Zoological Studies **49**(1): 132-151 (2010)

Zoological Studies

Phylogenetic Analyses of MHC Class II Genes in Bottlenose Dolphins and Their Terrestrial Relatives Reveal Pathogen-Driven Directional Selection

Wei-Cheng Yang, Jer-Ming Hu, and Lien-Siang Chou*

Institute of Ecology and Evolutionary Biology, National Taiwan University, 1 Roosevelt Rd., Sec. 4, Taipei 106, Taiwan

(Accepted June 3, 2009)

Wei-Cheng Yang, Jer-Ming Hu, and Lien-Siang Chou (2010) Phylogenetic analyses of MHC class II genes in bottlenose dolphins and their terrestrial relatives reveal pathogen-driven directional selection. *Zoological Studies* **49**(1): 132-151. The mammalian order Cetacea is believed to have made the challenging evolutionary transition from land to sea early in the Eocene, > 50 million years ago (Mya). With the shift in habitat, cetaceans had to evolve a range of specializations, including immune response. The major histocompatibility complex (MHC) multigene families have proven to be ideal candidates for research on the timing of selection and the spatial heterogeneity of selection pressures experienced by the ancestors of cetaceans because MHC multigene families consist of a mixture of divergent genes and play key roles in immune responses. Herein, we report the evolutionary relationship and estimates of divergence times of cetartiodactyls using MHC class II β-chain genes. In both the *DQB* and *DRB* phylogenies, cetaceans (*Tursiops truncatus* and *T. aduncus*) and artiodactyls (pig, hippo, and ruminants) formed 2 distinct clades, and the estimated divergence time was about 60 Mya or slightly earlier. Furthermore, our results showed that *T. truncatus* and *T. aduncus* diverged about 24 Mya, which greatly predates the emergence of the oldest dolphin (approximately 11 Mya). These findings are explained by postulating gene duplications in MHC genes and pathogen-driven directional selection during cetacean evolution. This research provides new information for studying the dynamics of MHC-pathogen co-evolution. http://zoolstud.sinica.edu.tw/Journals/49.1/132.pdf

Key words: Major histocompatibility complex, Bottlenose dolphins, Pathogen-driven directional selection.

Recent molecular and morphological studies suggested that the order Cetacea may be more closely related to the order Artiodactyla than to other orders of ungulates (Kumar and Hedges 1998, Arnason et al. 2000, Murphy et al. 2001, Boisserie et al. 2005), supporting a hypothesized order Cetartiodactyla (Montgelard et al. 1997). In addition, cetaceans and the hippopotamus (*Hippopotamus amphibious*) form a monophyletic group deeply nested within the Cetartiodactyla, while camels and pigs are basal to this order (Gatesy et al. 1996, Gatesy 1997, Nikaido et al. 1999, Boisserie et al. 2005). The molecular clock estimate for the divergence of the artiodactyls and cetaceans is about 60 million years ago (Mya) (Arnason and Gullberg 1996), a time slightly predating the oldest known cetacean, *Himalayacetus*, dated in the early Eocene, at 53.5 Mya (Bajpai and Gingerich 1998). It is believed that early cetaceans initially lived in freshwater habitats as terrestrial quadrupeds and were partially dependent on fresh water at some stages of their life before they gradually adapted to the marine environment and ultimately became fully aquatic marine mammals (Thewissen and Williams 2002). Adaptations from the land to sea life in cetaceans have been studied by comparative analyses of morphological traits (Thewissen et al.

*To whom correspondence and reprint requests should be addressed. Jer-Ming Hu and Lien-Siang Chou contributed equally to this work. Tel: 886-2-33662468. Fax: 886-2-23639902. E-mail:chouls@ntu.edu.tw

2001, Bejder and Hall 2002, Spoor et al. 2002, Nummela et al. 2004) as well as protein function (Naylor and Gerstein 2000, McClellan et al. 2005) among cetaceans and closely related species. However, these adaptations have never been investigated in terms of evolution of immune genes in cetaceans. This issue nonetheless had to be critical for cetaceans during their move from land to water, which is an enormous shift in habitat environments. Since major qualitative differences in microorganisms and infectious diseases are believed to exist between marine and terrestrial environments (McCallum et al. 2004), the immune genes of primitive cetaceans are supposed to have adapted for defending against distinct pathogens in aquatic environments.

Among immune genes, several major histocompatibility complex (MHC) genes are suggested to be under strong pathogen-driven selective forces (Hughes and Yeager 1998). The mammalian MHC related to peptide expression is divided into two main regions, class I and II gene regions (Hughes and Yeager 1998). Both class I and class II MHC gene families include a large number of loci that encode heterodimeric peptide-binding proteins consisting of extracellular domains (including the peptide-binding region; PBR), a connecting portion, a transmembrane portion, and a cytoplasmic tail. These multigene families consist of mixtures of divergent genes including large numbers of functionally closely related genes and pseudogenes resulting from birth-and-death evolution (Nei et al. 1997). In this evolutionary process, new genes are created by repeated gene duplications, while some genes may later become pseudogenes or even be deleted from the genome. The evolutionary longevity of new genes can vary from gene to gene in the MHC. MHC class II loci undergo a slower rate of birth-and-death evolution than do class I loci, and the longevity of class II genes is much greater than that of class I genes (Klein and Figueroa 1986). Consequently, many class II homologous loci are shared by different orders of mammals, while it is difficult to determine the homologous relationships of class I genes among different orders of mammals (Hughes and Nei 1990). These characteristics make MHC class II loci more-appropriate subjects for comparative evolutionary studies and for estimating divergence times of various loci (Klein and Sato 1998, Takahashi et al. 2000).

Divergence times between different gene clusters of mammalian class II genes were estimated (Klein and Figueroa 1986, Hughes and Nei 1990, Takahashi et al. 2000). Estimates by Takahashi et al. (2000) are more reliable because many more gene sequences were available at that time. However, only 3 species from the Cetartiodactyla (pig, cattle, and sheep) were studied by Takahashi et al. while evolutionary relationships of MHC class II genes among the Cetartiodactyla remain unresolved. Because the problem is important for understanding how immune system genes evolved in the Cetartiodactyla, herein we constructed phylogenetic trees and estimated the divergence times of clades using complementary (c)DNA sequences of MHC class II β-chain genes from bottlenose dolphins (*Tursiops truncatus* and *T. aduncus*), the hippo (*Hippopotamus amphibious*), other mammals, chicken, frog, and fish. By interpreting phylogenetic relationships and divergence time estimates of MHC class II genes in bottlenose dolphins and their closely related terrestrial species, we attempted to reveal how and when the habitat shift, accompanied by changes in foreign antigens, affected MHC evolution in cetaceans and other closely related mammals when they moved from land to water.

MATERIALS AND METHODS

Sample preparation and rapid amplification of cDNA ends (RACE) polymerase chain reaction (PCR) amplification

Tissue sample from 1 captive hippopotamus was collected and stored at -70°C. RNA was isolated by silica-based gel membranes supplied in the RNeasy Fibrous Tissue Mini Kit (Qiagen, Valencia, CA, USA). The isolated RNA was also stored at -70°C prior to the RACE cDNA synthesis. A cDNA population was constructed using SMART RACE cDNA amplification kits (Clontech, Mountain View, CA, USA) in order to facilitate the amplification of full-length gene transcripts. In brief, RNA samples from the hippopotamus were used as templates for cDNA synthesis. Adaptor-like sequences were added to either the 5' or 3' end of the cDNA fragments in 2 separate reactions. These modified cDNAs were generated from cellular RNA by Moloney murine leukemia virus (MMLV) reverse transcriptase-driven 1st-strand synthesis using lock-docking oligo(dT) primers and the SMART II oligonucleotide. The resulting 5'- and 3'-modified cDNA fragments were used

as templates for the subsequent PCR and RACE PCR analyses.

We used published degenerate oligonucleotide primers (Bowen et al. 2002 2004) which recognize conserved regions of each of 2 MHC class II genes, *DQB* and *DRB*. *DQB* and *DRB* genes were shown to be polymorphic in some terrestrial carnivores and domestic animals (Schook and Lamont 1996, Yuhki and O'Brien 1997, Wagner et al. 1999). PCR amplifications using these degenerate class II primers were performed on 20 ng of each RACE cDNA library in 50 ml volumes containing 20 pmol of each primer (either *DQB*-U172 and *DQB*-L769, or *DRB*-U182 and *DRB*-L729), 40 mm Tris-KOH (pH 8.3), 15 mm KOAc, 3.5 mm $Mg(OAc)_2$, 200 mM each dNTP, and 5 U of Advantage 2 *Taq* polymerase (Clontech). For the 5' gene transcript amplifications, the reactions contained 20 pmol of either DRB-L729 (for *DRB* gene products) or DQB-L769 (for *DQB* gene products), along with 1-5 pmol of the universal primer mix (UPM, SMART RACE cDNA amplification kit; Clontech). For 3' gene transcript amplifications, the reactions contained 20 pmol of either DRB-U182 (for *DRB* gene products) or DQB-U172 (for *DQB* gene products), along with 1-5 pmol of the UPM (SMART RACE cDNA amplification kit; Clontech). The PCR was performed on a thermal cycler and consisted of 30 cycles at 94°C for 30 s, 68°C for 30 s and 72°C for 2 min, followed by a single extension step of 72°C for 10 min. The products of these reactions were electrophoresed on 1.5% agarose gels and visualized by ethidium bromide staining. Bands representing PCR products of the predicted size were excised from the gel, extracted, and purified using a commercially available nucleic acid-binding resin (Qiaex II gel extraction kit, Qiagen). These isolated RACE fragments were then ligated into a T/A-type cloning vector (pGEM-T Easy Vector System, Promega, San Luis Obispo, CA, USA). Following transformation, growth, and blue-white selection in DH5α-competent cells, the DNA from positive clones was isolated. The nucleotide sequences of both strands were determined by dideoxynucleotide methodology using an automated sequencer. The nucleotide sequences of these amplicons were compared using MEGA vers. 3.0 (Kumar et al. 2004). However, we could not obtain the predicted product of the 5' gene transcript amplifications using 5'-modified cDNA fragments from the hippopotamus, possibly due to RNA degradation during tissue preservation. Therefore, we only used sequences from the 3' gene transcript amplifications, the nucleotide number of the coding region of which was around 660, for the subsequent phylogenetic analysis.

The full-length nucleotide sequences of *DQB* and *DRB* genes of bottlenose dolphins (*T. truncatus* and *T. aduncus*) (Yang et al. 2007) used in this study were obtained from blood samples by similar methods described above.

Phylogenetic analysis of MHC class II β-chain genes in vertebrates

We used nucleotide sequences of the class II β-chain genes consisting of the extracellular domain, connecting peptide, transmembrane, and part of the cytoplasmic tail obtained from RACE product sequences and GenBank (Table 1). Several different species including placental mammals, chicken, frog (*Xenopus*), and bony fish (zebrafish *Danio rerio*) were studied. *Tursiops truncatus* and *T. aduncus* were representatives of the Cetacea, and *H. amphibious* was representative of the Hippopotamidae. The choice of other species and loci followed Takahashi et al. (2000) with modifications. Pseudogenes were not used. When allelic variants were found at a locus, only a single typical allele was used to represent that locus. The nucleotide sequences obtained were aligned using the CLUSTAL W method (Thompson et al. 1994), taking into account the deduced amino acid sequences. Minor adjustments were made to the alignments after visual inspection. The number of nucleotides was 616 (205 codons). We identified positions 1-218 as the segment including the PBR and positions 219-616 as the non-PBR segment following Bowen et al. (2002 2004). A test for recombination events between sequences was conducted with the program GENECONV 1.81 (Sawyer 1999). The program runs 10,000 permutations of the original data in global and pairwise tests for the occurrence of possible inner and outer fragments involved in recombination or gene-conversion events. We used all codon positions and only the 1st + 2nd codon positions in the following analysis. The appropriate model was selected with Modeltest (Posada and Crandall 1998), using the Akaike information criterion (AIC) (Posada and Buckley 2004) with a parsimonious tree chosen as the basis for Modeltest. The best model was GTR + Γ + I (Rodríguez et al. 1990, Yang 1994). A phylogenetic analysis was carried out using the Bayesian inference (BI) implemented in MRBAYES 3.1 (Ronquist and Huelsenbeck

2003) with the following settings. The maximum likelihood model employed 6 substitution types (nst = 6), with base frequencies estimated from the data. Rate variations across sites were modeled using a γ-distribution (rates = invgamma). The prior probability distribution on branch lengths was specified as the birth-death clock model. A Markov chain Monte Carlo (MCMC) search was run with 4 chains for 10^6 generations, sampling the Markov chain every 100 generations, and the sample points of the 1st 2500 generations were discarded as "burn-in" after which the chain reached stationarity. We used the Shimodaira-Hasegawa (SH) test (Shimodaira and Hasegawa 1999), which assumes a resampling estimated log-likelihood (RELL) approximation, to compare the likelihood scores of potential topologies derived from the BI tree. To test for topological consistency, we used 3 more methods to construct phylogenetic trees: Neighbor-joining (NJ), maximum-parsimony (MP), and maximum-likelihood (ML) using PAUP* 4.0b10 (Swofford 2003). Parsimony analyses were performed using 100 replicates with random addition of taxa, tree-bisection-reconnection branch swapping, and a transversion: transition weighting of 2: 1. The ML analysis was performed using the GTR+ Γ + I model with estimated parameters. The NJ analysis was performed using the Kimura-2-parameter model of Takahashi et al. (2000). We performed the SH test to compare scores of trees resulting from the above searches. We also conducted Tajima's relative rate test (Tajima 1993) and likelihood-ratio test as described by Swofford et al. (1996) to test the molecular clock hypothesis of this dataset.

Divergence time estimation

To estimate the divergence time, zebrafish sequences were used as outgroups in the trees obtained from 4 inferences. We estimated the divergence times of MHC class II gene clades of the BI ultrametric tree by performing age calibration in r8s vers. 1.7 (Sanderson 2003) using the bird-mammal divergence time estimate (310 Mya) (Hedges et al. 1996, Kumar and Hedges 1998, Wang et al. 1999a) as a calibration point. For comparison, we used a linearized tree method under the assumption of a molecular clock to estimate the times of divergence of different clades in the NJ tree as described by Takahashi et al. (2000). However, we used 310 Mya as the calibration point in the linearized tree method, while Takahashi et al. (2000) used 300 Mya.

Table 1. Major histocompatibility complex (MHC) class II β-chain genes used in this study with respective GenBank accession numbers

Species	Gene [accession no.]	Species	Gene [accession no.]	Species	Gene [accession no.]
Zebrafish	DAB1 [L04805]	Dolphin (Tt)*	DRB [EF017817]	Mouse	DOB [NM010389]
	DAB4 [U08870]		DQB [EF017815]		DQB(AB1) [M13538]
	DCB [U08873]	Dolphin (Ta)*	DRB [EF017818]		DRB(EB1) [K01145]
	DEB [U08874]		DQB [EF017816]	Macaque	DRB4 [M60062]
Frog	B(1) [D13685]	Hippo	DRB [EF017820]		DRB1 [M60059]
	B(2) [D13684]		DQB [EF017819]	Gorilla	DRB3 [M77153]
Chicken	BL(1) [M26306]	Sheep	DRB [M93432]		DRB1 [M77154]
	BL(2) [M29763]		DQB [L08792]		DRB5 [M77152]
	BL(3) [M26307]	Cattle	DOB [AB117946]	Human	DRB1 [M11161]
Rabbit	DPB [M21468]		DRB [U77067]		DRB3 [NM022555]
Cat	DRB [U51575]		DQB [M30003]		DRB4 [NM021983]
Sea lion*	DRB [AY491464]	Goat	DRB [AB008346]		DRB5 [NM002125]
	DQB [AF503406]	Pig	DRB [M55165]		DPB1 [M57466]
Dog	DRB [M29611]		DQB [AY459301]		DOB [X03066]
	DQB [AF043908]	Mole rat	DPB [M16685]	Chimpanzee	DOB [M24358]
Horse	DQB [L33910]	Rat	DQB(BB) [X56596]	Tamarin	DRB [M76488]
			DRB1(EB1) [M12382]		

Tt, *Tursiops truncatus*; Ta, *Tursiops* aduncus; Sea lion, California sea lion *Zalophus californianus*.

RESULTS

Identification of MHC genes in the hippo and bottlenose dolphins

3' RACE cDNA clones were obtained from muscle tissue of a captive hippo. These partial-length transcripts were characterized as hippo-*DQB* (647 bp) or hippo-*DRB* (653 bp) based on comparisons with bovine lymphocyte antigen (BoLA) and human lymphocyte antigen (HLA) *DQB* and *DRB* sequences. The numbers of the deduced amino acids of hippo-*DQB* and *DRB* gene products were 213 and 216, respectively. No sequence compatible with a class II MHC pseudogene was identified.

Clones containing full-length bottlenose dolphin MHC class II sequences were obtained from the RACE cDNA products of 2 *T. truncatus* and 2 *T. aduncus* (Yang et al. 2007). The nucleotide sequence of the 780 (*DQB*-primer derived) and 801 bp (*DRB*-primer derived) products were typical of transcripts from mammalian class II genes. These transcripts were characterized as *Tt-DQB*, *Ta-DQB*, *Tt-DRB*, or *Ta-DRB* based on comparisons with HLA and BoLA *DQB* and *DRB* sequences. No sequence compatible with a class II MHC pseudogene was identified.

Phylogenetic analysis of MHC class II β-chain genes in vertebrates

No evidence for genetic-recombination or gene-conversion events between the aligned sequences was found by using the program GENECONV 1.81 with gscale values of 0, 1 and 2. The phylogenetic trees were obtained by BI using 3 data sets: 1-218 (PBR), 219-616 (non-PBR), and 1-616 nt positions of MHC class II β-chain genes from various vertebrate species (Fig. 1). The BI trees were constructed under the GTR + Γ + I model incorporating the birth-death clock model, and all 3 codon positions were used. When 616 nt was used, the tree revealed that all genes from mammalian species formed a monophyletic group separated from the clades of chicken, frog, and zebrafish (Fig. 2a). In the mammal group, there were 4 clades of genes corresponding to *DRB*, *DPB*, *DQB*, and *DOB*, respectively. The *DOB* clade was the sistergroup of the much-larger *DRB*/*DPB*/*DQB* clades. The *DPB* and *DQB* clades were more closely related to each other than to the other clades. The *DRB* and *DQB* clades both showed well-defined clustering of sequences from major mammalian lineages, such as rodent, primate, carnivore, and ungulate clades. Comparable clustering patterns were observed in the ML, MP, and NJ trees using all codon positions (Fig. 3a). The SH test showed that there were no significant differences among trees obtained from the BI, ML, MP, and NJ algorithms ($p > 0.5$). When only the 1st and 2nd codon positions were used, the topologies of the trees obtained by each method slightly differed from those using all codon positions, but the evolutionary relationships among the major clades remained the same (results not shown). Sequences from ungulates and dolphins formed a monophyletic group that was separated from other mammals with a posterior probability of 1.00 in the *DQB* and *DRB* clades (Fig. 2a). In addition, dolphins diverged from ungulates with moderately high (0.82 for *DQB*) or high (0.98 for *DRB*) posterior probabilities. Ruminants (cattle, sheep, and goat) were grouped with high posterior probabilities in both the *DQB* (0.92) and *DRB* (1.00) clades. However, relationships among the pig, horse, hippo, and ruminants were not conclusively resolved due to low posterior probabilities (0.50-0.69). The alternative hypothesis that dolphins represent a sistergroup of the hippo in both the *DQB* and *DRB* genes was rejected by the SH test ($p < 0.001$). Nevertheless, a basal position of the horse among the dolphin/ungulate group in the *DQB* clade was not rejected in the tree constructed using either the original data matrix ($p > 0.3$) or only *DQB* sequences ($p > 0.3$).

The BI tree based on PBR sequences only (positions 1-218) (Fig. 2b) showed that β-chain genes of mammals diverged from those of non-mammalian species with a low posterior probability (0.53). However, the *DPB*, *DQB*, and *DRB* clades formed a monophyletic group which was separated from the *DOB* clade with a posterior probability of 1.00. The *DQB* and *DRB* sequences of dolphins were sistergroups within the clade containing all other *DRB* sequences with high posterior probability (0.97). The *DQB* and *DRB* sequences of the California sea lion *Zalophus californianus* were sistergroups and clustered with the dog *DQB* sequence with a high posterior probability (0.96). *DQB* sequences of ungulates were grouped with rodent *DQB* sequences with a high posterior probability (0.98), but their relationships were not conclusively resolved. In the *DRB* clade, the hippo was grouped with ruminants with a low posterior probability (0.68), while the pig was grouped with rodents and tamarin although the posterior probability was low as well (0.68).

(A)

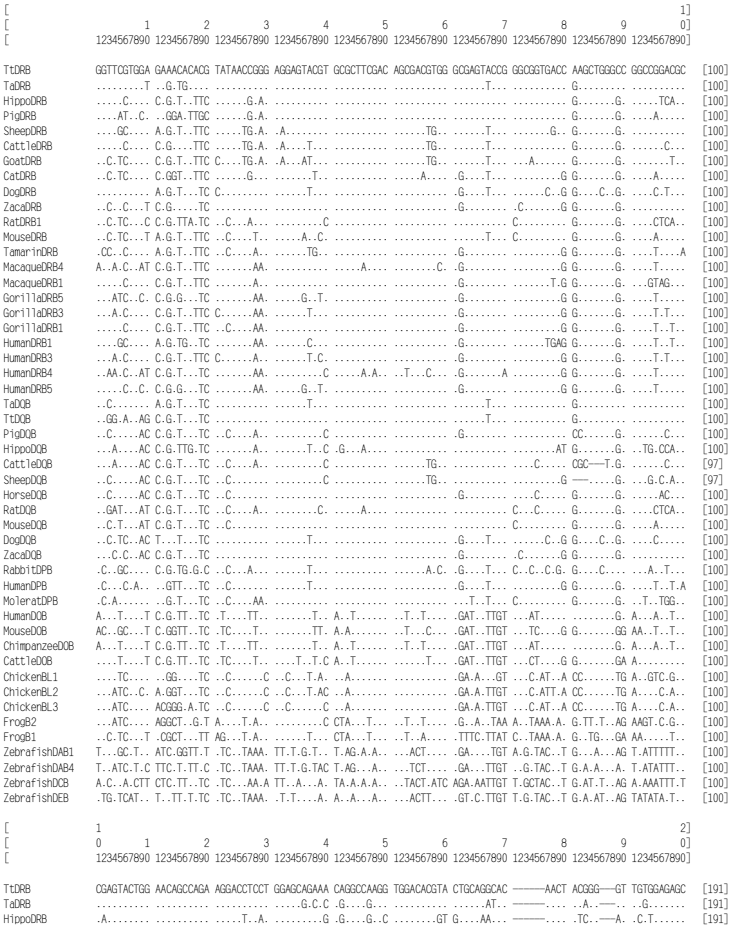


Fig. 1. Nucleotide sequence identity between *Tursiops truncatus DRB* sequence (Tt-*DRB*) and other major histocompatibility complex (MHC) class II β-chain sequences from *T. aduncus*, artiodactyls, other mammals, chicken, frog, and fishes. Tt, *Tursiops truncatus*; Ta, *T. aduncus*; Zaca, California sea lion *Zalophus californianus*.

(B)

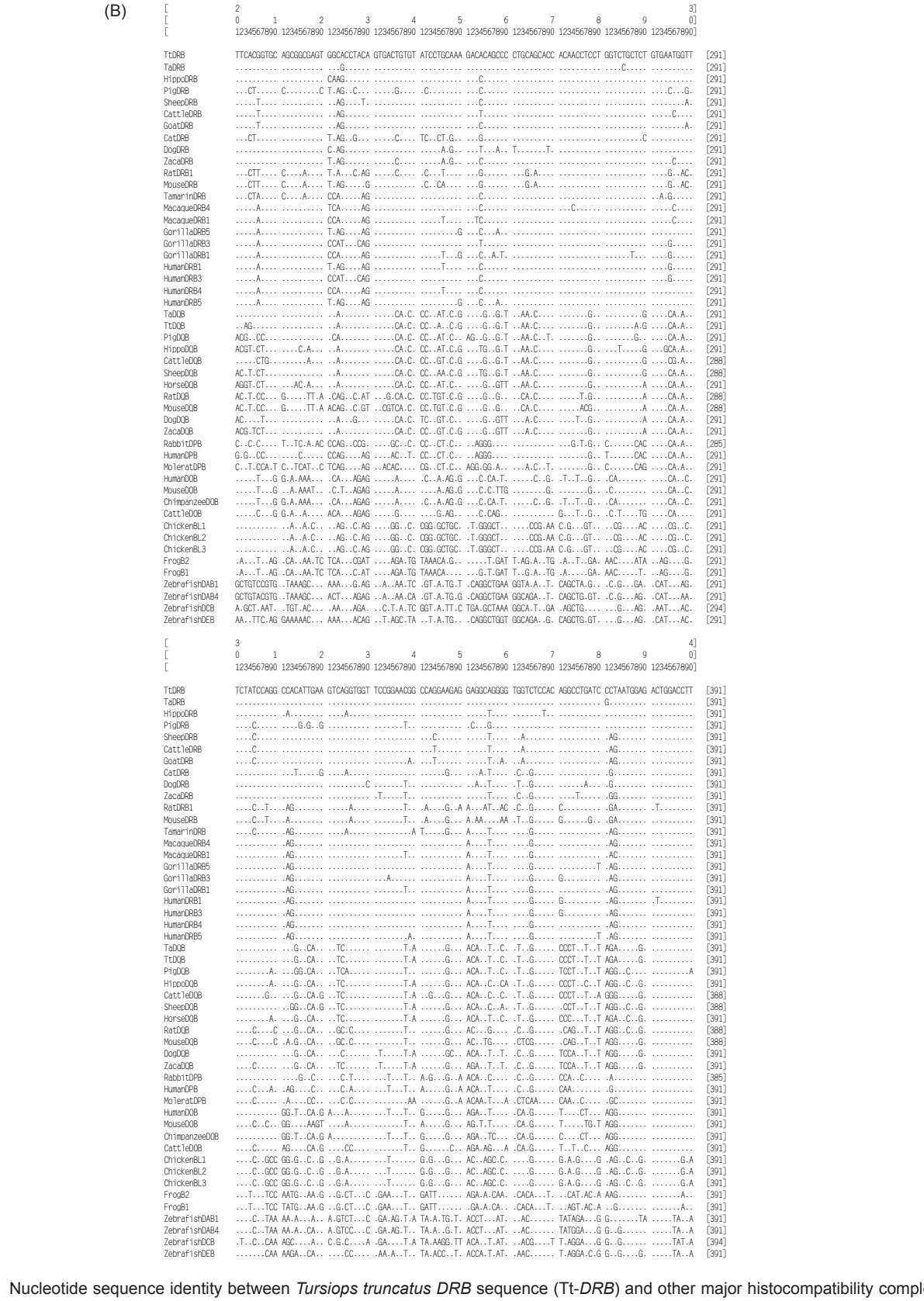


Fig. 1. Nucleotide sequence identity between *Tursiops truncatus DRB* sequence (Tt-*DRB*) and other major histocompatibility complex (MHC) class II β-chain sequences from *T. aduncus*, artiodactyls, other mammals, chicken, frog, and fishes. Tt, *Tursiops truncatus*; Ta, *T. aduncus*; Zaca, California sea lion *Zalophus californianus*.

(C)

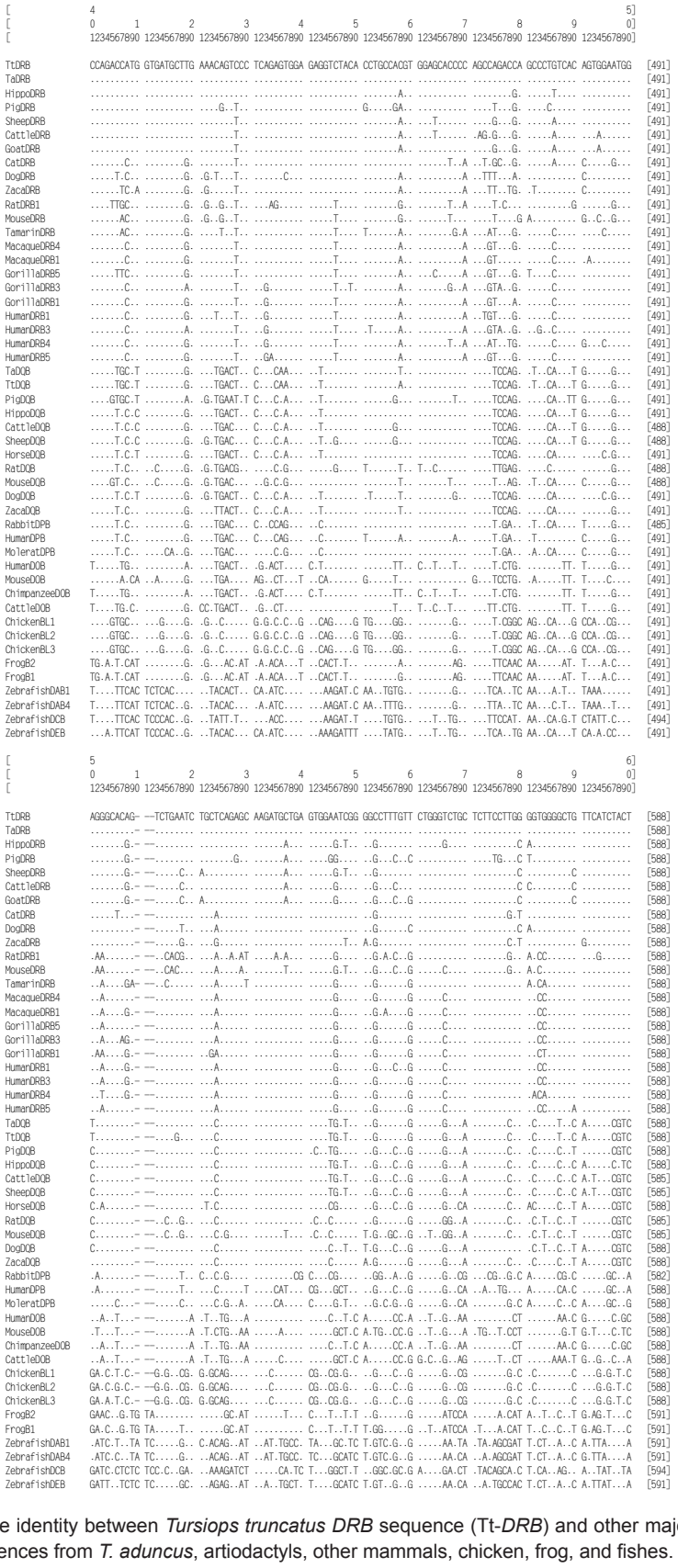


Fig. 1. Nucleotide sequence identity between *Tursiops truncatus DRB* sequence (Tt-*DRB*) and other major histocompatibility complex (MHC) class II β-chain sequences from *T. aduncus*, artiodactyls, other mammals, chicken, frog, and fishes. Tt, *Tursiops truncatus*; Ta, *T. aduncus*; Zaca, California sea lion *Zalophus californianus*.

DRB sequences of primates were grouped with a posterior probability of 1.00. Clades with relatively high posterior probabilities (> 0.90) in the BI tree were also significantly supported by bootstrap values on the MP and NJ trees (Fig. 3b).

The phylogenetic relationships of the BI tree based on non-PBR sequences (position 219-616) (Fig. 2c) were nearly identical to those of the tree using 616 nt (Fig. 2a) except that the horse was placed outside cetartiodactyls in the *DQB* clade with a moderate posterior probability (0.82). Furthermore, the artiodactyl clade had moderate posterior probabilities in both the *DQB* and *DRB* clades (0.85 and 0.88, respectively). The ML, MP, and NJ trees showed essentially the same clustering patterns compared to the BI tree, although the bootstrap values of the cetartiodactyl group in the *DQB* clade and the artiodactyl group in the *DQB* and *DRB* clades ware low (< 70%) under the MP and NJ criteria (Fig. 3c). The phylogenetic relationships of non-PBR sequences should be more reflective of the actual evolutionary history of the mammalian MHC genes since they are governed by purifying selection. For this reason, in figure 2a, it is advisable to regard the horse as a sistergroup of the cetartiodactyls in the *DQB* phylogeny. This postulation was further supported by the result of our SH tests on the position of the horse as mentioned above. Therefore, we used this postulated position of the horse *DQB* gene in the following estimation of divergence times.

(D)

```
[              6              ]
[              0       1      ]
[              1234567890 123456]

TtDRB          TCAGGAATCA GAAAGG   [604]
TaDRB          .......... ......   [604]
HippoDRB       .......... ......   [604]
PigDRB         ...A...... ......   [604]
SheepDRB       .......... ......   [604]
CattleDRB      .......... ......   [604]
GoatDRB        .......... ......   [604]
CatDRB         ..C....... ......   [604]
DogDRB         ....A..... ...G..   [604]
ZacaDRB        .......... ...G..   [604]
RatDRB1        ......GC.. ......   [604]
MouseDRB       .......... ......   [604]
TamarinDRB     .......... ......   [604]
MacaqueDRB4    .......... ......   [604]
MacaqueDRB1    .......... ......   [604]
GorillaDRB5    ...A...... ......   [604]
GorillaDRB3    ......G... ......   [604]
GorillaDRB1    .......... ......   [604]
HumanDRB1      .......... ......   [604]
HumanDRB3      .......... ......   [604]
HumanDRB4      .......... ......   [604]
HumanDRB5      ...A...... ......   [604]
TaDQB          A.....GC.. ...G..   [604]
TtDQB          A.....GC.. ...G..   [604]
PigDQB         A.....G... ...G..   [604]
HippoDQB       A.....GC.. ...G..   [604]
CattleDQB      A.....GC.. ...G..   [601]
SheepDQB       A.....GC.. ...G..   [601]
HorseDQB       A.....GCA. ...G..   [604]
RatDQB         A..A.CG... ......   [601]
MouseDQB       A.....G... ......   [601]
DogDQB         A.....GC.. ......   [604]
ZacaDQB        A.....GC.. ......   [604]
RabbitDPB      GG....GCA. .....C   [598]
HumanDPB       GG....GCA. .....T   [604]
MoleratDPB     .G....GCA. .....C   [604]
HumanDOB       .A...GC... ......   [604]
MouseDOB       ...A.GC... .....C   [604]
ChimpanzeeDOB  .A...GC... ......   [604]
CattleDOB      .....GCCTG ......   [604]
ChickenBL1     .GC.CGG... ......   [604]
ChickenBL2     .GC.CGG... ......   [604]
ChickenBL3     .GC.CGG... ......   [604]
FrogB2         .AC...G.A. ...GAC   [607]
FrogB1         ..C...G.A. A..GAC   [607]
ZebrafishDAB1  AG.A...ATC A.C...   [607]
ZebrafishDAB4  AG.A...ATC A.C...   [607]
ZebrafishDCB   AA..A..... A.C...   [610]
ZebrafishDEB   AG.A...ATC ATC...   [607]
```

Fig. 1. Nucleotide sequence identity between *Tursiops truncatus DRB* sequence (Tt-*DRB*) and other major histocompatibility complex (MHC) class II β-chain sequences from *T. aduncus*, artiodactyls, other mammals, chicken, frog, and fishes. Tt, *Tursiops truncatus*; Ta, *T. aduncus*; Zaca, California sea lion *Zalophus californianus*.

Divergence time estimations

The likelihood scores of the BI trees using 616 nt with and without a molecular clock did not significantly differ (p = 0.374) in the SH test; therefore the null hypothesis that the rate of evolution was homogeneous among all branches in the phylogeny was not rejected. The results of Tajima's relative rate test showed that there was no significant evolutionary rate difference between the 2 species of bottlenose dolphins, between the dolphin/ungulate pair, or between the *DQB*/*DRB* pair. By using the birth-death clock model in the BI tree (Fig. 2a), we estimated the divergence times between clades under the assumption that mammals and chicken diverged 310 Mya. We obtained an estimate of 249.36 Mya for the divergence between the *DOB* and other clades (Table 2). The estimated time of divergence between the *DRB* and *DQB*/*DPB* clades was close to that between the *DQB* and *DPB* clades, as expected from the tree in figure 2a. Divergence times between several major mammalian lineages (rodents, primates, carnivores, and cetartiodactyls) were also estimated. Rodents and other mammals diverged 109-122 Mya in the *DOB*, *DPB*, *DQB*, and *DRB* clades; primates and fereuungulates (carnivores + cetartiodactyls) diverged around 98 and 88 Mya in the *DOB* and *DRB* clades, respectively; carnivores and cetartiodactyls diverged around 78 and 76 Mya in the *DQB* and *DRB* clades, respectively. When cetartiodactyl *DQB* sequences were constrained as a monophyletic group, the estimated divergence time between dolphins and artiodactyls was 60.03 Mya, which was similar to that in the *DRB* clade. The 2 dolphin species (*T. truncatus* and *T. aduncus*) were separated by 26 and 22 Mya in the *DQB* and *DRB* clades, respectively. The BI tree using position 219-616 was also validated as being

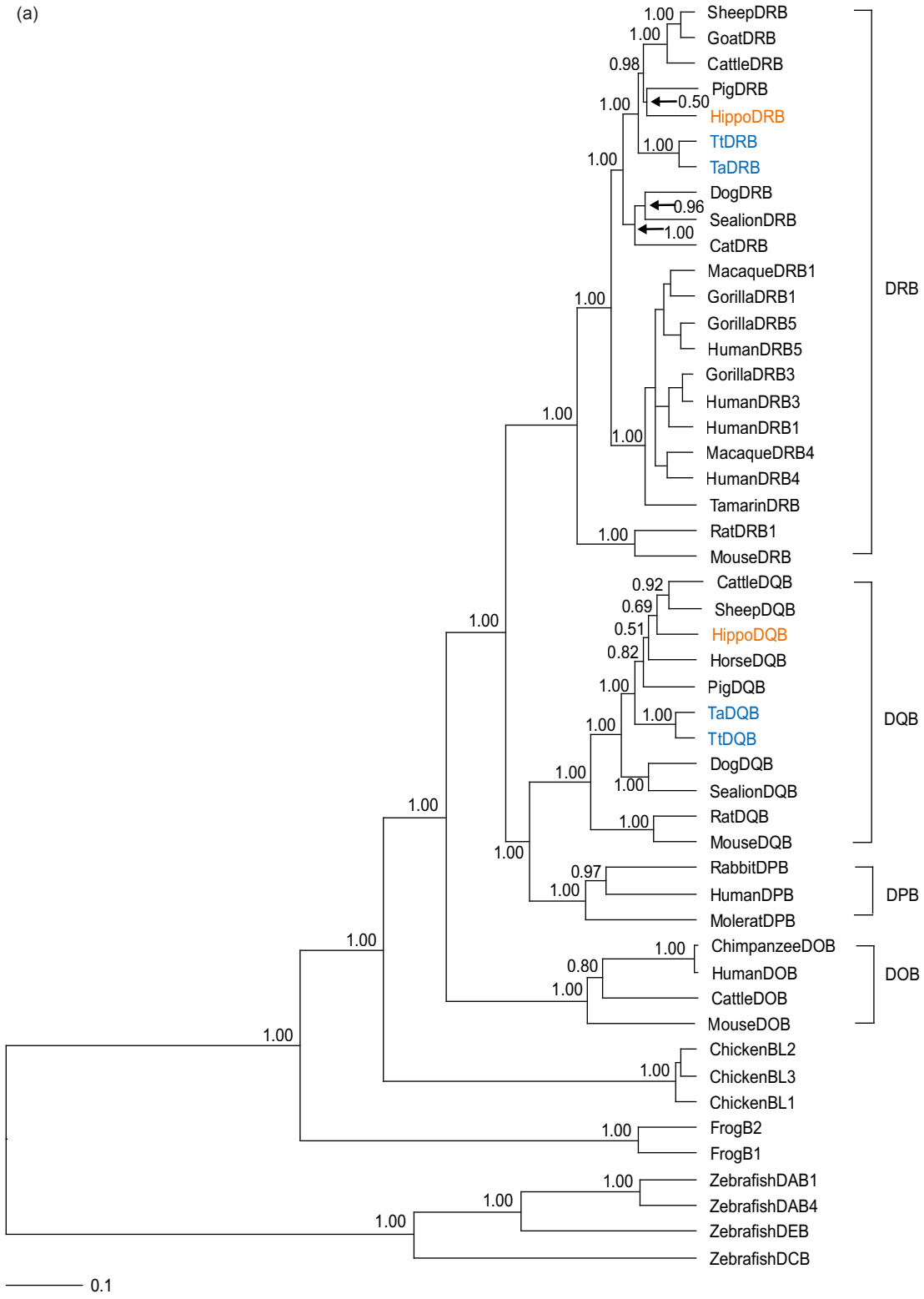


Fig. 2. Bayesian trees for major histocompatibility complex (MHC) class II β-chain genes using all codon positions with the GTR + Γ + I model and birth-death clock model. The nucleotide positions for tree construction were (a) positions 1-616, (b) 1-218, and (c) 219-616. The posterior probabilities are shown in parentheses for the nodes of interest. Tt, *Tursiops truncatus*; Ta, *T. aduncus*; Sealion, California sea lion *Zalophus californianus*.

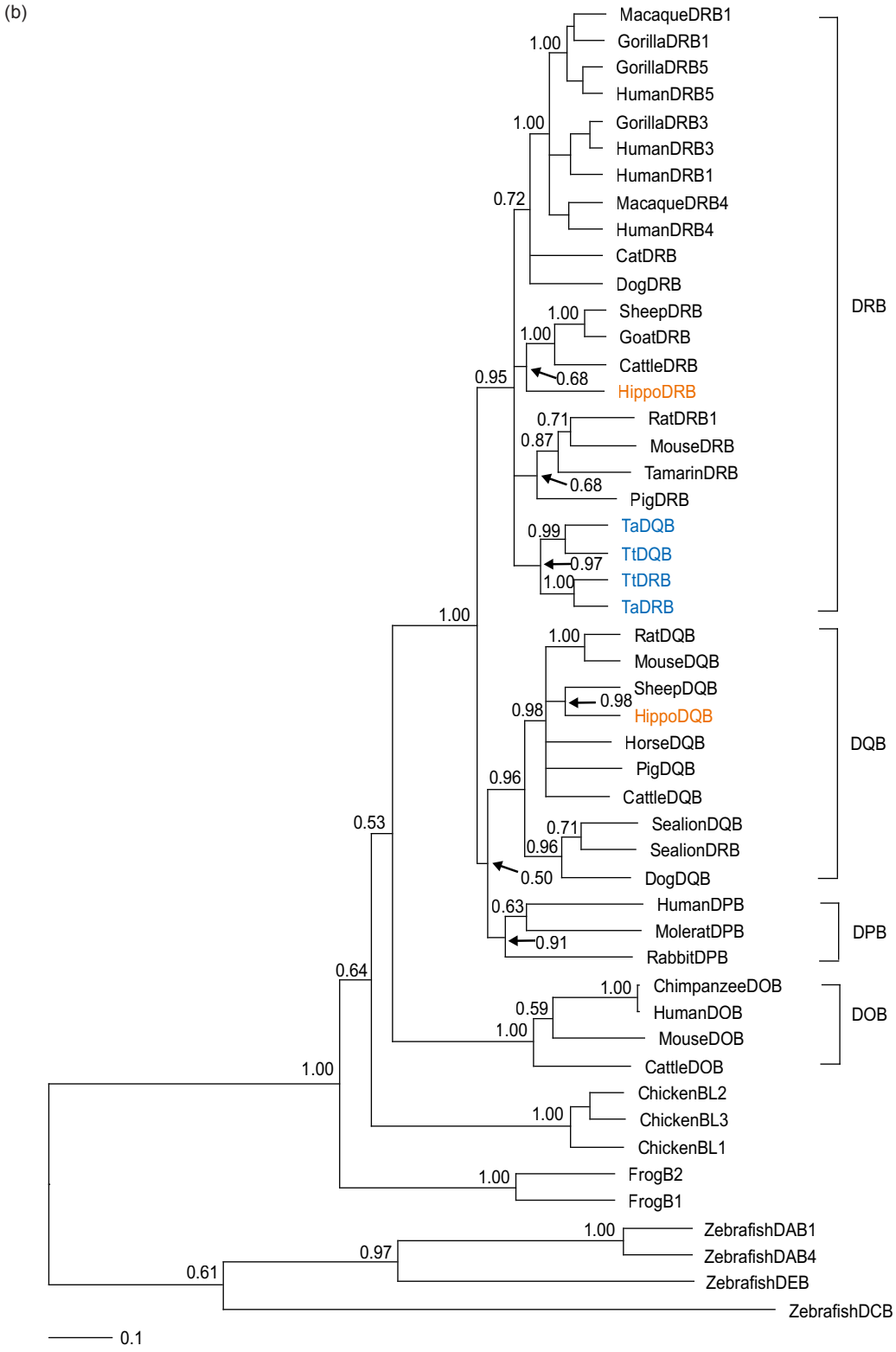


Fig. 2. Bayesian trees for major histocompatibility complex (MHC) class II β-chain genes using all codon positions with the GTR + Γ + I model and birth-death clock model. The nucleotide positions for tree construction were (a) positions 1-616, (b) 1-218, and (c) 219-616. The posterior probabilities are shown in parentheses for the nodes of interest. Tt, *Tursiops truncatus*; Ta, *T. aduncus*; Sealion, California sea lion *Zalophus californianus*.

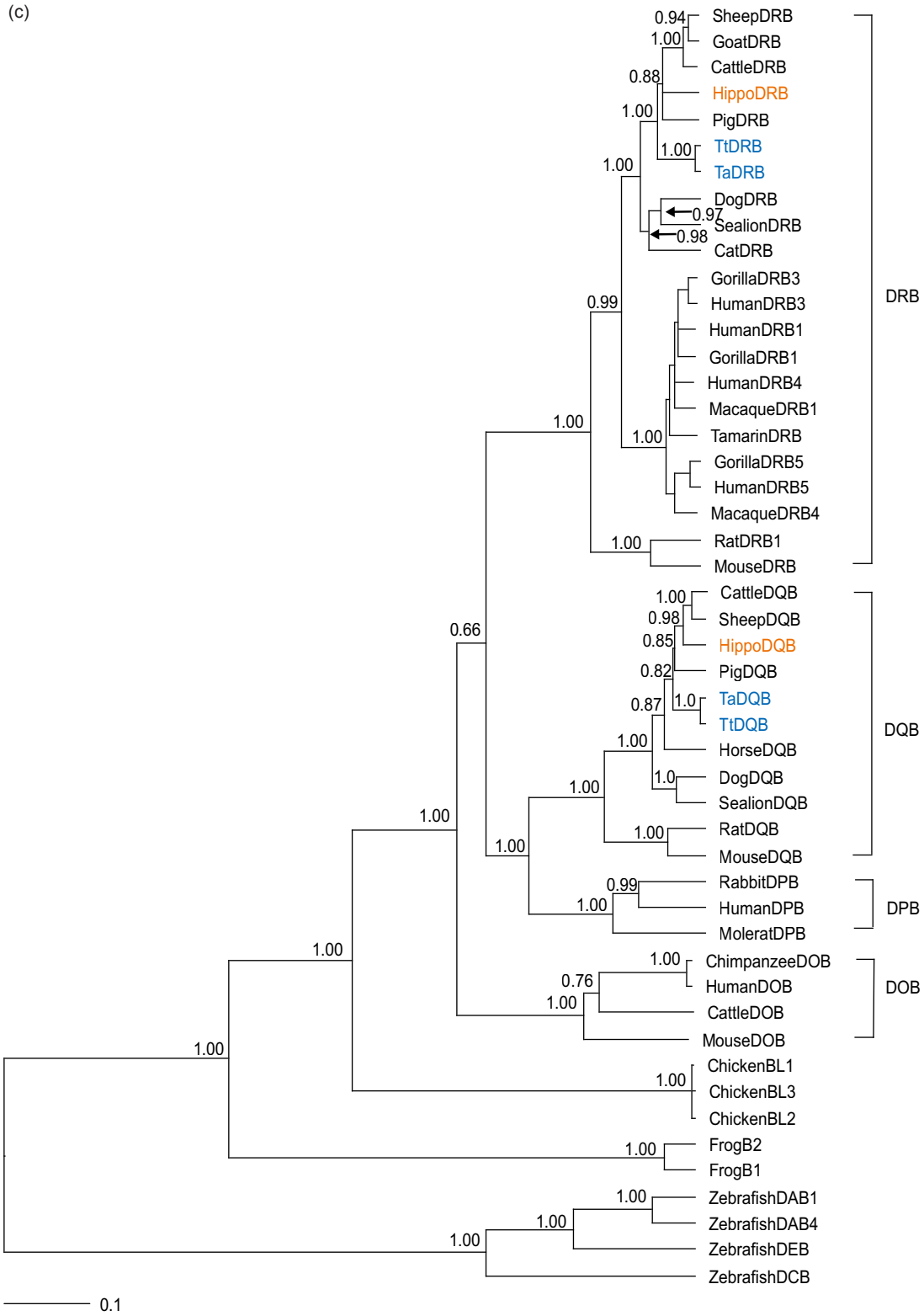


Fig. 2. Bayesian trees for major histocompatibility complex (MHC) class II β-chain genes using all codon positions with the GTR + Γ + I model and birth-death clock model. The nucleotide positions for tree construction were (a) positions 1-616, (b) 1-218, and (c) 219-616. The posterior probabilities are shown in parentheses for the nodes of interest. Tt, *Tursiops truncatus*; Ta, *T. aduncus*; Sealion, California sea lion *Zalophus californianus*.

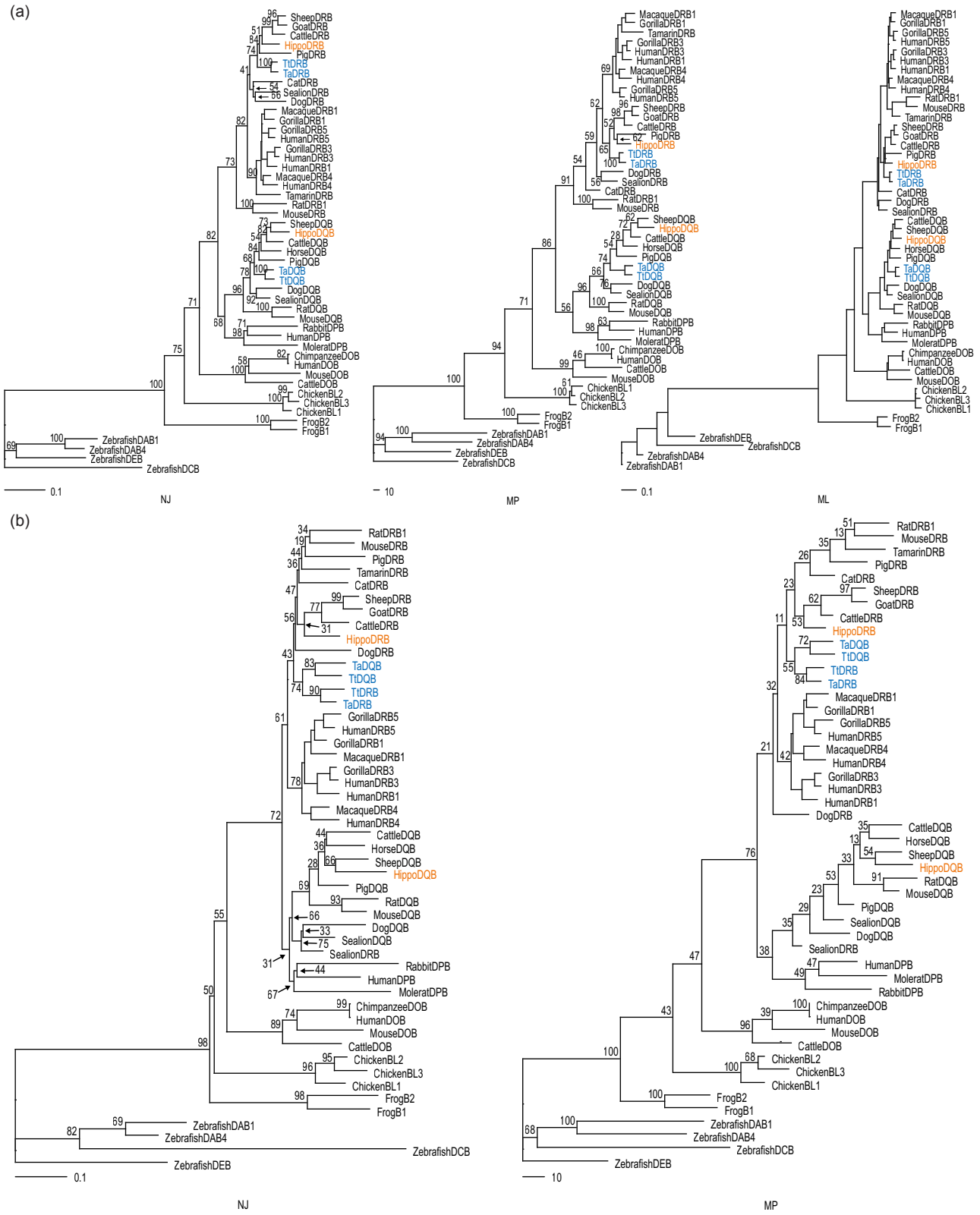


Fig. 3. Neighbor-joining (NJ), maximum-parsimony (MP), and maximum-likelihood (ML) trees for major histocompatibility complex (MHC) class II β-chain genes using all codon positions. The nucleotide positions for tree construction were (a) positions 1-616, (b) 1-218, and (c) 219-616. The ML tree was constructed only using nucleotide positions 1-616. The bootstrap values are shown in parentheses for the nodes of interest. Tt, *Tursiops truncatus*; Ta, *T. aduncus*; Sealion, California sea lion *Zalophus californianus*.

clock-like by performing the SH test and Tajima's relative rate test as described above, and we obtained estimates of divergence times of several nodes that were the same as those in the BI tree using all 616 nt (Table 2). The estimates were notably younger than those in the BI tree using 616 nt, except those for *DRB*/(*DQB*, *DPB*) and primate/fereuungulates in the *DOB* clade.

The Kimura distance was then employed to construct linearized trees using positions 1-616 and 219-616 to estimate divergence times of the same nodes in the BI trees. The bird/mammal divergence time (310 Mya) was also used as a calibration point. Similarly to the scenario in the BI trees, most of the divergence times in the NJ tree using position 219-616 were younger than those in the NJ tree using position 1-616 (Table 2). Moreover, in trees using the same nucleotide positions (1-616 or 219-616), estimates of the divergence time between the *DOB* and other mammal clades and between the *DRB* and *DQB*/*DPB* clades in the NJ tree ware close to those in the BI tree, but those of *DQB*/*DPB*, rodents/other mammals, primates/fereuungulates, and carnivores/cetartiodactyls in the NJ tree were older than those in the BI tree. In addition, estimates of the dolphin/artiodactyl divergence time for the *DQB* and *DRB* genes in the NJ tree were slightly older than those in the BI tree. For the *T. truncatus*/*T. aduncus* divergence time, estimates for the *DQB* and *DRB* genes in the NJ and BI trees were nearly the same.

DISCUSSION

Klein and O'hUigin (1994) demonstrated that substitutions of PBR sequences of MHC genes

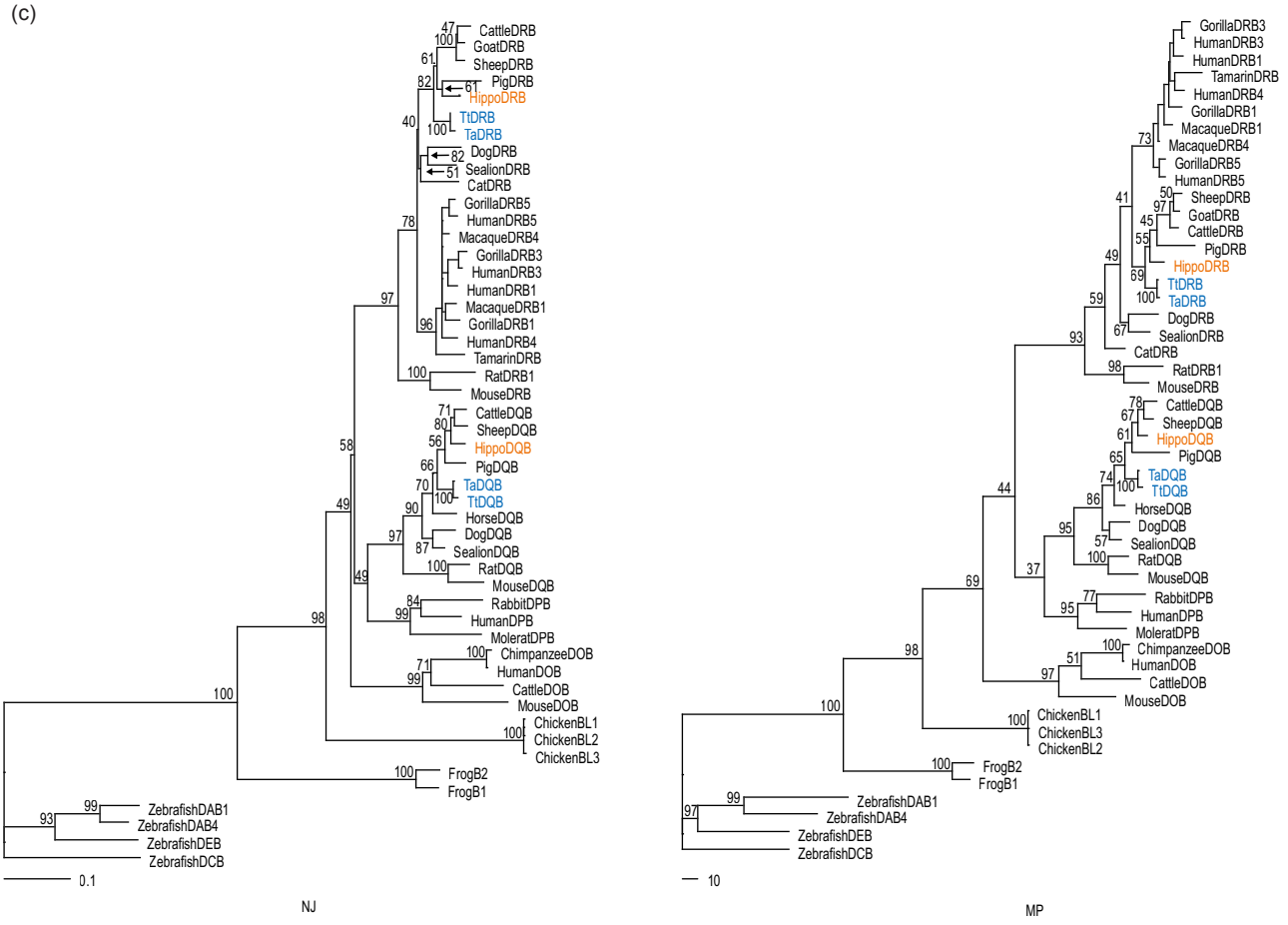


Fig. 3. Neighbor-joining (NJ), maximum-parsimony (MP), and maximum-likelihood (ML) trees for major histocompatibility complex (MHC) class II β-chain genes using all codon positions. The nucleotide positions for tree construction were (a) positions 1-616, (b) 1-218, and (c) 219-616. The ML tree was constructed only using nucleotide positions 1-616. The bootstrap values are shown in parentheses for the nodes of interest. Tt, *Tursiops truncatus*; Ta, *T. aduncus*; Sealion, California sea lion *Zalophus californianus*.

occur in a trans-specific manner. They suggested that closely related species often share some of the same pathogens which only differ in the strain, and some pathogens have co-evolved trans-specifically in parallel with their hosts. In our study, a similar scenario was generally observed for phylogenetic relationships based on PBR sequences (Fig. 2b). However, the distinct grouping pattern of dolphin *DQB* and *DRB* genes from those of other artiodactyls was noted despite dolphins and artiodactyls being closely related. Possible explanations for this pattern include analytical artifacts and an evolutionary hypothesis of convergent evolution. Analytical artifacts were considered unlikely because the expected homologous relationship of the *DQB* and *DRB* sequences of other mammals was shown in the same phylogenetic tree. Convergent evolution may

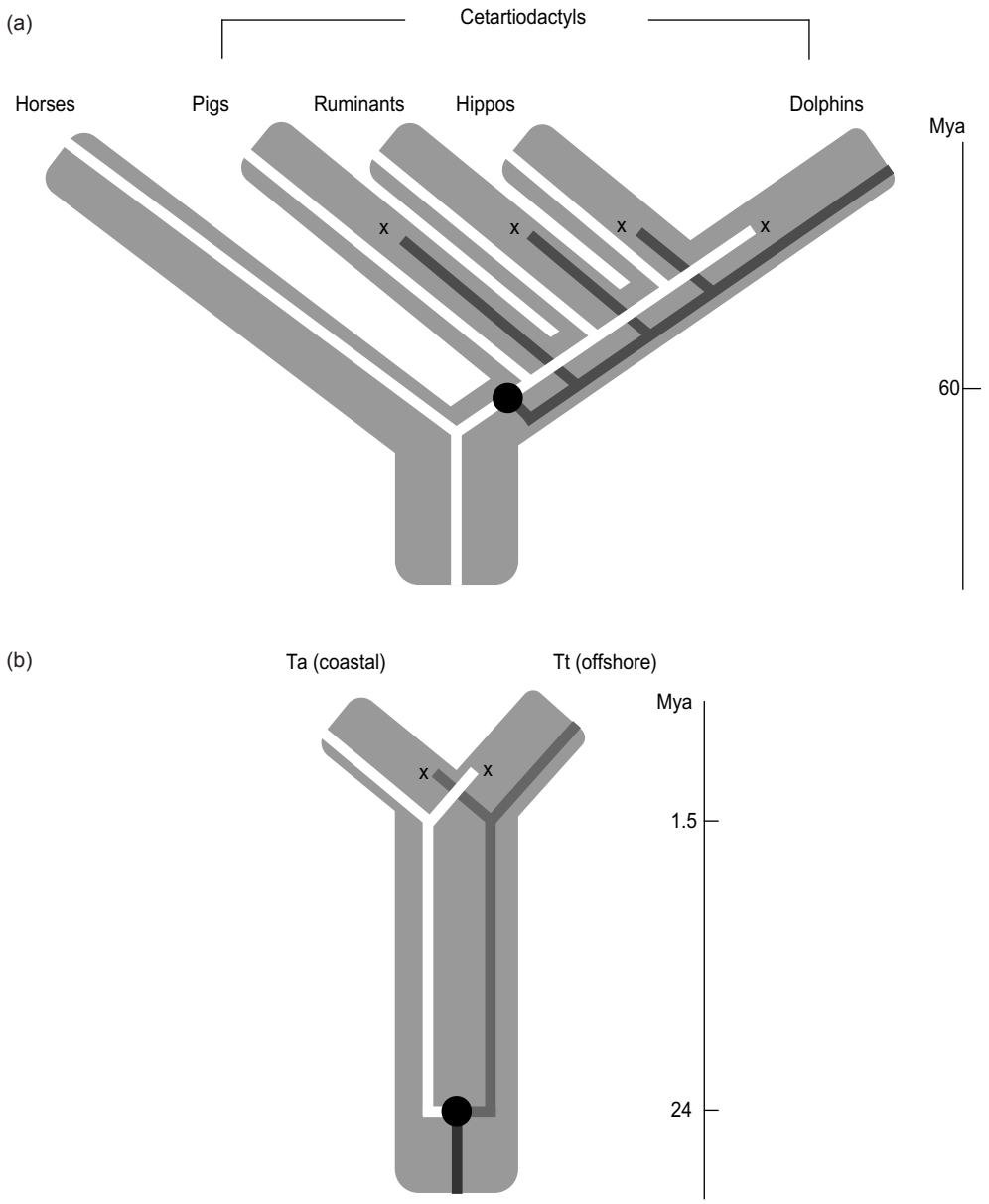


Fig. 4. Hypothesized birth-and-death process of major histocompatibility complex (MHC) class II *DQB* and *DRB* genes in (a) cetartiodactyls and (b) bottlenose dolphins. The broad trees represent the species tree, and the thin trees represent the MHC gene tree. The times were taken from summary times in table 2 except the time of speciation of these 2 bottlenose dolphins which was estimated by Wang et al. (1999b). •, gene duplication; X, gene was lost or became a pseudogene. Tt, *Tursiops truncatus*; Ta, *T. aduncus*; Mya, million years ago.

be relatively common in the MHC for which several mechanisms are potentially responsible: point mutations, large-scale intergenic recombinations, small-scale intergenic conversions, and strong selection (Yeager and Hughes 1999). The last mechanism is the most likely because it was shown that MHC alleles observed in populations have been selectively favored (Hughes et al. 1993), and no evidence for genetic recombination or gene conversion events between sequences was found in our dataset using the program GENECONV. This unusual grouping pattern of dolphin *DQB* and *DRB* sequences is probably due to the requirement that both genes defend against marine pathogens, which is evidence for permanent adaptations to marine environments by cetaceans. In sum, we showed the homoplasy of the PBR between *DQB* and *DRB* genes, which can best be explained by convergent evolution. A similar conclusion was reached in the humpback whale (Baker et al. 2006) and canids (Seddon and Ellegren 2002). Therefore, the PBR alone might not be a good indicator of the phylogenetic history of MHC genes because it is prone to homoplasy in different loci and species. The evolutionary history of the PBR in cetartiodactyl MHC genes could be a good example of pathogen-driven directional selection, which infers the varying selective advantages of MHC alleles in accordance to the diversity and abundance of pathogens in different environments (Bernatchez and Landry 2003). Potential homoplasy was noted in sequences from the California sea lion (Fig. 2b) although the support was moderate (0.71). Since the unique grouping pattern of PBR sequences was only found in marine mammals in this study, more sequences from other species are needed for further comparative investigations of PBR evolution in marine mammals.

We compared estimates of divergence times obtained using either non-PBR sites or all sites in the BI and NJ trees with previous MHC and multigene studies (Kumar and Hedges 1998, Takahashi et al. 2000). It was suggested that reliable estimates of divergence times for MHC genes can be obtained by using all sites with a birth-death clock model in the BI tree. Employing the birth-death clock model is likely more appropriate for estimating divergence times of MHC genes in view of the fact that MHC genes have evolved through a birth-and-death process (Nei et al. 1997). Furthermore, many studies on phylogenies and divergence times of mammalian MHC class I or II genes used longer nucleotide sequences, including the PBR and other regions, for phylogenetic reconstruction (Klein and Figueroa 1986, Hughes and Nei 1990, Takahashi et al. 2000, Piontkivska and Nei 2003). Inclusion of the PBR in these analyses and our study still provided plausible estimates of divergence times, in support of the previous suggestion for the use of as many nucleotide sites as possible (Nei and Kumar 2000).

The BI trees using all sites or non-PBR sites both showed that phylogenetic relationships in the respective cetartiodactyl group in the *DQB* and *DRB* clades did not correspond to those in previously accepted species trees. It was striking to observe that cetaceans (bottlenose dolphins) and artiodactyls (pig, hippo, and ruminants) formed 2 distinct clades in both the *DQB* and *DRB* phylogenies, rather than being of the same clade with the hippo and dolphin as the closest relatives (Gatesy and O'Leary 2001). The general lack of a match between the gene and species trees was also reported for MHC class I genes of cetartiodactyls (Holmes et al. 2003), although no

Table 2. Estimates of the times of divergence between different β-chain gene clades

	DOB/ (*DRB*, *DQB*, *DPB*)	*DRB*/ (*DQB*, *DPB*)	*DQB*/*DPB*	Rodents/other mammals (four gene clades)	Primates/ ferungulates (*DOB*; *DRB*)	Carnivores/ cetartiodactyls (*DQB*; *DRB*)	Dolphins/ artiodactyls (*DQB*; *DRB*)	Tt/Ta* (*DQB*; *DRB*)
Approach (nucleotide positions)								
BI (1-616)	249.36	191.79	168.26	109.30 - 122.03	97.53; 88.43	78.14; 76.45	60.03; 62.54	26.47; 21.58
BI (219-616)	220.04	193.52	157.31	83.43 - 111.22	97.41; 77.97	62.74; 51.09	38.42; 44.99	7.38; 9.31
NJ (1-616)	263 ± 12	195 ± 11	190 ± 17	126 - 147	138 ± 11; 92 ± 9	84 ± 9; 91 ± 8	67 ± 6; 68 ± 4	25 ± 3; 22 ± 3
NJ (219-616)	207 ± 8	194 ± 12	185 ± 13	118 - 145	120 ± 10; 75 ± 9	61 ± 3; 76 ± 5	47 ± 7; 49 ± 3	7 ± 2; 6 ± 2
Takahashi et al.•	254 ± 16	189 ± 12	179 ± 12	N/A	N/A	N/A	N/A	N/A
Kumar and Hedges◦	N/A	N/A	N/A	114	92	79	58	N/A

The numbers after the ± signs are standard errors. Tt, *Tursiops truncatus*; Ta*, *Tursiops aduncus*. •From Table 3 in Takahashi et al. 2000 ◦From figure 2 and 3 in Kumar and Hedges 1998.

sequence was included from cetaceans or hippo in that study. Possible sources contributing to the incongruence between the MHC gene tree and species tree include (1) an incorrect model of sequence evolution, (2) different evolutionary rates across the tree, and (3) utilization of paralogous sequences (Page and Holmes 1998). The 1st source is unlikely in our case because we obtained very similar phylogenetic relationships using different models with different algorithms. The possibility of the 2nd source can be excluded since our dataset was clock-like based on the likelihood-ratio and relative-rate tests. The 3rd source is very likely to be important because the gene tree can be discordant with the species tree if the species relationships are based on mixed orthologous and paralogous sequences (Page and Holmes 1998), and it has not yet been confirmed whether the respective *DQB* and *DRB* sequences of cetartiodactyls are paralogous or orthologous. While orthologous genes are separated by speciation events, paralogous genes are separated by gene duplication events, which were proposed as being a major force in MHC evolution (Klein et al. 1998). Gene duplications were observed in the genetic organization of MHC genes in many mammals. In bovine MHC class II genes, for example, 2 *DQB* genes and 9 *DRB* genes were detected, with eight of the *DRB* genes likely being pseudogenes (Ellis and Ballingall 1999). In conclusion, we presumed that the sequences of cetaceans and artiodactyls were paralogous in *DQB* and *DRB* genes, respectively. We explained the paralogy between sequences of cetaceans and artiodactyls by postulating gene duplications in MHC genes and pathogen-driven directional selection during cetacean evolution.

Gene-duplication events are hypothesized to have occurred at least once in both the *DQB* and *DRB* genes, prior to the cetacean/artiodactyl divergence (Fig. 4a). Gene duplication gave rise to 2 sets of paralogous genes. Given the 2 sets named A and B sets, the A sets of all artiodactyls have survived to the present day, while the B sets were lost or became pseudogenes; in contrast, the B sets in cetaceans were preserved, while their A sets were lost or became pseudogenes. Through this process, we obtained the phylogenetic relationship that cetaceans and artiodactyls are sistergroups in the *DQB* and *DRB* phylogenies. According to the estimates in table 2, the duplications took place 60 Mya or slightly earlier, which is comparable to the 1st appearances of fossil cetaceans at around 53.5 Mya, artiodactyls at 55 Mya, and another molecular estimate of the divergence time of cetaceans/artiodactyls at 60 Mya (Arnason and Gullberg 1996, Arnason et al. 2000, Arnason et al. 2004, Theodor 2004). As already pointed out, gene duplication is a common feature of MHC genes, and duplicated MHC loci are usually governed by diversifying selection, which may favor their specialization (Hughes 1999). Since natural selective pressures of infectious diseases between terrestrial and aquatic (especially marine) environments differ (McCallum et al. 2004), pathogen-driven directional selection is a plausible driving force for preserving or losing MHC genes by cetaceans and their terrestrial relatives in the period during which cetaceans progressively moved from land to water. Several other mammal groups also made the evolutionary transition from land to sea, such as pinnipeds, sea otters, polar bears, and sirenians. Studying MHC genes of these marine mammals and their terrestrial relatives should provide further insights into the evolution of MHC genes.

We selected *T. truncatus* and *T. aduncus*, which are closely related but inhabit different environments, to investigate MHC gene evolution in toothed whales. *Tursiops truncatus* (Montagu 1821) is a larger form of bottlenose dolphin, which is generally found in deep, offshore waters. The smaller form, *T. aduncus*, inhabits shallow, tropical, coastal waters (Ehrenberg 1832). Shallow waters along the coast may be influenced by terrestrial runoff, and the pathogen diversity and abundance in coastal waters may differ from those in oceanic areas. An investigation of stomach contents of stranded and by-caught specimens in Taiwan revealed that the diet and microflora presumably differ between *T. truncatus* and *T. aduncus* (Wang 2003), indicating different selective pressures from pathogens between these 2 species. Yang et al. (2007) reported the molecular characterization of *DQB* and *DRB* genes in bottlenose dolphins and found a clear phylogenetic division between *T. truncatus* and *T. aduncus*. This was an intriguing result compared to the general trans-specific pattern of evolution observed for MHC loci (Hughes 1999), i.e., different species and different genera are clustered within the same allelic lineages. It was assumed that the species-specific MHC allelic lineages in *Tursiops* resulted from different selective pressures from pathogens which exist in oceanic (*T. truncatus*) and coastal (*T. aduncus*) environments. Nonetheless, it was still not clear when the species-specific allelic lineages emerged and the subsequently evolutionary processes

occurred. The general evolutionary understanding is that species pairs of closely related delphinids (true dolphins), such as *T. truncatus* and *T. aduncus*, likely evolved during Pleistocene glaciation events (< 2 Mya; Wang et al. 1999b). However, our estimated divergence time of *T. truncatus* and *T. aduncus* according to the *DQB* and *DRB* genes (>20 Mya) was much earlier than the separation date of these 2 species, and the earliest fossils identifiable as *Tursiops* date to only 4-7 Mya (Barnes 1990), as well as the emergence of the oldest delphinid which is possible 11 Mya in the late Miocene (Barnes 1977). After conducting Tajima's relative rate test, we ruled out a possible error for the time estimate caused by accelerating substitution rates of *Tursiops*. The significant time gap between speciation and MHC gene estimates can be explained by postulating a birth-and-death process (Fig. 4b), similar to that in the cetartiodactyl clade we described above. These allelic lineages of *Tursiops* MHC genes may have emerged by gene duplication during the period of early radiation of small toothed whales (from the late Oligocene to early Miocene, ~22 Mya (Arnason et al. 2004). It was shown that MHC alleles can persist over extremely long time periods by balancing selection (Bernatchez and Landry 2003). Long persistence times allow the accumulation of multiple substitutions between allelic lineages of MHC genes, which may enable them to possess different assignments for pathogen defense. Baker et al. (2006) suggested the scenario that baleen whales and early divergent toothed whales having duplicate *DQB* genes is an ancestral condition shared with ruminants, while the more-derived cetaceans such as true dolphins (of the family Delphinidae) lost this condition and only have 1 *DQB* locus. The ancestor of *Tursiops* was possibly equipped with these 2 ancestral lineages at different loci, and these lineages were maintained for a long evolutionary period through speciation events. *Tursiops truncatus* and *T. aduncus* consequently kept only one of the 2 lineages expressed. Pathogen-driven directional selection may have been responsible for *T. truncatus* and *T. aduncus* maintaining different expressed lineages due to a heterogeneity of natural selective pressures by infectious diseases found in oceanic and coastal waters. Since the extant cetacean fauna consists of more than 80 species which live in various habitats, such as oceans, estuaries, polar regions, and rivers, the evolution of cetacean MHC genes may be more complicated than previously realized. It would be interesting to elucidate this by obtaining more sequences and loci from a variety of cetacean species.

Our phylogenetic analysis and divergence time estimates in cetartiodactyl MHC class II genes provide a new aspect for discussing the history of co-evolution between MHC genes and pathogens when cetaceans moved from land to water. To the best of our knowledge, this is the first study on the evolutionary history of cetartiodactyls using MHC class II genes. More sequences and loci from other cetacean species and closely related mammals should enable us to further study the birth-and-death process of cetartiodactyl MHC genes. Comparisons with MHC evolution of other marine mammals would also help elucidate the factors affecting the immune defense system in relation to evolutionary change of foreign antigens. Furthermore, our results show that cetacean MHC genes are adapted to marine environments. Their ability to defend against terrestrial pathogens needs investigation and close monitoring, especially in these times when there are potential risks of epidemics in cetaceans, when they have more occasions to encounter terrestrial pathogens through human exploitation of marine environments or, directly, by keeping dolphins in captivity.

Acknowledgments: This study could not be conducted without blood samples of bottlenose dolphins provided by Hong Kong Ocean Park and Hualien Ocean Park and tissue samples from a hippopotamus provided by Leofoo Village Theme Park and Taipei City Zoo. We appreciate colleagues in the laboratories of JMH and LSC for useful comments and technical support. We are grateful to many scholars at the Fifth Conference on Secondary Adaptation of Tetrapods to Life in Water held in Tokyo in 2008 for providing constructive comments on this paper. This study was funded by grants to LSC from the Council of Agriculture of Taiwan (94AS-9.1.7-FB-e1(8) and 95AS-11.1.3-FB-e1(10)) and Mission Biotech Corp. of Taiwan.

REFERENCES

Arnason U, A Gullberg. 1996. Cytochrome *b* nucleotide sequences and the identification of five primary lineages of extant cetaceans. Mol. Biol. Evol. **13:** 407-417.

Arnason U, A Gullberg, S Gretarsdottir, B Ursing, A Janke. 2000. The mitochondrial genome of the sperm whale and a new molecular reference for estimating eutherian divergence dates. J. Mol. Evol. **50:** 569-578.

Arnason U, A Gullberg, A Janke. 2004. Mitogenomic analyses provide new insights into cetacean origin and evolution. Gene **333:** 27-34.

Bajpai S, PD Gingerich. 1998. A new Eocene archaeocete (Mammalia, Cetacea) from India and the time of origin of whales. Proc. Natl. Acad. Sci. USA **95:** 15464-15468.

Baker CS, MD Vant, ML Dalebout, GM Lento, SJ O'Brien, N Yuhki. 2006. Diversity and duplication of *DQB* and *DRB*-like genes of the MHC in baleen whales (suborder: Mysticeti). Immunogenetics **58:** 283-296.

Barnes L. 1977. Outline of eastern north Pacific cetacean assemblages. Syst. Zool. **25:** 321-343.

Barnes L. 1990. The fossil record and evolutionary relationships of the genus *Tursiops*. *In* S Leatherwood, R Reeves, eds. The bottlenose dolphin. San Diego, CA: Academic Press, pp. 3-26.

Bejder L, BK Hall. 2002. Limbs in whales and limblessness in other vertebrates: mechanisms of evolutionary and developmental transformation and loss. Evol. Develop. **4:** 445-458.

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

Boisserie JR, F Lihoreau, M Brunet. 2005. The position of Hippopotamidae within Cetartiodactyla. Proc. Natl. Acad. Sci. USA **102:** 1537-1541.

Bowen L, BM Aldridge, F Gulland, W Van Bonn, R DeLong, S Melin, LJ Lowenstine, JL Stott, ML Johnson. 2004. Class II multiformity generated by variable MHC-*DRB* region configurations in the California sea lion (*Zalophus californianus*). Immunogenetics **56:** 12-27.

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

Ellis SA, KT Ballingall. 1999. Cattle MHC: evolution in action? Immunol. Rev. **167:** 159-168.

Gatesy J. 1997. More DNA support for a Cetacea/Hippopotamidae clade: the blood-clotting protein gene gamma-fibrinogen. Mol. Biol. Evol. **14:** 537-543.

Gatesy J, C Hayashi, MA Cronin, P Arctander. 1996. Evidence from milk casein genes that cetaceans are close relatives of hippopotamid artiodactyls. Mol. Biol. Evol. **13:** 954-963.

Gatesy J, MA O'Leary. 2001. Deciphering whale origins with molecules and fossils. Trends Ecol. Evol. **16:** 562-570.

Hedges SB, PH Parker, CG Sibley, S Kumar. 1996. Continental breakup and the ordinal diversification of birds and mammals. Nature **381:** 226-229.

Holmes EC, AFC Roberts, KA Staines, SA Ellis. 2003. Evolution of major histocompatibility complex class I genes in Cetartiodactyls. Immunogenetics **55:** 193-202.

Hughes AL. 1999. Adaptive evolution of genes and genomes. Oxford, UK: Oxford Univ. Press.

Hughes AL, M Hughes, D Watkins. 1993. Contrasting roles of interallelic recombination at the HLA-A and HLA-B loci. Genetics **133:** 669-680.

Hughes AL, M Nei. 1990. Evolutionary relationships of class II major-histocompatibility-complex genes in mammals. Mol. Biol. Evol. **7:** 491-514.

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

Klein J, F Figueroa. 1986. Evolution of the major histocompatibility complex. Crit. Rev. Immunol. **6:** 295-386.

Klein J, C O'hUigin. 1994. MHC polymorphism and parasites. Phil. Trans. R. Soc. Lond. B Biol. Sci. **346:** 351-357.

Klein J, A Sato. 1998. Birth of the major histocompatibility complex. Scand. J. Immunol. **47:** 199-209.

Klein J, A Sato, C O'hUigin. 1998. Evolution by gene duplication in the major histocompatibility complex. Cytogenet. Cell Genet. **80:** 123-127.

Kumar S, SB Hedges. 1998. A molecular timescale for vertebrate evolution. Nature **392:** 917-920.

Kumar S, K Tamura, M Nei. 2004. MEGA3: integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Brief. Bioinform. **5:** 150-163.

Kundu S, CG Faulkes. 2004. Patterns of MHC selection in African mole-rats, family Bathyergidae: the effects of sociality and habitat. Proc. Biol. Sci. **271:** 273-278.

McCallum H, A Kuris, C Harvell, K Lafferty, G Smith, J Porter. 2004. Does terrestrial epidemiology apply to marine systems? Trends Ecol. Evol. **19:** 585-591.

McClellan DA, EJ Palfreyman, MJ Smith, JL Moss, RG Christensen, JK Sailsbery. 2005. Physicochemical evolution and molecular adaptation of the cetacean and artiodactyl cytochrome *b* proteins. Mol. Biol. Evol. **22:** 437-455.

Montgelard C, FM Catzeflis, E Douzery. 1997. Phylogenetic relationships of artiodactyls and cetaceans as deduced from the comparison of cytochrome *b* and 12S rRNA mitochondrial sequences. Mol. Biol. Evol. **14:** 550-559.

Murphy WJ, E Eizirik, WE Johnson, YP Zhang, OA Ryder, SJ O'Brien. 2001. Molecular phylogenetics and the origins of placental mammals. Nature **409:** 614-618.

Naylor GJ, M Gerstein. 2000. Measuring shifts in function and evolutionary opportunity using variability profiles: a case study of the globins. J. Mol. Evol. **51:** 223-233.

Nei M, X Gu, T Sitnikova. 1997. Evolution by the birth-and-death process in multigene families of the vertebrate immune system. Proc. Natl. Acad. Sci. USA **94:** 7799-7806.

Nei M, S Kumar. 2000. Molecular evolution and phylogenetics. Oxford, UK: Oxford Univ. Press.

Nikaido M, AP Rooney, N Okada. 1999. Phylogenetic relationships among cetartiodactyls based on insertions of short and long interpersed elements: hippopotamuses are the closest extant relatives of whales. Proc. Natl. Acad. Sci. USA **96:** 10261-10266.

Nummela S, JG Thewissen, S Bajpai, ST Hussain, K Kumar. 2004. Eocene evolution of whale hearing. Nature **430:** 776-778.

Page RDM, EC Holmes. 1998. Molecular evolution: a phylogenetic approach. Oxford, UK: Blackwell Science.

Piontkivska H, M Nei. 2003. Birth-and-death evolution in primate MHC class I genes: divergence time estimates. Mol. Biol. Evol. **20:** 601-609.

Posada D, TR Buckley. 2004. Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. Syst. Biol. **53:** 793-808.

Posada D, KA Crandall. 1998. Modeltest: testing the model of DNA substitution. Bioinformatics **14:** 817-818.

Rodríguez F, JF Oliver, A MarÍn, JR Medina. 1990. The general stochastic model of nucleotide substitution. J. Theor. Biol. **142:** 485-501.

Ronquist F, JP Huelsenbeck. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models.

Bioinformatics **19:** 1572-1574.
Sanderson MJ. 2003. r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. Bioinformatics **19:** 301-302.
Sawyer S. 1999. GENECONV: a computer package for the statistical detection of gene conversion. St. Louis, MO: Department of Mathematics, Washington Univ.
Schook L, S Lamont. 1996. The major histocompatibility complex region of domestic animal species. Boca Raton, FL: CRC Press.
Seddon JM, H Ellegren. 2002. MHC class II genes in European wolves: a comparison with dogs. Immunogenetics **54:** 490-500.
Shimodaira H, M Hasegawa. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. **16:** 1114-1116.
Spoor F, S Bajpai, ST Hussain, K Kumar, JG Thewissen. 2002. Vestibular evidence for the evolution of aquatic behaviour in early cetaceans. Nature **417:** 163-166.
Swofford DL. 2003. PAUP*. Phylogenetic analysis using parsimony (*and other methods). Sunderland, MA: Sinauer Associates.
Swofford DL, GJ Olsen, PJ Waddell, DM Hillis. 1996. Phylogenetic inference. *In* DM Hillis, C Moritz, BK Mable, eds. Molecular systematics. Sunderland, MA: Sinauer Associates, pp. 407-514.
Tajima F. 1993. Simple methods for testing molecular clock hypothesis. Genetics **135:** 599-607.
Takahashi K, AP Rooney, M Nei. 2000. Origins and divergence times of mammalian class II MHC gene clusters. J. Hered. **91:** 198-204.
Theodor JM. 2004. Molecular clock divergence estimates and the fossil record of Cetartiodactyla. J. Paleontol. **78:** 39-44.
Thewissen JGM, EM Williams. 2002. The early radiations of Cetacea (Mammalia): evolutionary pattern and developmental correlations. Annu. Rev. Ecol. Syst. **33:** 73-90.
Thewissen JGM, EM Williams, LJ Roe, ST Hussain. 2001. Skeletons of terrestrial cetaceans and the relationship of whales to artiodactyls. Nature **413:** 277-281.
Thompson JD, DG Higgins, TJ Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22:** 4673-4680.
Wagner JL, RC Burnett, R Storb. 1999. Organization of the canine major histocompatibility complex: current perspectives. J. Hered. **90:** 35-38.
Wang DY, S Kumar, SB Hedges. 1999a. Divergence time estimates for the early history of animal phyla and the origin of plants, animals and fungi. Proc. Biol. Sci. **266:** 163-171.
Wang JY, LS Chou, BN White. 1999b. Mitochondrial DNA analysis of sympatric morphotypes of bottlenose dolphins (genus: *Tursiops*) in Chinese waters. Mol. Ecol. **8:** 1603-1612.
Wang MC. 2003. Feeding habits, food resource partitioning and guild structure of odontocetes in Taiwanese waters. PhD dissertation. Institute of Zoology, National Taiwan University, Taipei, Taiwan, pp. 148.
Yang WC, LS Chou, JM Hu. 2007. Molecular characterization of major histocompatibility complex class II *DQB* and *DRB* genes in bottlenose dolphins (*Tursiops truncatus* and *T. aduncus*) from the western Pacific. Zool. Stud. **46:** 664-679.
Yang Z. 1994. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. J. Mol. Evol. **39:** 306-314.
Yeager M, A Hughes. 1999. Evolution of the mammalian MHC: natural selection, recombination, and convergent evolution. Immunol. Rev. **167:** 45-58.
Yuhki N, SJ O'Brien. 1997. Nature and origin of polymorphism in feline MHC class II DRA and *DRB* genes. J. Immunol. **158:** 2822-2833.