



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**  
POSGRADO EN CIENCIAS DEL MAR Y LIMNOLOGÍA  
(BIOLOGÍA MARINA)

“EVOLUCIÓN DEL EXÓN 2 DEL GEN *DQB* EN EL  
COMPLEJO MAYOR DE HISTOCOMPATIBILIDAD  
DE MAMÍFEROS MARINOS Y TERRESTRES”.

**T E S I S**

QUE PARA OPTAR POR EL GRADO DE:  
DOCTORA EN CIENCIAS DEL MAR Y LIMNOLOGÍA

**PRESENTA:**

**MARÍA JOSÉ VILLANUEVA NORIEGA**

**TUTOR:**

**Dr. LUIS MEDRANO GONZÁLEZ**  
FACULTAD DE CIENCIAS, UNAM

**COMITÉ TUTOR:**

Dra. MARÍA LUISA FANJUL PEÑA  
FACULTAD DE CIENCIAS, UNAM

Dr. MANUEL URIBE ALCOCER  
INSTITUTO DE CIENCIAS DEL MAR Y LIMNOLOGÍA

Dr. EDGAR ZENTENO GALINDO  
FACULTAD DE MEDICINA, UNAM

Dr. FEDERICO ALBERTO ABREU GROBOIS  
INSTITUTO DE CIENCIAS DEL MAR Y LIMNOLOGÍA

MÉXICO, D. F. ENERO, 2013



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



**Posgrado en Ciencias del Mar y Limnología**

**Universidad Nacional Autónoma de México**



**EVOLUCIÓN DEL EXÓN 2 DEL GEN *DQB* EN EL  
COMPLEJO MAYOR DE HISTOCOMPATIBILIDAD  
DE MAMÍFEROS MARINOS Y TERRESTRES**

**T E S I S**

que para obtener el grado académico de  
Doctora en Ciencias del Mar y Limnología  
en el Área de Biología Marina

p r e s e n t a

**BIÓL. MARÍA JOSÉ VILLANUEVA NORIEGA**

**DIRECTOR DE TESIS**

**DR. LUIS MEDRANO GONZÁLEZ**

**COMITÉ TUTORAL Y JURADO**

**DRA. MARÍA LUISA FANJUL PEÑA**

**DR. MANUEL URIBE ALCOCER**

**DR. EDGAR ZENTENO GALINDO**

**DR. FEDERICO ALBERTO ABREU GROBOIS**

México D.F., Enero, 2013

*Para empezar un gran proyecto, hace falta valentía. Para terminar un gran proyecto, hace falta perseverancia (Anónimo).*

Esta tesis va dedicada a mis padres, Rosa y Jesús, quienes a través de los años me han enseñado a enfrentar la vida con una sonrisa, a jamás rendirme, a siempre luchar por mis sueños y a saber que la familia es la posesión más valiosa con la que contamos; a mi hermana Gabriela, quien es la voz en mi cabeza, mi más objetiva consejera, compañera de risas y mi mejor amiga; a mi hermano Jesús, quien me enseña todos los días a ser mejor persona con su gran corazón, con quien comparto anhelos, ideas, y siempre nuevos proyectos; a mi cuñado Sergio, quien más que cuñado es un hermano, compañero de vida, de experiencias y emociones; a mi sobrino Elías, quien llegó a este mundo a darnos alegría y fuerza para luchar por un mundo mejor y más justo; y finalmente a mi esposo Andrew, a quien admiro por su incansable dedicación a su profesión y su interminable amor por la naturaleza, porque todos los días aprendo cosas nuevas a su lado, porque es mi hogar y el motivo para despertar todas las mañanas.

Este logro no se podría celebrar sin el apoyo de todos ustedes. Desde el fondo de mi corazón, GRACIAS.

# Contenido

<b>Agradecimientos</b> .....	1
<b>Resumen</b> .....	2
<b>Abstract</b> .....	3
<b>Capítulo 1.</b> Biología del Complejo Mayor de Histocompatibilidad y de los procesos de enfermedad y mortalidad de los mamíferos	
1.1 <i>Introducción</i> .....	4
1.2 <i>El Complejo Mayor de Histocompatibilidad (MHC)</i> .....	5-8
1.3 <i>Sistemática breve de los mamíferos marinos</i> .....	8-10
1.4 <i>Los mamíferos en el ambiente marino</i> .....	10-12
1.5 <i>La mortalidad en los mamíferos</i> .....	13-18
1.6 <i>Mortalidad y MHC</i> .....	18-20
1.7 <i>Epidemiología en el ambiente marino</i> .....	20-21
1.8 <i>El MHC y los mamíferos marinos</i> .....	21-24
1.9 <i>Planteamiento de la investigación</i> .....	24-25
1.10 <i>Objetivos</i> .....	25
1.11 <i>Hipótesis</i> .....	26
<b>Capítulo 2.</b> Evolution of the MHC-DQB exon 2 in marine and terrestrial mammals	
<i>Publicación en la revista Immunogenetics</i> .....	27-41
<b>Capítulo 3.</b> Variación nucleotídica y peptídica del exón 2 del gen <i>DQB</i> en mamíferos marinos y terrestres	
3.1 <i>Introducción</i> .....	42-44
3.2 <i>Métodos</i> .....	44-47
3.3 <i>Resultados</i> .....	47-54
3.4 <i>Discusión</i> .....	55
<b>Capítulo 4.</b> Conclusiones generales y recomendaciones para futuras investigaciones	
4.1 <i>Conclusiones generales y recomendaciones para futuras investigaciones</i> .....	56-58
<b>Literatura citada</b> .....	59-66

## Agradecimientos

Agradecemos a todos quienes, en el laboratorio o en el mar, contribuyeron al desarrollo de este trabajo. En particular, agradecemos el apoyo técnico de M. Dalebout, J. Murrell, M.R. Robles, y D. Steel, así como la revisión y el consejo de M.L. Fanjul, D. Heimeier, y M. Uribe, E. Zenteno, F.A. Abreu Grobois y la orientación académica de G. Vilaclara. Apreciamos la asesoría de J. Zúñiga en las pruebas estadísticas, así como los comentarios de dos revisores anónimos de la revista *Immunogenetics* quienes significativamente mejoraron este trabajo. Agradecemos el apoyo institucional, académico, legal y financiero del Posgrado en Ciencias del Mar y Limnología-Universidad Nacional Autónoma de México, Consejo Nacional de Ciencia y Tecnología, Facultad de Ciencias-Universidad Nacional Autónoma de México, The Marsden Foundation, University of Auckland, Secretaría del Medio Ambiente y Recursos Naturales, Instituto Nacional de Ecología, y la Convención sobre el Comercio Internacional de Especies Amenazadas de Fauna y Flora Silvestres (CITES por sus siglas en inglés).



## Resumen

Basados en la existencia de un polimorfismo en general modesto, varios estudios sugieren que la selección balanceadora en el complejo mayor de histocompatibilidad (MHC) clase II, es más débil en los mamíferos marinos que en los terrestres. En este trabajo investigamos esta selección diferencial del exón 2 del gen *DQB* del MHC entre los taxones Cetacea, Artiodactyla y Primates contrastando indicadores de evolución molecular como presencia de polimorfismos transespecíficos, patrones de longitud de ramas filogenéticas por posición de codón, tasas de sustituciones no sinónimas y sinónimas así como la acumulación de sitios variables en las muestras de alelos analizadas. Estos indicadores se compararon entre el *DQB* y el gen mitocondrial del citocromo b como una referencia de expectativas de neutralidad y de diferencias entre relojes moleculares debidas a la historia de vida y la demografía histórica. Así mismo, se analizaron los perfiles de hidropatía y la estructura secundaria resultante de cada secuencia de *DQB* en el estudio. Todos los indicadores de evolución molecular mostraron que la selección balanceadora sobre el *DQB* es más variable y en general más débil en los cetáceos. En nuestro muestreo, los zífidos, el cachalote, los monodóntidos y la marsopa sin aleta formaron un grupo con bajo polimorfismo del *DQB* mientras que los mysticetos mostraron una alta variación del *DQB* similar a la de mamíferos terrestres, así como una alta presencia de polimorfismos transespecíficos. Distintos delfines aparecieron en ambos grupos. Una mayor variación en la selección sobre el *DQB* de los cetáceos puede relacionarse a una mayor estocasticidad en su demografía histórica y por lo tanto, a una mayor complejidad de la ecología general y los procesos de enfermedad de estos animales. En cuanto al análisis de la estructura secundaria del *DQB*, aunque no se observa un patrón distintivo entre los perfiles de hidropatía entre mamíferos marinos y terrestres, existe una mayor variación en los marinos. La conservación de la estructura secundaria del *DQB* observada en todos los mamíferos analizados sugiere que, a pesar de que algunas especies presenten menores niveles de polimorfismo, la funcionalidad en el PBR parece mantenerse.

## Abstract

On the basis of a general low polymorphism, several studies suggest that balancing selection in the class II major histocompatibility complex (MHC) is weaker in marine mammals as compared with terrestrial mammals. We investigated such differential selection among Cetacea, Artiodactyla, and Primates at exon 2 of MHC-*DQB* gene by contrasting indicators of molecular evolution such as occurrence of transpecific polymorphisms, patterns of phylogenetic branch lengths by codon position, rates of non-synonymous and synonymous substitutions as well as accumulation of variable sites on the sampling of alleles. These indicators were compared between the *DQB* and the mitochondrial cytochrome b gene (*cytb*) as a reference of neutral expectations and differences between molecular clocks resulting from life history and historical demography. Also, *DQB* hydropathy profiles and secondary structure were analyzed for all sequences in the study. All indicators of molecular evolution showed that the influence of balancing selection on the *DQB* is more variable and overall weaker for cetaceans. In our sampling, ziphiids, the sperm whale, monodontids and the finless porpoise formed a group with lower *DQB* polymorphism, while mysticetes exhibited a higher *DQB* variation similar to that of terrestrial mammals as well as higher occurrence of transpecific polymorphisms. Different dolphins appeared in the two groups. Larger variation of selection on the cetacean *DQB* could be related to greater stochasticity in their historical demography and thus, to a greater complexity of the general ecology and disease processes of these animals. Regarding the secondary structure analysis of the *DQB*, although no distinctive patterns were observed between marine and terrestrial mammals' hydropathy profiles, marine mammals showed more variation. The conservation of the *DQB* secondary structure in all mammals analyzed suggests that; in spite some species have lower levels of polymorphism, PBR functionality seems to be maintained.



# Capítulo 1

Biología del Complejo Mayor de  
Histocompatibilidad y de los procesos de  
enfermedad y mortalidad de los mamíferos



## 1.1 Introducción

La constante evolución del sistema inmune ha permitido el desarrollo de defensas adecuadas contra agentes invasores, que se encuentran cambiando continuamente. La respuesta inmune adaptativa involucra a varias células, dentro de las cuales destacan las células presentadoras de antígeno (CPA) y las células T, que controlan las acciones contra un amplio espectro de dichos agentes. Las CPA presentan moléculas del Complejo Mayor de Histocompatibilidad (MHC, por sus siglas en inglés) de manera que las células T las reconozcan como antígeno y puedan llevar a cabo la respuesta inmune (Čemerski y Shaw, 2006). Los genes clase I y II del MHC presentan los loci codificantes conocidos más polimórficos que se han reportado en el grupo de los vertebrados, en algunos casos con más de 200 alelos reportados para algunas especies (Potts y Wakeland, 1990).

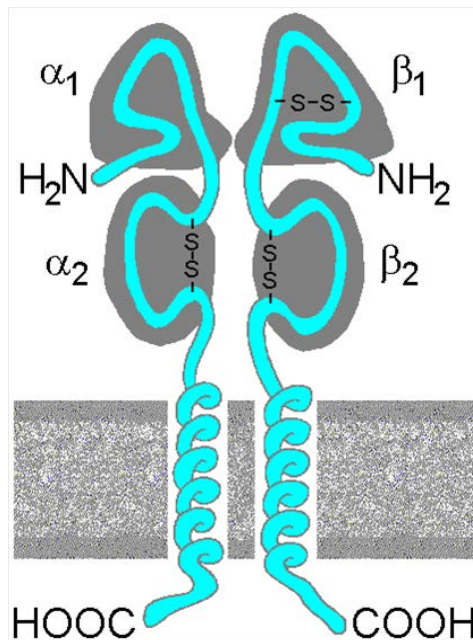
Debido a que las moléculas del MHC actúan como estructuras presentadoras de antígenos, el repertorio particular de moléculas de MHC de un individuo determina el repertorio de antígenos al cual los linfocitos T pueden responder (Goldsby et al. 2003). Por esta razón, se dice que el MHC determina en parte la respuesta inmune de un individuo hacia antígenos de organismos infecciosos, y por tanto, se ha asociado a la susceptibilidad a enfermedades y en el desarrollo autoinmune (Potts y Wakeland, 1990; Edwards y Hedrick, 1998; Goldsby et al. 2003). El MHC ha sido estudiado principalmente en humanos (*Homo sapiens*; complejo – HLA) y en ratones (*Mus musculus*; complejo H – 2) (Goldsby et al. 2003), sin embargo, el creciente interés por este complejo a raíz de las implicaciones que este puede tener sobre la viabilidad de algunas poblaciones silvestres (O'Brien et al. 1985), ha estimulado un aumento en el número de trabajos sobre especies silvestres (Bernatchez y Landry, 2003).

## 1.2 El Complejo Mayor de Histocompatibilidad (MHC)

Las moléculas del MHC son miembros de la superfamilia de las inmunoglobulinas. Los componentes principales extracelulares de estas moléculas son dos: un “pedúnculo” (dominio de la inmunoglobulina) que ancla la molécula a la membrana de la célula, y un receptor, con forma de canastilla, llamado la región de unión del péptido (PBR – peptide binding region) (Edwards y Hedrick, 1998). Las moléculas de clase I consisten en una larga cadena de una glicoproteína con 3 dominios extracelulares, un segmento transmembranal, y una  $\beta 2$  – microglobulina, la cual es una proteína con un solo dominio. Las moléculas de clase II están compuestas por dos glicoproteínas asociadas no – covalentemente, las cadenas  $\alpha$  y  $\beta$ , codificadas por diferentes genes (Figura 1) (Goldsby et al. 2003). La regulación transcripcional de las moléculas clase I y II se encuentra relacionada, directa o indirectamente, con la funciones defensivas del sistema inmune (Roitt y Delves, 2001).

Las moléculas de las clases I y II presentan antígenos a las células T. Las moléculas de la clase I presentan antígenos procesados endógenamente a células T CD8 y son expresadas por la mayoría de las células somáticas. Las moléculas de la clase II presentan antígenos procesados exógenamente a células T CD4 y son presentadas únicamente por células B, macrófagos, y células dendríticas (Goldsby et al. 2003).

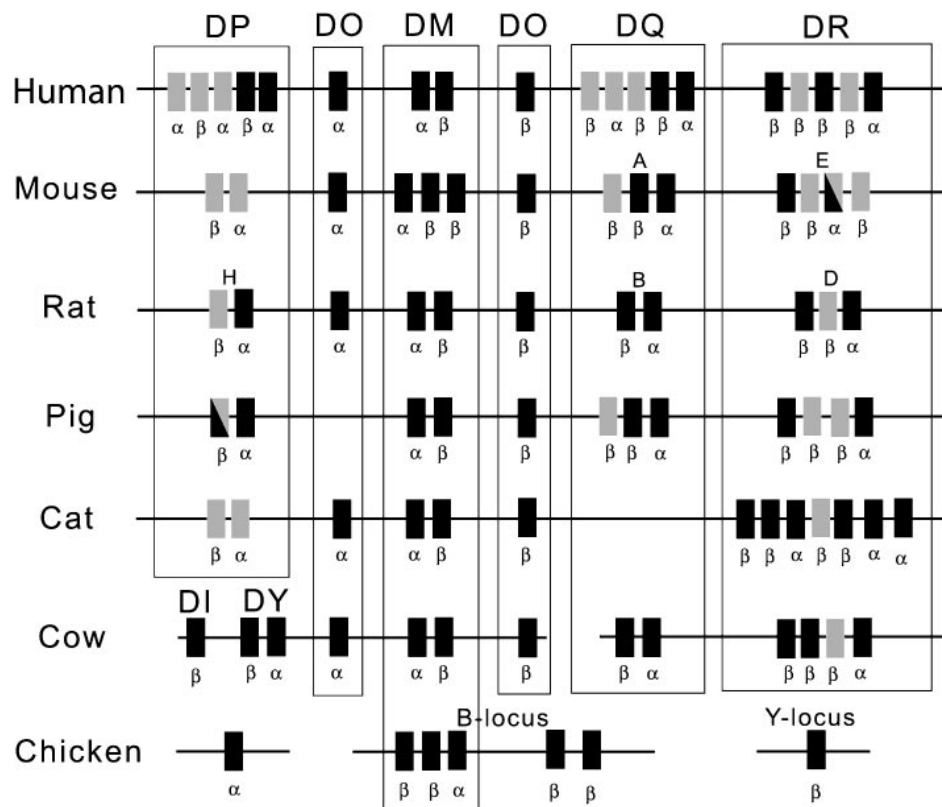
El MHC de clase I y II se encuentra separado en diferentes regiones o loci (Figura 2), la nomenclatura que se utiliza para definir los loci del MHC en especies silvestres es generalmente la misma del complejo HLA, donde las moléculas de clase I se dividen en *B*, *C*, y *A*; mientras que las moléculas de clase II se dividen en *DP*, *DQ* y *DR*, donde cada región se divide a su vez en cadenas  $\alpha$  y  $\beta$  (Roitt y Delves, 2001). Las proteínas de cada región de las clases I y II, son codificadas por exones separados. El juego de alelos de cada loci forma un haplotipo específico, parte del cual es transmitido por cada uno de los padres. En poblaciones con apareamiento aleatorio, la descendencia puede ser heteróciga en muchos loci con la expresión de los alelos de ambos padres (codominancia) (Goldsby et al. 2003).



**Fig. 1** Estructura de la molécula MHC clase II. La molécula de MHC clase II está compuesta por dos cadenas polipeptídicas diferentes (alfa y beta) las cuales están asociadas de manera no covalente. Los extremos alfa 1 y beta 1 del extremo amino terminal de las moléculas de clase II interaccionan y forman la hendidura de unión al péptido, estructuralmente similar a la de clase I. En las moléculas de clase II, los extremos de la hendidura de unión al péptido están abiertos, de manera que pueden unirse péptidos de 10 a 20 residuos.

Los sitios PBR son altamente polimórficos (Roitt y Delves, 2001) y se han identificado cientos de variantes alélicas para las clases I y II en los humanos, sin embargo se debe apuntar que un individuo sólo expresa una pequeña parte de dicha variabilidad, hasta 6 alelos de la clase I y 12 alelos de la clase II. La unión entre el péptido y la molécula de MHC se considera promiscua, ya que una molécula de MHC se puede unir a numerosos péptidos y algunos péptidos se pueden unir a varias moléculas de MHC. Los ligandos peptídicos son sostenidos por una larga y extendida cadena que corre a lo largo de un surco, las moléculas de clase I presentan un surco cerrado en ambos extremos, por lo que reconocen ligandos de 8 a 10 aminoácidos; mientras las moléculas de clase II presentan un surco abierto que les permite reconocer ligandos de 13 a 18 aminoácidos. La variación a lo largo de la cadena polipeptídica del MHC no se encuentra distribuida aleatoriamente, sino que se encuentra agrupada en algunas zonas de poca longitud, por lo general en los

dominios  $\alpha 1$  y  $\alpha 2$  de la clase I; y en los dominios  $\alpha 1$  y  $\beta 2$  de la clase II, estas zonas coinciden con los PBR, por lo tanto el exón dos del gen *DQB* de las moléculas de clase II del MHC es considerado como un PBR. La localización de tantos amino ácidos polimórficos dentro de los sitios de unión a los antígenos procesados sugiere fuertemente que la diferencias alélicas contribuyen a las diferencias observadas en la capacidad de las moléculas de MHC a interactuar con un péptido antigénico específico (Goldsby et al. 2003). Los loci de la cadena beta de los alelos de la clase II del MHC, son extremadamente polimórficos, lo que les confiere una característica única a este conjunto de moléculas (Lundberg y McDevitt, 1992).



**Fig. 2** Genes ortólogos clase II en mamíferos. En general, los genes clase II DP, DO, DM, DQ, y DR se conservan entre los mamíferos, pero el número de genes alfa y beta en estos loci es altamente variable debido a duplicaciones locales intraespecíficas. Los genes son representados por cajas negras; los pseudogenes están en gris. Las letras que se muestran arriba de algunos grupos indican la nomenclatura específica para ciertas especies, por ejemplo, A representa al H2-A y E representa al H2-E. Tomado de Kumánovics et al. (2003).

Por lo hasta lo que ahora se conoce en los mamíferos placentados, los genes del MHC clase II DP, DN, DM, DO, DQ, y DR codifican glicoproteínas de unión a péptidos. Los genes DP, DM, DQ y DR usualmente codifican para una cadena alfa y una cadena beta. En algunos artiodáctilos, los genes DP se sustituyen por los genes DI y DY (Figura 2). El arreglo general de los genes del MHC clase II se conserva entre los grupos de mamíferos aunque en algunas especies algunos genes se han perdido y el número de genes alfa y beta es variable (Beck y Trowsdale 1999; Takahashi et al. 2000; Goldsby et al. 2003; Kumánovics et al. 2003).

### **1.3 Sistemática breve de los mamíferos marinos**

Los mamíferos marinos son un grupo polifilético (Figura 3) en donde cuatro clados de mamíferos placentados desarrollaron adaptaciones para la vida en los océanos de manera independiente. Los mamíferos marinos incluyen a los pinnípedos (lobos, focas y morsas), cetáceos (ballenas y delfines), sirenios (manatíes y dugongos) y los extintos desmostilios (Rice, 2009). Actualmente se reconocen alrededor de 126 especies de mamíferos marinos: 86 cetáceos, 36 pinnípedos, 4 sirenios. Algunos autores incluyen en este grupo, dos especies de nutrias, el oso polar e incluso el murciélagos pescador (Jefferson et al. 2008) debido a que son organismos que su vida y alimento dependen enteramente del medio marino. A continuación se presenta una breve descripción de cada uno de los grupos de mamíferos marinos.

#### **Cetáceos**

El orden Cetacea tradicionalmente incluye tres subórdenes: el extinto Archaeoceti y los recientes Odontoceti (cetáceos con dientes) y Mysticeti (cetáceos con barbas). Se cree comúnmente que la evolución de los cetáceos ancestrales a la vida terrestre a la acuática ocurrió a partir de rápidas y drásticas adaptaciones. Las marcadas disimilitudes entre odontocetos y misticetos hicieron difícil para los primeros cetólogos establecer el orden Cetacea como un grupo monofilético (Árnason et al. 2004), Sin embargo, estudios morfológicos y moleculares han permitido constatar el origen monofilético de los cetáceos,

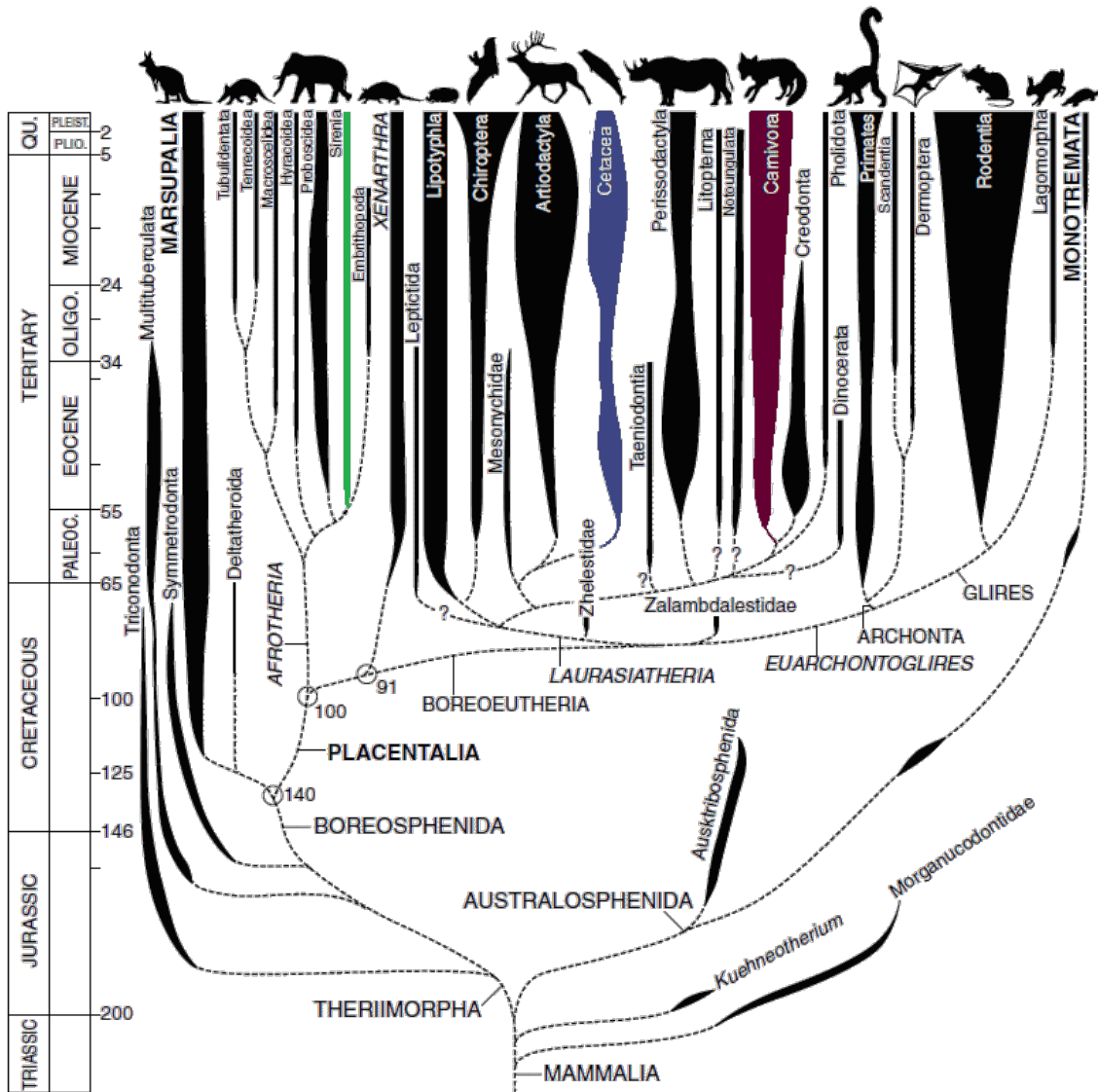
así como revelar que los artiodáctilos (mamíferos de dedos- pares con pezuñas) son sus parientes vivos más cercanos (Irwin y Arnanson, 1994; Árnason y Gullberg, 1994). Algunos autores colocan a los cetáceos dentro del mismo orden Artiodactyla como grupo hermano de Hippopotamidae bajo un supraorden denominado Cetartiodactyla (Rice, 2009). Esta relación ha sido a su vez apoyada por otros estudios con genes nucleares (Gatesy, 1997) y estadísticamente por medio del análisis de genomas mitocondriales completos (Ursing y Árnason, 1998). Análisis con estos genomas mitocondriales, sugieren que los cetáceos actuales se originaron en un radiación evolutiva hace 30 – 35 millones de años atrás (Árnason et al. 2004), contando actualmente con 13 familias que incluyen: ballenas francas, rorcuales, cachalotes, zifios, delfines (marinos y de río), marsopas, entre otras (Rice, 2009).

### **Pinnípedos**

En el pasado, los pinnípedos se consideraban como un orden distinto a los carnívoros terrestres (Carnivora), denominado Pinnipedia. Recientes análisis moleculares y morfológicos indican que los pinnípedos son un grupo monofilético pertenecientes al suborden Caniformia del orden Carnivora (Rice, 2009; Berta et al. 2006). La mayoría de los autores considera como los parientes más cercanos a este grupo a los carnívoros arctoides, con la mayoría de la evidencia apoyando la unión de los pinnípedos y de los úrsidos (osos y sus parientes extintos), aunque algunos estudios revelan que este grupo podría estar emparentado con la familia Mustelidae (Berta et al, 2006). Este grupo consta de 36 especies clasificadas dentro de tres familias: Otarioidea (lobos marinos y lobos finos), Odobenidae (morsas) y Phocoidea (focas) (Rice, 2009).

### **Sirenios**

Durante muchos años, debido a su similar apariencia a los cetáceos, los sirenios fueron considerados “cetáceos herbívoros”. Sin embargo, estudios modernos han revelado que los sirenios, junto con los extintos Desmostyilia, son miembros marinos del superorden Tethytheria, el cual también abarca a los elefantes (orden Proboscidea) y algunos otros grupos extintos (Berta et al. 2006). La monofilia de los sirenios es ampliamente aceptada y el grupo de los elefantes se consideran sus parientes vivos más cercanos (Rice, 2009).



**Fig. 3** Filogenia de mamíferos, mostrando la posición de los tres órdenes de mamíferos marinos. Tomado de (Benton, 2005)

## 1.4 Los mamíferos en el ambiente marino

### Distribución

Los mamíferos marinos han colonizado una vasta variedad de ambientes acuáticos que van desde ríos y áreas costeras hasta las aguas abisales y del ecuador hasta los polos. Con excepción del oso polar (*Ursus maritimus*), el manatí del Amazonas (*Trichechus manatus*)



*manatus*) y algunos misticetos, que pueden pasar largos periodos sin consumir alimento, los mamíferos marinos necesitan comer a lo largo de todo el año. Por lo tanto, su distribución se encuentra cercanamente relacionada con la de su comida. Algunas especies tienen requerimientos muy específicos de su hábitat, por ejemplo, en el caso de los pinnípedos y nutrias que requieren de tierra firme o hielo para reproducirse.

Existen especies de mamíferos marinos que presentan distribución *cosmopolita*, esto significa que se encuentran en todo el mundo como la ballena jorobada (*Megaptera novaengliae*) y la orca (*Orcinus orca*). Algunas especies aunque no se encuentren en todos los océanos, se pueden considerar de distribución amplia, ya que se encuentran en alguna cuenca oceánica, como es el caso de algunas especies del género *Stenella*. Algunas otras presentan distribución *circumpolar*, que se refiere a la distribución a altas latitudes en alguno de los dos polos como el narval (*Monodon monoceros*), la beluga (*Delphinapterus leucas*), la ballena franca del hemisferio sur (*Eubalaena australis*), la foca de Weddell (*Leptonychotes weddelli*) y la foca elefante del hemisferio sur (*Mirounga leonina*). Algunas especies tienen distribución circumpolar estacionalmente como la orca, la ballena jorobada, el rorcual de aleta (*Balaenoptera physalus*), y cachalotes machos (*Physeter macrocephalus*).

Otras especies como la ballena de Bryde (*Balaenoptera edeni*), se encuentran en aguas tropicales a ambos lados del ecuador, distribución denominada *pantropical*. Así mismo todas las especies del orden Sirenia (manatíes y dugongos) y las focas monje del género *Monachus* presentan distribuciones tropicales restringidas a localidades costeras.

La distribución costera suele incluir a las áreas de aguas adyacentes a la costa sobre la plataforma continental. Algunas de las especies que presentan esta distribución pueden llegar a ser tan especializadas que incluso se les puede llegar a considerar de distribución fluvial, como es el caso de los delfines de río y los manatíes. Algunas especies con distribución costera que presentan especialización “anfibia” ya que pasan alguna parte de su vida en tierra, como es el caso de los pinnípedos, nutrias y osos polares.

Las especies que viven en mar abierto, ya sea por una temporada o por toda su vida, son denominados como *pelágicos*. Los cachalotes y muchas especies de zífios como el de Cuvier (*Ziphius cavirostris*), presentan este tipo de distribución, sin embargo, también muchas especies de pinnípedos como las focas elefante del género *Mirounga* pasan una buena parte de su vida en mar abierto. Las especies *pagofílicas* (ama – hielo) que comprenden muchas especies de pinnípedos como la foca de Weddell, así como algunos cetáceos como la beluga y el narval, se distribuyen exclusivamente en regiones de hielos perpetuos (Reeves et al. 2002).

La alta diversidad de ambientes que ocupan los mamíferos marinos, aunado a la mezcla e intercambio de estos ambientes, ya sea de manera estacional o por estadio de vida (especies migratorias, anfibias, fluviales, etc.), forman una amplia gama de factores y presiones ambientales que pueden estar alterando la evolución de este grupo.

### **Alimentación**

Los mamíferos marinos tienen formas de alimentación y dietas muy diversas con variaciones considerables incluso dentro de una misma especie. Existen especies herbívoras como los manatíes, que se alimentan de pastos marinos que se encuentran cerca de la costa. Los misticetos al desarrollar barbas en vez de dientes, se han tornado hacia la alimentación de tipo planctófaga, consumiendo enormes cantidades de crustáceos llamados krill en zonas de alta productividad. El consumo de organismos que se encuentran en el necton es llevado a cabo principalmente por delfines, zífios, y pinnípedos que consumen peces, calamares y otros organismos. A los mamíferos marinos que se alimentan principalmente de moluscos, como las nutrias y morsas (*Odobenus rosmarus*) se les conoce como malacófagos.. Finalmente las especies sarcófagas son aquellas que consumen carne roja, de las cuales la orca es su representante más destacado, y el oso polar y algunas especies de fócidos también basan su ingesta en este tipo de alimento (Reeves et al. 2002).

## 1.5 La mortalidad en los mamíferos

La mortalidad, ya sea por patógenos o por otros agentes, juega un importante papel en los patrones poblacionales de los mamíferos marinos, y puede ser causada, de acuerdo con Geraci y Lounsbury (2009), por:

### **Fallas reproductivas**

Estas se refieren a la muerte de individuos recién nacidos por abortos y nacimientos prematuros, el riesgo que aumenta en madres primerizas y por los estados nutricional y de salud deficiente de las madres. Los pinnípedos, osos polares y nutrias son altamente susceptibles a este tipo de muerte por su dependencia de presas específicas y por la incidencia de patógenos como el *morbillivirus* que puede llegar a provocar abortos espontáneos y una alta tasa de muertes de crías. Otras causas de mortalidad en neonatos son defectos congénitos, abandono, depredación, lesiones, perturbación e inexperiencia.

### **Inanición**

Es una de las causas más comunes de mortalidad en mamíferos marinos jóvenes (Geraci et al. 1999). Factores como la edad, la reserva de grasa, la tasa metabólica, las demandas, energéticas o la salud en general, determinan cuánto puede sobrevivir un animal sin comida. Animales grandes, con bajas tasas metabólicas, como los misticetos (ballenas con barbas), sobreviven más tiempo que aquellos con altas demandas energéticas, como las nutrias marinas. Las crías recién nacidas dependen del estado nutricional de la madre, por lo que la disponibilidad de alimento influye en el adecuado crecimiento de las crías. La muerte por inanición se debe a hipotermia, agotamiento, enfermedades causadas por desnutrición, infección de parásitos o invasión de patógenos oportunistas.

### **Efectos ambientales directos**

La mortandad puede ser ocasionada por condiciones climáticas extremas, como largos periodos de bajas temperaturas, que han ocasionado la muerte de muchos manatíes (*T. manatus latirostris*); por tormentas, que llegan a arrojar a los pinnípedos contra la costa; o

bien por cambios en la cobertura de hielo, que pueden provocar el aprisionamiento de algunos cetáceos.

### **Trauma**

Las lesiones que tormentas, depredadores, encuentros agresivos o actividades humanas ocasionan pueden conllevar a la muerte de algunos organismos. En algunas zonas de reproducción de pinnípedos, muchas crías mueren a causa de aplastamientos y lesiones por el tránsito de individuos adultos. Una de las mayores causas de mortalidad en mamíferos marinos, en general, fue y es la caza ballenera la cual ha reducido considerablemente los tamaños poblacionales de la mayoría de las especies. Las lesiones a los manatíes, producto de colisión con embarcaciones causan la muerte de muchos individuos al año. El enmallamiento de pinnípedos y cetáceos costeros en redes pesqueras son una causa importante de mortalidad.

### **Depredación**

Otra causa de mortalidad en mamíferos marinos es la depredación de pinnípedos y crías de cetáceos por tiburones, otros cetáceos (orcas. *O. orca*) u otros pinnípedos (focas leopardo. *Hydrurga leptonyx*).

### **Parásitos**

Después del destete, prácticamente todos los mamíferos marinos presentan algún tipo de parásito. Existen parásitos dañinos que pueden llegar a causar la muerte de pinnípedos como algunos gusanos (*Uncinaria lucasi*); y algunos nemátodos del género *Crasicauda* y tremátodos del género *Nasitrema* pueden afectar a algunos cetáceos. Estos efectos dañinos se pueden acentuar por efectos de estrés, desnutrición y lesión.

### **Desórdenes metabólicos**

Las alteraciones en la secreción de la aldosterona por la corteza adrenal provocadas por estrés, desnutrición o enfermedades pueden provocar la muerte por pérdida de apetito y desorientación.

## **Tumores**

Pueden ser provocados por desbalances hormonales, virus, defectos hereditarios o agentes físicos y químicos. Las belugas (*D. leucas*) del Golfo de San Lorenzo, presentan una alta incidencia de tumores debido a la gran cantidad de desechos de tóxicos arrojados al mar.

## **Varamientos**

Se definen como la aparición en tierra de animales acuáticos enfermos, débiles, o simplemente perdidos. La mayoría de estos animales muere rápidamente. Algunos varamientos individuales son producto de alguna enfermedad o ingestión de basura marina. El grupo de los cetáceos es el grupo más susceptible a este tipo de mortalidad. Los varamientos masivos se ocurren cuando dos o más animales aparecen en tierra al mismo tiempo y en el mismo lugar. Estos eventos ocurren con mucha frecuencia en especies de odontocetos con estructuras sociales complejas como los calderones (*Globicephala spp*), los cachalotes (*P. macrocephalus*), pseudorcas (*Pseudorca crassidens*) y delfines de costados blancos del Atlántico (*Sagmatias acutus*). Los animales generalmente mueren por sofocamiento, insolación, fatiga respiratoria, falta de circulación sanguínea o colapso vascular.

## **Biotoxinas**

La ingestión de peces e invertebrados con altas concentraciones de toxinas acumuladas por florecimientos de dinoflagelados y algas, como la marea roja (*Gymnodinium brevis*), ha sido la causa de muertes masivas de toninas y manatíes en el Golfo de México.

## **Alteración del hábitat.**

El impacto que tienen las actividades antropogénicas sobre el ambiente marino se debe de tomar en cuenta cuando se evalúan los riesgos a nivel inmunológico, como es el caso de la nutria marina (*E. lutris nereis*) que comienza a presentar cada vez más casos de encefalitis a causa de infecciones por *Toxoplasma gondii* a causa de la creciente población de gatos portadores de la enfermedad (Miller et al. 2002), o del aumento en la cantidad de desechos arrojados al océano que contengan químicos inmunotóxicos que facilitan la aparición de enfermedades en niveles tróficos altos, como es el caso de los mamíferos marinos (Ross et

al. 1996; Ross, 2002). Los derrames petroleros han causado la muerte de muchos mamíferos marinos ya sea por ingesta del crudo o por destrucción, aunado a la pérdida de la capacidad de algunas especies de pinnípedos y nutrias para regular a través de su pelaje.

### **Patógenos**

*Bacterias.* Los cachorros de varias especies de pinnípedos son altamente susceptibles a infecciones de *Streptococcus spp* y *Corynebacterium spp.*, durante la época de muda. La reciente invasión de cepas de *Mycobacterium tuberculosis* en poblaciones silvestres del lobo marino de Nueva Zelanda, a partir de poblaciones de lobos en cautiverio, ha tenido efectos severos sobre los tamaños poblacionales.

La mayoría de los patógenos reportados para las orcas son de baja virulencia o presentan un bajo riesgo de generar una epizootia, mientras que para el resto de los odontocetos, el número es un poco mayor (Gaydos et al. 2004). Por otra parte se ha informado que en los mamíferos marinos existen dos cepas de bacterias Gram–negativas de origen terrestre del género *Brucella* que los infectan, una se identificó como *B. cetaceae* que infecta cetáceos y la otra se identificó como *B. pinnipediae* que infecta a pinnípedos. Esta bacteria altera el nivel reproductivo por medio de la esterilización de sus portadores, eliminando individuos reproductivos de las poblaciones naturales (Cloeckert et al. 2003).

*Hongos.* Son los patógenos que tienen el menor efecto sobre los mamíferos marinos, algunas cepas que causan infección pertenecen a los géneros *Candida*, *Aspergillus*, *Coccidioides*, *Blastomyces*, *Histoplasma*, *Fusarium*, *Nocardia* y *Loboa spp.* Esta última provoca infecciones lobomicóticas en toninas (*Tursiops truncatus*) y tucuxis (*Sotalia fluviatilis*) asociadas a la gente que habita en zonas costeras de América central y del sur, pero aun no se conoce su medio de transmisión.

### **Muertes masivas causadas por epizootias.**

Existe un buen número de informes de muertes masivas de mamíferos marinos a causa de algún brote epizootico como el caso de la *Phoca vitulina* en Europa (Harwood y Hall, 1990), y de *T. manatus* (Bossart et al. 1998). Otros casos de epizootias provocadas por el morbilivirus canino, CVD son: la muerte de 5,000 – 10,000 Focas del Lago Baikal (*Pusa sibirica*) (Grachev et al. 1989) en 1987 – 1988; 10,000 Focas del Mar Caspio (*Pusa caspica*) en 2000 (Kennedy et al. 2000); y probablemente de cerca de 2,500 Focas cangrejas (*Lobodon carcinophagus*) en la Antártida en 1955 (Laws y Taylor, 1957; Bengston et al. 1991).

Se han reportado eventos de mortalidad masiva en cetáceos provocada por un virus análogo al CDV, específico para especies de delfines llamado el morbilivirus de delfines (*DMV = Dolphin Morbillivirus*) el cual ha desencadenado la muerte de varios miles de delfines listados (*Stenella coeruleoalba*) en el Mar Mediterráneo de 1990 – 1992 (Domingo et al. 1990) que posteriormente se expandió e infectó a una población Delfines comunes (*Delphinus delphis*) del Mar Negro en 1994 (Birkun et al. 1999). Este virus ha sido encontrado también en tejidos de Toninas (*T. truncatus*) provenientes del varamiento de alrededor de 750 animales a lo largo de la costa este del los Estados Unidos entre 1987 – 1988 (Lipscomb et al. 1994).

Un cuarto tipo es el morbilivirus de las marsopas (*PMV = Porpoise Morbillivirus*) el cual ha sido encontrado en Toninas muertas del Golfo de México en 1993 y en un varamiento de Marsopa común (*Phocoena phocoena*) en el Mar Irlandés en 1994 (Kennedy et al. 1988; Taubenberger et al. 1996). Los virus DMV y PMV, son antigénicamente similares y representan diferentes cadenas del morbillivirus cetáceo (CMV) el cual se ha reportado en varamientos infectando algunos individuos en varias especies de cetáceos, sin embargo, sus efectos en términos de la mortalidad de las poblaciones naturales no han sido bien comprendidos (Reynolds III et al. 2005).

Las bacterias también pueden tener efectos dañinos en las poblaciones naturales de mamíferos marinos, la bacteria *Leptospira pomona* fue detectada desde 1970 en

poblaciones del Lobo Marino de California (*Zalophus californianus*) y desde entonces se han reportado eventos de mortalidades masivas periódicas a lo largo de la costa norte de California cada 4 años (Vedros et al. 1971; Gulland et al. 1996). A pesar de esto, la población de lobos marinos de California ha mantenido un aumento de cerca del 10% anual en su tamaño poblacional (Forney et al. 2000). Otro caso es una bacteria gram – negativa pleiomórfica que infecta a cachorros del Lobo marino de Nueva Zelanda (*Phocartos hookeri*) y que mató a casi el 60% de la camada de 1998 reduciendo a su vez el 5% de la camada del año siguiente (Reynolds III et al. 2005).

## 1.6 Mortalidad y MHC

La dependencia de la respuesta inmune de las moléculas de la clase II del MHC refleja el papel central que estas moléculas juegan al presentar antígenos a las células T. Los múltiples factores genéticos y ambientales tienen varios papeles en el desarrollo de una enfermedad donde el MHC juega un papel importante más no exclusivo, donde la combinación de varios genes son los que generan la respuesta inmune. Las diferencias alélicas pueden tener efectos sobre la susceptibilidad a enfermedades de dos formas: por la capacidad de la molécula de MHC de presentar un antígeno; o por la capacidad de las células T de reconocer a las APC (Goldsby et al. 2003). El genotipo del MHC puede tener una influencia baja en la susceptibilidad hacia enfermedades específicas de individuos, pero la suma de varias influencias modestas puede llevar hacia la selección de una clase genotípica sobre otra. Diferencias aparentemente pequeñas en el vigor ocasionadas por enfermedades, pueden ser ampliadas a grandes diferencias en la adecuación generadas por competencia intraespecífica (Potts y Wakeland, 1990).

Las enfermedades asociadas a alelos del MHC particulares incluyen los desórdenes autoinmunes, algunas enfermedades virales, desórdenes en el sistema de complemento, algunos desórdenes neurológicos y algunas alergias. El fracaso general para encontrar asociaciones entre la presentación exógena de alelos del MHC y la susceptibilidad a enfermedades en poblaciones naturales, ha llevado a algunos autores a sugerir que los



alelos del MHC son selectivamente neutrales o bien mantenidos por mecanismos no basados en enfermedades. De acuerdo con Penn (2002), una población polimórfica para el MHC tendrá menos probabilidad de extinguirse que una monomórfica, ya que la presentación de un polimorfismo extremo en las moléculas del MHC asegura que, por lo menos, algunos miembros de la población o especie sean capaces de responder y resistir a nuevas enfermedades; aunque la existencia de algunas poblaciones monomórficas del MHC que parecen ser viables (Goldsby et al. 2003; Potts y Wakeland, 1990) parece contradecir dicho argumento. Sin embargo, si el polimorfismo del MHC se encuentra asociado a presiones selectivas dependientes de enfermedades, éste puede afectar directamente la adecuación individual a través de la supervivencia de su prole a enfermedades infecciosas (Bernatchez y Landry, 2003). Se ha propuesto que la antigenicidad de los parásitos, será seleccionada para explotar los defectos en la respuesta inmune mediada por el MHC de los genotipos más comunes (Potts y Wakeland, 1990), esta selección dependiente de la frecuencia produce un aumento en la selección hacia la adaptación del patógeno hacia los alelos más comunes de MHC, lo que provoca que estos alelos pierdan ventaja sobre los alelos raros, favoreciendo el aumento de la frecuencia de éstos en las poblaciones (Penn, 2002).

Existen algunos trabajos recientes que asocian polimorfismos del MHC con resistencia a enfermedades en vertebrados silvestres, como en algunas especies de peces (Langefors et al. 2001; Wegner et al. 2003) y mamíferos (Paterson et al. 1998; Traul et al. 2005). Sin embargo, en algunos de éstos la asociación no es completamente clara. Se ha reportado en poblaciones de borregos ferales (*Ovis aries L.*) que la variación alélica del MHC se encuentra significativamente asociada a diferencias en la supervivencia de juveniles y resistencia a infecciones de nemátodos intestinales (Paterson et al. 1998). En el caso particular de los mamíferos marinos, Bowen y colaboradores (2006), encontraron diferenciación genética entre los haplotipos (combinación de alelos) del MHC de dos poblaciones de lobo marino de California (*Z. californianus*), probablemente por efecto de presiones selectivas diferenciales de patógenos en el medio (Acevedo – Whitehouse et al. 2006). Otros estudios en peces (Arkuch et al. 2002), aves (Hedrick et al. 2001) y mamíferos

(Sauermann et al. 2001) muestran una mayor adecuación dependiente del MHC en individuos heterocigos que en homocigos.

El bajo polimorfismo del MHC de poblaciones de borregos cimarrones (*Ovis canadensis*) parece no estar relacionado con la alta susceptibilidad a enfermedades transmitidas por el ganado doméstico (Gutiérrez – Espeleta et al. 2001). La presencia de alelos individuales, más que la heterocigosis, en el haplotipo del MHC parece ser el factor crítico que determina la mortalidad en algunas especies de bóvidos (Paterson et al. 1998), lo que puede explicar por qué, a pesar de presentar un polimorfismo considerable en el MHC, las poblaciones aún sean altamente susceptibles a enfermedades (Gutiérrez – Espeleta et al. 2001).

Se han estudiado poblaciones con bajos niveles de polimorfismo en el MHC como los guepardos (O'Brien et al. 1985; Castro-Prieto et al. 2011) y castores europeos (Ellegren et al. 1993) que parecen presentar poblaciones viables (Potts y Wakeland, 1990), por lo que los niveles de polimorfismo en el MHC pueden no ser los únicos componentes actuando sobre la respuesta inmune de las poblaciones silvestres (Terio et al. 2005).

## **1.7 Epidemiología en el ambiente marino**

La mayoría de las teorías epidemiológicas han sido desarrolladas para ambientes terrestres, sin embargo se está comenzando a reconocer la importancia del conocimiento de las enfermedades que ocurren en los océanos. Aparentemente para algunas especies marinas, las enfermedades y su impacto han ido en aumento a través de los últimos 30 años, como las de tortugas, corales, moluscos y de algunos mamíferos marinos (McCallum et al. 2004).

McCallum y colaboradores (2004), atribuyen las principales diferencias para la caracterización del riesgo epidemiológico entre el ambiente marino y el terrestre como: (1) Una mayor diversidad taxonómica de huéspedes, así como parásitos, en ambientes marinos; (2) Diferencias en historias de vida entre organismos acuáticos y terrestres; (3) Diferencias en el modo de transmisión de parásitos en ambientes marinos y terrestres; y (4) Diferencias

en impactos humanos en ambientes marinos y terrestres; y 5) Diferencias en el potencial de controlar enfermedades infecciosas en ambientes marinos y terrestre.. También explican que los casos de fuerte virulencia en organismos marinos podrían ser ocasionados por una baja variabilidad genética, como en el caso de la mayoría de especies de mamíferos marinos que han sufrido cuellos de botella muy intensos.

Algunos autores han reportado que los patógenos principales que afectan a cetáceos odontocetos son menos en los individuos de vida libre que en organismos en cautiverio. La mayoría de los patógenos reportados que afectan a los cetáceos en vida silvestre son de baja virulencia o presentan un bajo riesgo de generar una epizootia, mientras que los reportados para especies cautivas son potencialmente más virulento y pueden generar epizootias (Gaydos et al. 2004). Por otra parte se tiene reportado que para los mamíferos marinos existen, al menos, dos cepas de bacterias Gram–negativas de origen terrestre del género *Brucella* que infectan mamíferos marinos, una se identificó como *B. cetaceae* que infecta cetáceos y la otra se identificó como *B. pinnipediae* que infecta a pinnípedos, bacteria que tiene repercusiones a nivel reproductivo esterilizando a sus portadores y eliminando con ello a individuos reproductivos de las poblaciones naturales (Cloeckaert et al. 2003).

## **1.8 El MHC y los mamíferos marinos**

El notable alto grado de polimorfismo genético que presenta el MHC (Edwards y Hedrick, 1998; Klein y Figueroa, 1986; Takahata, 1990; Goldsby et al. 2003) es una de las razones por las cuales este conjunto de genes representa una herramienta clave para la comprensión de los mecanismos que mantienen la variación genética en las poblaciones silvestres (Takahata et al, 1992). A diferencia del sistema de inmoglobulinas, donde la variabilidad se logra individualmente a través de un sistema multigénico, el MHC ha evolucionado en términos de variabilidad entre individuos a través de un sistema polimórfico basado en alelos múltiples, por ejemplo, genes alternativos en cada locus (Roitt y Delves, 2001). De acuerdo con Potts y Wakeland (1990), la diversidad en el MHC tiene cuatro propiedades; (1) es producto de alguna forma de selección balanceadora que opera en muchos loci, (2) la

selección opera en poblaciones contemporáneas y no por episodios (con largos periodos de neutralidad), (3) la selección opera directamente sobre los sitios de unión al antígeno (Hughes y Nei, 1988; Gu y Nei, 1999), y (4) la selección puede llegar a ser muy fuerte ( $s \geq 0.1$ ), por lo menos en especies como ratones. Gu y Nei (1999), también postulan que la fuente de polimorfismo del MHC es la selección balanceadora y la mutación.

En la selección balanceadora existe la sobredominancia (cuando existe una superioridad en adecuación para el heterocigoto), la selección dependiente de la frecuencia, o la diversificación de la selección que favorece diferentes genotipos en diferentes ambientes (Dobzhansky, 1971). La expresión codominante de los alelos del MHC da ventaja al heterocigoto al transcribir los alelos de ambos padres, lo que incrementa su capacidad de respuesta inmune con respecto a los homocigotos (Potts y Wakeland, 1990). El mantenimiento de la diversidad del MHC puede también ser producto de una selección dirigida por patógenos y por mecanismos de evasión de endogamia (Apanius et al. 1997). Se han reportado mecanismos fisiológicos – conductuales que favorecen los apareamientos entre individuos con diferentes combinaciones de MHC, a partir de la discriminación quimiosensible especialmente de tipo olfativa, observadas en ratones (Potts y Wakeland, 1990).

A pesar del alto grado de polimorfismo que muestra el MHC en la mayoría de los vertebrados terrestres (Edwards y Hedrick, 1998; Klein y Figueroa, 1986; Takahata, 1990; Potts y Wakeland, 1990; Goldsby et al. 2003), muchas de las especies de mamíferos marinos hasta hoy examinados presentan niveles de polimorfismo bajos, como algunos cetáceos (Trowsdale et al, 1989; Murray et al. 1995; Hayashi et al. 2003) y algunos pinnípedos (Slade, 1992). Estudios recientes de los loci *DQB* y *DRB* del MHC clase II, han reportado niveles de variación alélica considerables (Hoelzel et al. 1999; Baker et al. 2006; Xu et al. 2007). Se han reportado duplicaciones en algunos loci de cetáceos (Yang et al. 2005; Baker et al. 2006; Xu et al. 2007) aumentando en el número de alelos que un individuo puede expresar, por lo tanto, incrementando el espectro de antígenos reconocibles y mejorando la respuesta inmune (Goldsby et al. 2003). Para el gen *DQB*, se han reportado

por lo menos 23 alelos para las ballenas jorobadas (*M. novaengliae*) y 43 alelos para el recientemente extinto baiji (*Lipotes vexillifer*).

En un estudio realizado por Trowsdale y colaboradores (1989) se caracterizó la molécula de MHC de dos especies de mysticetos (*B. physalus* y *B. borealis*), mediante sondas de DNA humano, con excepción del loci DP y algunos genes de la clase I. El gen *DQB* en cetáceos es homólogo al gene HLA – *DQB1* del humano y al *DQB* de las vacas (*Bos taurus*) (Hayashi et al. 2003); así mismo, los loci DRB2 y DRB3 de los mysticetos son similares a los de las vacas (Baker et al. 2006).

Existen varias explicaciones para el bajo polimorfismo que presentan algunas especies de mamíferos marinos, como el efecto fundador, cuellos de botella evolutivos, baja presión inmunológica producto de vivir en un ambiente marino, o estilos de vida poco comunes (Klein y Figueroa, 1986; Trowsdale et al, 1989; y Slade, 1992). La resistencia a enfermedades podría ser unas de las fuerzas que se encuentre actuando sobre el polimorfismo en los mamíferos marinos, ya que han existido reducciones poblacionales considerables en algunas especies, como las focas, que ha presentado casos de muertes masivas a raíz de un brote epizoótico causado por el morbilivirus fócido, análogo del moquillo canino CVD (Harwood y Hall, 1990).

Slade (1992) argumenta que debido a que los mamíferos marinos se encuentran expuestos a relativamente menos patógenos en el medio marino, existe una presión de selección menor sobre la diversidad, a nivel poblacional, de los loci que codifican para respuestas inmunes; por esto las poblaciones son más susceptibles a mortalidades más altas en ocasiones cuando los patógenos tienen el suficiente potencial de transmisión para invadir y persistir. También argumenta que los cetáceos y los sirenios tendrán una menor diversidad en el MHC con respecto a los pinnípedos, ya que los últimos tienen un mayor contacto con el ambiente terrestre, ya sea para la reproducción y/o el descanso. En los lobos marinos de California (*Z. californianus*), se ha observado variación geográfica en los alelos presentados por dos poblaciones, se cree que esta variación se debe a diferencias en la intensidad de selección a nivel inmunológico en cada zona (Bowen et al. 2006).

Otras especies de animales que presentan niveles de polimorfismo bajos son el pingüino de las Galápagos (*Spheniscus mendiculus*) (Bollmer et al. 2007), el panda gigante (*Ailuropoda melanoleucas*) (Wan et al. 2006), el guepardo (*Acinonyx jubatus*) (O'Brien et al. 1985; Castro-Prieto et al. 2011) y el borrego cimarrón (*O. canadiensis*) (Gutiérrez-Espeleta et al. 2001), incluyendo algunas especies aparentemente monomórficas para todos sus loci como el castor europeo (*Castor fiber*) (Ellegren et al. 1993). La mayoría de estas especies tienen poblaciones fuertemente reducidas en sus tamaños poblacionales (IUCN, 2012), lo cual puede tener efecto sobre la variabilidad del MHC. El principal riesgo que genera un bajo polimorfismo en el MHC es la calidad de la respuesta inmune hacia patógenos en el medio (Goldsby et al. 2003) teniendo repercusiones importantes en la adecuación de los individuos a corto plazo (Amos y Balmford, 2001).

## **1.9 Planteamiento de la investigación**

Se ha observado que las poblaciones de mamíferos marinos están sujetas a eventos ocasionales e impredecibles de mortalidad masiva como el caso de epizootias (Harwood y Hall 1990), los cuales afectan planes y programas de conservación de estas especies en tanto que los modelos poblacionales normalmente utilizados en esta tarea no son aplicables a largo plazo o se modifican sensiblemente por una fluctuación intensa en la tasa de mortalidad. Más recientemente se ha encontrado que los mamíferos marinos tienen en general niveles de diversidad relativamente bajos del complejo mayor de histocompatibilidad (MHC) y se ha sugerido que esto deriva de un régimen de selección más relajado que la del medio terrestre, ante la menor densidad de patógenos específicos para mamíferos marinos en el mar (Slade, 1992). Sin embargo, no es claro si la aparente menor diversidad genética del MHC deriva de un régimen diferente de selección en el medio marino o se desprende de procesos de deriva génica asociados a procesos demográficos. Aún si la mortalidad asociada a patógenos fuese diferente en el medio marino con respecto al medio terrestre, el régimen de selección del MHC no sería necesariamente distinto, por lo tanto, la aparente menor diversidad genética en el MHC

derivaría del efecto de la mortalidad infringida por patógenos en el medio marino, apoyando el efecto de procesos demográficos más que de formas diferentes de selección.

## 1.10 Objetivos

### Objetivo general

Determinar el modo de evolución del MHC en los mamíferos marinos y terrestres distinguiendo los efectos de la demografía a través de la variación de la secuencia y estructura del exón 2 del gen *DQB* que codifica el surco de reconocimiento de antígenos en las moléculas DQ (clase II) las cuales participan en la respuesta inmunológica exógena.

### Objetivos específicos

- Comparar el número y nivel taxonómico de los eventos de polimorfismo transtaxón que se presentan en los mamíferos marinos y terrestres a través de un análisis filogenético.
- Determinar la fuerza de selección existente sobre la secuencia completa del gen *DQB* y en cada posición de codón en mamíferos marinos y terrestres a través de la existencia de polimorfismos transespecíficos y distancia genética, determinada mediante un análisis filogenético.
- Comparar el modo de selección del gen *DQB* a partir de la proporción de sustituciones sinónimas y no sinónimas entre mamíferos marinos y terrestres.
- Describir diferencias en el número de sitios variables del gen *DQB* y del gen mitocondrial citocromo b (*cytb*).
- Reconocer las diferencias estructurales y en la reactividad inmunológica que presenta el péptido resultante de la transcripción del exón 2 del gen *DQB* entre mamíferos marinos y terrestres.

## 1.11 Hipótesis

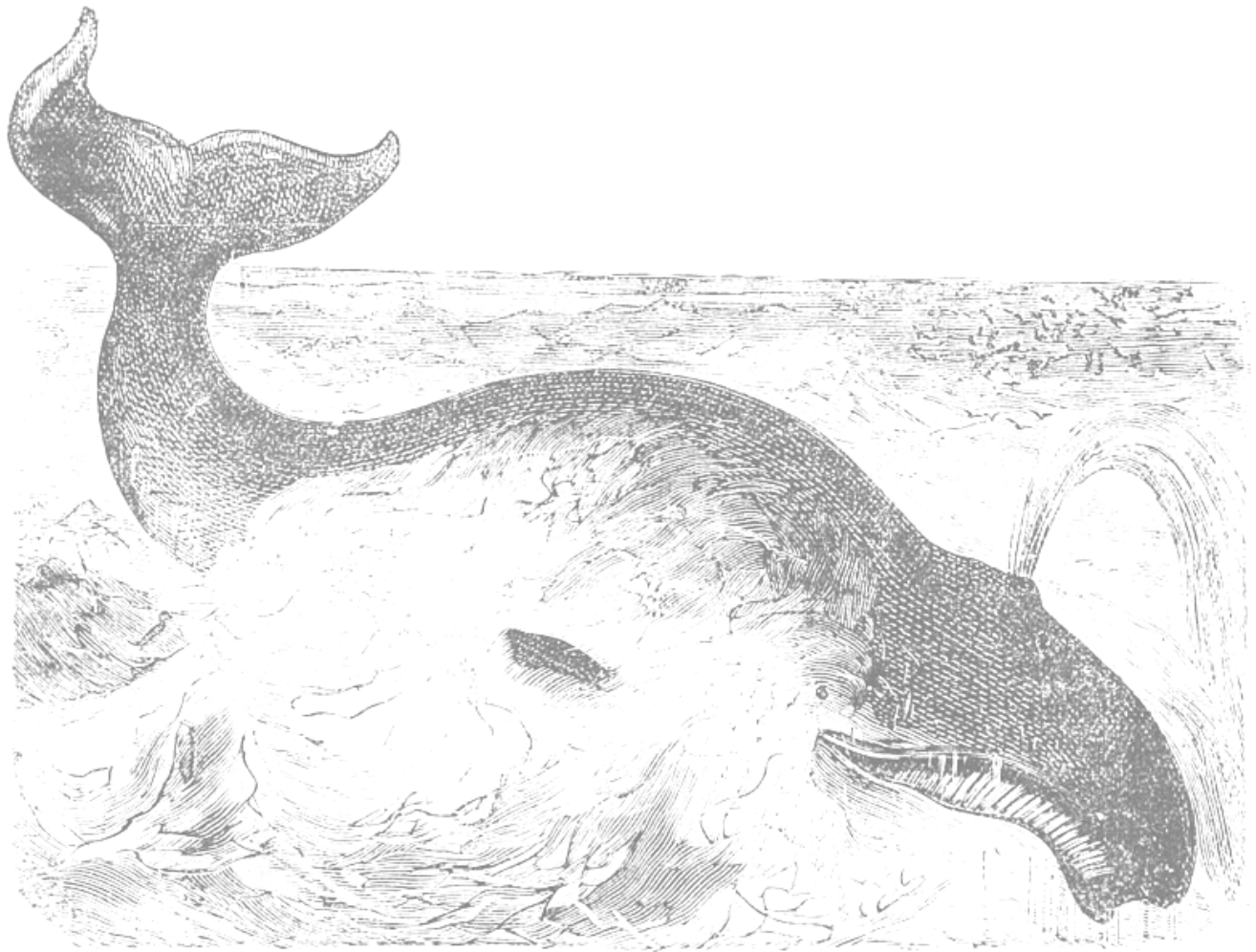
La aparente menor diversidad genética en el MHC de los mamíferos marinos puede derivar de causas diferentes a una menor intensidad de selección en el medio marino las cuales pueden ser resultado de historias de vida y estructuras poblacionales diferentes. La distinción entre los efectos de la selección y la demografía en la variación del MHC puede hacerse comparando la variación de estos genes y la de un marcador neutral.

Si el régimen de selección del MHC en mamíferos marinos y terrestres es distinto, deben observarse diferencias en la existencia de polimorfismos transespecíficos (persistencia de linajes), los patrones de sustitución sinónima y no sinónima en las posiciones de codón con referencia a un marcador neutral y/ o en las propiedades estructurales y de reactividad inmunológica de la región de unión a antígenos de las moléculas de clase II que son las relacionadas con las respuestas a antígenos exógenos.



## Capítulo 2

Evolution of the MHC-*DQB* exon 2 in  
marine and terrestrial mammals



# Evolution of the MHC-DQB exon 2 in marine and terrestrial mammals

María José Villanueva-Noriega · Charles Scott Baker ·  
Luis Medrano-González

Received: 25 May 2012 / Accepted: 13 August 2012 / Published online: 13 October 2012  
© Springer-Verlag 2012

**Abstract** On the basis of a general low polymorphism, several studies suggest that balancing selection in the class II major histocompatibility complex (MHC) is weaker in marine mammals as compared with terrestrial mammals. We investigated such differential selection among Cetacea, Artiodactyla, and Primates at exon 2 of MHC-*DQB* gene by contrasting indicators of molecular evolution such as occurrence of transpecific polymorphisms, patterns of phylogenetic branch lengths by codon position, rates of non-synonymous and synonymous substitutions as well as accumulation of variable sites on the sampling of alleles. These indicators were compared between the *DQB* and the mitochondrial cytochrome *b* gene (*cytb*) as a reference of neutral expectations and differences between molecular clocks resulting from life history and historical demography.

All indicators showed that the influence of balancing selection on the *DQB* is more variable and overall weaker for cetaceans. In our sampling, ziphiids, the sperm whale, monodontids and the finless porpoise formed a group with lower *DQB* polymorphism, while mysticetes exhibited a higher *DQB* variation similar to that of terrestrial mammals as well as higher occurrence of transpecific polymorphisms. Different dolphins appeared in the two groups. Larger variation of selection on the cetacean *DQB* could be related to greater stochasticity in their historical demography and thus, to a greater complexity of the general ecology and disease processes of these animals.

**Keywords** Major histocompatibility complex · *DQB* · Cytochrome *b* · Nonsynonymous nucleotide substitution · Transpecific polymorphism · Marine mammals

**Electronic supplementary material** The online version of this article (doi:10.1007/s00251-012-0647-8) contains supplementary material, which is available to authorized users.

M. J. Villanueva-Noriega · L. Medrano-González (✉)  
Departamento de Biología Evolutiva, Facultad de Ciencias,  
Universidad Nacional Autónoma de México,  
Avenida Universidad 3000, Ciudad Universitaria,  
México 04510, México  
e-mail: medranol@ciencias.unam.mx

C. S. Baker · L. Medrano-González  
School of Biological Sciences, University of Auckland,  
Private Bag 92019, Auckland, New Zealand

M. J. Villanueva-Noriega  
Posgrado en Ciencias del Mar y Limnología,  
Universidad Nacional Autónoma de México,  
Avenida Universidad 3000, Ciudad Universitaria,  
México 04510, México

C. S. Baker  
Marine Mammal Institute, Oregon State University,  
2030 SE Marine Science Drive,  
Newport, OR 97365, USA

## Introduction

Mammals possess a cluster of genes, the major histocompatibility complex (MHC), which codes for proteins that participate in humoral and cellular immune responses. MHC molecules are subdivided into three classes, each having different functions. Class I molecules are expressed in virtually all cells of the organism and serve as recognition elements for natural killer cells. Class II molecules are expressed primarily in antigen-presenting cells which present antigen peptides to CD4<sup>+</sup> T cells. Class III genes encode a heterogeneous set of proteins which includes complement components, steroid 21-hydroxylases, heat shock proteins, and cytokines (Goldsby et al. 2003). As far as is known, in placental mammals, the MHC class II genes *DP*, *DN*, *DM*, *DO*, *DQ*, and *DR* code for peptide-binding glycoproteins. Genes *DP*, *DM*, *DQ*, and *DR* usually code for one alpha chain and one beta chain, while gene *DN* code for an alpha

chain and gene *DO* code for a beta chain. In some artiodactyls, *DP* genes are substituted by *DI* and *DY* genes. The general arrangement of MHC class II genes is conserved among mammals although some genes are missing in some species and the number of alpha and beta genes is variable (Beck and Trowsdale 1999; Takahashi et al. 2000; Goldsby et al. 2003; Kumánovics et al. 2003). Genes of the MHC are highly polymorphic in most mammals (Klein and Figueroa 1986; Edwards and Hedrick 1998; Gu and Nei 1999) presumably because of balancing selection acting over functional regions of the MHC proteins such as the peptide-binding regions (PBR) (Hughes and Nei 1988; Gu and Nei 1999). This balancing selection is thought to be driven by the load of pathogens in the environment and by inbreeding avoidance (Apanius et al. 1997).

Until recent years, research on MHC variation was conducted mainly on humans (*Homo sapiens*, HLA Complex), mice (*Mus musculus*, H-2 Complex), and a few wild and domesticated ungulates (Sigurdardóttir et al. 1992; Schwaiger et al. 1994; Swarbrick and Crawford 1997; Amills et al. 1998; Smith et al. 2005) with the view that MHC variation may be relevant for the viability of populations especially of those severely depleted (Bernatchez and Landry 2003; O'Brien et al. 1985; Potts and Wakeland 1990; Ellegren et al. 1993; Gutierrez-Espeleta et al. 2001). However, information linking MHC variability with disease susceptibility is still poor because the whole immunogenetic complex is not yet examined in relation with disease mechanisms (Acevedo-Whitehouse and Cunningham 2006). Also, the presence of particular alleles within the MHC genotype, rather than levels of heterozygosity and other population attributes, appears to be related with morbidity and mortality in animals (Paterson et al. 1998; Hutchings 2009). A similar situation has been found for the association between the human gene *DRD4* and mental disorders (Tovo-Rodrigues et al. 2011) and could explain why some populations are highly susceptible to disease despite exhibiting a high degree of polymorphism at target MHC loci (Gutierrez-Espeleta et al. 2001) and also why populations with low levels of variation at MHC appear viable (Ellegren et al. 1993). Recent investigations using high-resolution methods as well as extensive samplings in species such as cheetahs (*Acionyx jubatus*) are also showing that MHC variation and population health status are actually higher than the traditional view of genetic depletion associated with high disease susceptibility (Castro-Prieto et al. 2011).

Investigation on MHC genes of marine mammals, mainly on *DR* and *DQ* genes, shows low levels of polymorphism for several species; this includes cetaceans (Trowsdale et al. 1989; Murray et al. 1995; Hayashi et al. 2003; Munguía-Vega et al. 2007; Nigenda-Morales et al. 2008) and pinnipeds (Slade 1992; Hoelzel et al. 1999). In other marine mammal species, MHC shows a high degree of variation

(Baker et al. 2006; Bowen et al. 2006; Xu et al. 2007). Founder effect, population bottlenecks, and a low immunologic selective pressure in the marine environment have been invoked to explain the low levels of MHC polymorphism found in some marine mammals (Trowsdale et al. 1989; Slade 1992). Resistance to disease has been proposed to be a major factor driving the evolution of marine mammal MHC judging from the observation of mass mortalities caused by epizooties and naturally occurring toxins (Harwood and Hall 1990). However, it is not clear whether the apparent low diversity in MHC of some marine mammals is derived from a different regime of selection in the aquatic environment or from genetic drift associated with demographic fluctuations, or even low rates of mutational substitutions (e.g., Jackson et al. 2009). Although mortality in the marine environment is associated with a pattern of pathogenic load different to that of the terrestrial environment (McCallum et al. 2004), the selective regime over the MHC on both realms is not necessarily different, since the low diversity of MHC in some marine mammals might result from a different pattern of mortality.

In order to distinguish a differential regime of selection in the MHC between marine and terrestrial mammals from the effects of different life history and historical demography on the MHC molecular variation, here we examined indicators of molecular evolution at exon 2 of the *DQB* gene which codes for the PBR of class II MHC (Beck and Trowsdale 1999; Takahashi et al. 2000; Goldsby et al. 2003). Such indicators were examined in the set of related marine and terrestrial mammals Cetacea and Artiodactyla as well as in Primates as outgroup. We chose the *DQB* gene because it is present in at least three of the four main placental clades, i.e., Afrotheria, Euarchontoglires, and Laurasiatheria, available data cover a broader spectrum of terrestrial and marine mammal species to compare, and also the *DQB* evolutionary phenomenology relies mainly on nucleotide substitutions among orthologous genes of large phylogenetic longevity (Takahashi et al. 2000).

Balancing selection acting over the evolutionary history of the *DQB* gene can result in the retention of alleles among species for longer than expected under neutral evolution (Cutrera and Lacey 2007). Balancing selection may, therefore, be characterized by the occurrence of transpecific polymorphisms along the phylogeny of MHC genes (Klein and Figueroa 1986; Takahata 1990; Takahata et al. 1992). Maintenance of high variation in the MHC driven by balancing selection can also be shown through the analysis of substitution rates by codon position. This can be evaluated by phylogenetic branch lengths as well as by the rates of synonymous ( $d_S$ ) and nonsynonymous ( $d_N$ ) substitutions where  $d_N$  has proved to be greater than  $d_S$  in the PBR in comparison with non-PBR sites (Hughes and Nei 1988). We also explored the pattern of occurrence of variable sites at

the PBR sites throughout the accumulation of alleles sampled. These molecular analyses represent the complementary phylogenetic and distance approaches.

An analysis of absolute molecular rates is restricted by differences among taxa in molecular clocks resulting from life history and demography (e.g., Jackson et al. 2009) by the lack of complete time calibration data and also because such analysis excludes examination of intraspecific variation. Since the mitochondrial cytochrome *b* gene (*cytb*) is a relatively conserved molecular marker which is under neutral evolution, its variation should reflect approximate phylogenetic distances among species in relation with their particular life histories and demographies (Irwin and Árnason 1994), giving, thus, a reference for contrast with variation at DQB (Takahata 1990). We have, thus, compared the patterns of molecular evolution listed above at the DQB with the *cytb* as a reference of molecular changes assumed to reflect neutral molecular variation as well as changes in life history and demography along history independent of balancing selection on the MHC.

## Methods

### DNA sequence data

We compiled from the GenBank (Benson et al. 2005) a total of 260 nucleotide sequences of the second exon in the *DQB* gene (171 bp), which codes for the PBR of the class II MHC antigen receptors. We compiled as well 248 sequences of the *cytb* gene (1,140 bp) for a total of 33 marine and terrestrial mammal species of the orders Cetacea, Artiodactyla, and Primates. We generated 20 of the analyzed DQB sequences from ten cetacean species and one artiodactyl (Table 1). These DQB data were obtained from tissue samples in the stocks of authors CSB and LMG by standard PCR amplification and cloning (Baker et al. 2006).

### Phylogeny and occurrence of ancestral polymorphisms

Separate phylogenies of the DQB and *cytb* sequences were reconstructed using the program MrBayes3.1.2 (Ronquist and Huelsenbeck 2003) partitioning sequence data in the three codon positions for which their own GTR + G models of nucleotide substitution were determined during the search as the ones selected by the program Modeltest 3.7 of Posada and Crandall (1998). Trees were built by four MC<sup>3</sup> chains with temperature 0.5 in 10 million generations sampling trees every 10,000. Given the convergence of the likelihood values, the first 300 trees were discarded in the DQB analysis and the first 150 trees were discarded for the *cytb*. Consensus trees were finally determined using the majority rule of the program PAUP\* 4.0 (Swofford 2001).

Each DQB sequence was identified with a vector of characters describing its classification at the levels of order, family, genus, and species. This character series was then mapped into the DQB phylogeny to identify unambiguous synapomorphies (i.e., shared derived characters) using the program MacClade 4 (Maddison and Maddison 2000). Synapomorphies of the classification vectors were interpreted as transpecific polymorphisms which result from the long-term persistence of MHC alleles beyond speciation events and which may even survive beyond divergence between genera and families. Since transpecific changes identified in the DQB tree may imply different taxonomical ranks above the species level (*trans*-order, *trans*-familial, and *trans*-genus), we refer to them in general as *trans*-taxon changes. A rate of *trans*-taxon polymorphism in DQB for the taxonomic levels ( $R_l$ ) species within genera, genera within families and families within orders was then defined as follows:

$$R_l = \frac{T_l}{(L_{\text{cyt}bl})(b_{\text{cyt}bl})}$$

where  $T_l$  stands for all the *trans*-taxon changes at DQB accumulated at taxonomic level  $l$ ,  $L_{\text{cyt}bl}$  is the mean length of branches at *cytb* phylogeny separating taxa at level  $l$  and  $b_{\text{cyt}bl}$  is the number of accumulated branches at level  $l$  on the *cytb* phylogeny.  $R_l$  is, thus, the number of *trans*-taxon changes per unitary branch length sampled at the *cytb* phylogeny.  $R_l$  was profiled for Cetacea, Artiodactyla, and Primates at the different levels  $l$  measured with the mean length of branches at *cytb* phylogeny separating taxa at level  $l$  ( $L_{\text{cyt}bl}$ ). The number of accumulated *trans*-taxon changes at each taxonomic level was examined between paired taxa with the  $\chi^2$  test after showing that *cytb* branch lengths are not significantly different among the taxa by the Student's  $t$  test. This, in turn, was made after examining the normality of data with the Kolmogorov–Smirnov  $d$  test.

### Patterns of phylogenetic branch lengths

For the DQB and *cytb* phylogenies, Bayesian tree lengths between all sequences among all species pairs within Cetacea, Artiodactyla, and Primates were averaged. Therefore, for each species pair mean, the proportion DQB tree length/*cytb* tree length ( $L_{\text{DQB}}/L_{\text{cyt}b}$ ), as a reference of DQB selection signal, was plotted against the *cytb* tree length ( $L_{\text{cyt}b}$ ) as a measure of neutral phylogenetic divergence. The plot obtained is a decaying curve asymptotic to the proportion  $L_{\text{DQB}}/L_{\text{cyt}b} = 1$  and appearing as a power law in log–log representation; this is,  $L_{\text{DQB}}/L_{\text{cyt}b} = AL_{\text{cyt}b}^b$  where  $A$  ( $>1$ ) is the antilogarithm of the intercept ( $a$ ) and  $b$  ( $<0$ ) is the slope of the linear regression of the  $\log(L_{\text{DQB}}/L_{\text{cyt}b})$  vs.  $\log(L_{\text{cyt}b})$  plot. In order to discriminate variation due to specific neutral rates among taxa, this comparison was

**Table 1** Families and species of the three mammal orders examined including GenBank accessions as well as the number of sequences (*n*) for *DQB* and *cytb* genes

Order/family	Species	DQB GenBank accession <sup>a</sup>	DQB <i>n</i>	<i>cytb</i> GenBank accession	<i>Cytb</i> <i>n</i>
Cetacea					
Delphinidae	<i>Cephalorhynchus hectori</i>	DQ354629, DQ354628, Own (1)	3	EF093033, AF084071	2
	<i>Globicephala macrorhynchus</i>	AB164228, AB164227, AB164226	3	AF084055, AF084054	2
	<i>Sagmatias obliquidens</i>	AB164225, AB164224, Own (1)	3	EF093041, EF093040, EF093039, EF093038, EF093037, EF093036, AF084067	7
	<i>Tursiops truncatus</i>	AB164221, Own (1)	2	DQ466029, DQ466028, DQ466027, DQ466026, DQ466025, EF093029, X92526, AF084095, AF084094, AF084093	10
Monodontidae	<i>Delphinapterus leucas</i>	U16990, U16989, U16988, U16987, U16986, Own (1)	6	DLU72037, X92531	2
	<i>Monodon monoceros</i>	U16991, Own (1)	2	AJ554062, NC_005279, MMU72038, X92532	4
Phocoenidae	<i>Neophocaena phocaenoides</i>	AB164212, AB164214, AB164215, AB164216, AB164217, AB164218, AB164219	7	EF203442, DQ364692, DQ364691, AF334489, NPU09680	5
Balaenopteridae	<i>Balaenoptera bonaerensis</i>	AB164202, AB164203, AB164204, AB164205, AB164206, AB164207	6	X75581, NC_006926, AP006466	3
Ziphiidae	<i>Mesoplodon europaeus</i>	DQ354641, DQ354640	2	X92537	1
	<i>Mesoplodon grayi</i>	DQ354639, Own (1)	2	AY579546, AY579545	2
	<i>Mesoplodon stejnegeri</i>	AB164209, DQ354638	2	AY579554, AY579553	2
Physeteridae	<i>Physeter macrocephalus</i>	AB164208, Own (1)	2	X75589, AF304073, AJ277029, NC_002503	4
	<i>Balaenoptera physalus</i>	AB164199, DQ300263, DQ300262, DQ300261, DQ354627, DQ354626, Own (1)	7	X61145, NC_001321	2
	<i>Megaptera novaeangliae</i>	DQ354664, DQ354663, DQ354662, DQ354661, DQ354660, DQ354659, DQ354658, DQ354657, DQ354656, DQ354655, DQ354654, DQ354653, DQ354652, DQ354651, DQ354650, DQ354649, DQ354648, DQ354647, DQ354646, DQ354645, DQ354644, DQ354643, DQ354642, Own (10)	33	X75584, NC_006927, AP006467	3
Eschrichtiidae	<i>Eschrichtius robustus</i>	DQ354636, DQ354635	2	AJ554053, AP006471, NC_005270	3
Balaenidae	<i>Balaena mysticetus</i>	DQ354625, DQ354624, DQ354623	3	AJ554051, AP006472, NC_005268, X75588	4
	<i>Eubalaena australis</i>	AB164198, DQ354634, DQ354633, DQ354632, DQ354631, DQ354630, Own (1)	7	DQ095153, DQ095152, NC_006930, AP006473	4
<b>Subtotal</b>	<b>17</b>		<b>92</b>		<b>60</b>
Artiodactyla					
Hippopotamidae	<i>Hippopotamus amphibius</i>	Own (1)	1	Y08813, U07565, AJ010957, NC_000889	4
Bovidae	<i>Bos indicus</i>	AJ421636, AJ249896, AJ249897, AJ249898, AJ249718, AJ249717, AJ249716, X79348, X79352	9	AF531473, NC_005971, AF492350, AY126697	4
	<i>Bos taurus</i>	U77795, U77796, U77798, AJ421635, AJ421634, AJ421633, AJ421632, AJ421631, AF376814, AF376808, Z48206, Z48205, NM_001034668, XM_585873, D37954, D37952, D37953, S43263, AY730728, AY730727, U77800, U77799, U77794, U77793, U77792, U77791, U77790, U77789, U77788	28	AY526085, AY521038, AY521037, AY521036, AY521035, AY521034, AY521033, AY521032, AY521031, AY521030, AY521063, AY521062, AY521061, AY521060, AY521059, AY521058, AY521057, AY521056, AY521055, AY521054, AY521053, AY521052, AY521051, AY521050, AY521049, AY521048, AY521047, AY521046, AY521045, AY521044,	35

**Table 1** (continued)

Order/family	Species	DQB GenBank accession <sup>a</sup>	DQB <i>n</i>	<i>cytb</i> GenBank accession	<i>Cytb</i> <i>n</i>
Cervidae	<i>Bubalus bubalis</i>	AY699887, AY699886, AY699885, AY699884, AY699883, AY699882, AY699881, AY699880, AY699879, AY699878, AY699877, AY699876	12	AY521043, AY521042, AY521041, AY521040, AY521039, EF409942, EF409941, EF409940, EF409939, AY702618, AY488491, NC_006295	7
	<i>Ovibos moschatus</i>	AY077796, AY077795, AY077794, AY077793, AY077792, AY077791, AY077790, AY077789, AY077788	9	AY669322, OMU90303, OMU90302, OMU90301, OMU90300, OMU17862	6
	<i>Cervus elaphus</i>	U39052, U39058, U39057, U39056, U39055, U39054, U39053, U39051, AF119786, AF119785, AF119784, AF119783, AF119782, AF119780, AF119779, AF119778, AF119777, AF119776, AF119775, AF119774	20	AJ000021, AY244490, AY044859, AY035871, AY347753, AY347752, AB021099, AB021098, AB021097, AB021096, AY148966, AY142327, AY142326, AY118199, AY118198, AY118197, AY070226, AY070225, AY070224, AY070223, AY070222, AY070221, AY044861, AY044860, AY044858, AY044856	26
	<i>Cervus nippon</i>	AY679483, AY679482, AY679481, AY679480, AY679479, AY679478, AY679477, AY679476, AY679475, AY679474, AY679473, AY679472, AY679471, AY679470, AY679468, AY679467, AY679466, AY679465, AY679464, AY679463, AY679462, AY679461	22	AB160860, AY035876, AB021095, AB021094, AB021093, AB021092, AB021091, AB021090, NC_008462, EF058308, DQ985076, NC_007179, AB218689, NC_006993, NC_006973, AB211429, AB210267	17
	<i>Alces alces</i>	AY077784, AY077785, AY077786, AY077787	4	AJ000026, AY090104, AY090103, AY090102, AY090101, AY090100, AY090099	7
Suidae	<i>Sus scrofa</i>	AY135572, AY135571, AY135570, AY135569, AY135568	5	EF061505, EF061504, EF061503, EF061502, EF061501, EF061500, EF061499, EF061498, EF061497, EF061496, EF061495, EF061494, EF061493, EF061492, EF061491, EF061490, EF061489, EF061488, EF061487, EF061486, EF061485, EF061484, EF061483	23
<b>Subtotal</b>	<b>9</b>		<b>110</b>		<b>129</b>
Primates					
Hominidae	<i>Homo sapiens</i>	NM_002123, M24364	2	EF459669, EF458167, EF459670, EF452294, EF452293, EF452295, EF449507, EF449506, EF061150, EF061159, EF061158, EF061157, EF061156, EF061155, EF061154, EF061153, EF061152, EF061151, EF061149, EF061148, EF061147, EF061146, EF061145, EF061144, AY495156, AY495324, AY495325, AY495326, AY495327, AY495328	30
	<i>Gorilla gorilla</i>	M81265, M81266, M81267	3	X93347, NC_001645, GORMTCD38114	3
	<i>Pan troglodytes</i>	AJ308040, AJ308039, AJ308038, M81260, M81262, M81263, M81264, M81259	8	NC_001643, AY585843, AY585842, AY585841, AY585840, AY585839, AY585838, AY585836, AY585835, AY585834, AY585833, X93338, X93339	13
	<i>Pongo pygmaeus</i>	AJ308043, AJ308042, AJ308041, M81270, M81273, M81271, M81272	7	NC_001646, NC_002083	2
Cercopithecoidea	<i>Macaca arctoides</i>	AJ308053, AJ308052, AJ308051	3		7

**Table 1** (continued)

Order/family	Species	DQB GenBank accession <sup>a</sup>	DQB <i>n</i>	<i>cytb</i> GenBank accession	<i>Cytb</i> <i>n</i>
				AY738634, AY685835, AY685834, AY685833, AY685832, AY685831, AY685830	
	<i>Macaca mulatta</i>	AJ308050, AJ308049, AJ308048, AJ308047, AJ308046, AJ308045, AJ308044, M81269, M81301, M81299, M81292, M81294, M81286, M81287, M81289, M81290, M81291, M81293, M81296, M81297, M81300, M81305	22	NC_005943, AY612638	2
Aotidae	<i>Aotus nancymae</i>	AF213642, AF213641, AF213640, AF213639, AF213638, AF213637, AF213636, AF213635, AF213634, AF213633, AF213632, AF213631, AF213630	13	AJ489745, AJ489746	2
<b>Subtotal</b>		<b>7</b>	<b>58</b>		<b>59</b>
<b>Total</b>		<b>33</b>	<b>260</b>		<b>248</b>

Bold numbers indicate subtotals of the number of species as well as DQB and *cytb* sequences examined for Cetacea, Artiodactyla and Primates

<sup>a</sup> Own sequences were deposited in the Dryad Repository: <http://dx.doi.org/10.5061/dryad.b33m7>

partitioned by codon position, since  $L_{DQB}/L_{cytb}$  values are expected to be higher in the first and second codon positions for balancing selection and higher at the third codon position for general neutrality. Parameters  $a$ ,  $b$ , and the regression index ( $r^2$ ) were compared in the three main taxa by codon position ( $i$ ) with reference to the whole sequence as  $a_i - a_{all}$ ,  $b_i/b_{all}$  and  $r_i^2/r_{all}^2$ . The regression parameters were also compared for each taxon  $j$  with reference to Primates as  $a_j - a_{Primates}$ ,  $b_j/b_{Primates}$ , and  $r_j^2/r_{Primates}^2$ . Statistical significance for the differences of  $L_{DQB}/L_{cytb}$  among taxa in the whole sequences as well as in the three codon positions was tested with a general linear model (GLM) dependent upon taxa and  $L_{cytb}$  using the program Statistica 7. GLM included the Levene's test for homogeneity of variances.

#### Patterns of nonsynonymous and synonymous substitutions

The nonsynonymous ( $d_N$ ) and synonymous ( $d_S$ ) substitution distances of Nei and Gojobori (1986) were calculated for all DQB sequence pairs using the program MEGA (Kumar et al. 2008). We obtained average  $d_N$  and  $d_S$  distances from pairwise comparisons between sequences belonging to different species within the three taxa groups, since the interspecific  $d_N/d_S$  ratio adequately measures selection (Kryazhimskiy and Plotkin 2008). The parameter for the selection coefficient  $4Ns$  was then estimated from Eq. 3 of Nielsen and Yang (2003) by means of interpolation. The interspecific  $d_N/d_S$  ratio is related with the selection coefficient ( $s$ ) as follows:

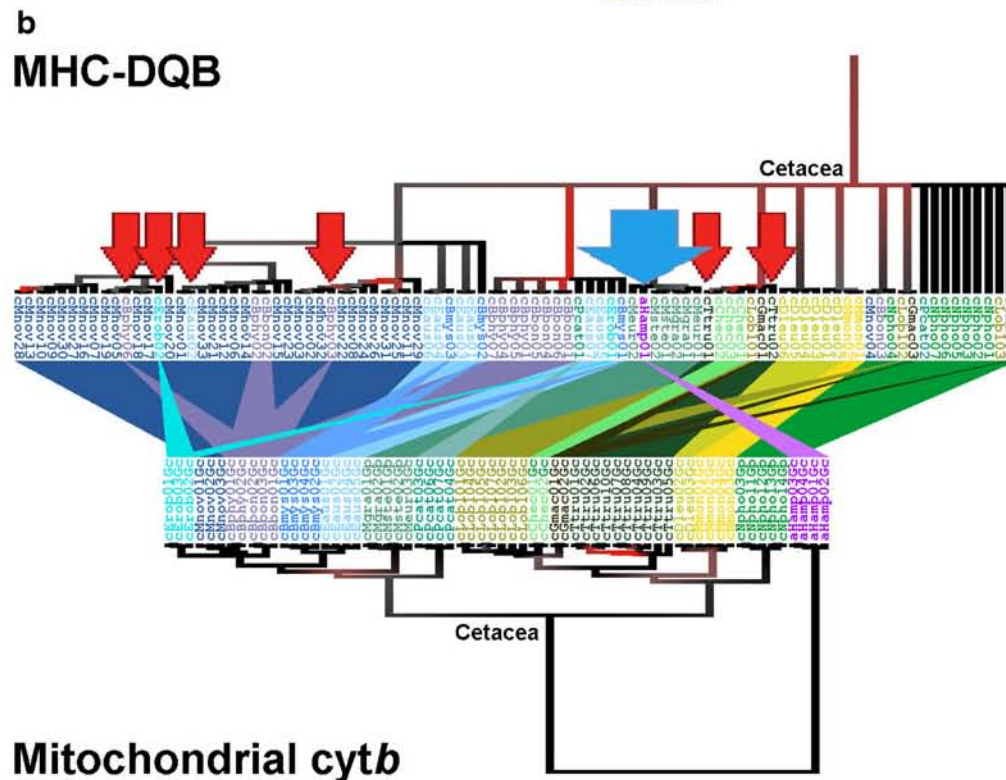
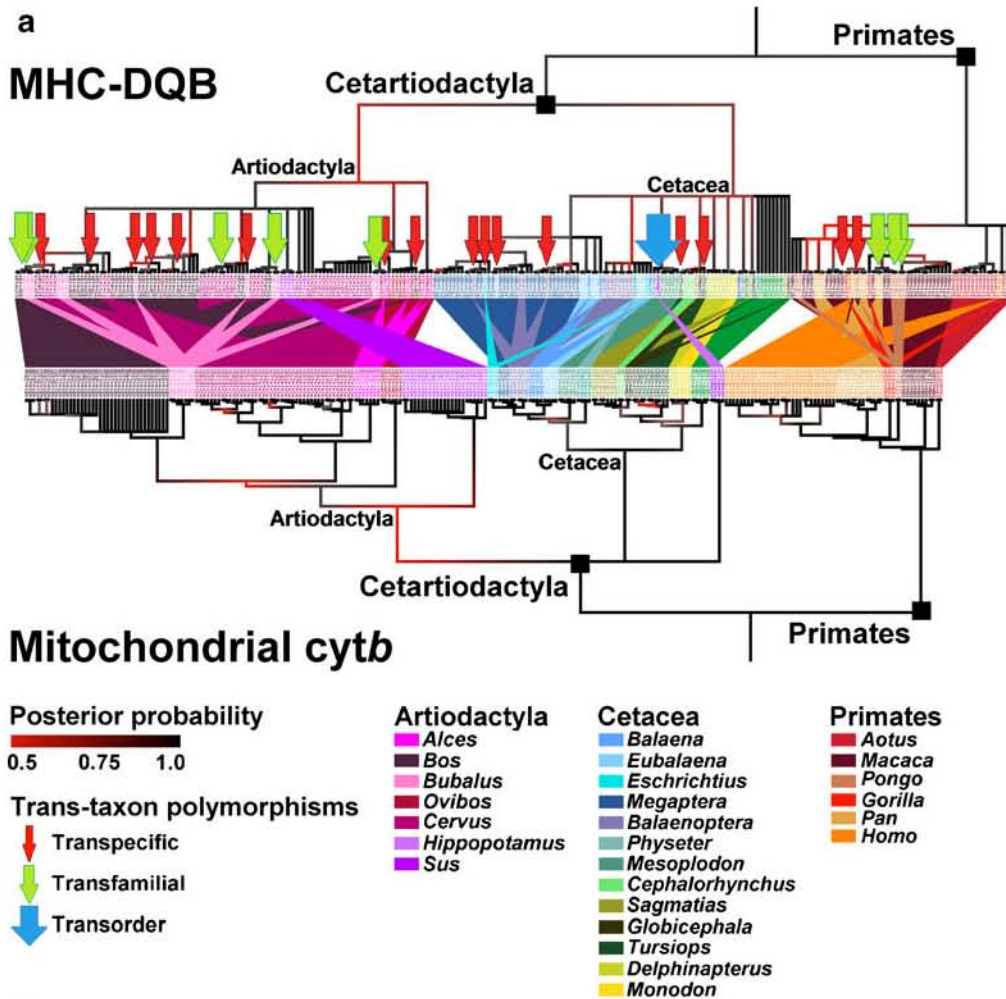
$$\frac{d_N}{d_S} = \frac{4Ns}{1 - e^{-4Ns}}$$

where  $N$  is the long-term effective population size. Notice that  $d_N/d_S$  is nearly equal to  $4Ns$  at large positive values of both (e.g.,  $>3$ ). When  $d_N=1$  and  $d_S=0$ , values were replaced with the highest determined positive  $d_N/d_S$  ratio found in the comparison set. The values of  $4Ns$  for each interspecies comparison of DQB within the three studied mammal orders were plotted against their respective average DQB Jukes–Cantor distance. Statistical significance for the differences of  $4Ns$  among taxa was tested with a GLM dependent upon the DQB Jukes–Cantor distance and included the Levene's test for variance homogeneity using the program Statistica 7. Because the Jukes–Cantor distance of DQB is distributed differently among the three main taxa and also because the  $d_N/d_S$  ratio mismeasures selection within species (Kryazhimskiy and Plotkin 2008), the distribution of intraspecific  $d_N$  values in the tree main taxa was examined with a Student's  $t$  test after demonstrating the normality of data with the Kolmogorov–Smirnov  $d$  test.

#### Accumulation of variable sites

The number of variable sites in sequence alignments, as additional evidence of selection-driven polymorphism, was determined for the DQB and *cytb* sequences for each species of Cetacea, Artiodactyla, and Primates. Thereafter, the

**Fig. 1 a** Bayesian majority rule consensus tree for the mitochondrial *cytb* gene (900 trees; LnL=-23,323.304) and the MHC-DQB gene (600 trees; LnL=-8634.446). DQB unambiguous *trans*-taxon events are indicated with arrows. **b** Detail of the cetacean trees. Sequences are abbreviated with one character indicating the order, one character indicating genus, three characters indicating specific name and a serial number





proportion of variable sites per allele and nucleotide was determined and partitioned by codon position. After testing for normality of these data within the main taxa with the Kolmogorov–Smirnov  $d$  test, statistical significance for the differences between taxa were determined with the Student's  $t$  test. A profile for the accumulation of variable sites ( $v_s$ ) in dependence of the alleles sampled ( $s$ ) for each group of species was built and fitted to the following *ad hoc* asymptotic–hyperbolic function:

$$v_s = v_{\max} \frac{s - 1}{k + s - 1}$$

where  $s-1$  stands for the accumulation start of  $v_s=0$  at  $s=1$ ,  $v_{\max}$  is the asymptotic number of variable sites at an infinite number of alleles sampled and  $k$  is the number of alleles minus one at which half of the variable sites are observed. Parameters  $v_{\max}$  and  $k$  and their standard errors were estimated for the three mammal orders examined with the least squares procedure available in the program SigmaPlot 8. Fitting was not practicable for the *cytb* sequences.

## Results

### Phylogeny and occurrence of ancestral polymorphisms

Separate phylogenies of the DQB and *cytb* were built for 260 and 248 sequences, respectively, belonging to Cetacea, Artiodactyla, and Primates (Table 1). The DQB phylogeny revealed several *trans*-taxon changes for both marine and terrestrial mammals (Fig. 1) suggesting the retention of allele diversity among divergent species in the three mammal orders. The *cytb* provided a phylogeny consistent with the known organismal phylogeny on which cetaceans and artiodactyls are linked in the clade Cetartiodactyla (Irwin and Árnason 1994; Montgelard et al. 1997; Murphy et al. 2001; Fig. 1a).

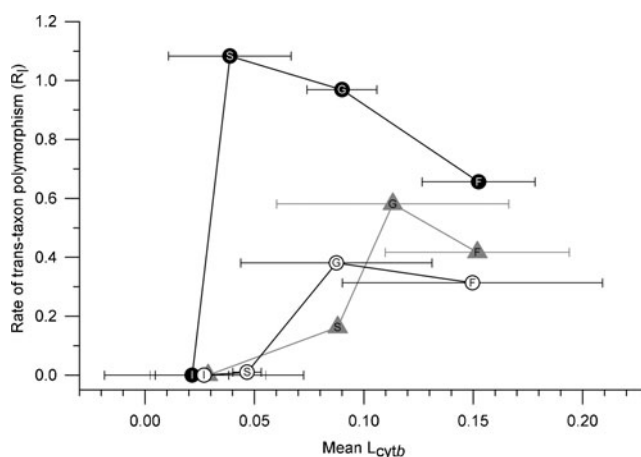
A mapping of species classification into our Bayesian phylogeny of the DQB, revealed a total of 25 unambiguous *trans*-taxon changes for marine and terrestrial mammal species of which 22 were transspecific or transgenus, two were transfamilial, and one was transorder linking the hippopotamus (*Hippopotamus amphibius*) DQB sequence with the Ziphiidae clade at a Jukes–Cantor distance as low as 0.06 for the comparison between the hippopotamus and the Gervais' beaked whale (*Mesoplodon europaeus*; Fig. 1b). Mysticetes showed DQB transspecific polymorphisms particularly among the humpback whale (*Megaptera novaeangliae*) and the fin whale (*Balaenoptera physalus*) and a transfamilial polymorphism was observed among the southern right whale (*Eubalaena australis*) and the gray whale (*Eschrichtius robustus*) (Fig. 1b). We found no transfamilial polymorphisms within Delphinoidea but found two transgenus polymorphisms between the bottlenose dolphin

(*Tursiops truncatus*) and the short-finned pilot whale (*Globicephala macrorhynchus*) (Fig. 1b). Sequences from Bovidae and Cervidae families from the Artiodactyla were found much intermingled in the DQB tree as reported previously (Hassanin and Douzery 2003). There was also a consistent relationship between DQB sequences from moose (*Alces alces*) and musk ox (*Ovibos moschatus*). As for Primates, we found persistence of DQB alleles among the Catarrhini such as the stump-tailed macaque (*Macaca arctoides*) and the chimpanzee (*Pan troglodytes*). DQB sequences from night monkeys (*Aotus nancymae*; Platyrrhini) remained apart from the rest of the Primates examined (Fig. 1a).

Accumulated occurrence of *trans*-taxon polymorphisms profiled along the accumulated branch length of the *cytb* gene was lower for cetaceans despite the large number of species sampled (Table 1). Among the three analyzed mammal orders, artiodactyls showed the largest rate of accumulated *trans*-taxon variations (Fig. 2). No statistically significant differences were observed between the  $L_{\text{cytb}}$  distances of the three main taxa in the three taxonomic levels examined (interval of  $p$  0.14–0.48). According to the  $\chi^2$  test, no significant differences in the occurrence of *trans*-taxon polymorphisms for the three taxonomic levels were observed between cetaceans and primates ( $p>0.31$ ), but significant or marginally significant higher occurrence of *trans*-taxon variations were observed for artiodactyls compared with cetaceans (interval of  $p$  0.02–0.11) and primates (interval of  $p$  0.06–0.10).

### Patterns of phylogenetic branch lengths

In all taxa and codon positions, the  $L_{\text{DQB}}/L_{\text{cytb}}$  ratio showed a decaying relationship with the branch length of the *cytb*

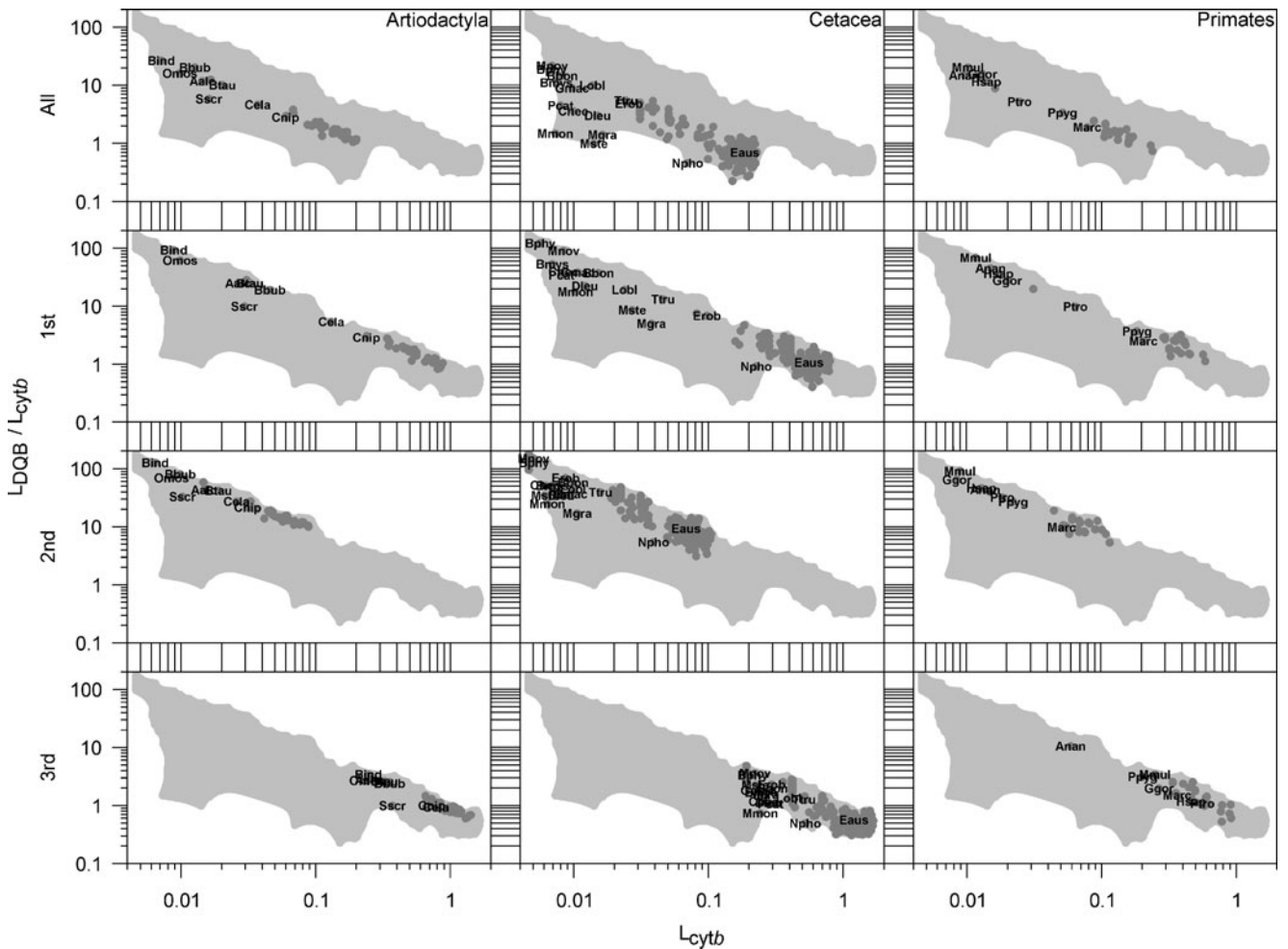


**Fig. 2** Rate of *trans*-taxon changes at the taxonomic levels of individuals within species (*I*), species within genera (*S*), genera within families (*G*), and families within orders (*F*). Black circles stand for Artiodactyla, open circles for Cetacea and gray triangles for Primates. Taxonomic levels are measured with the mean branch length at *cytb* phylogeny separating taxa ( $L_{\text{cytb}}$ ). Horizontal bars indicate standard deviation of the average *cytb* lengths

gene ( $L_{cytb}$ ) which seems asymptotic to one (Fig. 3). Intercept values ( $a$ ) in artiodactyls and primates ( $-0.59$  and  $-0.67$ , respectively) were higher as compared with cetaceans ( $-0.80$ ) (Table 2). This can be interpreted as a higher selection signal over the DQB sequences of terrestrial mammals. Higher values of the regression parameter  $a$  are also observed in the first and second codon positions as compared with the third position for all mammal orders examined. Lower  $r^2$  values indicate more variation in selection signal in cetaceans (0.72) as compared with artiodactyls and primates (0.96 and 0.97, respectively). For the whole sequence as well as for the three codon positions, the GLM identified a significant dependence of the  $L_{DQB}/L_{cytb}$  ratio on  $L_{cytb}$  ( $p < 0.01$ ). The estimated  $L_{DQB}/L_{cytb}$  means were lower for cetaceans in the whole sequence and in the first codon position, while estimated means and standard errors of  $L_{cytb}$  were almost equal among the three main taxa in the entire sequence and at each codon position. For the whole

sequence, the  $L_{DQB}/L_{cytb}$  ratio was significantly different in the pairwise comparisons of cetaceans vs. artiodactyls and primates ( $p < 0.01$ ) but no significant difference was found between artiodactyls and primates ( $p = 0.56$ ). In the whole sequence, the Levene's test showed significant heterogeneity of variances for the  $L_{DQB}/L_{cytb}$  ratio and for  $L_{cytb}$  between cetaceans and the other two main taxa (interval of  $p$  0.00–0.09) but not between artiodactyls and primates ( $p_{L_{cytb}} = 0.77$  and  $p_{L_{DQB}/L_{cytb}} = 0.76$ ).

Values of the  $L_{DQB}/L_{cytb}$  ratio in the third codon position were the lowest and the most similar among the three major taxa examined. Comparisons of parameters  $a$ ,  $b$ , and  $r^2$  with reference to primates show that variation in cetaceans is higher ( $r_j^2/r_{Primates}^2 = 0.75$ ) and that selective signal is lower ( $a_j - a_{Primates} = -0.13$ ). Artiodactyls showed a pattern similar to primates ( $r_j^2/r_{Primates}^2 > 0.99$ ) with a slightly higher selective signal ( $a_j - a_{Primates} = 0.09$ ) (Table 2; Fig. 3). For all codon positions, the slope parameter ( $b$ ) is lower in



**Fig. 3** Comparisons of branch length distances between the *cytb* gene ( $L_{cytb}$ ) and the ratio between DQB and *cytb* lengths ( $L_{DQB}/L_{cytb}$ ) for Artiodactyla, Cetacea, and Primates in the whole sequence and the three codon positions. Intraspecific comparisons are indicated with the

abbreviated species name using the first character of the genus name and the first three characters of the species name. Interspecific distances are shown in dark gray circles. The light gray area in the plots ground is the envelope of all points in all graphs

**Table 2** Log–log regression values as well as GLM means and standard errors (SE) for branch length comparisons among *cytb* and DQB sequences shown in Fig. 3

	Artiodactyla	Cetacea	Primates
All			
<i>A</i>	−0.59	−0.80	−0.67
<i>B</i>	−0.90	−0.84	−0.93
$r^2$	0.96	0.72	0.97
$L_{DQB}/L_{cytb}$ mean [SE]	3.81 [0.81]	1.87 [0.24]	3.86 [0.96]
$L_{cytb}$ mean [SE]	0.11 [0.01]	0.13 [0.01]	0.11 [0.01]
1st			
<i>a</i>	−0.07	−0.18	−0.11
<i>b</i>	−0.94	−0.88	−0.94
$r^2$	0.98	0.88	0.96
$L_{DQB}/L_{cytb}$ mean [SE]	7.24 [2.55]	4.61 [1.12]	9.01 [3.02]
$L_{cytb}$ mean [SE]	0.49 [0.04]	0.45 [0.02]	0.29 [0.03]
2nd			
<i>a</i>	0.10	−0.137	−0.029
<i>b</i>	−0.85	−0.915	−0.892
$r^2$	0.94	0.819	0.906
$L_{DQB}/L_{cytb}$ mean [SE]	23.2 [3.88]	21.2 [2.40]	19.6 [3.72]
$L_{cytb}$ mean [SE]	0.05 [0.00]	0.06 [0.00]	0.06 [0.01]
3rd			
<i>A</i>	−0.07	−0.20	−0.16
<i>B</i>	−0.84	−0.89	−0.99
$r^2$	0.88	0.73	0.84
$L_{DQB}/L_{cytb}$ mean [SE]	1.16 [0.10]	0.97 [0.07]	1.88 [0.34]
$L_{cytb}$ mean [SE]	0.87 [0.05]	0.95 [0.04]	0.51 [0.04]

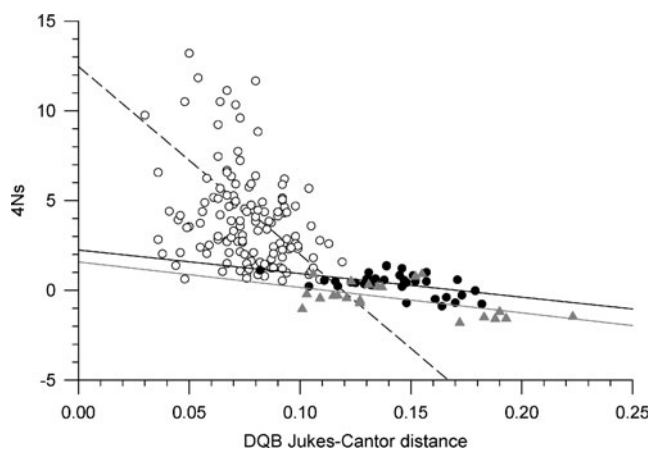
cetaceans than that of artiodactyls and primates with noticeable exception of the second codon position for which the cetacean slope is higher. A slower decrease of the  $L_{DQB}/L_{cytb}$  ratio with reference to  $L_{cytb}$ , indicated by the slope parameter in cetaceans, may reflect the lower selection signal on the DQB as inferred from the values of the intercept parameter (*a*). The GLM mean value of  $L_{cytb}$  was higher for the third codon position and lower for the second codon position, while the GLM mean  $L_{DQB}/L_{cytb}$  ratio was higher for the second codon position and lower for the third codon position (Table 2). The  $L_{DQB}/L_{cytb}$  ratio was similar and not significantly different between artiodactyls, cetaceans, and primates in the second codon position. For the first and third codon positions, primates exhibited the highest  $L_{DQB}/L_{cytb}$  ratio and cetaceans, the lowest. Differences were significant or marginally significant in the third codon position for all comparisons (interval of  $p$  0.00–0.10) and significant only between cetaceans and primates at the first codon position ( $p=0.06$ ).

Artiodactyl species from the Bovidae family showed the highest intraspecific selective signal in this group, whereas species from the Cervidae family showed the lowest. All mysticetes showed consistent high selective signal, while

ziphiids, monodontids, the sperm whale (*Physeter macrocephalus*), and the finless porpoise (*Neophocaena phocaenoides*) revealed the lowest selective pressure in the cetacean group. No evident pattern among species was observed in the Delphinidae and primates (Fig. 3).

#### Patterns of nonsynonymous and synonymous substitutions

In the three mammal orders examined, the  $d_N/d_S$  ratio expressed as  $4N_s$  decreases in parallel to the DQB sequence distance between species (Fig. 4). For the paired taxa comparisons and for the global comparison among the three taxa, the GLM showed a significant dependence of  $4N_s$  on the DQB sequence distance ( $p<0.01$ ). However, this distance was not equally distributed among taxa; cetaceans showed the highest values of  $4N_s$  and the lower DQB sequence distance, while primates showed the lower  $4N_s$  values and the higher DQB sequence distance. Mean  $4N_s$  value was 4.62 for cetaceans, 0.37 for artiodactyls, and  $-0.45$  for primates, while the mean DQB distances were 0.07, 0.14, and 0.14, respectively (Fig. 4). After testing the normality of the DQB sequence distances with the Kolmogorov–Smirnov  $d$  test ( $p>0.10$ ), the Student's  $t$  test showed a significant difference between cetaceans and the two terrestrial orders of mammals ( $p<0.02$ ) as well as a marginally significant difference between artiodactyls and primates ( $p=0.10$ ). This pattern was similar in the *cytb* gene and precluded application of the GLM, since the three main taxa cannot be compared on the same levels of variation of DQB or *cytb* sequences. Notice, however, that linear regressions for the plot  $4N_s$  vs. DQB distance indicate an expected  $4N_s$  value for cetaceans of  $-2.2$  at the mean DQB distance of 0.14 in artiodactyls and primates which corresponds to  $4N_s$  values of 0.40 and  $-0.40$ , respectively (Fig. 4).



**Fig. 4** Selection parameter  $4N_s$  on the MHC–DQB gene in reference of the DQB Jukes–Cantor distance. *Black circles and line* stand for values and linear regression of Artiodactyla, *open circles and dashed line* for the corresponding Cetacea results and *gray triangles and line* for the Primates data

Therefore, we compared the distribution of intraspecific  $d_N$  values in the three taxa. Within species, synonymous and nonsynonymous sequence distances of the DQB are higher and more variable than the distances of the *cytb*. For the DQB also, nonsynonymous distance is larger than the synonymous distance in the three mammal orders examined. Mean intraspecific distance of the *cytb* gene is lower for cetaceans (0.004) and

higher for primates (0.017). Nonsynonymous distance within species of the DQB is similar for artiodactyls and primates (0.11 and 0.10, respectively) and lower for cetaceans (0.05) (Table 3). After testing the normality of data, the Student's *t* test showed significant differences comparing the DQB distances between cetaceans and the two terrestrial groups ( $p < 0.01$ ) and significant or marginally significant differences for the *cytb*

**Table 3** Average Jukes–Cantor distance ( $d$ ), nonsynonymous ( $d_N$ ) and synonymous ( $d_S$ ) distances and their standard errors (SE) of *cytb* and DQB sequences of the examined species. Abbreviation *nc* stands for not calculated values due to the lack of data

	<i>cytb</i> alleles	DQB alleles	<i>cytb</i> $d$ [SE]	DQB $d$ [SE]	DQB $d_N$ [SE]	DQB $d_S$ [SE]
<b>Artiodactyla</b>						
<i>A. alces</i>	7	4	0.009 [0.003]	0.104 [0.019]	0.117 [0.031]	0.069 [0.027]
<i>B. bubalis</i>	7	12	0.014 [0.003]	0.168 [0.020]	0.174 [0.033]	0.152 [0.038]
<i>B. indicus</i>	4	9	0.001 [0.001]	0.114 [0.016]	0.122 [0.024]	0.094 [0.031]
<i>B. taurus</i>	35	28	0.006 [0.003]	0.108 [0.013]	0.115 [0.023]	0.091 [0.020]
<i>C. elaphus</i>	26	20	0.037 [0.003]	0.102 [0.015]	0.101 [0.023]	0.109 [0.028]
<i>C. nippon</i>	17	22	0.025 [0.003]	0.100 [0.015]	0.104 [0.024]	0.089 [0.028]
<i>H. amphibius</i>	4	1	0.009 [0.002]	nc	nc	nc
<i>O. moschatus</i>	6	9	0.002 [0.001]	0.075 [0.013]	0.085 [0.020]	0.049 [0.020]
<i>S. scrofa</i>	23	5	0.009 [0.001]	0.068 [0.015]	0.079 [0.023]	0.038 [0.024]
<b>Mean</b>	<b>14</b>	<b>12</b>	<b>0.012 [0.004]</b>	<b>0.105 [0.011]</b>	<b>0.112 [0.010]</b>	<b>0.086 [0.013]</b>
<b>Cetacea</b>						
<i>B. bonaerensis</i>	3	6	0.004 [0.001]	0.071 [0.014]	0.084 [0.025]	0.035 [0.020]
<i>B. mysticetus</i>	4	3	0.001 [0.001]	0.049 [0.014]	0.062 [0.028]	0.012 [0.012]
<i>B. physalus</i>	2	7	0.000 [0.000]	0.071 [0.013]	0.081 [0.021]	0.030 [0.015]
<i>C. hectori</i>	2	3	0.002 [0.001]	0.012 [0.007]	0.016 [0.009]	0.000 [0.000]
<i>D. leucas</i>	2	6	0.004 [0.002]	0.031 [0.009]	0.039 [0.016]	0.008 [0.008]
<i>E. australis</i>	4	7	0.000 [0.000]	0.072 [0.015]	0.094 [0.025]	0.014 [0.015]
<i>E. robustus</i>	3	2	0.001 [0.001]	0.055 [0.017]	0.066 [0.033]	0.021 [0.018]
<i>G. macrorhynchus</i>	2	3	0.003 [0.002]	0.085 [0.019]	0.106 [0.035]	0.022 [0.017]
<i>L. obliquidens</i>	7	3	0.003 [0.001]	0.071 [0.017]	0.086 [0.026]	0.023 [0.020]
<i>M. europaeus</i>	1	2	nc	0.019 [0.011]	0.026 [0.015]	0.000 [0.000]
<i>M. grayi</i>	2	2	0.010 [0.005]	0.013 [0.009]	0.000 [0.000]	0.000 [0.000]
<i>M. monoceros</i>	4	2	0.002 [0.001]	0.000 [0.000]	0.000 [0.000]	0.000 [0.000]
<i>M. novaeangliae</i>	3	33	0.002 [0.001]	0.078 [0.012]	0.082 [0.021]	0.056 [0.022]
<i>M. stejnegeri</i>	2	2	0.003 [0.003]	0.000 [0.000]	0.000 [0.000]	0.000 [0.000]
<i>N. phocaenoides</i>	5	7	0.013 [0.004]	0.026 [0.008]	0.035 [0.013]	0.000 [0.000]
<i>P. macrocephalus</i>	4	2	0.002 [0.001]	0.007 [0.007]	0.010 [0.010]	0.000 [0.000]
<i>T. truncatus</i>	10	2	0.011 [0.002]	0.039 [0.017]	0.034 [0.026]	0.036 [0.039]
<b>Mean</b>	<b>4</b>	<b>5</b>	<b>0.004 [0.001]</b>	<b>0.041 [0.007]</b>	<b>0.048 [0.009]</b>	<b>0.015 [0.004]</b>
<b>Primates</b>						
<i>A. nancymae</i>	2	13	0.003 [0.001]	0.093 [0.014]	0.099 [0.023]	0.075 [0.024]
<i>G. gorilla</i>	3	3	0.008 [0.002]	0.124 [0.023]	0.132 [0.032]	0.102 [0.040]
<i>H. sapiens</i>	30	2	0.003 [0.001]	0.141 [0.031]	0.152 [0.040]	0.111 [0.053]
<i>M. arctoides</i>	7	3	0.009 [0.003]	0.098 [0.020]	0.094 [0.026]	0.109 [0.046]
<i>M. mulatta</i>	2	22	0.000 [0.000]	0.114 [0.015]	0.110 [0.024]	0.126 [0.025]
<i>P. pygmaeus</i>	2	7	0.079 [0.009]	0.103 [0.018]	0.085 [0.023]	0.162 [0.048]
<i>P. troglodytes</i>	13	8	0.018 [0.003]	0.069 [0.014]	0.060 [0.018]	0.098 [0.033]
<b>Mean</b>	<b>8</b>	<b>8</b>	<b>0.017 [0.011]</b>	<b>0.106 [0.009]</b>	<b>0.105 [0.011]</b>	<b>0.112 [0.010]</b>

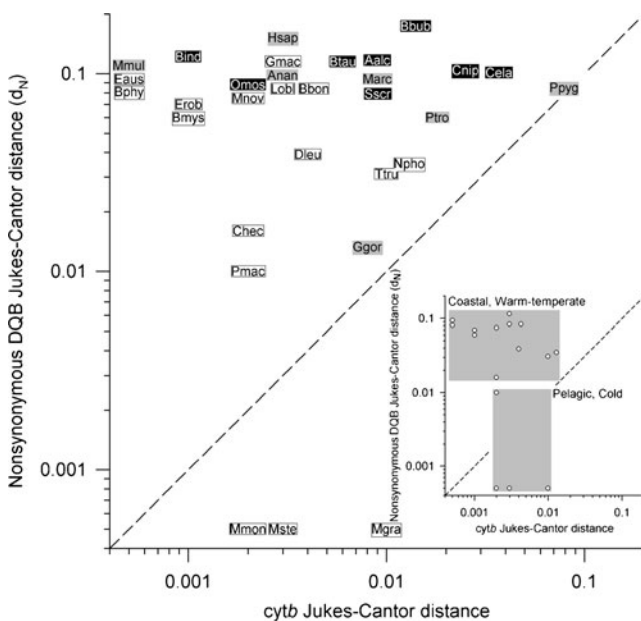
Bold numbers highlights mean values of Cetacea, Artiodactyla and Primates

distance (interval of  $p$  0.01–0.07). The difference between artiodactyls and primates was not significant either for the DQB distance or for the *cytb* distance (interval of  $p$  0.63–0.65).

The results above indicated that statistical comparisons of the rate of nonsynonymous substitution at the DQB between cetaceans and terrestrial mammals depend on differences between taxa affecting molecular clocks which are slower in cetaceans for the two genes examined. However, plotting of the *cytb* and DQB $d_N$  distances showed a pattern of variation noticeably different for cetaceans in which at least two groups of  $d_N$  are observed. A group of high values of DQB $d_N$  (>0.06) is formed by all mysticetes examined together with the short-finned pilot whale and the North Pacific white-sided dolphin (*Sagmatias obliquoidens*), while a group of low  $d_N$  values (<0.04) is formed by ziphiids, the sperm whale, monodontids, the Hector's dolphin (*Cephalorhynchus hectori*), the bottlenose dolphin, and the finless porpoise (Table 3; Fig. 5). This pattern is the same as the one observed in the analysis of phylogenetic branch lengths (Fig. 3). The plot of the *cytb* and DQB $d_N$  distances also suggests that in spite of the differential variation of the *cytb* distance, the DQB $d_N$  of cetaceans is, in general, lower than its values for artiodactyls and primates (Fig. 5).

#### Accumulation of variable sites

DQB sequences of terrestrial mammals showed a higher proportion of variable sites compared to cetacean sequences.



**Fig. 5** Jukes–Cantor distance of the *cytb* gene and the nonsynonymous DQB Jukes–Cantor distance ( $d_N$ ) within species of Artiodactyla (black rectangles), Cetacea (white rectangles) and Primates (gray rectangles) in Table 3. The dashed line is the diagonal. Species names are indicated with abbreviations as in Fig. 3. The inner graph is the same plot showing only cetaceans with distinction of two groups having different levels of DQB $d_N$  which correspond to different marine habitats

Cetaceans, which had the largest number of species analyzed in this investigation (17 species, 92 alleles), had on average less than 8 % of variable sites (13/171), while artiodactyls, which had fewer species and a similar amount of sequences analyzed (nine species, 110 alleles), had 25 % (43/171) of variable sites (Table 4). Both taxa had a comparable sampling of divergence degree as measured by the *cytb* (Fig. 2). Mean proportion of variable sites on *cytb* sequences was also lower in cetaceans (0.5 %) as compared with artiodactyls and primates (3 % and 2 %, respectively) and thus, the general proportion of variable sites at the *cytb* gene was lower than it is in the DQB. The Student's  $t$  test showed significant differences in the proportion of variable sites per allele and nucleotide between cetaceans and the two groups of terrestrial mammals. These differences were found for the DQB and the *cytb* sets ( $p < 0.02$ ). The proportion of variable sites per allele and nucleotide was not statistically different between artiodactyls and primates for the DQB ( $p = 0.50$ ) or for the *cytb* ( $p = 0.66$ ).

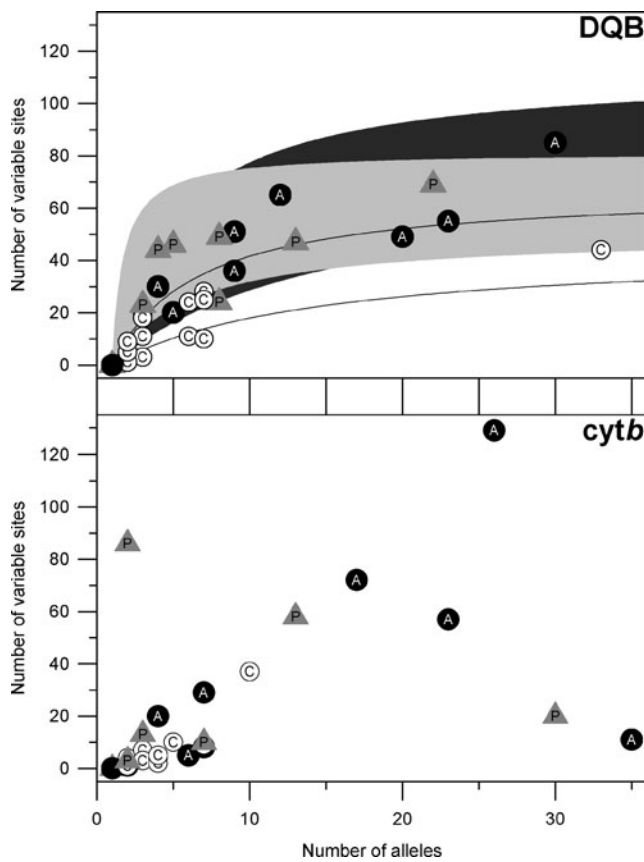
Estimation of the asymptotic number of variable sites at the DQB sequence yielded  $56 \pm 11$  (SE;  $r^2 = 0.79$ ) for cetaceans,  $93 \pm 20$  (SE;  $r^2 = 0.79$ ) for artiodactyls, and  $66 \pm 15$  (SE;  $r^2 = 0.72$ ) for primates (Fig. 6). Distribution of variable sites among codon positions in the *cytb* gene was similar in the three mammal orders occurring more in the third position and less in the second position. Variable sites in the DQB gene in both terrestrial mammal orders were almost equally distributed among the three codon positions, while cetaceans showed more occurrences of variable sites in the first position and less in the third position.

#### Discussion

All indicators of molecular evolution examined pointed to weaker balancing selection in MHC-DQB of cetaceans in comparison with artiodactyls and primates. The analyses on

**Table 4** Sampling of species, alleles, and variable sites for the *cytb* and DQB of the mammal orders examined. Parameters  $v_{max}$  and  $k$ , their standard errors (SE), and the fit regression index for the estimation of total variable sites at DQB are also shown

	Artiodactyla	Cetacea	Primates
Species	9	17	7
DQB alleles	110	92	58
<i>cytb</i> alleles	129	60	59
DQB variable sites (171 bp)	43	13	43
<i>cytb</i> variable sites (1,140 bp)	37	6	27
DQB $v_{max}$ [SE]	92.7 [19.9]	55.9 [11.5]	65.8 [15.4]
DQB $k$ [SE]	8.96 [4.84]	9.42 [3.58]	2.95 [2.22]
$r^2$	0.86	0.79	0.72



**Fig. 6** Accumulation of variable sites on DQB and *cytb* alleles sampled. Each point represent a species among the examined Artiodactyla (black circles, A), Cetacea (open circles, C), and Primates (gray triangles, P). The maximum and minimum accumulation profiles of a fitted hyperbolic function are plotted for the DQB using the function parameters  $v_{\max}$  and  $k$  and their standard errors, SE (Table 4;  $v_{\max} + S E_{v_{\max}}$  and  $k - S E_k$  for upper bound as well as  $v_{\max} - S E_{v_{\max}}$  and  $k + S E_k$  for lower bound). Fitted function intervals are dark gray for Artiodactyla, open space between black lines for Cetacea, and light gray for Primates

nonsynonymous substitutions and the accumulation of variable sites impeded statistical examination of selection on the DQB given the general lower rates of substitution in cetaceans. However, the general weaker balancing selection observed in this taxon seemed not related to differences in sampling of species and alleles among taxa. Analyses of the  $L_{DQB}/L_{cytb}$  and  $d_N/d_S$  ratios also showed a higher variation in the strength of balancing selection for the DQB in cetaceans which was unrelated to sampling differences. Because molecular clocks differ between taxa, phylogenetic approaches showed more clearly a different selection mode on the cetacean DQB than distance approaches.

We found apparent transorder sharing of an allele between the hippopotamus and a beaked whale with a sequence similarity of 0.94. Although the phylogenetic relationship between Cetacea and Hippopotamidae is not surprising given other molecular evidence (Gatesy 1997; Nikaido et al. 1999; Boissier et al. 2005), the preservation

of this allelic similarity for over 50 Mya (Thewissen et al. 2007) seems remarkable.

Variable degrees of polymorphism have been described for the different MHC genes in several mammal species (Amills et al. 1998; Baker et al. 2006), each having a particular role in the immune response to specific pathogens (Acevedo-Whitehouse and Cunningham 2006). Also, duplication of MHC genes allows the expression of alternative MHC proteins in one individual. In the California sea lion (*Zalophus californianus*), for example, the DRB is a family of eight loci, each of which exhibits limited variability but which may have different loci configurations among individuals (two to eight loci for a total of 40 configurations). This yields a highly variable set of possible DR proteins within and between individuals (Bowen et al. 2006). Duplication of DQB has been reported in some cetacean species (Baker et al. 2006; Xu et al. 2007) and is widely reported in artiodactyls showing, in some cases, high variation between duplicated loci (Sigurdardóttir et al. 1992; Amills et al. 1998). MHC duplication and all its multigenic features comprise a genotypic level of polymorphism different to the variation in nucleotide sequence that we have examined for the *DQB* gene. MHC function depends on sequence variation among individuals within loci, variation of loci configurations (genotypes) within and between individuals, and expression patterns of genes on cells. These three levels of genetic structure and function may be subjected to different modes and strengths of natural selection. Further research on polymorphism at different MHC loci as well as on the patterns of MHC expression on cells against different pathogens are needed to properly assess MHC evolution in marine and terrestrial environments (Acevedo-Whitehouse and Cunningham 2006).

The apparent low levels of polymorphism previously reported for some species of marine mammals does not occur for all species (Hoelzel et al. 1999; Baker et al. 2006; Xu et al. 2007). For example, at least 23 different DQB alleles have been reported in humpback whales (Baker et al. 2006). As is in our sampling, the ziphiids, the sperm whale, monodontids, and the finless porpoise formed a group with low DQB polymorphism, while mysticetes exhibited a relatively high DQB variation similar to that of terrestrial mammals which is concordant with a higher occurrence of transpecific polymorphisms. Different dolphin species appeared in the two groups; the bottlenose dolphin and the Hector's dolphin in the low DQB variation group and the short-finned pilot whale and the North Pacific white-sided dolphin in the high DQB variation group. The distinction of these two groups among cetaceans suggests that life differences between coastal and pelagic realms as well as between warm and cold environments might influence the selection regime on the DQB.

The attributes of epidemic processes in the marine environments, mostly high dispersal and openness (McCallum et

al. 2004), seem involved in mass mortalities and other stochastic phenomena which are relevant for the life history and demography of marine mammals (Harwood and Hall 1990). The stronger and less variable balancing selection found in terrestrial mammals might, thus, result from a lower stochasticity in their demography part of which could be related with domestication in artiodactyls of our sampling such as bovines, pigs, and even deer. However, the pattern of stronger and less variable balancing selection of the DQB occurs also in primates for which life history and demography are different. This suggests that the evolutionary pattern of DQB in artiodactyls and primates might be general for terrestrial mammals. The differences of epidemiology among marine and terrestrial environments seem to have the same basis on dispersal, habitat isolation and population openness (McCallum et al. 2004) than the contrasting rates of chromosomal evolution and speciation between cetaceans, artiodactyls, and primates (Árnason 1972; Bush et al. 1977). Therefore, the particular DQB evolution of cetaceans might be part of a general stochastic phenomenology of high dispersal in habitats at once highly continuous and highly contrasting.

In conclusion, our results show that DQB evolution in cetaceans follows the general pattern of balancing selection observed in other mammals (Takahata 1990). However, we found a clear indication of a general weaker balancing selection for the DQB of cetaceans and a larger variation on its strength. Such evolutionary variation for cetaceans also seems, in our results, associated to differences between marine habitats. Further comparative studies on local selective pressures occurring in subpopulations or restricted populations (e.g., Vassilakos et al. 2009) are needed to disentangle environmental, demographic, and phylogenetic influences on the evolution of mammal MHC and their consequences for conservation.

**Acknowledgments** We are grateful to all people who, at the labs and/or in the sea, have contributed to this work. In particular, we acknowledge the technical work of M. Dalebout, J. Murrell, M.R. Robles, and D. Steel as well as the review and advice from M.L. Fanjul, D. Heimeier, and M. Uribe, and the scholar orientation by G. Vilaclara. We appreciate the assistance of J. Zúñiga on statistical tests as well as the comments of two anonymous reviewers who greatly improved this article. Institutional, academic, legal, and funding supports were received from The Marsden Foundation, University of Auckland, Consejo Nacional de Ciencia y Tecnología, Facultad de Ciencias-Universidad Nacional Autónoma de México, Posgrado en Ciencias del Mar y Limnología-Universidad Nacional Autónoma de México, Secretaría del Medio Ambiente y Recursos Naturales, Instituto Nacional de Ecología, and Convention on International Trade in Endangered Species of Wild Fauna and Flora.

## References

- Acevedo-Whitehouse K, Cunningham AA (2006) Is MHC enough for understanding wildlife immunogenetics? *Trends Ecol Evol* 21(8):433–438
- Amills M, Ramiya V, Norimine J, Lewin HA (1998) The major histocompatibility complex of ruminants. *Rev Sci Tech Off Int Epiz* 17(1):108–120
- Apanius V, Penn D, Slev PR, Ruff LR, Potts WK (1997) The nature of selection on the major histocompatibility complex. *Crit Rev Immunol* 17:179–224
- Árnason Ú (1972) The role of chromosomal rearrangement in mammalian speciation with special reference to Cetacea and Pinnipedia. *Hereditas* 70:113–118
- Baker CS, Vant MD, Dalebout ML, Lento GM, O'Brien SJ, Yuhki N (2006) Diversity and duplication of DQB and DRB-like genes of the MHC in baleen whales (suborder: Mysticeti). *Immunogenetics* 58:283–296
- Beck S, Trowsdale J (1999) Sequence organization of the class II region of the human MHC. *Immunol Rev* 167:201–210
- Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL (2005) GenBank. *Nucleic Acids Res* 33:34–38
- 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
- Boissierie JR, Lihoreau F, Brunet M (2005) The position of Hippopotamidae within Cetartiodactyla. *Proc Natl Acad Sci USA* 102:1531–1537
- Bowen L, Aldridge BM, DeLong R, Melin S, Godinez C, Zavala A, Gulland F, Lowenstine L, Stott JL, Johnson ML (2006) MHC gene configuration variation in geographically disparate populations of California sea lions (*Zalophus californianus*). *Mol Ecol* 15:529–533
- Bush GL, Case SM, Wilson AC, Patton JL (1977) Rapid speciation and chromosomal evolution in mammals. *Proc Natl Acad Sci USA* 74:3942–3946
- Castro-Prieto A, Watcher B, Sommer S (2011) Cheetah paradigm revisited: MHC diversity in the world's largest free-ranging population. *Mol Biol Evol* 28(4):1455–1468
- Cutrer AP, Lacey EA (2007) Trans-species polymorphism and evidence of selection on class II MHC in tuco-tucos (Rodentia: Ctenomyidae). *Immunogenetics* 59(12):937–948
- Edwards SV, Hedrick PW (1998) Evolution ecology of MHC molecules: from genomics to sexual selection. *Trends Ecol Evol* 13(8):305–311
- Ellegren H, Hartaman G, Johansson M, Andersson L (1993) Major histocompatibility complex monomorphism low levels of DNA fingerprinting variability in a reintroduced rapidly expanding population of beavers. *Proc Natl Acad Sci USA* 90:8150–8153
- Gatesy J (1997) More DNA support for a Cetacea/Hippopotamidae clade: the blood-clotting protein gene  $\gamma$ -fibrinogen. *Mol Biol Evol* 14:537–543
- Goldsbey RA, Kindt TJ, Osborne BA, Kuby J (2003) *Immunology*, 5th edn. Freeman, New York
- Gu X, Nei M (1999) Locus specificity of polymorphic alleles evolution by a birth-and-death process in mammalian MHC genes. *Mol Biol Evol* 16(2):147–156
- Gutierrez-Espeleta G, Hedrick PW, Kalinowski ST, Garrigan D, Óbice WM (2001) Is the decline of the desert bighorn sheep from infectious disease the result of low MHC variation? *Heredity* 86:439–445
- Harwood J, Hall A (1990) Mass mortality in marine mammals: its implications for population dynamics genetics. *Trends Ecol Evol* 5(8):254–257
- Hassanin A, Douzery EJP (2003) Molecular and morphological phylogenies of Ruminantia and the alternative position of the Moschidae. *Syst Biol* 52(2):206–228
- Hayashi K, Nishida S, Yoshida H, Goto M, Pastene LA, Koike H (2003) Sequence variation of the DQB allele in the cetacean MHC. *Mamm Stud* 28(2):89–96
- Hoelzel AR, Claiborne Stephens J, O'Brien SJ (1999) Molecular genetic diversity: evolution at the MHC DQB locus in four species of pinnipeds. *Mol Biol Evol* 16(5):611–618

- Hughes AL, Nei M (1988) Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. *Nature* 335:167–170
- Hutchings K (2009) Parasite-mediated selection in an island endemic, the Seychelles warbler (*Acrocephalus sechellensis*). PhD thesis, University of East Anglia, London
- Irwin DM, Arnason Ú (1994) Cytochrome b gene of marine mammals: phylogeny and evolution. *J Mamm Evol* 2:37–55
- Jackson JA, Baker CS, Vant M, Steel DJ, Medrano-González L, Palumbi SR (2009) Big and slow: phylogenetic estimates of molecular evolution in baleen whales (Suborder Mysticeti). *Mol Biol Evol* 26(11):2427–2440
- Klein J, Figueroa F (1986) Evolution of the major histocompatibility complex. *Crit Rev Immunol* 6:295–386
- Kryazhimskiy S, Plotkin JB (2008) The population genetics of dN/dS. *PLoS Genet* 4(12):e1000304
- Kumánovics A, Takada T, Fischer Lindahl K (2003) Genomic organization of the mammalian MHC. *Annu Rev Immunol* 21:629–657
- Kumar S, Nei M, Dudley J, Tamura K (2008) MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. *Brief Bioinform* 9(4):299–306
- Maddison DR, Maddison WP (2000) MacClade 4: analysis of phylogeny and character evolution. Version 4.0. Sinauer, Sunderland
- McCallum HI, Kuris A, Harvell CD, Lafferty KD, Smith GW, Porter J (2004) Does terrestrialepidemiology apply to marine systems? *Trends Ecol Evol* 19(11):585–591
- Montgelard C, Catzeflis FM, Douzery E (1997) Phylogenetic relationships of artiodactyls and cetaceans as deduced from the comparison of cytochrome b and 12S rRNA mitochondrial sequences. *Mol Biol Evol* 14(5):550–559
- Munguía-Vega A, Esquer-Garrigós Y, Rojas-Bracho L, Vázquez-Juárez R, Castro-Prieto A, Flores-Ramírez S (2007) Genetic drift vs. natural selection in a long-term small isolated population: major histocompatibility complex class II variation in the Gulf of California endemic porpoise *Phocoena sinus*. *Mol Ecol* 16:4051–4065
- Murphy WJ, Elzirik E, Johnson WE, Zhang YP, Ryder OA, O'Brien SJ (2001) Molecular phylogenetics and the origins of placental mammals. *Nature* 409:614–618
- Murray BW, Malik S, White BN (1995) Sequence variation at the major histocompatibility complex locus DQβ in beluga whales (*Delphinapterus leucas*). *Mol Biol Evol* 12(4):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
- Nielsen R, Yang Z (2003) Estimating the distribution of selection coefficients from phylogenetic data with applications to mitochondrial and viral DNA. *Mol Biol Evol* 20(8):1231–1239
- Nigenda-Morales S, Flores-Ramírez S, Urbán-R J, Vázquez-Juárez R (2008) MHC *DQB-1* polymorphism in the Gulf of California fin whale (*Balaenoptera physalus*) population. *J Hered* 99(1):14–21
- Nikaido M, Rooney AP, Okada N (1999) Phylogenetic relationships among cetartiodactyls based on insertions of short and long interspaced elements: hippopotamuses are the closest extant relative of whales. *Proc Natl Acad Sci USA* 96:10261–10266
- O'Brien SJ, Roelke ME, Marker L, Newman A, Winkler CA, Meltzer D, Colly L, Evermann JF, Bush M, Wildt DE (1985) Genetic basis for species vulnerability in the cheetah. *Science* 227(4693):1428–1434
- 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* L.). *Proc Natl Acad Sci USA* 95:3714–3719
- Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14(9):817–818
- Potts WK, Wakeland EK (1990) Evolution of diversity at the major histocompatibility complex. *Trends Ecol Evol* 5:181–187
- Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574
- Schwaiger F-W, Weyers E, Buitkamp J, Ede AJ, Crawford A, Epplen JT (1994) Interdependent *MHC-DRB* exon-plus-intron evolution in artiodactyls. *Mol Biol Evol* 11(2):239–249
- Sigurdardóttir S, Borsch C, Gustafsson K, Andersson L (1992) Gene duplications and sequence polymorphism of bovine class II DQB genes. *Immunogenetics* 35:205–213
- Slade RW (1992) Limited MHC polymorphism in the southern elephant seal: implications for MHC evolution marine mammal population biology. *Proc R Soc Lond B* 249(1325):163–171
- Smith DM, Lunney JK, Ho CS, Martens GW, Ando A, Lee JH, Schook L, Renard C, Chardon P (2005) Nomenclature for factors of the swine leukocyte antigen class II system 2005. *Tissue Antigens* 66:623–639
- Swarbrick PA, Crawford AM (1997) The red deer (*Cervus elaphus*) contains two expressed major histocompatibility complex class II DQB genes. *Anim Genet* 28:49–51
- Swofford DL (2001) PAUP\*: phylogenetic analysis using parsimony (and other methods). Versión 4.0. Sinauer, Sunderland
- Takahashi K, Rooney AP, Nei M (2000) Origins and divergence times of mammalian class II MHC gene clusters. *J Hered* 91:198–204
- Takahata N (1990) A simple genealogical structure of strongly balanced allelic lines and trans-species evolution of polymorphism. *Proc Natl Acad Sci USA* 87:2419–2423
- Takahata N, Satta Y, Klein J (1992) Polymorphism and balancing selection at major histocompatibility complex loci. *Genetics* 130:925–938
- Thewissen JGM, Cooper LN, Clementz MT, Bajpai S, Tiwari BN (2007) Whales originated from aquatic artiodactyls in the Eocene epoch of India. *Nature* 450:1190–1194
- Tovo-Rodrigues L, Rhode LA, Roman T, Schmitz M, Polanczyk G, Zeni C, Marques FZC, Contini V, Grevet EH, Belmonte-de-Abreu P, Bau CHD, Hutz MH (2011) Is there a role for rare variants in *DRD4* gene in the susceptibility for ADHD? Searching for an effect of allelic heterogeneity. *Mol Psychiatry* 17(5):520–526
- Trowsdale J, Groves V, Arnason A (1989) Limited MHC polymorphism in whales. *Immunogenetics* 29:19–24
- Vassilakos D, Natoli A, Dahlheim M, Hoelzel AR (2009) Balancing and directional selection at exon-2 of the MHC DQB1 among populations of odontocete cetaceans. *Mol Biol Evol* 26(3):681–689
- Xu S, Sun P, Zhou K, Yang G (2007) Sequence variability at three MHC loci of finless porpoises (*Neophocaena phocaenoides*). *Immunogenetics* 59:581–592



## Capítulo 3

Variación nucleotídica y peptídica del exón 2 del gen *DQB* en mamíferos marinos y terrestres



### 3. 1 Introducción

La predicción precisa de la estructura secundaria a partir de las abundantes fuentes de secuencias de proteínas y genes que existen actualmente, permite entender cuestiones de diferentes ramas de la ciencia experimental y teórica, como es el diseño de medicinas, las bases genéticas de una enfermedad, el rol de la estructura proteica en funciones enzimáticas, etc. (Munson et al. 1994). Los modelos de estructura secundaria representan un intento para añadir realismo biológico al análisis de genes para productos de RNA que presenten una estructura secundaria. Estos modelos pueden explicar la dependencia entre los sitios de nucleótidos, donde la selección favorecerá sustituciones que no afecten la estabilidad de esta estructura (Lewis, 2001).

Los elementos fundamentales de la estructura secundaria de las proteínas con las  $\alpha$ - hélices, láminas  $\beta$ , espirales y giros, debido principalmente a los puentes de hidrógeno que se forman entre el esqueleto carbonil y los grupos NH. En promedio, las proteínas contienen aproximadamente 30% de estructura  $\alpha$  - hélices (H), cerca de 20% de láminas  $\beta$  (E) y alrededor de 50% de estructuras espirales (C). El espiral es la estructura más fácil de predecir, mientras que las láminas  $\beta$  son las más difíciles, dada su rareza relativa y su naturaleza irregular en zonas de contacto, contrario a las  $\alpha$ - hélices, donde las zonas de contacto se encuentran bien localizadas (Kloskowski et al. 2002).

#### **Método cuadrático – logístico para la predicción de la estructura secundaria de proteínas.**

Existen métodos como el GOR (Garnier – Osguthorpe – Robson) (Garnier et al. 1996) que predicen la estructura de secundaria de las proteínas a partir de secuencias de amino ácidos basándose en la teoría de información y redes neuronales. Así mismo, existen métodos estadísticos que obtienen resultados con una tasa de predicción comparable e incluso mayor a aquellos que utilizan la teoría de la información y redes neuronales.

El método cuadrático – logístico (QL method por su siglas en inglés) se basa en un enfoque puramente estadístico, viendo la situación como un problema de regresión o clasificación

con un número desconocido de variables. Este acercamiento es un híbrido de un modelo estrictamente paramétrico (logística lineal y cuadrática) con técnicas no – paramétricas (o modelo – libre). Este acercamiento permite la incorporación de variables de importancia conocida mientras permite el surgimiento de una estructura del modelo muy general a partir de los datos. En este método la complejidad es controlada ya sea por medio de la adición o remoción de términos en el modelo de regresión, o por medio del ajuste del “parámetro de suavizamiento” que efectivamente permite el continuo ajuste en el número de parámetros en el modelo. De esta manera se puede encontrar el modelo óptimo dentro de un gran universo de modelos complejos (Munson et al. 1994).

La predicción de la estructura secundaria permite inferir la inmunoreactividad de moléculas como las del MHC, por lo tanto su funcionalidad. En algunos trabajos se ha observado que al alterar la estructura secundaria de ciertas proteínas en condiciones de almacenamiento y preparación, el contenido de estructuras  $\alpha$  hélice muestra una relación inversa a la inmunoreactividad y por tanto a la disponibilidad de sitios antigénicos (Rousset et al. 1983).

### **Predicción de determinantes antigénicos proteicos a partir de secuencias de amino ácidos.**

El método de Hopp y Woods (1981) se utiliza para localizar determinantes antigénicos proteicos por medio de secuencias de amino ácidos de manera que se puedan reconocer el punto de máxima hidropatía local. Esta es logrado a través de un valor numérico a cada amino ácido (valor hidropático) y promediando repetitivamente estos valores a través de la cadena peptídica. El punto de máxima hidropatía local promedio se encuentra localizado adyacente a un determinante antigénico o es un sitio antigénico. La tasa de éxito de esta predicción depende del promedio del grupo. Los valores que se les asignan a cada amino ácido se muestran a continuación en el Cuadro 1.

Esta escala fue desarrollada para predecir sitios antigénicos potenciales en proteínas globulares, las cuáles serán probablemente ricas en residuos polares con carga. Esta escala

es esencialmente un índice de hidrofilia, con asignación de valores negativos a residuos no polares.

**Cuadro 1.** Valores de hidropatía de Hopp y Woods (1981).

<b>Amino ácido</b>	<b>s,*kcal/mol</b>	<b>Valor de hidropatía</b>
Arginina	3.0	3.0
Ácido aspártico	2.5	3.0
Ácido glutámico	2.5	3.0
Lisina	3.0	3.0
Serina	0.3	0.3
Asparagina	0.2	0.2
Glutamina	0.2	0.2
Glicina	0.0	0.0
Prolina	-1.4	0.0
Treonina	-0.4	-0.4
Alanina	-0.5	-0.5
Histidina	-0.5	-0.5
Cisteina	-1.0	-1.0
Metionina	-1.3	-1.3
Valina	-1.5	-1.5
Isoleucina	-1.8	-1.8
Leucina	-1.8	-1.8
Tirosina	-2.3	-2.3
Fenilalanina	-2.5	-2.5
Triptofano	-3.4	-3.4

\* Valores de los parámetros de los solventes asignados por Levitt

### 3.2 Método

Se obtuvieron un total de 291 secuencias de nucleótidos del exón-2 del gen *DQB* (171 pb), el cual codifica para PBR de los receptores de antígeno de las moléculas clase II del MHC y 294 secuencias de citocromo b (*cytb*) mitocondrial (1,140 pb) pertenecientes a 41 especies de mamíferos marinos y terrestres de los órdenes: Cetácea, Carnívora, Artiodactyla y Primates (Cuadro 2) a través del banco de genes: Genbank (Benson et al. 2005). Se generaron otras 21 secuencias del gen *DQB* provenientes de diez especies de cetáceos, una de pinnípedo y una de artiodáctilo para completar el juego de 41 especies para ambos marcadores genéticos. Estas secuencias se obtuvieron por amplificación por

PCR utilizando los oligonucleótidos iniciadores de Murray et al. (1995). Algunas amplificaciones se secuenciaron directamente (secuencias genómicas) y otros fueron clonados en células competentes de *Escherichia coli* DH5 $\alpha$  transformándolas con el plásmido vector pGEM-T al cual se ligó el producto de la PCR. Las células con el plásmido con el producto de la PCR inserto se identificaron y separaron por el método de las colonias blancas y azules. Las colonias bacterianas con el plásmido con el inserto se amplificaron separadamente en cultivo. De estos cultivos se extrajeron los plásmidos con los productos amplificados insertos y se secuenciaron utilizando oligonucleótidos asentados en el plásmido (secuencias clonadas) (Baker et al. 2006).

**Cuadro 2.** Especies de mamíferos marinos y terrestres incluidos en el análisis de estructura secundaria del gen *DQB*

<b>Cetacea</b>	<b>Carnivora*</b>	<b>Artiodactyla</b>	<b>Primates</b>
<b>Delphinidae</b>	<b>Otaridae</b>	<b>Hippopotamidae</b>	<b>Hominidae</b>
<i>C. hectori</i>	<i>A. forsteri</i>	<i>H. amphibius</i>	<i>H. sapiens</i>
<i>G. macrorhynchus</i>	<i>A. gazella</i>	<b>Bovidae</b>	<i>G. gorilla</i>
<i>S. obliquidens</i>	<i>Z. californianus</i>	<i>B. indicus</i>	<i>P. troglodytes</i>
<i>T. truncatus</i>	<b>Phocidae</b>	<i>B. taurus</i>	<i>P. pygmaeus</i>
<b>Monodontidae</b>	<i>M. angustirostris</i>	<i>B. bubalis</i>	<b>Cercopithecidae</b>
<i>D. leucas</i>	<i>M. leonina</i>	<i>O. moschatus</i>	<i>M. arctoides</i>
<i>M. monoceros</i>	<i>M. schauinslandi</i>	<b>Cervidae</b>	<i>M. mulatta</i>
<b>Phocoenidae</b>	<b>Canidae</b>	<i>C. elaphus</i>	<b>Aotidae</b>
<i>N. phocaenoides</i>	<i>C. familiaris</i>	<i>C. nippon</i>	<i>A. nancymaae</i>
<b>Balaenopteridae</b>	<i>U. cinereoargenteus</i>	<i>A. alces</i>	
<i>B. bonaerensis</i>	<i>U. littoralis</i>	<b>Suidae</b>	
<i>B. physalus</i>		<i>S. scrofa</i>	
<i>M. novaengliae</i>			
<b>Eschrichtidae</b>			
<i>E. robustus</i>			
Balaenidae			
<i>B. mysticetus</i>			
<i>E. australis</i>			
<b>Ziphiidae and Physteridae</b>			
<i>M. europeaus</i>			
<i>M. grayi</i>			
<i>M. stejnegeri</i>			
<i>P. catodon</i>			

\*El orden Carnivora se subdividió en dos grupos, carnívoros terrestres y carnívoros marinos para poder realizar comparaciones de acuerdo a su hábitat.

### **Método *ad hoc* para analizar estructura *DQB***

Con el fin de analizar la variación de las secuencias con un sentido funcional, las secuencias nucleotídicas del gen *DQB* y del gen *cytb* se tradujeron a las secuencias de aminoácidos respectivas usando el programa Mega 4.1 software (Kumar et al. 2008) el cual considera los códigos genéticos universal (para el gen *DQB*) y mitocondrial de los vertebrados (para el gen *cytb*). En todas las secuencias peptídicas se determinaron los perfiles de carga eléctrica e hidropatía en la escala de Hopp y Woods la cual mide la inmunorreactividad peptídica. Para ello, se utilizó el programa MultiHQ7 desarrollado por L. Medrano (1990, 2007 en prep). Ya que la inmunorreactividad de los péptidos se asocia a su estructura secundaria (Hopp y Woods, 1981; Rousset et al. 1983), se determinó también la estructura secundaria con el método QL (Munson et al. 1994) que se encuentra disponible en línea (<http://abs.cit.nih.gov/>). De forma paralela a cómo las distancias nucleotídicas se compararon con respecto a la a variación del gen *cytb* como indicador de variación neutral, la variación de la estructura peptídica se normalizó con respecto a la distancia promedio Jukes-Cantor de las secuencias *cytb* dentro y entre especies utilizando el programa VARSECHQ para hidropatía y carga y el programa VARSECES para estructura secundaria desarrollados *ad hoc* por L. Medrano (2009 en prep). Dado un alineamiento de valores de hidropatía o carga eléctrica ( $V$ ) de  $m$  secuencias de  $s$  especies con  $k$  sitios, para cada sitio  $x$  se obtuvo el promedio de todas las comparaciones pareadas de las secuencias  $i$  e  $j$  respecto a la distancia promedio dentro o entre las especies de las secuencia  $i$  e  $j$  ( $dsp_{ij}$ ), esto es:

$$R_x = \frac{1}{n_{ij}} \sum_{ij} \frac{|V_{ix} - V_{jx}|}{dsp_{ij}}$$

donde  $n_{ij}$  es el número de comparaciones pareada  $i$  e  $j$  y  $R_x$  es la tasa promedio de cambio de la variable  $V$  en el residuo  $x$  con respecto a la distancia J-C del gen *cytb* entre las especies a las que pertenecen las especies  $i$  e  $j$ . Es importante mencionar que muchas de las distancias intraespecíficas ( $dsp_{ii}$ ) no son cero. Para cada sitio  $x$ , se determinó también la desviación estándar del valor  $R_x$ .

A diferencia de la hidropatía y la carga eléctrica que consisten de un solo parámetro, la estructura secundaria se describe con los valores de probabilidad para hélice alfa (pa), lámina beta (pb) y giro (coil; pc) y para los cuales se cumple la condición  $pa+pb+pc=1$ . Asimismo, debe considerarse que la tasa de cambio de la estructura secundaria depende de la incertidumbre de la predicción la cual puede caracterizarse con la diversidad de las probabilidades pa, pb y pc. Una alta diversidad (homogeneidad) de las tres probabilidades implica una mayor incertidumbre en la predicción de la estructura secundaria y por lo tanto, debe ponderarse como una menor tasa de cambio. Para caracterizar la incertidumbre, se eligió a la diversidad de Shannon-Wiener (H) que se calculó como sigue:

$$H = -\sum_{i=a}^c p_i \ln_2(p_i)$$

La incertidumbre en la predicción de la estructura secundaria se definió como  $I=H/H_{\max}$  donde  $H_{\max}$  es la diversidad máxima posible ( $H_{\max}=\ln_2(3)$ ). La tasa promedio de cambio  $R_x$  en la estructura secundaria se determinó entonces como sigue:

$$R_x = \frac{1}{n_{ij}} \sum_{ij} \frac{|p_{ix} - p_{jx}|_a + |p_{ix} - p_{jx}|_b + |p_{ix} - p_{jx}|_c}{|i|_j \text{dsp}_{ij}}$$

Igual que en el caso de la hidropatía y la carga eléctrica, se determinó la desviación estándar de  $R_x$  en la estructura secundaria.

### 3.3 Resultados

El análisis de hidropatía (HW) y carga eléctrica de Hopp – Woods dio como resultado perfiles similares y consistentes en todos los grupos de mamíferos analizados (Figuras 4 y 5). Todos los grupos mostraron valores altos de desviación estándar (Cuadro 3). En promedio, la hidropatía fue significativamente diferente entre la mayoría de los grupos de

mamíferos con excepción de los carnívoros terrestres y ambos grupos de mamíferos marinos. Los mamíferos marinos mostraron los valores más altos de hidropatía (0.528 cetáceos y 0.478 carnívoros marinos) sin diferencias significativas entre ellos. La hidropatía de los mamíferos terrestres parece ser más variable, principalmente en artiodáctilos y los primates. Con respecto a la carga eléctrica, no se presentaron diferencias significativas entre los mamíferos y se mostraron patrones similares de variación. Los primates son el único grupo que resultó de manera significativa menos negativamente cargados que el resto de los grupos de mamíferos.

**Cuadro. 3** Promedio y desviación estándar (DE) y valores p de prueba t -student del análisis de hidropatía (HW) y datos de carga (Q) de Hopp –Woods

	Promedio (HW)	DE (HW)	Promedio (Q)	DE (Q)	
Cetáceos	0.528	0.645	-0.019	0.202	
Artiodáctilos	0.442	0.690	-0.033	0.223	
Carnívoros marinos	0.478	0.609	-0.015	0.225	
Carnívoros terrestres	0.390	0.628	-0.028	0.203	
Primates	0.368	0.661	-0.012	0.204	
<i>valores p</i>	C	A	CM	CT	P
Cetáceos		0.2470	0.3970	0.1260	0.5550
Artiodáctilos	<b>0.0150</b>		0.2040	0.2540	<b>0.0190</b>
Carnívoros marinos	0.0940	<b>0.0030</b>		0.2250	<b>0.0020</b>
Carnívoros terrestres	<b>0.0500</b>	<b>0.0220</b>	0.5020		<b>0.0080</b>
Primates	<b>0.0000</b>	<b>0.0130</b>	<b>0.0000</b>	<b>0.0000</b>	

C = Cetáceos, A = Artiodáctilos, CM = Carnívoros Marinos, CT = Carnívoros Terrestres, P = Primates. Valores de hidropatía (matriz inferior) y valores de carga (matriz superior); Los valores significativos ( $p \geq 0.05$ ) se encuentran en negritas.

La comparación de los perfiles de hidropatía entre una molécula conservada como el *cytb* y una molécula altamente variable como el *DQB* del MHC, nos permitió observar que las zonas hidropáticas de encuentran muy conservadas en el *cytb* (Figuras 4 y 5). La predicción de la estructura secundaria del *cytb* no reveló cambios estructurales significativos entre los mamíferos para este marcador. La variación de dicha probabilidad fue muy poca entre secuencias, indicando que las predicciones fueron congruentes entre todos los grupos y



entre las mismas secuencias. La confiabilidad de dichas predicciones fue alta, principalmente en dos zonas hidropáticas localizadas entre los sitios 200 – 300 del *cytb*. La baja confiabilidad que se observa en las predicciones que se encuentran entre el sitio 75 – 130 en los carnívoros terrestres se le puede atribuir a que en algunas secuencias que se tienen para una de las dos especies de este grupo, los zorros grises, ésta zona no se encuentra bien secuenciada e incluso incompleta.

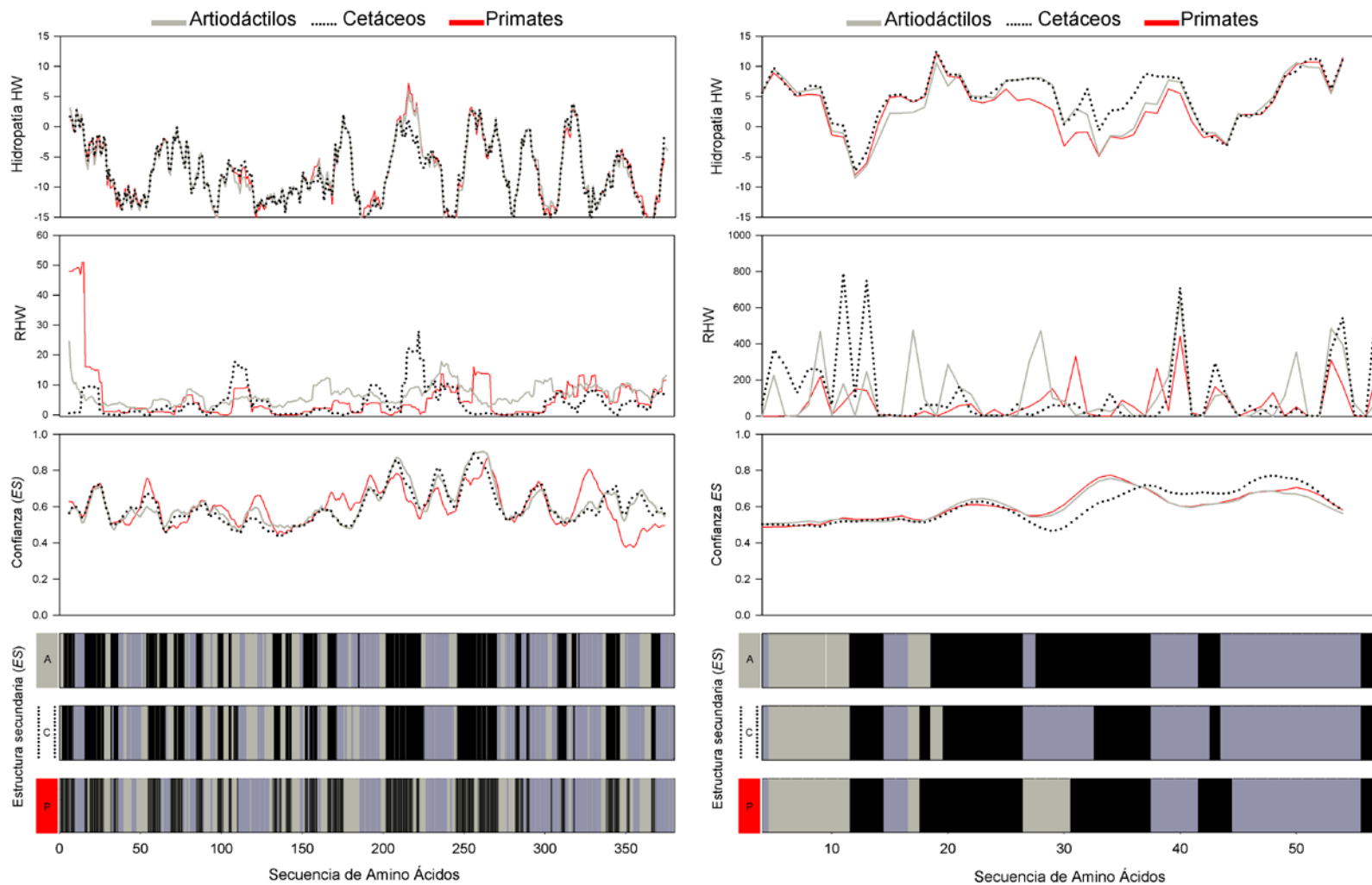
Es notable la diferencia entre los perfiles de hidropatía que existen entre ambos marcadores, ya que el *cytb* presenta zonas con niveles de hidropatía muy bajos, mientras que el gen *DQB* parece una molécula prácticamente hidropática en todas sus zonas.

Los perfiles de hidropatía del gen *DQB* son similares para todos los grupos de mamíferos analizados, sin embargo se observa que los primates presentan zonas menos hidropáticas que el resto de los grupos. Las zonas de alta hidropatía para este marcador corresponden a zonas con predicciones de estructura de hoja beta. Se puede observar una zona entre los sitios 27 – 32 donde existe mayor variación en las predicciones y que presentan altos valores de confiabilidad, esto puede indicar que para esta zona, la estructura es particular para cada uno de los diferentes grupos, aunque se debe de realizar un análisis más profundo para poder llegar a aseverar dicha sugerencia. Los carnívoros terrestres mostraron datos extremadamente variables de hidropatía a pesar de que la distancia genética fue muy corta, ya que sólo se constaba de dos especies diferentes, lo que pudiera indicar que para este grupo la hidropatía es un factor de variabilidad en esta molécula. Cabe destacar que las secuencias de amino ácidos de carnívoros marinos y terrestres presentaron sólo 45 – 50 residuos al traducirse, ya que las secuencias que se obtuvieron del banco de genes para dichos grupos únicamente se tenían secuenciadas alrededor de 100 pb de cada secuencia.

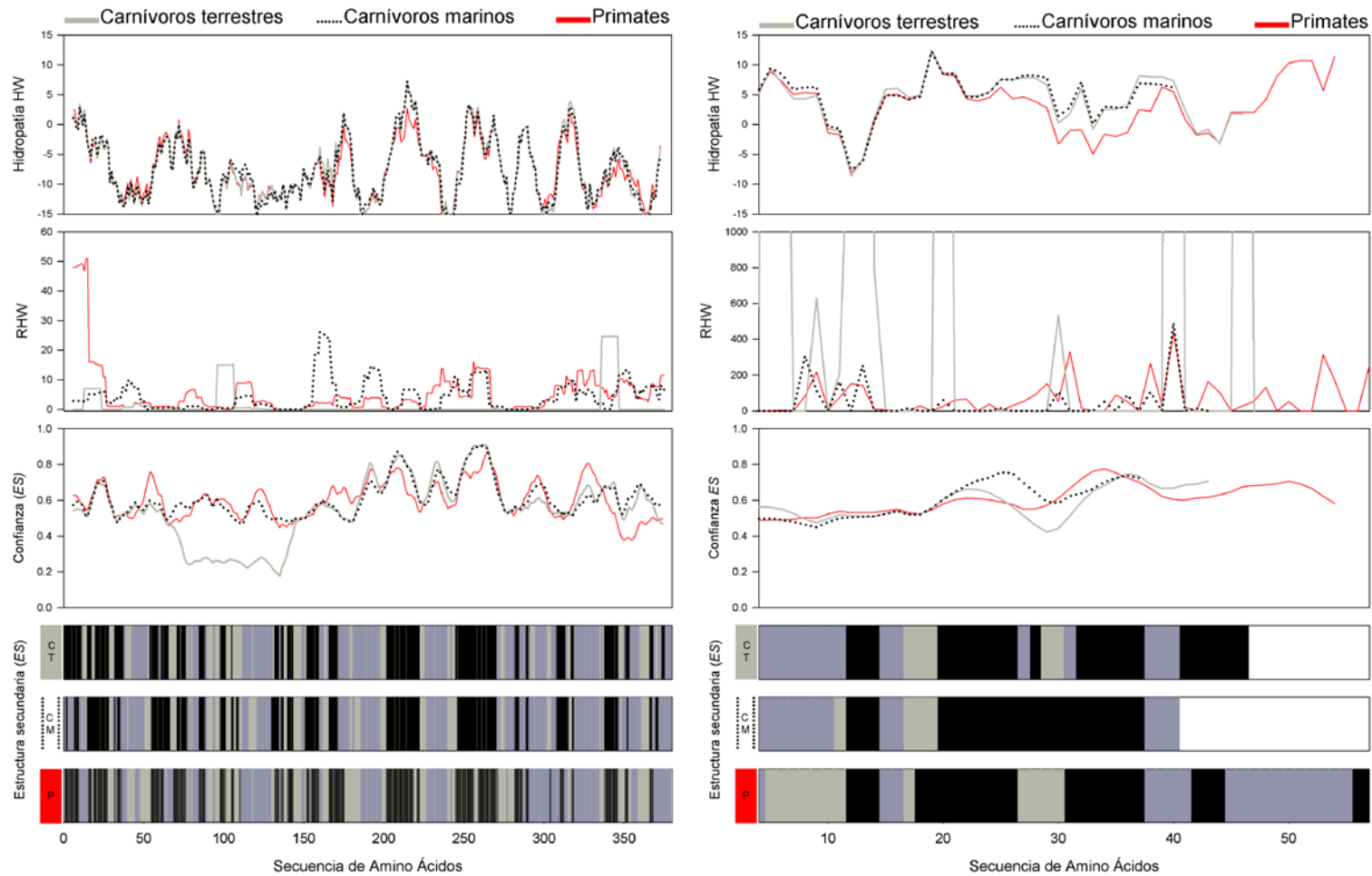
Analizamos la conservación de la estructura secundaria y la hidropatía (Rd) de ambos marcadores (Figuras 6 y 7 respectivamente) a partir de las distancias genéticas que obtuvimos del *cytb* que colocan a las secuencias en un contexto evolutivo de manera que podemos discernir si la variación observada en las moléculas es producto de un proceso evolutivo distinto o es más bien un efecto de la distancia acumulada a lo largo de la

distancia genética (VarSec.pas y VarHQ.pas, Medrano en prep.). Contrastamos valores de  $Rd$  de primates (grupo externo) contra el resto de los grupos de mamíferos y se contabilizaron por medio de un método de cuadrantes (Zonard.pas, Medrano en prep.) que nos permite observar la acumulación de datos en ciertas zonas para comparar el comportamiento de la variación en la estructura secundaria y en la hidropatía de los grupos de manera que se pudiera inferir si existía algún patrón que fuera exclusivamente marino o terrestre.

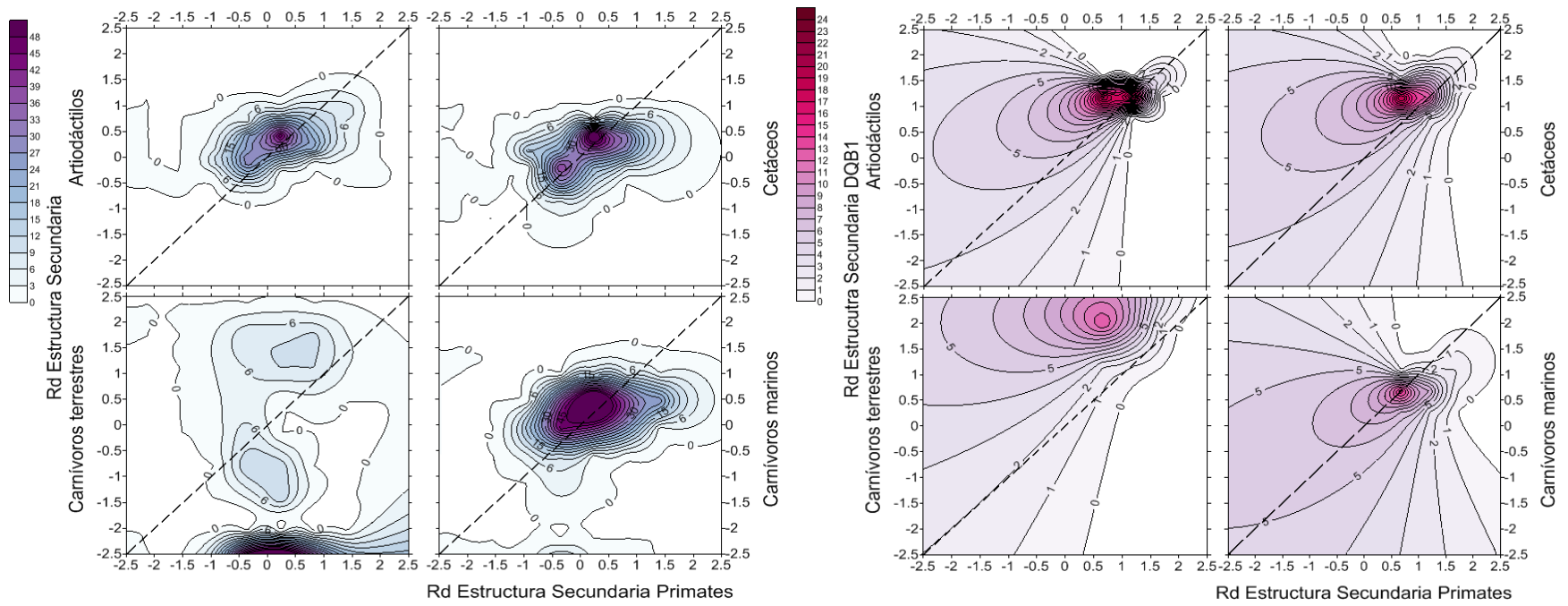
Sin embargo, los resultados de conservación en la estructura secundaria muestran que no existe un patrón distintivo entre mamíferos marinos y terrestres, en el caso de hidropatía parece haber una mayor variación en los mamíferos marinos, ya que parece que hay sitios que son más conservados que otros. En el caso de la estructura secundaria podemos observar que el gen *DQB* presenta una acumulación de datos menos restringida que el *cytb*, donde parece que hay sitios que presentan mayor variación que otros, mientras que en el *cytb*, podemos observar que los sitios tienen una variación similar entre los sitios.



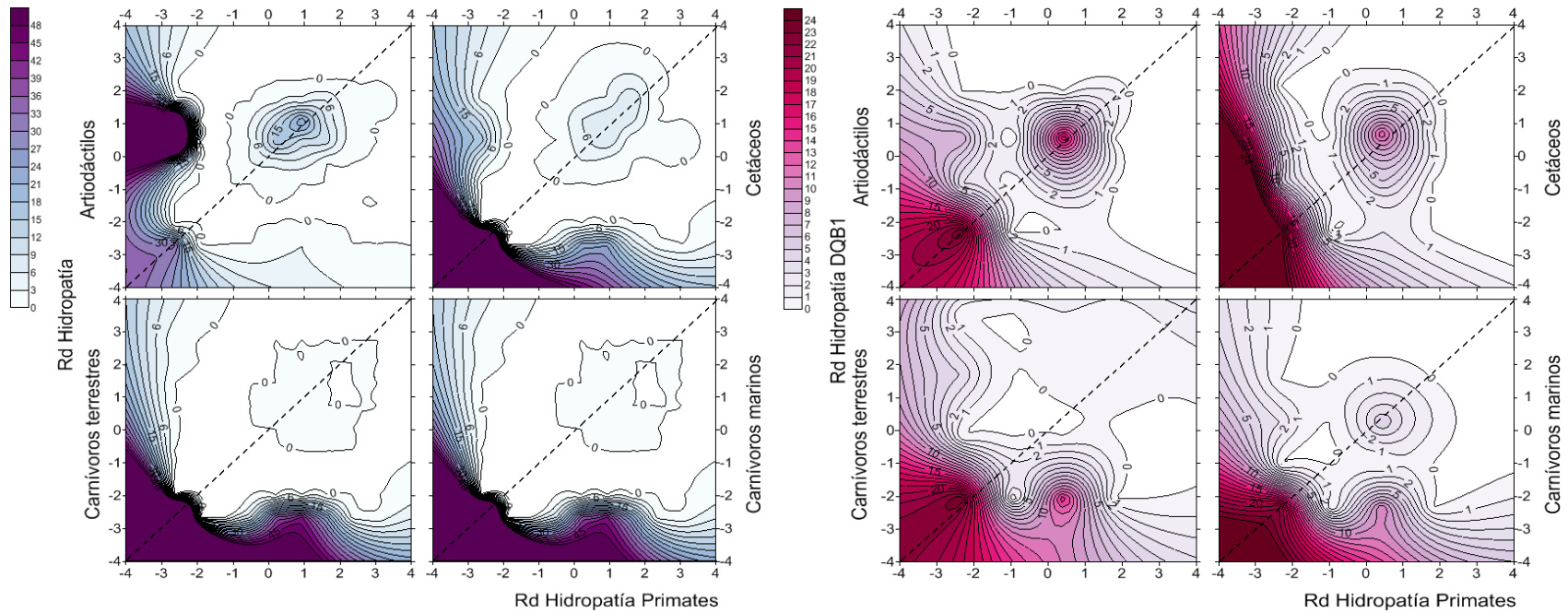
**Fig. 4** Perfiles de hidropatía (*HW*) de Hopp y Woods, variación de la hidropatía (*RHW*), confianza de la predicción de estructura secundaria (*ES*) y estructura secundaria promedio (gráficas de barras) del gen *cytb* (conjunto de gráficas a la izquierda) y gen *DQB* (conjuntos de gráficas a la derecha) en artiodáctilos, cetáceos y primates calculado con MultiHQ7 (Medrano – González, en prep.). *A* = artiodáctilos, *C* = cetáceos y *E* = Primates. Barras negras = predicción de giros; barras gris oscuro = predicción de hélices alfa; y barras gris claro = predicción de hojas beta.



**Fig. 5** Perfiles de hidropatía (*HW*) de Hopp y Woods, variación de la hidropatía (*RHW*), confianza de la predicción de estructura secundaria (*ES*) y estructura secundaria promedio (gráficas de barras) del gen *cytb* (conjunto de gráficas a la izquierda) y gen *DQB* (conjuntos de gráficas a la derecha) en carnívoros marinos y terrestres así como primates calculado con MultiHQ7 (Medrano – González, en prep.). *A* = artiodáctilos, *C* = cetáceos y *E* = Primates. Barras negras = predicción de giros; barras gris oscuro = predicción de hélices alfa; y barras gris claro = predicción de hojas beta. Las áreas en blanco indican que no existen datos para este grupo



**Fig. 6** Conservación de la estructura secundaria (RdES) en cada grupo de mamíferos. La acumulación de datos representa zonas con cierta variabilidad en la estructura secundaria de la proteína en cada uno de los sitios a lo largo de la cadena peptídica del gen *cytb* (conjunto de gráficas a la izquierda) y del gen *DQB* (conjunto de gráficas a la derecha). NOTA. Cuando la mayor densidad de datos se concentra en un lado de la línea punteada, indica que la variabilidad es mayor en alguno de los dos grupos que se contrastan en cada gráfica. Los primates siempre se colocaron siempre en el eje de las x actuando como grupo externo en la comparación.



**Fig. 7** Conservación de la hidropatía (RdH) en cada grupo de mamíferos. La acumulación de datos representa zonas con cierta variabilidad en la hidropatía en cada uno de los sitios a lo largo de la cadena peptídica del gen *cytb* (conjunto de gráficas a la izquierda) y del gen *DQB* (conjunto de gráficas a la derecha). NOTA. Cuando la mayor densidad de datos se concentra en un lado de la línea punteada, indica que la variabilidad es mayor en alguno de los dos grupos que se contrastan en cada gráfica. Los primates siempre se colocaron siempre en el eje de las x actuando como grupo externo en la comparación.

### 3.4 Discusión

Las diferencias significativas en la hidropatía de las secuencias de aminoácidos de los cetáceos podría sugerir que, a pesar de que algunas especies dentro de este grupo presenten menores niveles de polimorfismo, la funcionalidad de la estructura secundaria se conserva (Vassilakos et al. 2009). Esto sugiere que los bajos niveles de polimorfismo en loci de MHC no se traducen necesariamente en una menor capacidad en la respuesta inmune, sino que pareciera que los mamíferos marinos, a pesar de presentar un menor polimorfismo, llegan a traducir un PBR altamente hidropático y funcional, suficiente para responder a las amenazas al sistema inmune en el medio marino. Asimismo, la ausencia de cambios estructurales significativos entre las moléculas resultantes del gen *DQB* entre mamíferos terrestres y marinos nos lleva a pensar que es posible que la baja variabilidad en este complejo, al menos en este gen, no pone en peligro la respuesta inmune, ya que el resultado es una molécula con rasgos estructurales y funcionales similares. Incluso pareciera que en el caso de los mamíferos marinos existe una mayor variación en términos funcionales de hidropatía donde la variabilidad es mayor a pesar de que el número de alelos sea menor que en el resto de los grupos. Los carnívoros terrestres que en su mayoría son cánidos, presentan una variación mínima a pesar que presentan hasta 20 alelos, en el caso de los perros domésticos donde la estructura secundaria resultante y la hidropatía para cada alelo de gen *DQB* es prácticamente la misma.

La investigación a nivel funcional de la existencia de altos niveles de polimorfismo con una mayor adecuación en los individuos o especies de algún grupo en particular, es crucial para poder entender los mecanismos reales de la mantención de los asombrosos niveles de polimorfismo que exhiben algunas especies. Establecer el vínculo entre la supervivencia de organismos que presentan ciertas combinaciones de alelos de MHC es el siguiente peldaño que se debe escalar. De no poder establecer dicho vínculo, no se puede asegurar fehacientemente que la evolución del MHC sea únicamente producto de un efecto selectivo, descartando la posibilidad de que sea un proceso meramente demográfico que sigue patrones de permanencia de alelos dependiente del régimen de mortalidad particular de cada especie.

# Capítulo 4

Conclusiones generales y recomendaciones para futuras investigaciones



There between A + B. various  
type of relation. C + B. The  
first predation, B + D  
rather greater distinction  
then genus would be  
formed. - binary relation



## 4.1 Conclusiones generales y recomendaciones para futuras investigaciones

La presente investigación muestra algunos rasgos de la evolución del exón-2 del gen *DQB* del MHC en los mamíferos marinos y en los mamíferos terrestres en relación con las diferencias epidemiológicas y poblacionales entre ambos medios. Se encuentra que la visión tradicional de menor variación del gen *DQB* en el medio marino asociada a una menor presión de selección por patógenos es inexacta. Los principales puntos que se pueden concluir sobre esta investigación son:

1. Nuestros resultados muestran que la evolución del *DQB* en cetáceos sigue el patrón general de selección balanceadora observada en los mamíferos. Sin embargo, encontramos un claro indicio de una selección balanceadora en el *DQB* más débil y con una mayor variación en su fuerza para los cetáceos. Dicha variación en los cetáceos parece, en nuestros resultados, estar asociada a diferencias entre los hábitats marinos.
2. Los análisis de sustitución sinónimas/no-sinónimas y de acumulación de sitios variables impidió examinar estadísticamente la selección en el *DQB* debido a las menores tasas de sustitución encontradas en los cetáceos. Los indicadores filogenéticos permitieron probar estadísticamente las diferencias en la evolución del gen *DQB* entre los mamíferos marinos y terrestres.
3. Los análisis de las proporciones  $L_{DQB}/L_{cytb}$  y  $d_N/d_S$  mostraron una alta variación en la fuerza de selección balanceadora en el gen *DQB* entre las especies de cetáceos.
4. Se encontró evidencia de un polimorfismo transespecífico entre un hipopótamo y un zifio con una similitud en secuencias de 0.94. Esta es una retención notable que se ha conservado a través de 50 millones de años.

5. Los zifios, cachalotes, monodóntidos y la marsopa sin aleta formaron un grupo con bajos niveles de polimorfismo en el gen *DQB* mientras que los misticetos mostraron una alta variación similar a la observada en los mamíferos terrestres. El mayor polimorfismo en los misticetos y mamíferos terrestres coincide con una mayor presencia de polimorfismos transespecíficos en estos grupos.
6. Las especies de delfines se dividieron en los dos grupos: las toninas y el delfín de Héctor se encontraron en el grupo con baja variación en el *DQB* mientras que el calderón de aletas cortas y el delfín de costados blancos de Pacífico Norte se localizaron en el grupo de alta variación.
7. Esta conformación de grupos en los cetáceos sugiere que las diferencias entre vivir en ámbitos costeros o pelágicos, así como entre vivir en ambientes cálidos o fríos, pueden influenciar el régimen de selección en el *DQB*.
8. La selección balanceadora más fuerte y menos variable encontrada en los mamíferos terrestres estudiados puede resultar de una menor estocasticidad en su demografía lo cual podría estar relacionado con la domesticación de los artiodáctilos de nuestro muestreo (bovinos, cerdos e incluso venados). Sin embargo, el mismo patrón se observa en los primates sugiriendo que el patrón evolutivo encontrado para primates y artiodáctilos podría ser generalizado para los mamíferos terrestres.
9. El análisis de hidropatía (HW) y carga eléctrica de Hopp – Woods dio como resultado perfiles similares y consistentes en todos los grupos de mamíferos analizados. En general, los cetáceos muestran niveles de hidropatía significativamente mayores con respecto a los mamíferos terrestres.
10. Las zonas de alta hidropatía en el exón 2 del gen *DQB* corresponden a zonas con predicciones de estructura de hoja beta las cuales corresponden a su vez al surco de unión a antígenos.

11. La comparación entre los perfiles de hidropatía de los genes *cytb* y *DQB*, muestra zonas hidropáticas muy conservadas en el *cytb*.
12. El *cytb* presenta zonas con niveles de hidropatía bajos mientras que el segmento codificado por el exón 2 del gen *DQB* es casi todo hidropático.
13. No se observan cambios estructurales significativos entre las moléculas resultantes del gen *DQB* entre mamíferos marinos y terrestres. Esto sugiere que la variación del gen *DQB* está sujeta a un principio de conservación estructural y funcional que no parece ser diferente en medios ambientes distintos.
14. La metodología utilizada en este trabajo, representa una herramienta innovadora para el análisis de la evolución del MHC así como de otros genes sujetos a selección balanceadora. A nivel funcional (proteínas), es crucial encontrar el vínculo entre los altos niveles de polimorfismo y la adecuación (supervivencia) de los individuos de una especie al poder responder favorablemente ante las presiones inmunológicas de su entorno. La aplicación de esta metodología podría resultar en una evaluación más profunda de las presiones selectivas que inciden sobre este complejo de genes, tanto acotada a sub-poblaciones o poblaciones restringidas, como aplicada a un espectro más amplio de genes del MHC. Un enfoque poblacional permitirá discernir con mayor claridad las influencias ambientales, demográficas y filogenéticas que actúan sobre la evolución del MHC de los mamíferos así como las implicaciones que éstas tienen sobre la conservación de estos animales.

## Literatura citada

- Acevedo-Whitehouse K, Cunningham AA (2006) Is MHC enough for understanding wildlife immunogenetics? *Trends Ecol Evol* 21 (8):433-438
- Amills M, Ramiya V, Norimine J, Lewin HA (1998) The major histocompatibility complex of ruminants. *Rev Sci Tech Off Int Epiz* 17(1):108-120
- Amos W, Balmford A (2001) When does conservation genetics matter? *Heredity* 87:257-265
- Apanius V, Penn D, Slev PR, Ruff LR, Potts WK (1997) The nature of selection on the major histocompatibility complex. *Crit Rev Immunol* 17:179-224
- Arkuch KD, Giese AR, Mendonca HL, McBride AM, Marty GD, Hedrick PW (2002) Resistance to three pathogens in the endangered winter – run Chinook salmon (*Oncorhynchus tshawytscha*): effects of inbreeding and major histocompatibility complex genotypes. *Can J Fish Aquat Sci* 59:966-975
- Árnason Ú (1972) The role of chromosomal rearrangement in mammalian speciation with special reference to Cetacea and Pinnipedia. *Hereditas* 70:113-118
- Árnason Ú, Gullberg A (1994) Cytochrome b Nucleotide Sequences and the Identification of Five Primary Lineages of Extant Cetaceans. *Mol Biol Evol* 13:407-417
- Árnason Ú, Gullberg A, Janke A (2004) Mitogenomic analyses provide new insights into cetacean origin and evolution. *Gene* 333: 27-34
- Baker CS, Vant MD, Dalebout ML, Lento GM, O'Brien SJ, Yuhki N (2006) Diversity and duplication of *DQB* and *DRB* like genes of the MHC in baleen whales (suborder: Mysticeti). *Immunogenetics* 58:283-296
- Beck S, Trowsdale J (1999) Sequence organization of the class II region of the human MHC. *Immunol Rev* 167:201-210
- Bengston JL, Boveng P, Franzén U, Have P, Heide-Jørgensen MP, Härkönen TL (1991) Antibodies to Canine Distemper Virus in Antarctic Seals. *Mar Mammal Sci* 7(1):85-87
- Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL (2005) GenBank. *Nucleic Acids Res* 33:34-38
- Benton MJ (2005) *Vertebrate Paleontology*, 2nd edn. Blackwell Science Ltd
- 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

- Berta A, Sumich JL, Kovacs KM (2006) *Marine Mammals: Evolutionary Biology* 2nd edn. Academic Press, San Diego
- Birkun A, Kuiken T, Krivokhizhin S, Haines DM, Osterhaus ADME, van de Bildt MWG, Joiris CR, Siebert U (1999) Epizootic of morbilliviral disease in common dolphins (*Delphinus delphis ponticus*) from the Black Sea. *Vet Rec* 144:85-92
- Boisserie JR, Lihoreau F, Brunet M (2005) The position of Hippopotamidae within Cetartiodactyla. *Proc Natl Acad Sci USA* 102:1531-1537
- Bollmer JL, Vargas FH, Parker PG (2007) Low MHC variation in the endangered Galápagos penguin (*Spheniscus mendiculus*). *Immunogenetics* 59:593-602
- Bossart GD, Baden DG, Ewing RY, Roberts B, Wright SD (1998) Brevetoxicosis in Manatees (*Trichechus manatus latirostris*) from the 1996 Epizootic: Gross, Histologic and Immunohistochemical Features. *Toxicol Pathol* 26(2):276-282
- Bowen L, Aldridge BM, DeLong R, Melin S, Godinez C, Zavala A, Gulland F, Lowenstine L, Stott JL, Johnson ML (2006) MHC gene configuration variation in geographically disparate populations of California sea lions (*Zalophus californianus*). *Mol Ecol* 15:529-533
- Bush GL, Case SM, Wilson AC, Patton JL (1977) Rapid speciation and chromosomal evolution in mammals. *Proc Natl Acad Sci USA* 74:3942-3946
- Castro-Prieto A, Watcher B, Sommer S (2011) Cheetah paradigm revisited: MHC diversity in the world's largest free-ranging population. *Mol Biol Evol* 28(4):1455-1468
- Čemerski S, Shaw A (2006) Immune synapses in T-cell activation. *Curr Opin Immunol* 18:298-304
- CloECKAERT A, Grayon M, Grépinet O, Boumedine KS (2003) Classification of Brucella strains isolated from marine mammals by infrequent restriction site-PCR and development of specific PCR identification tests. *Microbes Infect* 5(7):593-602
- CutrerA AP, Lacey EA (2007) Trans-species polymorphism and evidence of selection on class II MHC in tuco-tucos (Rodentia: Ctenomyidae). *Immunogenetics* 59(12):937-948
- Dobzhansky TH (1971) *Genetics of the Evolutionary Process*. Columbia University Press, New York
- Domingo M, Ferrer L, Pumarola M, Marco A, Plana J, Kennedy S, Mcaliskey M, Rima BK. (1990) Morbillivirus in dolphins. *Nature (London)* 348(6296):21
- Edwards SV, Hedrick PW (1998) Evolution ecology of MHC molecules: from genomics to sexual selection. *Trends Ecol Evol* 13 (8):305-311
- Ellegren H, Hartaman G, Johansson M, Andersson L (1993) Major histocompatibility complex monomorphism low levels of DNA fingerprinting variability in a reintroduced rapidly expanding population of beavers. *Proc Natl Acad Sci USA* 90:8150-8153

- Forney KA, Barlow J, Muto MM, Lowry M, Baker J, Cameron G, Mobley J, Stinchcomb C, Carretta JV (2000) U.S. Pacific Marine Mammal Stock Assessments: 2000. NOAA-TM-NMFS-SWFSC-300
- Garnier J, Gibrat JF, Robson B (1996) GOR method for predicting proteins secondary structure from amino acid sequence. *Method Enzymol* 266:540-553
- Gatesy J (1997) More DNA support for a Cetacea/Hippopotamidae clade: the blood-clotting protein gene  $\gamma$ -fibrinogen. *Mol Biol Evol* 14:537-543
- Gaydos JK., Balcomb III KC, Osborne RW, Dierauf L (2004) Evaluating potential infectious disease threats for southern resident killer whales, *Orcinus orca*: a model for endangered species. *Biol Conserv* 117:253-262
- Geraci JR, Harwood J, Lounsbury VJ (1999) Marine Mammal Die – Offs: Causes, Investigations and Issues En: Conservation and Management of Marine Mammals. Twiss JR, Reeves RR (eds) Smithsonian Institute, US pp 367-396
- Geraci JR, Lounsbury VJ (2009) Health En: Encyclopedia of Marine Mammals. Perrin WF, Würsig B, Thewissen JGM (eds) 2nd edn. Elsevier Press, San Diego pp 562-569
- Goldsby RA, Kindt TJ, Osborne BA, Kuby J (2003) Immunology, 5<sup>th</sup> edn. Freeman, New York
- Grachev MA, Kumarev VP, Mamaev LV, Zorin VL, Baranova LV, Denikina NN, Belikov SI, Petrov EA, Kolesnik VS, Kolesnik RS, Dorofeev VM, Beim AM, Kudelin VN, Nagieva FG, Sidorov VN (1989) Distemper virus in Baikal seals. *Nature* 338:209
- Gu X, Nei M (1999) Locus specificity of polymorphic alleles evolution by a birth-and-death process in mammalian MHC genes. *Mol Biol Evol* 16(2):147-156
- Gulland FM, Koski M, Lowenstine LJ, Colagross A, Morgan L, Spraker T (1996) Leptospirosis in California Sea Lions (*Zalophus californianus*) stranded along the Central California Coast, 1981-1994. *J Wildlife Dis* 32(4): 572-580
- Gutiérrez-Espeleta G, Hedrick PW, Kalinowski ST, Garrigan D, Óbice WM (2001) Is the decline of the desert bighorn sheep from infectious disease the result of low MHC variation? *Heredity* 86:439-445
- Harwood J, Hall A (1990) Mass mortality in marine mammals: its implications for population dynamics genetics. *Trends Ecol Evol* 5(8):254-257
- Hassanin A, Douzery EJP (2003) Molecular and morphological phylogenies of Ruminantia and the alternative position of the Moschidae. *Syst Biol* 52(2):206-228
- Hayashi K, Nishida S, Yoshida H, Goto M, Pastene LA, Koike H (2003) Sequence variation of the *DQB* allele in the cetacean MHC. *Mamm Stud* 28(2):89-96

- Hedrick PW, Kim TJ, Parker KM (2001) Parasite resistance and genetic variation in the endangered Gila topminnow. *Anim Cons* 4(2):103-109
- Hoelzel AR (2002) *Marine mammal biology: An evolutionary approach* 1st edn. Blackwell Pub, Oxford
- Hoelzel AR, Claiborne Stephens J, O'brien SJ (1999) Molecular genetic diversity: evolution at the MHC *DQB* locus in four species of pinnipeds. *Mol Biol Evol* 16(5):611-618
- Hopp TP, Woods KR (1981) Prediction of protein in antigenic determinants from amino acid sequences. *Proc Natl Acad Sci USA* 78(6):3824-3828
- Hughes AL, Nei M (1988) Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. *Nature* 335:167-170
- Hutchings K (2009) Parasite-mediated selection in an island endemic, the Seychelles warbler (*Acrocephalus sechellensis*). PhD thesis, University of East Anglia, London
- Irwin DM, Árnason Ú (1994) Cytochrome b gene of marine mammals: phylogeny and evolution. *J Mamm Evol* 2:37-55
- IUCN (2012) The IUCN Red List of Threatened Species. Version 2012.1 <<http://www.iucnredlist.org>> Downloaded on 19 June 2012
- Jackson JA, Baker CS, Vant M, Steel DJ, Medrano-González L, Palumbi SR (2009) Big and slow: phylogenetic estimates of molecular evolution in baleen whales (Suborder Mysticeti). *Mol Biol Evol* 26(11):2427-2440
- Jefferson TA, Webber MA, Pitman RL (2008) *Marine Mammals of the World: A Comprehensive Guide to their Identification* 1st edn. Academic Press/Elsevier
- Kennedy S, Kuiken T, Jepson PD, Deaville R, Forsyth M, Barrett T, van de Bildt MWG, Osterhaus ADME, Eybatov T, Duck C, Kydyrmanov A, Mitrofanov I, Wilson S (2000) Mass die-off of Caspian Seals Caused by Canine Distemper Virus. *Emerg Infect Dis* 6(6):637-639
- Kennedy S, Smyth JA, McCullough SJ, Allan CM, McNeilly F, McQuaid S (1988) Confirmation of cause of recent seal deaths. *Nature* 335(6189):404
- Klein J, Figueroa F (1986) Evolution of the major histocompatibility complex. *Crit Rev Immunol* 6:295-386
- Kloskowski A, Pilarczyk M, Namiesnik J (2002) Denudation - A Convenient Method of Isolation and Enrichment of Analytes. *Crit Rev Anal Chem* 32(4):301-335
- Kryazhimskiy S, Plotkin JB (2008) The population genetics of dN/dS. *PLoS Genet* 4(12):e1000304
- Kumánovics A, Takada T, Fischer Lindahl K (2003) Genomic organization of the mammalian MHC. *Annu Rev Immunol* 21:629-657

- Kumar S, Nei M, Dudley J, Tamura K (2008) MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. *Brief Bioinform* 9(4):299-306
- Langefors Å, Lohm J, Grahn M, Andersen Ø, von Schantz T (2001) Association between major histocompatibility complex class IIB alleles and resistance to *Aeromonas salmonicida* in Atlantic salmon. *Proc R Soc Lond B* 268:479-485
- Laws RM, Taylor RJF (1957) A Mass Dying of Crabeater Seals, *Lobodon carcinophagus* (Gray). *Proc Zool Soc Lond* 129(3):315-324
- Lewis PO (2001) Phylogenetic systematics turns over a new leaf. *Trends Ecol Evol* 16(1):30-37
- Lipscomb TP, Schulman FY, Moffett D, Kennedy S (1994) Morbilliviral disease in Atlantic Bottlenose Dolphins (*Tursiops truncatus*) from the 1987–1988 Epizootic. *J Wildlife Dis* 30(4):567-571
- Lundberg AS, McDevitt HO (1992) Evolution of major histocompatibility complex class II allelic diversity: direct descent in mice and humans. *Proc Natl Acad Sci USA* 89:6545-6549
- Maddison DR, Maddison WP (2000) MacClade 4: analysis of phylogeny and character evolution. Version 4.0. Sinauer, Sunderland
- McCallum HI, Kuris A, Harvell CD, Lafferty KD, Smith GW, Porter J (2004) Does terrestrial epidemiology apply to marine systems? *Trends Ecol Evol* 19(11):585-591
- Miller MA, Gardner IA, Kreuder C, Paradies DM, Worcester KR, Jessup DA, Dodd E, Harris MD, Ames JA, Packham AE, Conrad PA (2002) Coastal freshwater runoff is a risk factor for *Toxoplasma gondii* infection of southern sea otters (*Enhydra lutris nereis*). *Int J Parasitol* 32(8):997-1006
- Montgelard C, Catzeflis FM, Douzery E (1997) Phylogenetic relationships of artiodactyls and cetaceans as deduced from the comparison of cytochrome b and 12S rRNA mitochondrial sequences. *Mol Biol Evol* 14(5):550-559
- Munguía-Vega A, Esquer-Garrigós Y, Rojas-Bracho L, Vázquez-Juárez R, Castro-Prieto A, Flores-Ramírez S (2007) Genetic drift vs. natural selection in a long-term small isolated population: major histocompatibility complex class II variation in the Gulf of California endemic porpoise *Phocoena sinus*. *Mol Ecol* 16:4051-4065
- Munson PJ, Di Francesco V, Porrelli R (1994) Protein Secondary Structure Prediction using Periodic-Quadratic-Logistic Models: Theoretical and Practical Issues. 27th Annual Hawaii International Conference on System Science 5:375-384
- Murphy WJ, Elzirik E, Johnson WE, Zhang YP, Ryder OA, O'Brien SJ (2001) Molecular phylogenetics and the origins of placental mammals. *Nature* 409:614-618



- 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(4):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
- Nielsen R, Yang Z (2003) Estimating the distribution of selection coefficients from phylogenetic data with applications to mitochondrial and viral DNA. *Mol Biol Evol* 20(8):1231-1239
- Nigenda-Morales S, Flores-Ramírez S, Urbán-R J, Vázquez-Juárez R (2008) MHC *DQB-1* polymorphism in the Gulf of California fin whale (*Balaenoptera physalus*) population. *J Hered* 99(1):14-21
- Nikaido M, Rooney AP, Okada N (1999) Phylogenetic relationships among cetartiodactyls based on insertions of short and long inter-spaced elements: hippopotamuses are the closest extant relative of whales. *Proc Natl Acad Sci USA* 96:10261-10266
- O'Brien SJ, Roelke ME, Marker L, Newman A, Winkler CA, Meltzer D, Colly L, Evermann JF, Bush M, Wildt DE (1985) Genetic basis for species vulnerability in the cheetah. *Science* 227(4693):1428-1434
- 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 L.*). *Proc Natl Acad Sci USA* 95:3714-3719
- Penn DJ (2002) The Scent of Genetic Compatibility: Sexual Selection and the Major Histocompatibility Complex. *Ethology* 108(1):1-21
- Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14(9):817-818
- Potts WK, Wakeland EK (1990) Evolution of diversity at the major histocompatibility complex. *Trends Ecol Evol* 5:181-187
- Reeves RR, Stewart BS, Clapham PJ, Powell JA (2002) National Audubon Society: Guide for Marine Mammals of the World. Alfred A Knopf, New York
- Reynolds III JE, Perrin WF, Reeves RR, Montgomery S, Ragen TJ (2005) Marine Mammal Research: Conservation beyond Crisis. The Johns Hopkins University Press, Baltimore pp 47-62
- Rice DW (2009) Classification (Overall). En: Perrin WF, Würsig B, Thewissen JGM (eds) *Encyclopedia of Marine Mammals*. 2nd edn. Elsevier Press, San Diego pp 234-238
- Roitt, IM, Delves P J (2001) *Roitt's Essential Immunology*. 10th edn. Blackwell Science, Oxford

- Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572-1574
- Ross PS (2002) The Role of Immunotoxic Environmental Contaminants in Facilitating the Emergence of Infectious Diseases in Marine Mammals. *Hum Ecol Risk Assess* 8(2):277-292
- Ross PS, De Swart R, Addison R, Van Loverend H, Vosd J, Osterhaus A (1996) Contaminant – induced immunotoxicity in harbour seals: Wildlife at risk? *Toxicology* 112(2):157-169
- Rousset B, Bernier-Valentin F, Wolff J, Roux B (1983) Alterations in tubulin immunoreactivity; relation to secondary structure. *Eur J Biochem* 131(1):31-39
- Sauermann U, Nürnberg P, Bercovitch FB, Berard JD, Trefilov A, Widding A, Kessler M, Schmidtke J, Krawczak M (2001) Increased reproductive success of MHC heterozygous males across free – ranging rhesus macaques. *Hum Genet* 108:249-254
- Schwaiger F-W, Weyers E, Buitkamp J, Ede AJ, Crawford A, Eppel JT (1994) Interdependent MHC-DRB exon-plus-intron evolution in artiodactyls. *Mol Biol Evol* 11(2):239-249
- Sigurdardóttir S, Borsch C, Gustafsson K, Andersson L (1992) Gene duplications and sequence polymorphism of bovine class II *DQB* genes. *Immunogenetics* 35:205-213
- Slade RW (1992) Limited MHC polymorphism in the southern elephant seal: implications for MHC evolution marine mammal population biology. *Proc R Soc Lond B* 249(1325):163-171
- Smith DM, Lunney JK, Ho CS, Martens GW, Ando A, Lee JH, Schook L, Renard C, Chardon P (2005) Nomenclature for factors of the swine leukocyte antigen class II system 2005. *Tissue Antigens* 66:623-639
- Swarbrick PA, Crawford AM (1997) The red deer (*Cervus elaphus*) contains two expressed major histocompatibility complex class II *DQB* genes. *Anim Genet* 28:49-51
- Swofford DL (2001) PAUP\*: phylogenetic analysis using parsimony (and other methods). Version 4.0. Sinauer, Sunderland
- Takahashi K, Rooney AP, Nei M (2000) Origins and divergence times of mammalian class II MHC gene clusters. *J Hered* 91:198-204
- Takahata N (1990) A simple genealogical structure of strongly balanced allelic lines and trans-species evolution of polymorphism. *Proc Natl Acad Sci USA* 87:2419-2423
- Takahata N, Satta Y, Klein J (1992) Polymorphism and balancing selection at major histocompatibility complex loci. *Genetics* 130:925-938
- Taubenberger JK, Tsai M., Kraft AE, Lichy JH, Reid AH, Schulman FY, Lipscomb DVM (1996) Two Morbilliviruses Implicates in Bottlenose Dolphin Epizootics. *Emerg Infect Dis* 2(3):213-216

- Terio KA, Munson L, Marker L, Aldridge BM, Solnick JV (2005) Comparison of *Helicobacter* spp. in Cheetahs (*Acinonyx jubatus*) with and without Gastritis. *J Clin Microbiol* 43(1):229-234
- Thewissen JGM, Cooper LN, Clementz MT, Bajpai S, Tiwari BN (2007) Whales originated from aquatic artiodactyls in the Eocene epoch of India. *Nature* 450:1190-1194
- Tovo-Rodrigues L, Rhode LA, Roman T, Schmitz M, Polanczyk G, Zeni C, Marques FZC, Contini V, Grevet EH, Belmonte-de-Abreu P, Bau CHD, Hutz MH (2011) Is there a role for rare variants in DRD4 gene in the susceptibility for ADHD? Searching for an effect of allelic heterogeneity. *Mol Psychiatry* 17(5):520-526
- Traul DL, Bhushan B, Eldridge JA, Crawford TB, Li H, Davies CJ (2005) Characterization of *Bison bison* major histocompatibility complex class IIa haplotypes. *Immunogenetics* 57(11):845-854
- Trowsdale J, Groves V, Arnason A (1989) Limited MHC polymorphism in whales. *Immunogenetics* 29:19-24
- Ursing BM, Arnason Ú (1998) Analysis of mitochondrial genomes strongly support hippopotamus-whale clade. *Proc R Soc Lond B* 265:2251-2255
- Vassilakos D, Natoli A, Dahlheim M, Hoelzel AR (2009) Balancing and directional selection at exon-2 of the MHC *DQB1* among populations of odontocete cetaceans. *Mol Biol Evol* 26(3):681-689
- Vedros NA, Smith AW, Schonweld J, Migaki G, Hubbard RC (1971) Leptospirosis Epizootic among California Sea Lions. *Science New Series* 172(3989):1250-1251
- Wan Q, Zhu L, Wu H, Fang S (2006) Major Histocompatibility complex class II variation in the giant panda (*Ailuropoda melanoleuca*). *Mol Ecol* 15(9):2441-2450
- Wegner KM, Reusch TBH, Kalbe M (2003) Multiple parasites are driving major histocompatibility complex polymorphism in the wild. *J Evol Biol* 16:224-232
- Xu S, Sun P, Zhou K, Yang G (2007) Sequence variability at three MHC loci of finless porpoises (*Neophocaena phocaenoides*). *Immunogenetics* 59:581-592
- Yang G, Yang J, Zhou K, Wei F (2005) Sequence Variation and Gene Duplication at the MHC *DQB* Loci of Baiji (*Lipotes vexillifer*), a Chinese River Dolphin. *J Hered* 96(4):310-317