



Université
de Toulouse

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par *l'Université Toulouse III - Paul Sabatier*

Discipline ou spécialité : *Ecologie et Evolution des populations et communautés*

Présentée et soutenue par *Matthias Macé*
Le 30 septembre 2009

Titre :

*Le Dauphin bleu et blanc (Stenella caeruleoalba) en Méditerranée :
de la vicariance à l'épisode à Morbillivirus*

JURY

M. Pierre Pontarotti	Directeur de Recherches (DR1) CNRS	Rapporteur
M. Michel André	Professeur – Universitat Politècnica de Catalunya	Rapporteur
M. Yves Lignereux	Professeur – Ecole Nationale Vétérinaire de Toulouse	Examineur
Mlle Morgane Gibert	Chargé de Recherches (CR1) CNRS	Examineur
M. Alexandre Ribéron	Maître de Conférences – Université Paul Sabatier	Examineur
Mme Brigitte Crouau-Roy	Professeur – Université Paul Sabatier	Directeur de thèse

Ecole doctorale : *Ecole Doctorale Sciences Ecologiques, Vétérinaires, Agronomiques et Bioingénieries*

Unité de recherche : *Evolution et Diversité Biologique, UMR5174*

Directeur(s) de Thèse : *Pr Brigitte Crouau-Roy*

Note à l'attention des lecteurs de la version PDF : des liens permettent la navigation au sein du document (en général, à la suite d'un « cf. » et en gras, ou les liaisons vers les articles dans l'avant-propos). Vous pourrez les visualiser en « survolant » certaines sections. En cliquant dessus, vous serez transféré à la section liée. De même, vous pourrez accéder à la référence bibliographique en cliquant sur la citation dans le texte. Des liens externes dans la liste bibliographique permettent d'accéder directement à l'article en ligne lorsque cela est possible.

SOMMAIRE

AVANT-PROPOS	7
CONTEXTE BIBLIOGRAPHIQUE	16
I. QUELQUES NOTIONS D'ÉCOLOGIE	18
I. A. DEFINITIONS GENERALES	18
A.1) Espèce, population	18
A.2) Dynamique d'une population	20
A.3) Le concept de niche écologique	21
A.4) Compétition entre génotypes, mutation-sélection	22
A.5) Equilibre de Hardy-Weinberg et dérive génétique	26
I. B. SPECIATION ALLOPATRIQUE, VICARIANCE ET DISPERSION	27
B.1) Différents modes de spéciation	27
B.2) Spéciation allopatrique	29
I. C. NOTION DE « SERVICES RENDUS » PAR LES ECOSYSTEMES ET DIMINUTION DES RESSOURCES	30
C.1) Notion de services écologiques	30
C.2) Place des Prédateurs dans les écosystèmes, conséquences de leur diminution, cas des Cétacés	33
II. LE MODELE BIOLOGIQUE : LE DAUPHIN BLEU ET BLANC, STENELLA CŒRULEOALBA (MEYEN, 1833)	36
II. A. SYSTEMATIQUE, TAXONOMIE, PALEONTOLOGIE, EVOLUTION	36
A.1) Résumé systématique	36
A.2) Les Cétacés, position relative des Cetartiodactyla	37
A.3) Les Delphinidæ et le clade Delphinus/Stenella/Tursiops	39
A.4) Le Dauphin bleu et blanc, taxonomie-synonymie	42
II. B. LE DAUPHIN BLEU ET BLANC : ASPECTS BIOLOGIQUES	45
B.1) Morphologie, détermination	45
B.2) Données physiologiques	46
B.3) Distribution, migrations	48
II. C. CONSERVATION DU DAUPHIN BLEU ET BLANC.	52
C.1) Pacifique Nord-ouest (IUCN, 1996).	52
C.2) Méditerranée.	53
III. LA ZONE D'ETUDE : LA MER MEDITERRANEE ET LE PROCHE ATLANTIQUE	54
III. A. DONNEES PHYSIQUES (DIVERS, 1999)	54
A.1) Géologie	54
A.2) Hydrographie	55
III. B. ECOLOGIE	56
B.1) Le talus continental est une zone trophique importante alors que les fonds sont désertiques	56
B.2) La biodiversité a dû s'adapter à des conditions changeantes	57
B.3) La Méditerranée héberge quelques Vertébrés remarquables	60
III. C. MENACES PESANT SUR L'ENVIRONNEMENT MEDITERRANEEN.	61
C.1) Une menace non ou marginalement anthropique : le phénomène de tropicalisation	61
C.2) Menaces d'origine anthropique	63
IV. LE COMPLEXE MAJEUR D'HISTOCOMPATIBILITE (C.M.H.)	65
IV. A. MISE EN EVIDENCE DU ROLE DE RECONNAISSANCE DU « SOI MODIFIE »	66
IV. B. HISTOIRE EVOLUTIVE DU COMPLEXE MAJEUR D'HISTOCOMPATIBILITE	67
B.1) Emergence du système immunitaire adaptatif	67

B.2) Histoire naturelle des gènes impliqués dans la présentation de l'antigène par le C.M.H. de classe I et II	70
IV. C. LA REGION DE CLASSE II	71
C.1) Description	71
C.2) Présentation de l'antigène	72
C.3) Autres fonctions	74
C.4) Conséquences de la présentation de l'antigène par les molécules de Classe II	74
C.5) La molécule DRB	75
V. L'ÉPIZOOTIE A DOLPHIN MORBILLIVIRUS (D.M.V., PARAMYXOVIRIDÆ)	79
V. A. EPIDEMIOLOGIE DESCRIPTIVE	79
V. B. TABLEAU ANATOMO-CLINIQUE	81
V. C. EPIDEMIOLOGIE ANALYTIQUE	82
V. D. ETIO-PATHOGENIE : EMERGENCE DES MORBILLIVIRUS DE MAMMIFERES MARINS	83
V. E. EVOLUTION DES MORBILLIVIRUS	84

RECHERCHE D'UN MARQUEUR DE SEQUENCE, POTENTIELLEMENT EQUIVALENT A LA REGION DE CONTROLE DE L'ADN MITOCHONDRIAL, SUR LE CHROMOSOME Y CHEZ LES CETARTIODACTYLES **85**

I. INTRODUCTION	87
I. A. UTILISATION DE MARQUEURS MOLECULAIRES EN ECOLOGIE, GENETIQUE DES POPULATIONS ET EVOLUTION	87
A.1) L'ADN mitochondrial	87
A.2) <i>Des marqueurs prometteurs portés par le chromosome Y</i>	87
I. B. MEDECINE LEGALE ET BIOTECHNOLOGIES	90
I. C. AMELOGENINE	91
C.1) Aspects physiologiques	91
C.2) Evolution	92
I. D. LE MODELE BIOLOGIQUE ETUDIE : LES <i>CETARTIODACTYLA</i>	92
II. METHODOLOGIE	94
II. A. LA CLADISTIQUE	94
II. B. LE CODAGE DES INDELS DANS LES ALIGNEMENTS DE SEQUENCE	97
B.1) Les algorithmes d'alignement	97
B.2) Codage des indels : la méthode MCIC	98
II. C. MODELES DE MUTATION	101
C.1) Qu'est-ce qu'un modèle de mutation ?	101
C.2) Caractéristiques, horloge moléculaire	102
C.3) Aspect mathématique des modèles de substitution	102
C.4) Modèles les plus courants d'évolution de l'ADN	105
II. D. INFERENCE PHYLOGENETIQUE	108
D.1) Principe	109
D.2) Les trois grandes familles de méthodes	109
III. MATERIEL ET METHODES	112
III. A. BIOLOGIE MOLECULAIRE	112
III. B. ANALYSES BIOINFORMATIQUES	112
IV. OBJECTIFS ET PRINCIPAUX RESULTATS	113
IV. A. OBJECTIFS	113
IV. B. PRINCIPAUX RESULTATS	114
V. PERSPECTIVES	124

ÉTUDE DE LA VARIABILITE GENETIQUE DU DAUPHIN BLEU ET BLANC EN MEDITERRANEE **125**

I. INTRODUCTION	127
I. A. QUELQUES OBSERVATIONS NATURALISTES	127

I. B.	TRAVAUX PRECEDENTS CONCERNANT LA MORPHOLOGIE ET LE DEVELOPPEMENT	127
I. C.	TRAVAUX PRECEDENTS DE GENETIQUE DES POPULATIONS	127
C.1)	Inbreeding et épizootie	127
C.2)	Structure génétique inter-océanique	129
I. D.	ETUDES PRELIMINAIRES	129
D.1)	Polymorphisme trans-spécifique chez les petits Delphinidæ	129
D.2)	Recherche de microsatellites utilisables chez les Cétacés et le Dauphin bleu et blanc en particulier	130
II.	MATERIEL	137
III.	VARIABILITE INTER-OCEANIQUE	137
IV.	POLYPHYLIE ET CRISE DE SALINITE DU MESSINIEN	145
IV. A.	METHODES	145
A.1)	Extraction	145
A.2)	Amplification et séquence	145
IV. B.	METHODOLOGIE	146
B.1)	Inférence statistique par Vraisemblance (L) et Critère d'Information d'Akaiké (AIC)	146
B.2)	Test hiérarchique de modèles de substitution (ModelTest)	149
B.3)	Inférence bayésienne	150
B.4)	Théorie de la Coalescence	152
B.5)	Une méthode d'inférence démographique non-paramétrique : le Bayesian Skyline Plot	155
IV. C.	ETUDE DE LA POLYPHYLIE DU DAUPHIN BLEU ET BLANC EN MEDITERRANEE	159
C.1)	Introduction et hypothèses testées	159
C.2)	Principaux résultats	160
C.3)	Conclusions	205

ÉTUDE DU DRB CHEZ LE DAUPHIN BLEU ET BLANC ET PERSPECTIVES GENERALES

207

I.	INTRODUCTION ET BUTS DE L'ETUDE	209
I. A.	DIVERSITE DU CMH ET <i>BALANCING SELECTION</i>	209
A.1)	Evolution non-neutre et polymorphisme élevé	209
A.2)	Le polymorphisme trans-spécifique	210
I. B.	EVOLUTION MEDIEE PAR LES PATHOGENES	210
B.1)	Cadre général	210
B.2)	Cas des <i>Paramyxoviridæ</i>	211
I. C.	LE DRB CHEZ LES CETACES/CETARTIODACTYLES	211
I. D.	OBJECTIFS	211
II.	MATERIEL ET METHODES	212
II. A.	MATERIEL BIOLOGIQUE ET <i>IN SILICO</i>	212
A.1)	Dauphins bleu et blanc échoués	212
A.2)	Séquences de DRB de Cétacés issus de GenBank	212
II. B.	ELECTROPHORESE SUR GEL A GRADIENT DE DENATURATION, <i>DGGE</i>	213
B.1)	Principe	213
B.2)	Mise en œuvre	214
III.	METHODOLOGIE : ANALYSE DES SIGNATURES DE SELECTION	217
III. A.	EVALUATION DE L'HETEROGENEITE DES MODELES DE SUBSTITUTION ENTRE LES 3 POSITIONS DE CODONS	217
A.1)	Principe	217
A.2)	Implémentation logicielle : BEAST	218
III. B.	ANALYSE CODON PAR CODON : VRAISEMBLANCE PAR EFFETS-FIXES, <i>FIXED-EFFECT LIKELIHOOD</i>	218
IV.	PRINCIPAUX RESULTATS	220

IV. A.	UNE VARIABILITE NOTABLE DU DRB CHEZ LE DAUPHIN BLEU ET BLANC	220
A.1)	Variabilité du nombre d'allèles	220
A.2)	Variabilité nucléotidique	221
A.3)	Phylogénie des allèles	221
IV. B.	UNE SIGNATURE DE SELECTION AU SEIN DE L'EXON 2	221
V.	PERSPECTIVES	241
DISCUSSION GENERALE		243
REFERENCES		263

Avant-Propos

L'origine de ce travail se situe dans le prolongement d'un investissement personnel et de longue date dans divers aspects de la biologie et la pathologie des Cétacés. Un détail avait attiré notre attention il y a quelques années : lors de manipulation d'ossements de Dauphins bleu et blanc issus d'échouages sur les côtes de Méditerranée, nous avons remarqué des variations dans les soudures des Vertèbres cervicales. Cela nous a fait émettre des doutes quant à l'homogénéité de la population méditerranéenne. L'épizootie à Morbillivirus qu'a subie cette espèce a aussi été un cadre de réflexion.

La Méditerranée renferme un nombre important de ressources biologiques caractérisées par un endémisme élevé et une relative fragilité vis-à-vis des phénomènes invasifs, qu'ils soient naturels ou artificiels. L'isolement presque total de cette actuelle extension continentale de l'Atlantique explique en grande partie ces caractéristiques mais il n'en a pas toujours été ainsi : en effet, la Méditerranée est le vestige de l'ancien océan pantropical des périodes Mésozoïque et Cénozoïque (jusqu'à la fin du Miocène) : la Téthys.

La Méditerranée est un « point chaud » reconnu de la diversité végétale terrestre mais elle est aussi, et cela est moins connu, une mer dont la diversité biologique marine est très significative : elle concentre 7% des espèces mondiales, dont 20-30% d'endémismes, dans 0,8% de la surface océanique (IUCN). Ceci est à rapprocher des données de densité en biodiversité évoquées précédemment. Cependant, malgré l'ancienneté des préoccupations pour l'environnement méditerranéen, dont les origines remontent à l'Antiquité par l'instauration d'espaces protégés, les mesures mises en œuvre sont très nettement insuffisantes : à peine 3,8% de la surface fait partie d'une Aire Marine Protégée et les pays riverains peinent à mettre en place une politique commune efficace. En outre, si l'on ne raisonne qu'en termes utilitaires, on peut affirmer que les écosystèmes marins (comme d'ailleurs tout écosystème) fournissent un nombre très important de « services » nécessaires à la survie de l'humanité : régulation dynamique des chaînes alimentaires, des processus sédimentaires, des flux de carbone (Holmlund & Hammer, 1999). De même les composants mobiles de ces écosystèmes

assurent le lien entre divers compartiments et écosystèmes.

Au sein de ce Pélagos, les Cétacés occupent une place particulière. En effet, il convient de rappeler qu'ils trouvent leur origine sur une grande bordure peu profonde de la Téthys qui a permis le lent passage de la vie terrestre à la vie aquatique. Ces Mammifères marins sont en outre situés à l'extrémité de la chaîne alimentaire, ce qui leur confère à la fois un rôle central régulateur au point de contact entre divers compartiments de l'écosystème benthique mais aussi une fragilité inhérente aux prédateurs en cas de crise démographique dans les étages « inférieurs » des chaînes trophiques. Il a aussi été démontré que leur surexploitation avait provoqué des processus d'effondrement en cascade de stocks de divers Vertébrés marins (**Springer et al., 2003**) ; et il n'est pas impossible que l'actuel collapsus des pêcheries mondiales trouve en partie son origine dans la gestion irresponsable de cette biomasse précieuse, les grands Vertébrés marins en général, tant sur le plan de la diversité que sur le plan quantitatif (cycle du carbone).

En Méditerranée, l'intérêt de l'Homme pour les Cétacés date de l'Antiquité : on retrouve des représentations et des récits mettant en scène en particulier le Dauphin commun et le Grand dauphin. Mais de nombreuses espèces sont actuellement présentes en Méditerranée, des grands Mysticètes comme le Rorqual commun aux Delphinidés en passant par les Baleines à bec (Ziphiidés) ou le Cachalot. Il est probable que dans un passé proche, la liste ait été plus longue, incluant en particulier des Mysticètes migrateurs, Baleine franche et Baleine grise (**Macé, 2003**).

Les menaces que l'Homme fait peser sur ces espèces en Méditerranée sont très importantes. Ces dangers sont de deux ordres : des perturbations directes (navigation, pollution, destruction d'habitats, plus rarement - aujourd'hui - destruction directe) et des perturbations indirectes par la surexploitation des ressources halieutiques ou les changements climatiques. A titre indicatif, 7% de la population mondiale vit dans l'espace méditerranéen et ce dernier reçoit annuellement 32% du tourisme global.

Le Dauphin bleu et blanc *Stenella coeruleoalba* est une espèce de Cétacé très

largement répandue à l'échelle mondiale. Il s'agit de l'espèce la plus communément rencontrée en Méditerranée mais son statut exact (population, endémisme ?) n'est pas évalué avec précision. Ce Dauphin a en outre été victime en 1991 d'une épizootie due au *Dolphin Morbillivirus* (DMV). Dans le cadre de cette thèse avec une double approche scientifique et vétérinaire, il nous a donc semblé intéressant de nous pencher sur ce modèle biologique.

L'écologie moléculaire est un concept proposé par différents auteurs parmi lesquels E.B. Ford dans ses travaux sur ce qu'il a nommé l'« écologie génétique ». Elle peut être considérée comme une branche de la biologie évolutive qui vise à appliquer les concepts de génétique des populations, phylogénie et génomique à des questions écologiques. La génétique a permis de caractériser de plus en plus finement les populations animales. Les Cétacés n'ont pas échappé à la règle, même si leur distribution, océanique pour la plupart, est un handicap pour l'échantillonnage. Ainsi, une source non négligeable de matériel biologique de ces espèces réside-t-elle dans les échouages d'animaux sur les côtes.

De nombreuses régions du génome ont fait l'objet d'une attention de la part des chercheurs en écologie mais certaines d'entre elles ont eu plus d'impact que d'autres. L'ADN mitochondrial est le support de très nombreuses études à la fois en génétique des populations, en particulier chez les Vertébrés à l'aide de sa partie non codante la région de contrôle, mais aussi en phylogénie sur les gènes codant les protéines du métabolisme énergétique. Son intérêt réside en plusieurs caractéristiques qui le différencient du reste du génome : il est haploïde, peu ou pas recombinant et hérité en règle générale par la lignée maternelle. Les deux premières propriétés le rendent intéressant tant au niveau technique qu'analytique, alors que la dernière permet de tester des hypothèses comportementales liées au sexe. L'utilisation de sa séquence permet une approche généalogique des processus évolutifs (coalescence). Les microsatellites ou motifs répétés en tandem (*tandem repeats*) sont aussi très largement utilisés en génétique des populations. Leur ubiquité à l'échelle du génome, leur grande variabilité, leur évolution présumée neutre et la facilité

de manipulation en ont fait des marqueurs de choix pendant plus d'une décennie. Enfin, le Complexe Majeur d'Histocompatibilité (C.M.H.) a d'abord été mis en évidence par une application « artificielle » : le rejet de greffes. Il est intéressant de constater que ce rôle artificiel comporte un pendant naturel chez les Urocordés où un proto-C.M.H. joue un rôle de barrière contre la création de chimères (équivalente du phénomène de greffe) entre plusieurs colonies par le biais d'un locus d'alloreconnaissance. Cet élément génomique apparaît tôt dans l'histoire des Vertébrés : en même temps que l'Immunité Adaptative Spécifique (IAS). L'immunité spécifique contre les pathogènes est un élément clef de la réussite adaptative des Tétrapodes et tout particulièrement des Mammifères.

Les questions que nous nous sommes posées peuvent être regroupées en trois catégories. La première, la plus générale et d'ordre méthodologique, est la recherche d'un marqueur de séquence du chromosome Y suffisamment variable pour pouvoir être utilisé à des fins de biologie évolutive. Le deuxième groupe de questions est relatif au statut populationnel du Dauphin bleu et blanc en Méditerranée. En particulier, nous avons abordé une hypothèse de différenciation géographique, puis l'éventualité de processus vicariants pouvant avoir une influence sur la structure génétique de l'espèce en Méditerranée. Enfin, dans un troisième temps, notre projet était de tester si une éventuelle structure dans l'échantillon méditerranéen (par exemple par le biais de l'*inbreeding* - ou consanguinité - et la plus grande fréquence d'allèles susceptibles) pouvait avoir une influence sur la susceptibilité au Morbillivirus ayant touché le Dauphin bleu et blanc en Méditerranée en 1991 et certainement en 2003. En raison de difficultés d'échantillonnage, nous n'avons fait qu'aborder cette question mais cela a permis de défricher un aspect peu abordé chez les Cétacés, la génomique du système immunitaire adaptatif, ainsi que de dégager des perspectives pour de futures investigations.

C'est dans un premier temps d'un point de vue bibliographique que nous aborderons diverses questions en relation avec le travail de thèse.

L'échantillonnage biologique a été effectué entre 1990 et 2005. Il est constitué d'une part de 138 individus de l'espèce Dauphin bleu et blanc dont 79 en provenance de la Méditerranée, 45 de l'Atlantique nord-est et 14 du Pacifique nord (cf. **Erreur ! Source du renvoi introuvable.**) et d'autre part d'individus des espèces suivantes : le Dauphin commun *Delphinus delphis*, le Grand dauphin *Tursiops truncatus*, le Dauphin de Risso *Grampus griseus*, le Cachalot *Physeter macrocephalus*, la Baleine à bec de Cuvier *Ziphius cavirostris*, le Rorqual commun *Balænoptera physalus*, le Rorqual à museau pointu *Balænoptera acutorostrata* et la Baleine grise *Eschrichtius robustus*.

Dans un premier temps, ce travail a consisté en l'étude d'un marqueur du chromosome Y, le locus situé allant de l'exon 4 à l'exon 5 du gène de l'amélogénine. Pour cela, nous avons cloné et séquencé cette région chez 10 Dauphins bleu et blanc ainsi que 5 autres espèces de Cétacés et rapporté ces séquences à d'autres séquences de *Cetartiodactyla* disponibles dans les bases de données. L'analyse phylogénique de ces données a montré comme principal résultat l'existence d'une insertion d'environ 500 paires de bases sur le chromosome Y, datant de la fin du Crétacé, pouvant être utilisée à la fois à des fins évolutives (phylogénétique et génétique des populations) et de sexage (zootecnie et biologie de la conservation). Ce travail est abordé dans le 0 et a fait l'objet d'une publication dans BMC Genetics :

Article n°1 : Macé, M. and B. Crouau-Roy (2008). A highly polymorphic insertion in the Y-chromosome amelogenin gene can be used for evolutionary biology, population genetics and sexing in Cetacea and Artiodactyla. *BMC Genetics* **9**: 64. <http://dx.doi.org/0.1186/1471-2156-9-64>

Dans un deuxième temps, nous avons abordé la biologie évolutive du Dauphin bleu et blanc en Méditerranée par le biais de l'ADN mitochondrial et des microsatellites. Les données de séquence et de génotypage ont subi à la fois une approche de génétique des populations, fréquentiste et bayésienne, ainsi que des approches phylogéniques visant à caractériser les relations évolutives entre allèles mais aussi à co-inférer des paramètres comme la datation des divergences et la démographie. Le principal résultat de ce travail est la démonstration de l'existence de deux groupes fortement différenciés au sein de la Méditerranée. Une partie importante de la discussion a porté sur le phénomène vicariant ayant pu mener à cette divergence ainsi que les facteurs ayant pu permettre la persistance

de cette différenciation. Ce travail est abordé dans le 0 et a fait l'objet de trois publications :

Article n°3 : Bourret V, **Macé M**, Crouau Roy B (2007) Genetic variation and population structure of western Mediterranean and northern Atlantic *Stenella coeruleoalba* populations inferred from microsatellite data. *Journal of the Marine Biological Association of the United Kingdom* **87** (1) **Special Issue**, 265-269. <http://dx.doi.org/10.1017/S0025315407054859>

Article n°2 : Bourret V, **Macé M**, Bonhomme M, Crouau-Roy B (2008) Microsatellites in Cetaceans: An Overview. *The Open Marine Biology Journal* 2, 38-42. <http://dx.doi.org/10.2174/1874450800802010038>

Article n°4 : **Macé M**, Bourret V, Crouau-Roy B « Polyphyly in the Mediterranean Striped dolphin (*Stenella coeruleoalba*) and the Messinian Salinity Crisis », en révision dans la revue PLoS ONE

Dans un dernier temps, nous avons caractérisé le séquence d'un gène du Complexe Majeur d'Histocompatibilité (C.M.H.) de Classe II, le DRB, chez le Dauphin bleu et blanc ainsi que chez quelques autres espèces de Cétacés. Pour cela, nous avons utilisé plus particulièrement la DGGE (Gel d'Electrophorèse par Gradient de Dénaturation, *cf.* 0). Les données ont subi une analyse phylogénique ainsi qu'une analyse des profils de sélection sur la globalité de la séquence ainsi que site par site. Ce travail est abordé dans le chapitre IV et fait l'objet d'une publication en cours d'achèvement :

Article n°5 en préparation : **Macé M** & Crouau-Roy B « Diversity and patterns of selection on the DRB of the Striped dolphin, *Stenella coeruleoalba* ».

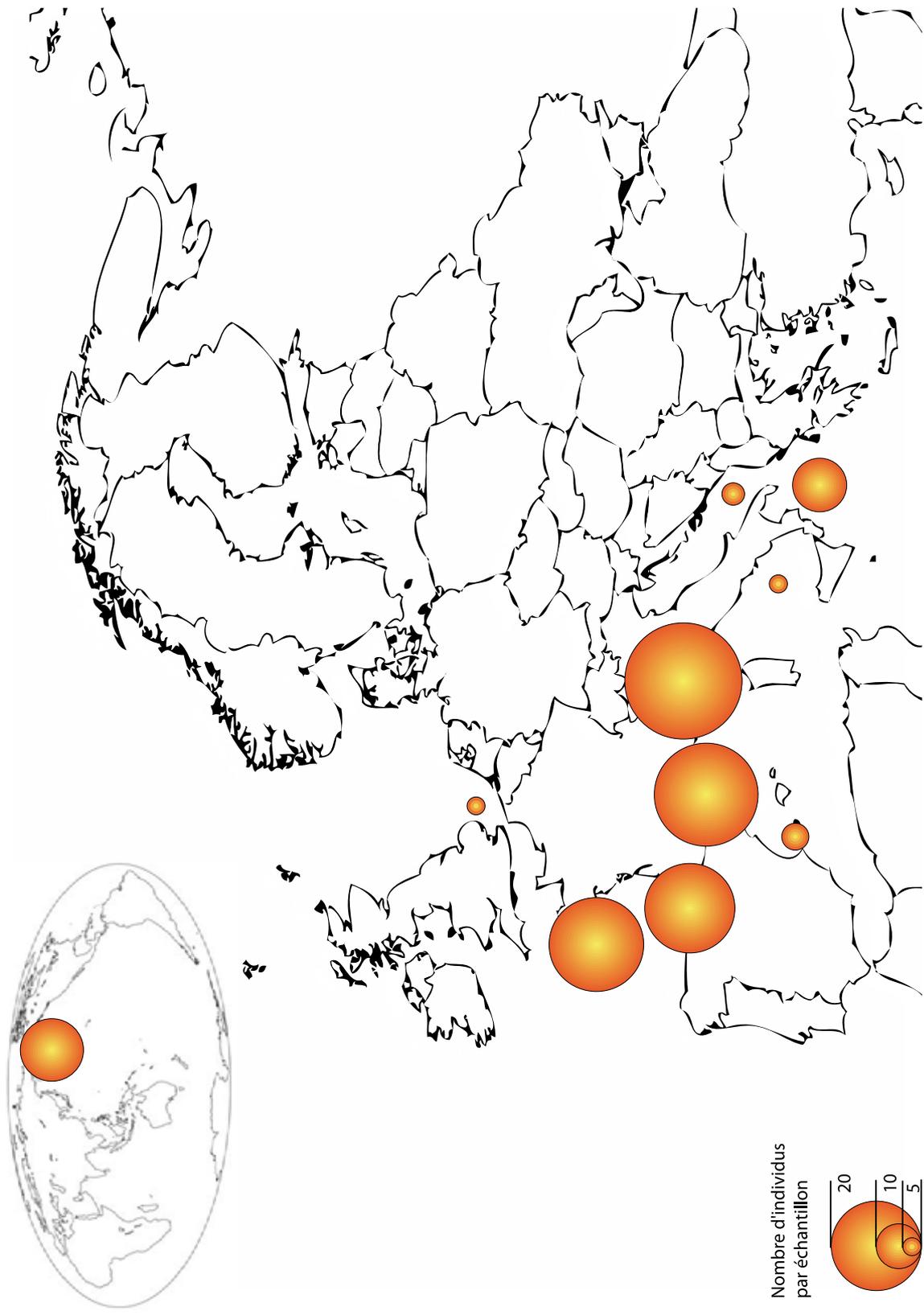


Figure 1. Répartition des échantillons de Dauphin bleu et blanc. Le médaillon en haut à gauche désigne l'échantillon Pacifique.

Chapitre I : Contexte Bibliographique

Afin de mieux positionner le problème et en raison de la diversité des approches, il nous a semblé utile de procéder à des rappels concernant l'écologie, le modèle animal, le cadre géographique, et les bases moléculaires de l'immunité. Nous espérons que ces rappels aideront les lecteurs issus d'horizons différents.

I. Quelques notions d'écologie

I. A. Définitions générales (Ridley, 2004)

A.1) Espèce, population

1(a) L'espèce

L'espèce est une notion difficile à appréhender et relativement fluctuante selon qu'on se réfère à un groupe d'êtres vivants ou à un autre. Comme toute création de l'esprit en vue d'interpréter la nature, elle est un calque de la culture de son créateur. On peut considérer dans une première approche qu'il s'agit de l'unité fondamentale du vivant.

Il existe plusieurs définitions de l'espèce. Sur un plan biologique au sens large (dynamique), la plus communément admise est la définition d'Ernst Mayr : *Species are groups of actually or potentially interbreeding natural populations, which are isolated from other such groups* (Mayr, 1942). Ainsi, on peut considérer qu'un postulat majeur de la notion d'espèce est l'isolement reproductif, c'est à dire l'isolement né des barrières qui vont s'opposer à ce que des êtres vivants appartenant à des groupes différents puissent engendrer une progéniture viable et féconde. De cette notion de groupe différent naît l'aspect circulaire de la réflexion autour de l'espèce : la présence de l'isolement reproductif est évaluée à partir de la définition de groupes mais les groupes sont eux-mêmes définis par cette barrière.

Les barrières peuvent être prézygotiques ou postzygotiques. Les premières seront d'ordre comportementales (parades par exemple), physiologiques (p. ex. modes de copulation,) ou morphologiques (p. ex. aspect) alors que les secondes empêcheront l'œuf de devenir adulte (p. ex. incompatibilité foëto-maternelle chez les Vivipares).

On peut aisément se rendre compte que l'espèce ne revêt pas du tout la même signification selon qu'on songe à un Procaryote, où les flux de gènes horizontaux

sont monnaie courante, ou à un Mammifère chez qui ces transferts ne peuvent avoir lieu que par l'intermédiaire de virus et dans une bien moins grande mesure.

1(b) La population

Avec les mêmes précautions que précédemment à propos de l'espèce, on peut dire qu'une population est un groupe d'individus partageant un même site où ils exploitent une même ressource et/ou s'y reproduisent entre eux.

A.2) Dynamique d'une population

On parle de dynamique de population à propos des changements au cours du temps de la taille de cette population. Etudier la dynamique d'une population revient à formuler l'ensemble des lois qui gouvernent ces changements.

Une première approche avait été formulée par Leonardo Pisano dit *Fibonacci* dans un de ses fameux « exercices de l'esprit » connu sous le nom des « lapins de Fibonacci » (**Pisano, 1202**). Il s'était posé la question de l'évolution du nombre de lapins si l'on part d'un couple, que chaque couple produit chaque mois un nouveau couple, lequel ne devient productif qu'après deux mois, et que les lapins sont éternels... Cela revient à écrire : $n_{t+2} = n_{t+1} + n_t$ avec $n_1 = 1, n_2 = 1$. Il était l'un des premiers à entrevoir que la croissance des populations isolées et sans contrainte suit une croissance multiplicative.

Bien plus tard, Malthus a décrit un modèle exponentiel défini par une suite géométrique (**Malthus, 1798**) :

$$n_t = n_0(1 + f - m)^t$$

$$\text{ou par intégration : } n(t) = n_0 e^{kt}$$

Deux orientations sont alors possibles pour la population décrite par ce modèle :

- soit la fécondité est supérieure à mortalité ($m < f \Leftrightarrow k > 0$) ce qui se traduit par une croissance exponentielle ;

- fécondité inférieure à mortalité ($m > f \Leftrightarrow k < 0$) ce qui se traduit par une décroissance exponentielle.

Ces modèles de croissance exponentielle restent théoriques puisque l'on constate expérimentalement que même lorsqu'il y a une phase de croissance exponentielle,

celle-ci est toujours suivie d'une phase au mieux stationnaire.

En effet, lorsque les individus d'une même population exploitent une ressource commune, la ressource s'épuise à mesure que la population augmente en taille. Cet épuisement a pour conséquence une diminution de la fécondité et on aboutit au mieux (pour la population !) à un infléchissement de la croissance, au pire à une décroissance. C'est ainsi qu'ont émergé des modèles de type logistique tel celui de Verulst (**Verhulst, 1838**) qui propose que la différence entre fécondité et mortalité décroît linéairement lorsque la taille de la population augmente :

$$\frac{1}{n} \frac{dn}{dt} = r \left(1 - \frac{n}{K} \right),$$

Cela correspond à une courbe d'allure sigmoïde : lorsque n est petit, la croissance est de type exponentiel et lorsque n se rapproche de K , la taille tend vers cette valeur (en croissant ou décroissant). La valeur K , dénommée capacité biotique, est la taille qu'atteint la population lorsqu'elle est à l'équilibre. Toutes les espèces ne suivent pas la même stratégie. Cela se traduit par des variations dans la taille des portées, temps de gestation, âge de maturité sexuelle, etc. On peut scinder schématiquement les espèces en deux groupes :

- celles qui investissent plus dans la fécondité (stratégie dite « r ») ;
- celles qui investissent plus dans la survie (stratégie dite « K »).

A.3) Le concept de niche écologique

La distinction « stratégie r » vs « stratégie K » ouvre la voie au concept de la niche écologique, concept sujet à controverses mais il est vrai que toute théorie ne prétend être qu'une approximation *a posteriori* de la réalité et ne s'avère que rarement prédictive.

La niche écologique est décrite comme l'ensemble des conditions de l'environnement qui permettent à une espèce de former des populations viables. Si la démographie de l'espèce considérée suit le modèle logistique de Verhulst, on peut dire que la niche écologique est l'ensemble des conditions telles que $r > 0$ (la fécondité excède la mortalité) et $K > 0$ (il y a assez de ressource pour la survie d'une population). Lors du Symposium sur la Biologie Quantitative de Cold Spring Harbor intitulé *Population*

Studies: Animal Ecology and Demography, la niche écologique a été décrite comme un « hypervolume » (**Hutchinson, 1957**), c'est à dire un espace multi-dimensionnel des ressources (intensité lumineuse, température, divers nutriments, structure,...) disponibles pour un groupe spécifique d'organismes (espèce, complexe d'espèces, communauté). On peut scinder ce concept en deux notions : la niche fondamentale (conditions théoriques nécessaires à l'existence d'un organisme) et la niche réalisée (niche proche de la niche fondamentale et occupée en pratique) (**Pidwirny, 2006**). On parle parfois d'enveloppe écologique, concept voisin décrivant la somme des conditions du milieu (biotope, climat,...) nécessaires à la vie d'une espèce. Dans certains cas, selon cette théorie, deux espèces différentes peuvent occuper un biotope commun mais dans des niches différentes. Celles-ci doivent différer d'au moins un paramètre et peuvent se recouvrir partiellement.

En raison de la limitation des ressources, et si l'on suppose l'action d'une sélection naturelle, soit l'espèce la « mieux adaptée » à la niche est favorisée, soit il y a coexistence. Trois conditions sont alors nécessaires : la niche fondamentale de chaque espèce doit être inférieure au point d'offre du milieu, la consommation des ressources doit être partagée - en particulier les facteurs limitants doivent être différents pour les deux espèces - et la compétition intraspécifique doit être complémentaire de la compétition interspécifique.

A.4) Compétition entre génotypes, mutation-sélection

4(a) Variance phénotypique et sélection directionnelle

On définit la variance phénotypique par $\sigma_P^2 = \frac{\sum_{i=1}^n (x_i - \hat{\mu}_X)^2}{n-1}$ où x_i est le phénotype des individus, $\hat{\mu}_X$ la moyenne des x_i et n le nombre d'individus mesurés. Cette variance peut être fractionnée en variance environnementale σ_E^2 - variance observée entre individus génétiquement identiques placés dans des conditions environnementales hétérogènes - et variance génétique σ_G^2 - variance au sein d'une population hétérogène d'individus placés dans des conditions environnementales homogènes (écotron par exemple). Ainsi, on a $\sigma_P^2 = \sigma_G^2 + \sigma_E^2$ (cf. **Figure 2** ci-après).

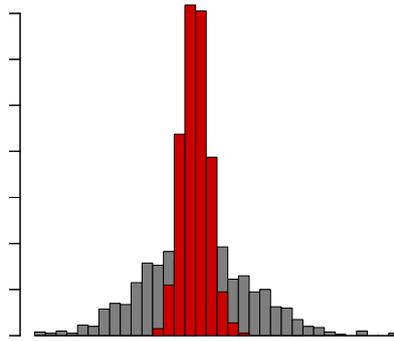


Figure 2. Exemple de représentation des parts relatives génétique (en rouge) et environnementale (en gris) dans la dispersion d'un phénotype.

Si $\sigma_E^2 = \sigma_G^2$ (donc $\sigma_P^2 = 0$), avec $\hat{\mu}_X(\text{parents}) = \hat{\mu}_X(\text{progéniture})$, il n'y a pas de différence observable entre les phénotypes des parents et de la progéniture). Après itération sur plusieurs générations, la distribution des phénotypes peut se décaler vers des valeurs de traits (sélection dite directionnelle). De même, entre la génération $n+1$ et la génération n , on considère que la sélection opère en éliminant les variants les moins avantageés et la variance génétique σ_G^2 diminue.

Lorsque $\sigma_E^2 \neq 0$ et $\sigma_G^2 \neq 0$, on définit l'héritabilité comme $h^2 = \frac{\sigma_G^2}{\sigma_P^2}$.

On voit alors que plus h^2 est élevée, plus la sélection aura une forte influence (cf. 0). Cela signifie que la moyenne $\hat{\mu}_X(\text{génération } n+1)$ sera d'autant plus proche de la moyenne du phénotype des individus qui auront d'une part survécu et d'autre part procréé.

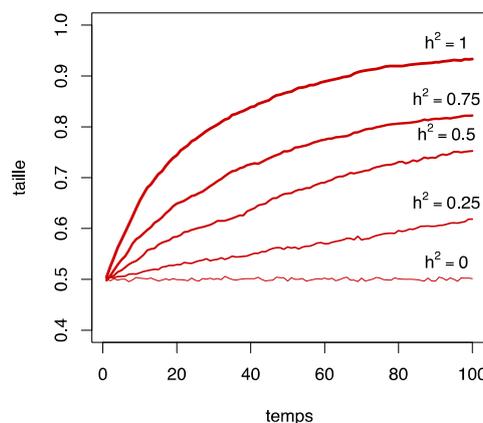


Figure 3. Evolution d'un caractère (ici la taille) au cours du temps (exprimé en générations) en fonction de l'héritabilité.

D'un point de vue expérimental, Wilhelm Johannsen (Johannsen, 1903; Johannsen,

1908), inventeur en particulier du mot « gène », a formalisé les effets de la sélection artificielle, opérée en particulier par les agriculteurs depuis le Néolithique, sur la variabilité génétique. En effet, dans un premier temps, la sélection de phénotypes extrêmes, intéressant l'agronome conduit à une évolution du phénotype des descendants dans le sens voulu mais cette évolution se tarit au bout d'un certain nombre de générations, traduisant l'appauvrissement concomitant de variabilité génétique. La sélection ne peut alors plus opérer car la moyenne du trait change selon la direction de la sélection et la variance diminue peu à peu. Cette diminution, si elle n'est pas trop rapide, est compensée par la dérive génétique.

4(b) Valeur adaptative ou Fitness

(Felsenstein, 2004a)

On définit le *Fitness* - ou « valeur sélective » ou « succès reproducteur » - comme une mesure de la capacité d'un génotype à se reproduire, en comparaison avec les autres génotypes. Il montre donc à quel point un génotype est favorisé par la sélection. Sa formulation la plus courante est donc :

$$W_{absolu} = \frac{N_{n+1}}{N_n} \text{ (fitness absolu)}$$

Mais il existe aussi la formulation d'un *fitness relatif*: si un des génotypes est normalisé à $w=1$, on quantifie le taux de survie de la progéniture d'individus porteurs d'un génotype en comparaison avec les porteurs de ce génotype de référence. Cette mesure est utilisée en raison de la difficulté d'accéder à une mesure absolue. Les deux mesures sont liées par la formulation suivante :

$$\frac{W_{absolu}}{\bar{W}_{absolu}} = \frac{W_{relatif}}{\bar{W}_{relatif}}$$

(où les *fitness* moyens sont pondérés par les fréquences des génotypes)

Dans le cas particulier d'un gène à deux allèles (A et a), on aura :

$$W_{relatif} = \frac{W_{AA}}{W_{max}}, \quad W_{relatif} = \frac{W_{Aa}}{W_{max}}, \quad W_{relatif} = \frac{W_{aa}}{W_{max}}$$

(*fitness* relatifs des génotypes AA, Aa et aa respectivement)

Le *fitness* varie entre 0 (génotype destiné à disparaître rapidement), et 1 (génotype destiné à devenir dominant rapidement).

Etant donné la versatilité expérimentale du problème de la mesure du *fitness*, une version probabiliste est donnée par l'assertion suivante :

« si x est plus adapté que y dans l'environnement E,
cela équivaut à ce que x ait une propension probabiliste supérieure à 0,5
de laisser plus de descendants que y » (Mills & Beatty, 1979).

Cela se traduit par : $A^*(O, E) = \sum P(Q_i^{OE})Q_i^{OE} - f(E, \sigma^2)$ où Q_i^{OE} serait la gamme du nombre possible de descendants à la génération i , $P(Q_i^{OE})$ la propension probabiliste de laisser Q_i^{OE} descendants à la génération i et $f(E, \sigma^2)$ une fonction de la variance du nombre de descendants.

On peut représenter le fitness en fonction de la fréquence des allèles, sous la forme d'un « paysage adaptatif » (Figure 4 ci-après).

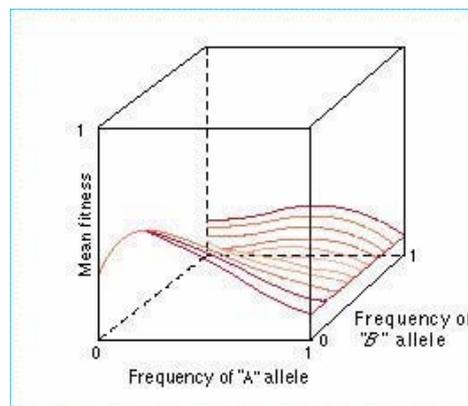


Figure 4. Paysage adaptatif : le fitness représenté comme une fonction des fréquences des allèles.

4(c) Différents types de sélection

(Felsenstein, 2004a)

Nous avons vu précédemment que la sélection directionnelle opère un déplacement de la moyenne du phénotype au cours des générations (cf. **Figure 5a** ci-après).

La sélection stabilisante, ou « purifiante », voit son intensité croître avec l'écart du phénotype par rapport à la moyenne. La moyenne du trait ne changera donc pas mais sa variance ira en diminuant rapidement (cf. **Figure 5b** ci-après).

La sélection disruptive favorise les individus situés aux extrémités de la distribution du trait (cf. **Figure 5c** ci-après). On pense qu'elle est la principale force en action dans la spéciation sympatrique (cf. B.1).

La sélection « balancée » (*balancing selection*) quant à elle tend à maintenir le polymorphisme génétique à l'intérieur d'une population. Ses deux principaux mécanismes sont l'avantage hétérozygote (*cf.* Anémie falciforme) et la sélection fréquence-dépendante (*cf.* choix du partenaire sexuel chez beaucoup d'espèces). Comme nous le verrons plus loin que c'est le principal mode en action sur les molécules du CMH.

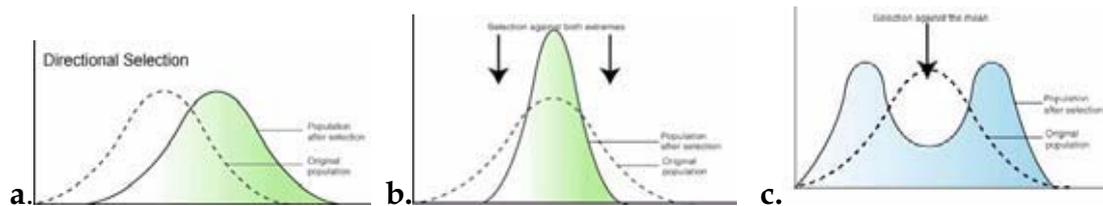


Figure 5. Schématisation des principaux types de sélection opérant sur un trait quantitatif (a. sélection directionnelle ; b. sélection purifiante ; c. sélection disruptive).

Cas extrême : sélection disruptive. La sélection favorise les phénotypes aux marges de la distribution (improprement appelé « phénotypes extrêmes ») par sélection autour de la moyenne du trait. Ce type de sélection est un des mécanismes principaux de la spéciation.

A.5) Équilibre de Hardy-Weinberg et dérive génétique

La dérive génétique ou dérive allélique est l'évolution des fréquences alléliques relatives d'une génération à l'autre en raison de l'échantillonnage au hasard des allèles parentaux transmis à la descendance.

Ce phénomène est d'autant plus important que la population est petite. En effet, selon le principe de l'équilibre de Hardy-Weinberg, les fréquences alléliques auront tendance à rester constantes d'une génération à l'autre pourvu que la population soit de grande taille, que les accouplements soient déterminés par le hasard, et qu'il n'y ait ni migration, ni sélection, ni mutation.

I. B. Spéciation allopatrique, vicariance et dispersion

B.1) Différents modes de spéciation

(Felsenstein, 1976; Felsenstein, 2004a; Munguia-Vega *et al.*, 2007; Ridley, 2004)

On considère schématiquement que la spéciation peut survenir selon quatre modes principaux (*cf.* Figure 6) :

- la **spéciation péripatrique**, au cours de laquelle de petits isolats en périphérie de distribution voient leurs échanges limités avec la population-mère. Ce type de spéciation est lié à l'effet fondateur et la dérive génétique et on l'observe tout particulièrement lors de goulets d'étranglement (*bottlenecks*) ;

- la **spéciation parapatric** : deux populations divergent mais il existe une zone hybride. Cependant, le fitness réduit des hybrides diminuerait peu à peu le flux de gènes jusqu'à la spéciation définitive. Ce type de spéciation pourrait expliquer l'existence des écotypes de Grand dauphin (Hoelzel *et al.*, 1998; Ross, 1977a; Ross & Cockroft, 1990) ainsi que d'autres Delphinidæ (*e.g.* (Escorza-Treviño *et al.*, 2005)) voire certaines populations d'Orque ;

- la **spéciation sympatrique** se produit lorsque qu'il y a divergence au sein d'un même habitat. L'exemple le plus communément cité est la spécialisation des Insectes en fonction de la plante hôte ou les Cichlidæ des Grands Lacs Africains, une des familles les plus diversifiées de Vertébrés (Salzburger *et al.*, 2005; Snoeks & Carvalho, 2004). Certaines populations d'Orque pourraient subir un tel processus (Hoelzel & Dover, 1991). Les mécanismes invoqués pour cette spéciation sont la polyploïdisation et la formation d'hybrides;

- la spéciation allopatrique que nous décrivons ci-après.

En outre, un processus appelé renforcement (*reinforcement*) - de la spéciation - augmente l'isolement reproductif par le biais de la sélection lors d'un contact secondaire entre deux populations d'une même espèce après un isolement. Si l'isolement reproductif est déjà avéré, il y a spéciation *de facto* et si cet isolement est partiel, les hybrides peuvent avoir un *fitness* moindre. Cela conduit à ce que l'on appelle aussi l'effet Wallace en l'honneur du naturaliste célèbre qui avait déjà évoqué cette possibilité au XIXe siècle (Wallace, 1889), par la suite corroboré par des études

(Ollerton, 2005).

La spéciation peut être aussi artificielle, c'est à dire induite par la sélection utilisée par l'Homme afin d'obtenir des organismes domestiques. En particulier, le Mouton domestique *Ovis aries*, *a priori* créé par hybridation ne peut plus se reproduire avec les différentes (sous-)espèces d'*Ovis orientalis* (Hiendleder *et al.*, 2002) espèces desquelles il dérive, en raison d'une grande instabilité chromosomique. En revanche, on ne peut considérer le Bœuf domestique *Bos taurus* comme une espèce différente des Bovidae d'Asie centrale (Yack, Gaur, etc.) puisque l'hybridation est la règle (Nowak, 1999).

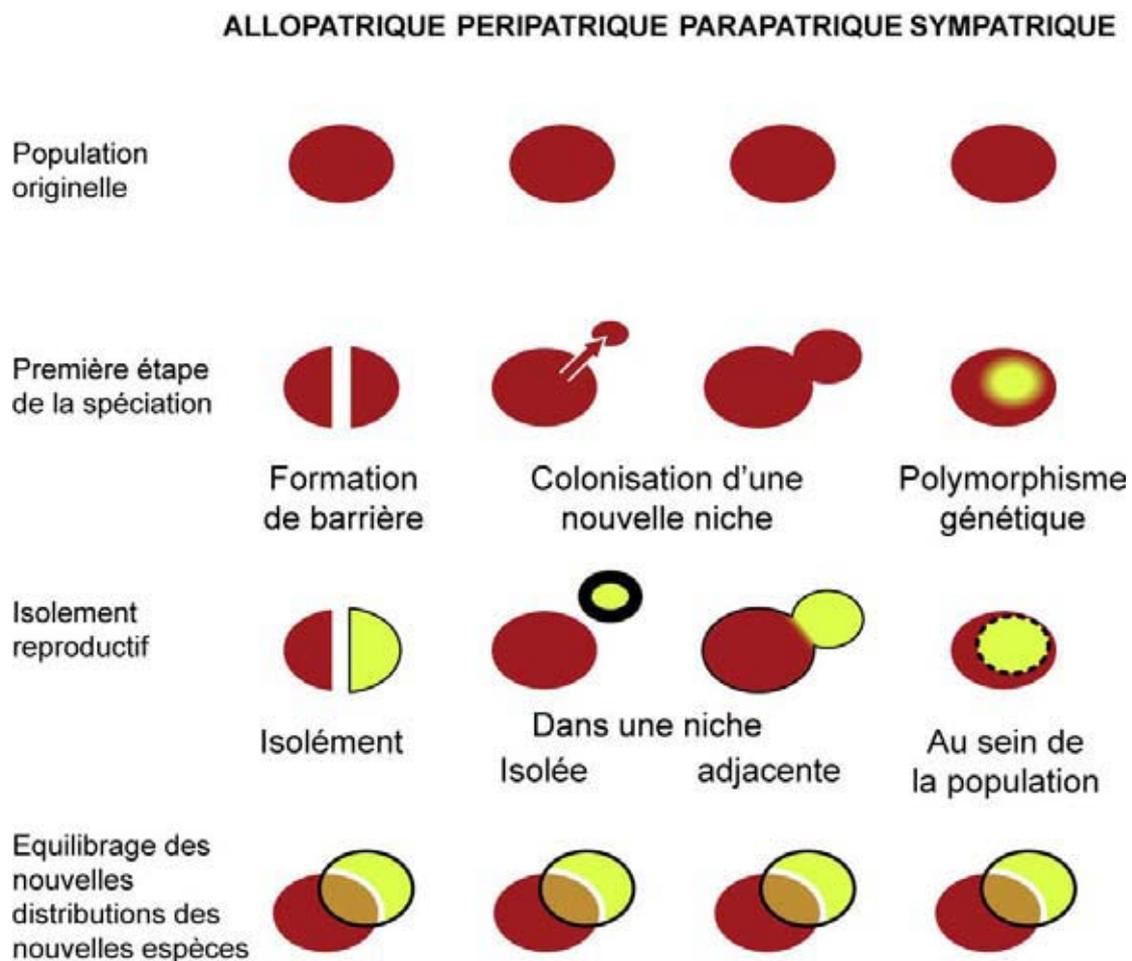


Figure 6. Principes et séquence de différents modes de spéciation (a. allopatrique, b. péripatrique, c. parapatric et d. sympatric).

D'un point de vue génétique, peu de gènes impliqués dans la spéciation ont été mis en évidence. Ils sont en général liés au « renforcement » de la spéciation, l'un d'entre

eux induisant par exemple la stérilité des hybrides entre sous-espèces liées (**Phadnis & Orr, 2009**). On suspecte maintenant que les transposons sont une des composantes de l'isolement reproductif (**Masly *et al.*, 2006**). En particulier, les rétrotransposons (motifs répétés transposables par l'intermédiaire des ARNs) empêchent les phénomènes de conversion génique (« recombinaison » entre portions paralogues du génome) en créant des zones de non-homologie entre des séquences au départ homologues. Ce phénomène permet l'émergence de nouveaux variants sans que ceux-ci ne soient éliminés par conversion avec les « gènes-parents ».

B.2) Spéciation allopatrique (Ridley, 2004)

L'allopatrie est un terme biogéographique qui désigne des organismes dont les répartitions sont entièrement disjointes. Si ces organismes sont des espèces-sœurs, cette distribution est généralement le fruit d'une spéciation dite allopatrique. Nous nous intéresserons à l'allopatrie dans notre étude, au Chapitre I.IV.

La spéciation allopatrique, aussi connue sous le nom de spéciation géographique, naît de l'existence (préalable ou *a posteriori*) d'une barrière extrinsèque (vicariance) et aboutit à des populations puis à des espèces plus ou moins non interfécondes. L'exemple le plus popularisé de ce type de spéciation est celui des Pinsons des Galápagos. Elle peut ainsi survenir lorsqu'une espèce est subdivisée en deux grandes populations par exemple par un événement géologique (tectonique, ou autre) ou lorsque un petit groupe d'individus colonise un nouvel habitat à la périphérie de l'aire de répartition de l'espèce. Dans le premier cas, en raison de la plus grande force des phénomènes sélectifs dans les grandes populations, l'évolution des génomes aboutissant à l'isolement reproductif s'opère par le biais des phénomènes adaptatifs. En revanche, cette évolution est moins bien comprise dans le cas d'une spéciation péripatrique qui pourrait s'opérer par le biais de la dérive génétique. Cependant, des cas de figure intermédiaires entre ces deux situations extrêmes semblent s'observer sur le terrain, comme l'isolement ou la fragmentation.

I. C. Notion de « services rendus » par les écosystèmes et diminution des ressources

C.1) Notion de services écologiques

1(a) Cadre général

Les écosystèmes de la planète sont caractérisés par des interactions complexes qui ont permis à la diversité biologique de s'épanouir en profitant au maximum que le permettent les réactions biochimiques de l'énergie entrant dans ces écosystèmes (énergie solaire principalement).

Une espèce est ainsi redevable par rapport à la communauté à laquelle elle appartient d'un certain nombre de services et réciproquement, elle en fournit elle-même. Un des cas les plus simples et évidents de cette dépendance réciproque est la zoochorie : l'animal se nourrit du fruit et transporte la graine et favorise ainsi la dispersion de la plante consommée. De cette interdépendance naît la notion de services et de biens « écologiques ».

Les services des écosystèmes (et les « biens écologiques » si l'on considère la nature selon une vision propre à Adam Smith) sont un aspect fondamental du fonctionnement du système terrestre support de la vie.

D'un point de vue anthropocentrique, non seulement ils contribuent au bien-être de l'humanité mais aussi participent directement des éléments nécessaires à sa survie.

On peut ainsi considérer qu'ils représentent une part de la valeur économique de la planète soit une valeur du même ordre que le P.I.B. de l'ensemble des pays - autour de 50.10^{12} \$ (Costanza *et al.*, 1997). A mesure que s'érode la biodiversité, tant qualitativement que quantitativement, on s'aperçoit de leur importance croissante voire qu'ils sont une condition *sine qua non* de l'existence de l'humanité. Ainsi, on peut considérer qu'une estimation monétaire n'est pas adaptée, même si elle a permis un début de prise de conscience de la part des acteurs économiques. On peut même remarquer que la carte montrant les valeurs que l'on peut attribuer à ces services et biens suit la distribution de la surexploitation des milieux (Figure 7).

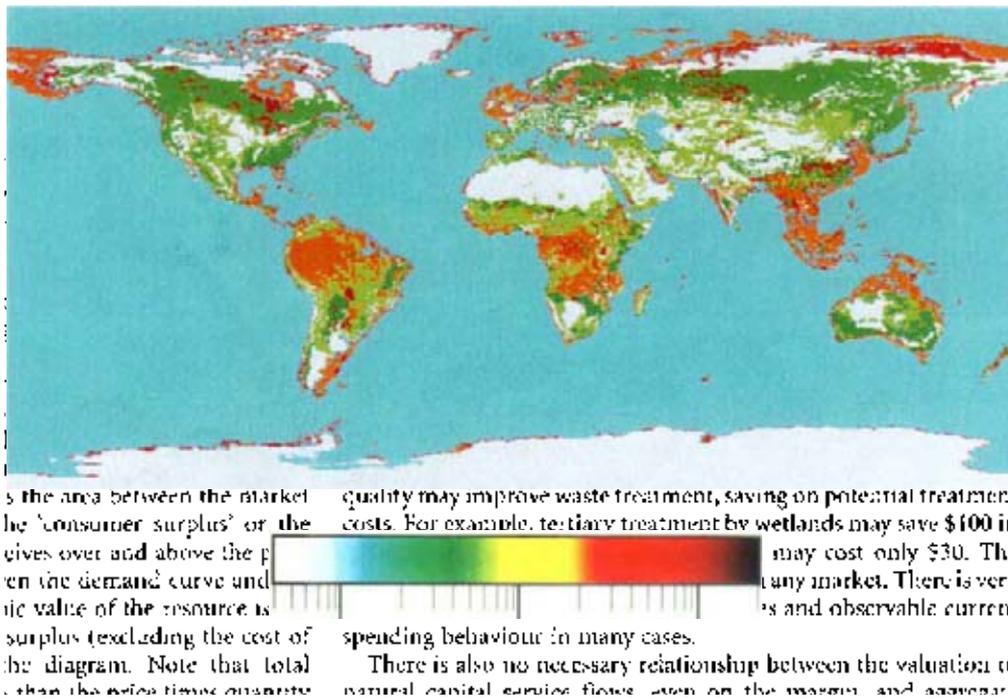


Figure 7. Carte globale de la valeur des services rendus par les écosystèmes. (d'après (Costanza *et al.*, 1997)).

Concrètement des exemples fondamentaux de biens sont l'air et l'eau et des exemples de services sont l'épuration de ces éléments, la décomposition des déchets et la création de la couche fertile du sol, ou encore l'ensemble des phénomènes biotiques et abiotiques concourant au contrôle des niveaux de gaz à effet de serre. Un autre service moins « visible » est le maintien de la biodiversité. Il peut être perçu sous deux angles :

- quantitatif, lié à l'énergie stockée sous forme de chaînes carbonées ;
- qualitatif, la complexité des réseaux et rétrocontrôles permettant le maintien d'une grande quantité de cette biomasse.

1(b) Milieu marin

Ceci est particulièrement vrai dans le milieu marin (Holmlund & Hammer, 1999) où l'on peut distinguer deux catégories de services : les services « fondamentaux », c'est à dire ceux dont la permanence est essentielle pour le fonctionnement des écosystèmes, et les services « à la demande ». L'existence de l'espèce humaine, comme celle d'autres espèces, est donc entièrement dépendante des premiers, que l'Homme sans que l'Homme s'en rende forcément compte : il ne sont liés à aucune espèce de

valeur ou marché. En revanche, les seconds sont liés à l'activité économique mais non nécessaires à la survie (p. ex. services « récréationnels »).

1(c) Services fondamentaux

Ces services sont résumés dans le **Tableau 1** ci-après.

Services d'ordre régulateur	Services « de liaison »
Régulation des chaînes trophiques	Liaison des compartiments océaniques
Résilience des écosystèmes	Liaisons écosystèmes terrestres <-> marins
Contrôle et maintien de la diversité génétique et métagénomique	
Recyclage des nutriments, redistribution des substrats (du fond par exemple), contrôle des processus de sédimentation	Transport de nutriments, carbone et minéraux
Régulation des flux de carbone eau/atmosphère	Transport d'énergie

Tableau 1. Principaux services fondamentaux des écosystèmes générés par les population marines et dulçaquicoles de Poissons

1(d) Services « à la demande »

Ces services sont résumés dans le **Tableau 2** ci-après.

Services d'ordre culturels	Services d'ordre informatif
Production alimentaire, aquaculture	Estimation des stress infligés aux écosystèmes, prise en compte ?
Production médicinale, contrôle de pathogènes	Estimation de la résilience des écosystèmes
Contrôle des algues et macrophytes	Informations d'ordre évolutif, historique et éducatif
Réduction des déchets	
Valeurs esthétiques et récréatives	

Tableau 2. Principaux services « à la demande » des écosystèmes générés par les population marines et dulçaquicoles de Poissons

C.2) Place des Prédateurs dans les écosystèmes, conséquences de leur diminution, cas des Cétacés

2(a) Menaces sur les services halieutiques

La surexploitation actuelle des ressources halieutiques globales est un sujet de préoccupation majeure pour les scientifiques, et encore trop peu pour la population et le milieu politique.

En 1995, la déplétion ou la surexploitation touchait 70% des populations mondiales de Poissons (Holmlund & Hammer, 1999). Actuellement, l'ensemble des espèces d'intérêts commercial a vu ses stocks plonger d'entre 20% et 80% selon les espèces (Worm *et al.*, 2006b). Cela n'affecte bien entendu pas seulement pêche récréative ou commerciale mais aussi services que nous avons précédemment évoqués.

2(b) Un cas d'école : l'écosystème dominé par les forêts de Kelp dans le Pacifique nord-est

Une des démonstrations les plus spectaculaires de l'existence de services de liaisons dont la détérioration d'origine humaine a eu pour conséquence une atteinte préjudiciable à des services encore plus fondamentaux (fourniture alimentaire) a été fournie par les travaux de James Estes et ses collaborateurs entre 1975 et 2005 dans le Pacifique nord.

Ces auteurs ont d'abord démontré que la prédation par la Loutre de mer *Enhydra*

lutris limite les populations d'Invertébrés épibenthiques, en particulier les Oursins *Strongylocentrotus polyacanthus*, entraînant en conséquence le développement de forêts d'algues (Kelp) (Estes *et al.*, 1978). En effet, les territoires caractérisés par une grande abondance de Loutres le sont aussi par la rareté et la petite taille des Oursins et la compétition semble réguler l'association entre Macroalgues (*Laminaria* spp. et *Agarum cribrosum* qui sont les algues contribuant le plus à la production primaire des milieux côtiers). En revanche, dans les territoires moins accessibles aux Loutres parce que plus profonds, les Oursins sont plus nombreux et plus gros et la compétition entre Macroalgues est moins intense. En milieu insulaire, en l'absence de Loutres, les Oursins prolifèrent et ont éliminé les Macroalgues. Par la suite, en examinant les associations entre forêts de Kelp, les herbivores et Mammifères marins, il sont arrivés à la conclusion que la distribution de ces algues était limitée aux eaux froides, plus particulièrement du Pacifique nord (Estes & Steinberg, 1988). En outre, le Kelp et les autres algues brunes « arborescentes » ont subi une radiation évolutive au Cénozoïque au moment où les *Mustelidæ* marins voyaient leur distribution limitée au Pacifique nord. L'interaction Mammifère/Herbivore/Kelp a ainsi pu favoriser une compartimentation de l'habitat selon l'intensité de l'herbivorie. Les auteurs ont par la suite vu ce paradigme confirmé par l'observation de différences dans la vitesse de récupération des forêts de Kelp après réintroduction de Loutres (Estes & Duggins, 1995). Ils ont enfin mis en évidence le rôle-clef des Orques dans la régulation de ce réseau par prédation des Loutres (Estes *et al.*, 1998), démontrant ainsi le lien entre écosystèmes côtiers et pélagiques et suspectant des modifications d'origine anthropique. Cette influence humaine a ensuite pu être vérifiée par l'observation d'un probable report de prédation des Orques depuis les grands Mysticètes (dont les populations dans le Pacifique nord ont été réduites de plus de 90% entre 1950 et aujourd'hui) sur des espèces de plus en plus petites de Mammifères marins (Baleines, Lions marins, Phoques puis Loutres), créant tout d'abord des effondrements en série de ces espèces (Springer *et al.*, 2003) puis, arrivé au maillon ultime, la Loutre, une augmentation des populations d'Oursins et une grande diminution des forêts de Kelp, siège de la reproduction d'un grand nombre d'alevins. L'effondrement des activités halieutiques a logiquement suivi.

La difficulté et le temps qu'a nécessité la démonstration des liens entre ces divers compartiments de l'écosystème (plus de 30 ans séparent les premières publications des dernières) sont à la mesure de la complexité des interactions au sein des écosystèmes.

II. Le modèle biologique : le Dauphin bleu et blanc, *Stenella caeruleoalba* (Meyen, 1833)

II. A. Systématique, Taxonomie, Paléontologie, Evolution

A.1) Résumé systématique

Il nous apparaît tout d'abord important de préciser la systématique de l'espèce étudiée. Dans le résumé ci-après, les termes en italique et gras proviennent de la classification phylogénétique et n'ont pas de rang.

Eukaryota (domaine) (**Whittaker & Margulis, 1978**)

Fungi/Metazoa

Metazoa (règne) (**Haeckel, 1874**)

Eumetazoa (sous-règne) (**Bütschli, 1910**)

Bilateria (sous-règne) (infra-règne) (**Haeckel, 1874**)

(Coelomata)

Deuterostomia (branche) (super-embranchement) (**Grobben, 1908**)

Chordata (phylum) (**Balfour, 1880**)

Craniata

Vertebrata (sous-phylum) (**Cuvier, 1812**)

Gnathostomata (infra-phylum) (**Zittel, 1879**)

Teleostomi (**Bonaparte, 1836**)

Euteleostomi

Sarcopterygii (Classe) (**Romer, 1955**)

Tetrapoda (super-classe) (**Broili, 1913**)

Synapsida

Amniota (série) (**Haeckel, 1866**)

Mammalia (classe) (**Linnæus, 1758**)

Theria (sous-classe) (**Parker & Haswell, 1897**)

Eutheria (infra-classe) (**Gill, 1872**)

Laurasiatheria (**Waddell et al., 1999**)

Cetartiodactyla (super-ordre) (**Montgelard et al., 1997**)

Cetacea (ordre) (**Brisson, 1762**)
Odontoceti (infra-ordre) (**Flower, 1867**)
Delphinoïdea (super-famille)
Delphinidae (famille) (**Gray, 1821**)
Delphininae (sous-famille (**Barnes, 1990**))
Stenella (**Gray, 1866**)/ genre (**Oliver, 1922**)
Stenella caeruleoalba (genre espèce)

A.2) Les Cétacés, position relative des Cetartiodactyla

2(a) Les Laurasiatheria au sein des Mammifères

Les *Laurasiatheria* sont un clade ayant rang de cohorte ou de super-ordre selon les classifications, comprenant les Echidnés, Musaraignes, Chauves-Souris, Cétacés, Ongulés (à nombre de doigts pair et à nombre de doigts impair), Carnivores et Pangolins (**Murphy et al., 2001**). Ils sont définis par des caractères binaires (présence-absence) moléculaires, les rétrotransposons (**Kordis et al., 2006**). Il s'agirait en outre du clade de Mammifères ayant subi leur radiation lors du Crétacé récent sur le supercontinent Laurasien, après sa scission d'avec le Gondwana (**Rose & Archibald, 2005**). Ce détail est important vu que la colonisation du milieu marin par les Cétacés s'est effectuée vraisemblablement sur la côte Laurasienne de la Téthys océan pantropical séparant les deux continents, ce qui correspond aux actuelles plaines d'Asie centrale. C'est donc le groupe-frère des *Euarchontoglires* (Primates, Rongeurs et Lagomorphes) et des *Afrotheria* (Damans, Oryctéropes, Tangués ou *Tenrecs*, Elephantules, Taupes dorées, Elephants et Lamantins).

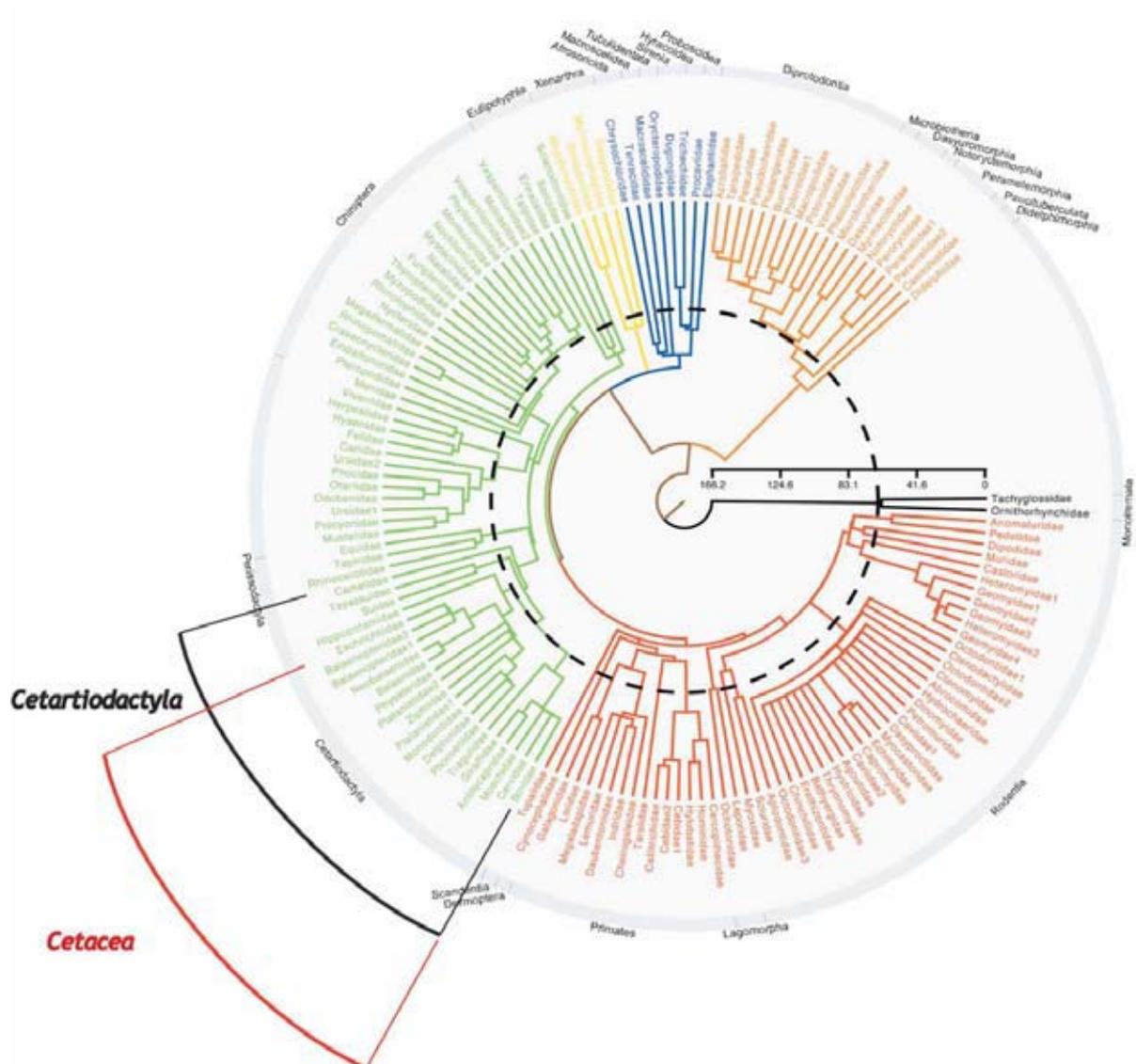


Figure 8. Super-arbre des Mammifères d'après (Bininda-Emonds et al., 2007).

2(b) Les Cetartiodactyla au sein des Laurasiatheria

Les Cétartiodactyles sont un regroupement récent des Cétacés et Artiodactyles en raison d'un certain nombre de synapomorphies fondamentales (Montgelard *et al.*, 1997) dont certaines d'ordre de l'écologie comportementale (Carrier *et al.*, 2002; Lusseau, 2003). Une description plus détaillée est fournie au I. D du chapitre II. Leur origine remonte à la fin du Crétacé, la divergence Perissodactyla-Cetartiodactyla étant datée d'environ 86 M.a (Rose & Archibald, 2005; Waddell *et al.*, 1999). La 0est un super-arbre montrant les relation des Cetartidactyla et Cétacés avec les autres Mammifères.

Il s'agit d'un groupe très intéressant pour l'Homme en raison du grand nombre

d'espèces domestiques. Il est possible que l'aspect grégaire qui caractérise un grand nombre de ces espèces ait facilité la domestication de certaines (**Clutton-Brock, 1999**). Au sein de ce groupe les *Tylopoda* (groupe des Chameaux *sensus lato*) et les *Suioidea* (groupes des Porcs *sensus lato*) bien qu'incontestablement des *Cetartiodactyla*, ont une place non clairement élucidée (pour plus de détails, cf. Chapitre I.I. D du chapitre II).

A.3) Les Delphinidæ et le clade Delphinus/Stenella/Tursiops

3(a) Les Delphinidæ

Cette famille regroupe l'ensemble des Dauphins océaniques soit environ 35-40 espèces. Il s'agit donc de la famille de Cétacés la plus diversifiée et elle comprend des espèces de tailles comprises entre 40 kg (Céphalorhynques) et 5 tonnes (Orque). Les caractéristiques communes du groupe sont résumées dans le tableau ci-après.

Synapomorphies des <i>Delphinidæ</i>	Caractères apomorphes partagés avec les autres Odontocètes	Caractères notables partagés avec les autres Cétartiodactyles
<p>Crâne dyssymétrique (os de la partie droite plus volumineux, en particulier l'incisif)</p> <p>Un sinus ptérygoïdien complexe (Muizon, 1988)</p> <p>La bulle tympanique, Pars tympanica du temporal, possède un processus caudal court à la différence en particulier des Ziphiidæ (Kasuya, 1973) et présentant de fines stries dont une médiane plus importante (ce dernier caractère est partagé avec les Monodontidæ)</p> <p>Sac nasal caudal absent, une délétion de 160 paires de bases dans un microsatellite (Árnason et al., 1984) (partagé avec les Phocœnidæ)</p> <p>(Heyning, 1997)</p>	<p>Echolocation à l'aide d'un organe spécifique</p>	<p>Grégarité</p> <p>Estomac pluriloculaire</p> <p>« coup de tête »</p> <p>pénis fibro-élastique et sigmoïde</p>

Tableau 3. Synapomorphies partagées par les Delphinidæ, Odontocètes et Cetartiodactyla.

Certains auteurs considèrent qu'il n'existe pas de genre fossile différent des genres actuels (**Muizon, 1988**) et que les genres rapportés précédemment (**Simpson, 1945**)

sont des *Kentriodontidæ*, considérés comme un taxon-frère des *Delphinidae*.

Cependant, d'autres auteurs (**Barnes et al., 1985; Slijper, 1936**) préfèrent rattacher les *Delphinidæ* aux *Kentriodontidæ* en se basant sur l'évolution des sacs aériens. Enfin, un dernier auteur (**Kasuya, 1973**) considère à partir des pétro-tympaniques que les *Kentriodontidæ* ne sont pas un groupe monophylétique et les répartit entre les divers groupes de Delphinoidea.

Par ailleurs, des fossiles du Miocène tardif du Japon attribués au genre *Stenella* (**Oishi & Hasegawa, 1994**) ainsi que d'autres (mandibules et pétro-tympaniques) attribués à *Stenella* ou *Delphinus* (**Barnes, 1976**) montrent que les relations entre ces genres sont questionnables.

Aussi, nous considérons (**Macé et al., 1999**) que les synapomorphies utilisées (**Muizon, 1988**) ne sont pas assez nombreuses ni suffisamment objectivables dans leur totalité comme en convient leur propre auteur considérant que les réversions et parallélismes sont probablement fréquentes chez les *Delphinidæ*. Il doit en être de même au niveau moléculaire.

3(b) Problème du clade *Stenella-Delphinus-Tursiops*

Une étude moléculaire a remis en question la monophylie des espèces précédemment définies au sein du complexe *Sousa-Tursiops-Lagenodelphis-Stenella-Delphinus* (**Leduc et al., 1999**). En effet, sur la base de la séquence du cytochrome b de l'ADN mitochondrial, les *Delphininæ* se retrouvent fortement supportés mais dans une acceptation différente des morphologistes (**Barnes et al., 1985; Barnes, 1990; Muizon, 1988**): le genre *Grampus* s'en retrouve exclu ainsi que les genres *Lagenorhynchus* (*Steno* et *Sotalia* nous sembleraient alors constituer des *incertae sedes*) comme l'indique la **Figure 9** ci-après.

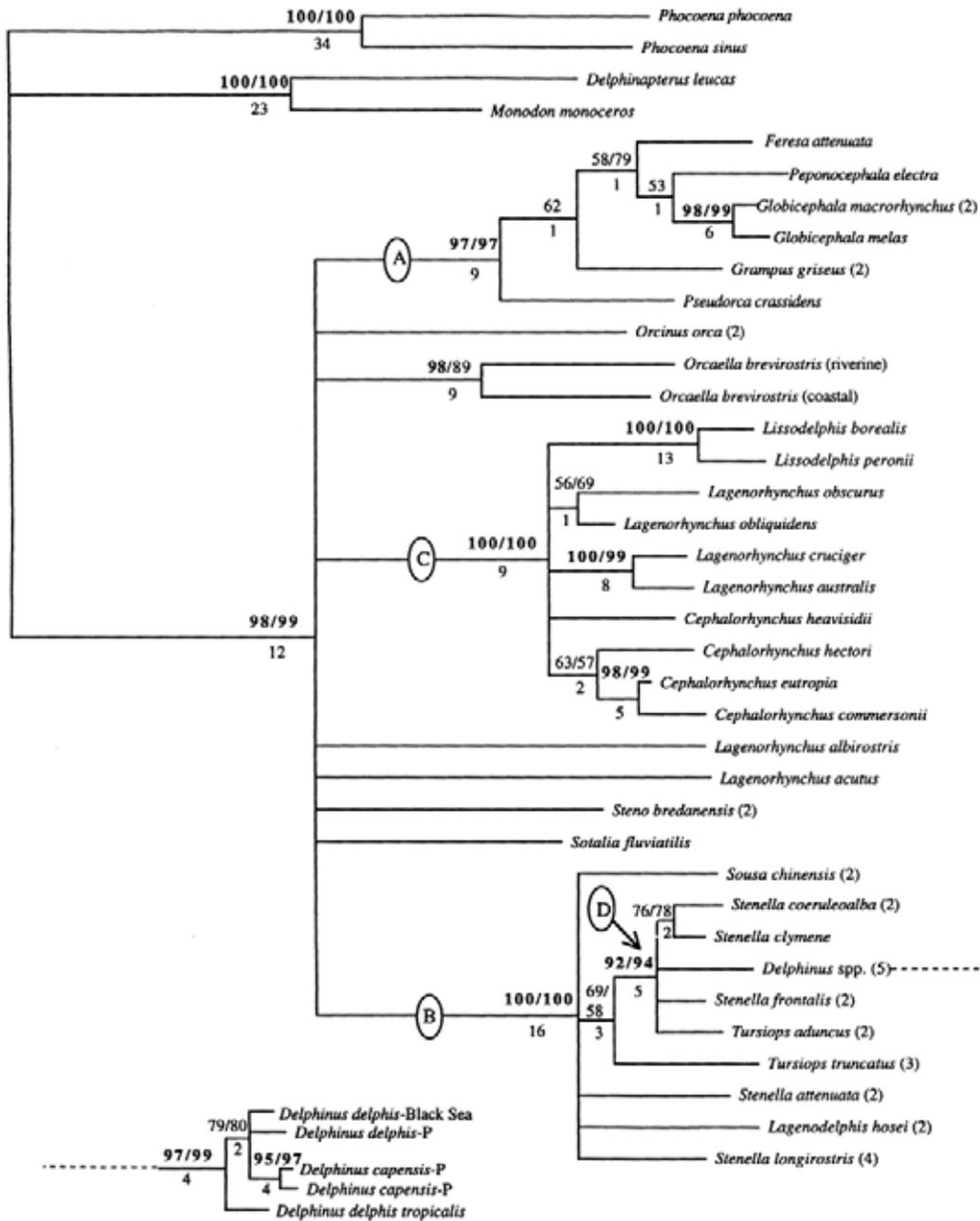


Figure 9. Phylogénie des Delphinidæ telle qu'inférée à partir de séquences du gène du Cytochrome B.

Les valeurs situées au-dessus des nœuds correspondent aux nombre de répliques de bootstrap concordant avec cette topologie.

Cette étude est la première faite à partir de marqueurs indiscutablement comparables et elle conduit à une remise en question de la taxonomie séculaire des genres et espèces de *Delphininae*. Elle démontre en substance que des les séquences issues d'échantillons appartenant à la même espèce peuvent s'avérer aussi divergentes entre elles, qu'envers d'autres séquences provenant de genres différents. Il nous semble en outre que le petit nombre d'échantillons par espèce dans cette étude ne

peut que conduire à une sous-estimation de l'ampleur de ce phénomène. Par ailleurs, la grande variabilité intraspécifique des patrons de colorations au sein des trois genres *Tursiops*, *Stenella* et *Delphinus* (**Perrin, 1997**) (et observation personnelle) va dans le même sens d'une polyphylie des espèces précédemment définies même si l'identité des spécimens des séquences déposées dans GenBank est discutée (**Ross et al., 2005**).

A.4) Le Dauphin bleu et blanc, taxonomie-synonymie

4(a) Taxonomie

Stenella est le diminutif latin du nom de genre *Steno* dont *Stenella* était considéré comme un sous-genre (**Gray, 1866**). Le nom spécifique *cæruleoalba* fait référence à la coloration bleue et blanche dont la signification évolutive avancée par certains auteurs serait la distinction face ventrale-face dorsale et le maintien de la cohésion intraspécifique (**Würsig et al., 1990**). Cette dernière peut s'avérer particulièrement importante dans un groupe d'espèces (genres *Stenella*, *Delphinus* et *Tursiops*) très proches et dont le rang évolutif reste encore très discutée.

Le genre *Stenella* contient 5 espèces actuelles (**Mead & Brownell, 1993**) :

- Le Dauphin tacheté pantropical, *Stenella attenuata* (**Gray, 1846**)
- Le Dauphin de Clymène, *Stenella clymene* (**Gray, 1850**)
- Le Dauphin tacheté atlantique, *Stenella frontalis* (**Cuvier & Valenciennes, 1828**)
- Le Dauphin longirostre, *Stenella longirostris* Gray, 1828
- Le Dauphin bleu et blanc, *Stenella cæruleoalba* (**Meyen, 1833**)

Ce genre est très probablement un assemblage artificiel de même que les genres-frères de *Delphinus* et *Tursiops* et leur taxonomie nécessite des travaux supplémentaires malgré une première mise au point relativement récente (**Leduc et al., 1999**). Par ailleurs, comme nous le verrons dans notre travail, l'histoire évolutive des espèces peut amener dans une même espèce des individus provenant d'horizons très différents surtout lorsqu'on a à faire avec des animaux dont la barrière d'espèce est parfois floue comme les Cétacés.

Après avoir été divisé en trois espèces (*S. styx* (**Gray, 1846**), *S. euphrasyne* Gray 1846 et *S. cæruleoalba* (**Meyen, 1833**)), ces dernières ont été réunies pour former *Stenella cæruleoalba* (**Okada, 1939; True, 1889**) malgré les réticences de Kellogg. La discussion

prit fin en 1970 (Fraser & Noble, 1970; Mitchell, 1970).

4(b) Synonymie

* *Clymene* Gray, 1864:237. Type : *Delphinus euphrosyne*. Sous-Genre de *Delphinus* Linnaeus. (≠ *Clymene* Lamarck, 1818 ou *Clymene* Savigny, 1822, Polychète)

* *Stenella* Gray, 1866b:213. Type d'espèce : *Steno attenuatus*. Sous-Genre de *Steno* Gray.

* *Euphrosyne* Gray, 1866b:214. Espèce-type : *Clymene euphrosyne*. Sous-Genre de *Clymene* Gray. (≠ *Euphrosyne* Miegen, 1800, Diptère).

* *Clymenia* Gray, 1868a:6. Espèce-type : *Delphinus longirostris*. (≠ *Clymenia* Savi, 1818 et Orsted, 1844, un genre de Vermes (vers indet.); ou Munster, 1834, Mollusque).

* *Micropia* Gray, 1868a:6. Espèce-type : *Clymenia stenorhyncha* (= *Delphinus microps* Gray).

* *Prodelphinus* Gervais, 1880:604. En substitution à *Clymenia* Gray.

* *Fretidelphis* Iredale & Troughton, 1934:65. Espèce-type : *Delphinus roseiventris* Wagner (= *Delphinus longirostris* Gray).

S. coeruleoalba ne comporte pas actuellement de sous-espèces. Les noms cités ci-dessus sont donc traités comme synonymes.

4(c) Spécimen-types

Delphinus coeruleo-albus Meyen, 1833:610. Localité-type : Rio de la Plata.

Delphinus Styx Gray, 1846:39. Localité-type : Afrique du Sud.

Delphinus Euphrosyne Gray, 1846:40. Localité-type : inconnue.

Delphinus Holbollii Nilsson, 1847:595. Localité-type : Groenland.

Delphinus lateralis Peale, 1848:35. Localité-type : 13°58' N, 161°22' W (Pacifique nord).

Delphinus tethyos Gervais, 1853:150. Localité-type : Valras, embouchure de l'Orb, Languedoc, France.

Delphinus marginatus Pucheran, 1856:545. Localité-type : Dieppe, France.

Delphinus mediterraneus Loche, 1860:475. Localité-type : Algérie.

Delphinus cesthenops Cope, 1865:200. Localité-type : inconnue.

Delphinus crotaphiscus Cope, 1865:200. Localité-type : inconnue.

Tursio Dorcides Gray, 1866a:400. Localité-type : inconnue.

Clymene dorides Gray, 1866b:214. Syn. *Tursio Dorcides* Gray.

Clymenia euphrosynoides Gray, 1868a:6. Syn. *Delphinus euphrosyne* Gray.

Clymene similis Gray, 1868b:146. Localité-type : Cape of Good Hope, Afrique du Sud.

Clymene Burmeisteri Malm, 1871:63. Localité-type : Brasil.

Clymenia novae-zelandiae Hector, 1873:159. Localité-type : Waikanae, New Zealand.

Prodelphinus Petersii Lutken, 1889:40. Localité-type : Océan Indien.

Delphinus amphitriteus Philippi, 1893:7. Localité-type : 29°15' N, Océan Atlantique.

4(d) Noms vernaculaires

Les noms suivants sont utilisés dans diverses langues pour désigner *S. coeruleoalba* : Anglais : Euphrosyne dolphin, Gray's dolphin, blue-white dolphin, Meyen's dolphin, longsnouted dolphin, Greek dolphin, harnessed dolphin, black-jawed dolphin, Gray's porpoise, Gray's longsnouted porpoise, streaker porpoise, streaker, whitebelly porpoise ; Français : dauphin bleu et blanc, dauphin rayé (inusité), dauphin euphrosyne ; Espagnol : delfín listado, delfín rayado, delfín a rayas, delfín de betas, estenela ; Catalan : dofí llistat, dofí blanc i blau ; Occitan : dofin lhístat, dofin blanc e blau ; Portugais : boto, golfinho estriado, golfinho riscado ; Italien : stenella striata, delfino dalle briglie, delfino eufrosine ; Grec : ζωνοδέλφινο (zonodélfino) ; Allemand : blauweisser Delfin or Delphin, Streifen-delfin; Néerlandais : doljijn van Gray, gestreepte dolfijn; Afrikaans: Streepdolfyn; Danois : stri- bet deljin; Suédois : strimmig deljin; Norvégien : Stripedelfin Polonais : Delfin pręgoboki ; Russe : polosatyy prodelfin ; Croate : prugasti dupin ; Serbe : pliskauika praua, dupin prau; Turc : çizgili yunus ; Japonais : スジイルカ (suji iruka ou suzi iruka) ; Chinois mandarin : 蓝白细吻海豚 ; Arabe : نيفلـد قرزا ضـيباو (delfin azraq wa abyad) ; Hébreu : הלגטס תספסופם (stenella mefuspeset).

II. B. Le Dauphin bleu et blanc : aspects biologiques

B.1) Morphologie, détermination

1(a) Aspect morphologique

La forme générale du corps du Dauphin bleu et blanc est similaire à celle de la plupart des autres petits *Delphinidae* : un corps fusiforme, un assez long bec nettement séparé du melon par un sillon très net, une nageoire dorsale falciforme et des nageoires pectorales effilées.

Malgré des caractéristiques assez précises (*cf.* ci-dessous), dans les zones où celle-ci cohabite avec d'autres, un œil non averti peut le confondre avec d'autres espèces de *Delphinidae* des Genre *Stenella*, *Delphinus*, *Tursiops*, *Lagenodelphis*,...

1(b) Détermination

Il s'agit d'un dauphin moyennement robuste (long. max. : 2,56 m, (Okada, 1939) et masse max. 156 kg (Ross, 1984) ; cf. Tableau 1 ci-après).

Stenella caeruleoalba est distingué des autres *Delphinidae* par patron de coloration composé d'une « flamme » de couleur claire et d'une rayure (d'où ses noms en anglais et castillan de *Striped dolphin* et *Delfin listado*) depuis l'œil jusqu'à l'anus, accompagnée de deux rayures accessoires dont une vers la nageoire pectorale.

Longueur totale moyenne (nombre d'individus)	Localisation	Références
220(♀)/236(♂)	Pacifique	(Kasuya, 1973; Kasuya, 1976; Miyazaki, 1984)
	Méditerranée (4-5 cm de différence entre le Nord et le Sud ?)	(Calzada & Aguilar, 1995)
5-8 cm de plus qu'en Méditerranée	Atlantique	(Calzada & Aguilar, 1995)

Tableau 4. Taille relative des Dauphins bleu et blanc selon leur origine géographique

B.2) Données physiologiques

2(a) Alimentation

La plupart des espèces de *Delphinidae* se nourrissent de Poissons et, dans des proportions variables, de Mollusques Céphalopodes et de petits Crustacés Décapodes. Ils possèdent entre 50 et 200 dents.

Le Dauphin bleu et blanc est un Cétacé opportuniste, c'est-à-dire qu'il s'alimente d'une grande variété de proies épi-, méso- et bathypélagiques qui forment de grandes agrégations. Ils consomment donc aussi bien de Poissons Téléostéens, Crustacés Décapodes et Mollusques pélagiques. Les Téléostéens (*Gadidae*, *Sparidae* & *Gonostomiidae*) représentent la majorité des éléments consommés mais il semble que les Mollusques (*Histioteuthidae*, *Ommastrephidae*, *Enoploteuthidae* & *Onychoteuthidae*) représentent la majorité de la masse consommée (Blanco *et al.*, 1995; Meotti & Podestà, 1997). Les Crustacés consommés sont des petites crevettes (*Decapoda* : *Euphausiacea*) : leur consommation est saisonnière et toujours minoritaire. La taille des proies consommées est inférieure à 130 mm pour les Poissons et à 200 mm pour les Céphalopodes (Wurtz & Marrale, 1991).

	IB	GZ	PC	PC	GC	MP	DP	IS	Niveau trophique
SC	0,05		0,20	0,15	0,05	0,3	0,25		4,2
Slong		0,20	0,20		0,40	0,20		4,3	
TT			0,20	0,05	0,15		0,60		4,2
Gm			0,05	0,05	0,1		0,40	0,40	4,5
BP		0,80	0,05		0,05	0,05	0,05		3,4
ZC	0,1		0,3	0,3		0,15	0,15		4,3
PM	0,05		0,1	0,6	0,05	0,05	0,15		4,4
DD			0,15	0,15	0,10	0,40	0,20		4,2
PP	0,05		0,10	0,10	0,30		0,45		4,1

Tableau 5. Composition relative de l'alimentation du Dauphin bleu et blanc (données compilées par Pauly *et al.* (1998), d'après Jefferson & Curry (2003); Leatherwood & Reeves (1983); Perrin *et al.* (1973); Santos *et al.* (1994); Wurtz *et al.* (1992))

SC = *Stenella coeruleoalba*, Slong = *Stenella longirostris*, TT = *Tursiops truncatus*, Gm = *Globicephala melæna*, BP = *Balænoptera physalus*, ZC = *Ziphius cavirostris*, PM = *Physeter macrocephalus*, DD = *Delphinus delphis*, PP = *Phocæna phocæna* ;

IB = Invertébrés benthiques, GZ = Zooplancton de grande taille, PC = Petits Calmars, GC = Grands Calmars, MP = Mésopélagiques, DP = Poissons Divers, IS = Invertébrés Supérieurs.

2(b) Reproduction-Croissance-Développement

La période œstrale a été estimée être en août selon la variation saisonnière de la taille des testicules (Calzada & Aguilar, 1996). On estime que la saison de mise-bas se situe entre la fin de l'été et l'automne (pour une durée de gestation en conséquence d'un peu plus de 12 mois), précèdent un pic de productivité, lors du mélange des eaux les plus chaudes avec les *upwellings* d'automne (Aguilar, 1991). Pendant la saison de reproduction, environ 25% des troupeaux comportent des petits et la mise bas semble se faire dans les troupeaux les plus importants (« reproductive schools »). Ce comportement est plus généralement celui de la plupart des espèces de petits *Delphinidae* (Jefferson & Curry, 2003; Miyazaki & Nishiwaki, 1978; Norris & Dohl, 1988). Le sex-ratio passerait d'une valeur supérieure à 1 (1,4 :1 ; n=12) à 1 en raison d'une mortalité précoce supérieure des mâles comme c'est généralement le cas chez d'autres Mammifères. La mortalité en lactation est la plus importante d'après le pic de mortalité observé en octobre-novembre (Aguilar, 1991)).

L'âge à la maturité sexuelle a été estimé à 12 ans/187 cm pour les femelles et 11,3 ans/190 cm for the males (Calzada & Aguilar, 1996) mais (Di Mèglio *et al.*, 1996) n'observent pas de dimorphisme sexuel de taille significatif. Ces valeurs sont bien

supérieures aux valeurs observées par (Kasuya, 1985) dans la population pacifique. En outre, (Calzada & Aguilar, 1995) observent des variations géographiques de la taille des Dauphins bleu et blanc à l'échelle de la Méditerranée occidentale.

2(c) Comportement

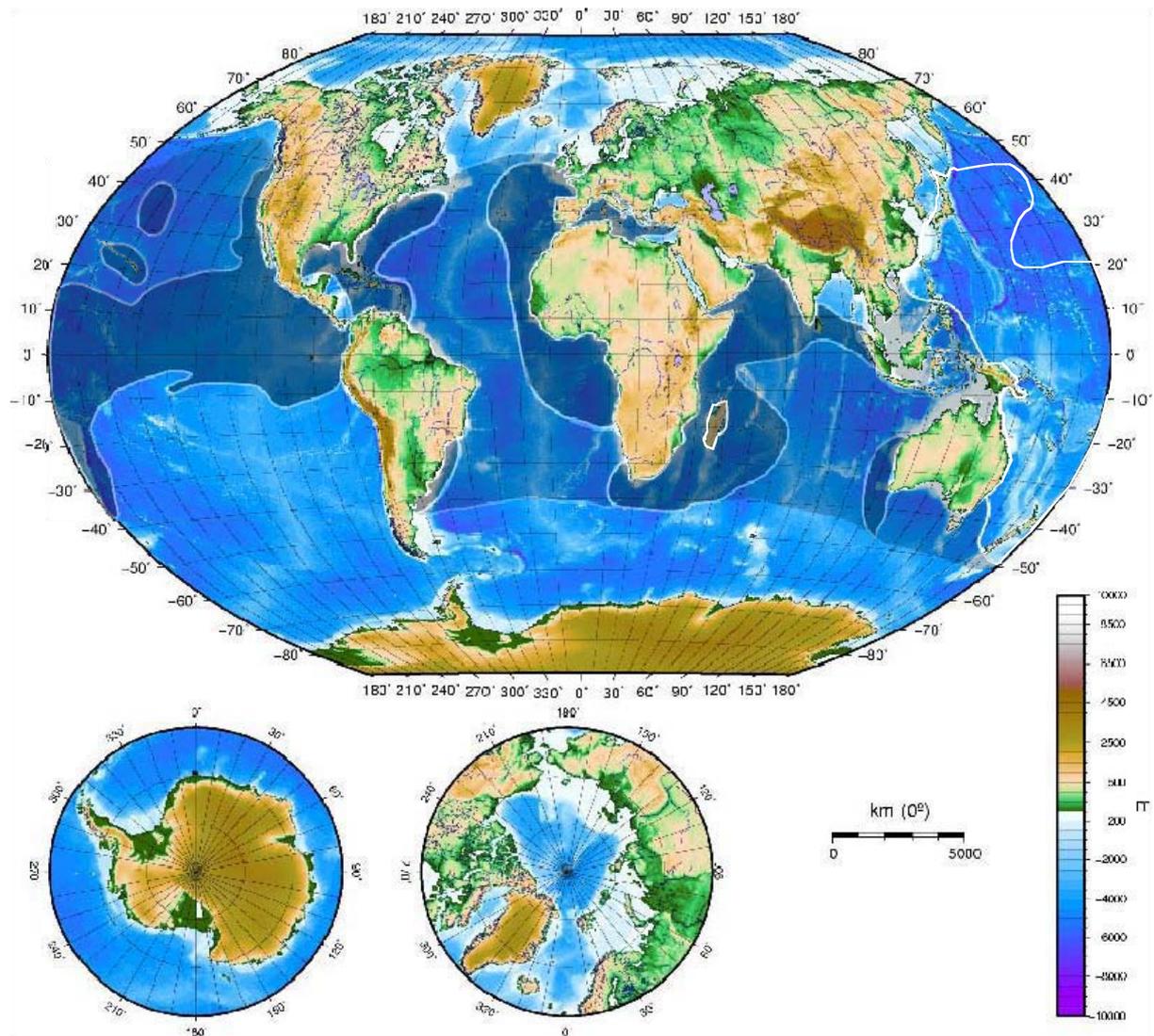
Le Dauphin bleu et blanc est régulièrement observé en troupes mixtes avec des Dauphins communs *Delphinus delphis* et/ou des dauphins de Risso *Grampus griseus* (Frantzis & Herzog, 2002). Ces regroupements donnent lieu à des interactions interspécifiques complexes incluant bien sûr des événements de prédation en groupe mais aussi des interactions sociales.

B.3) Distribution, migrations

3(a) Répartition géographique

a(i) Répartition mondiale

Le Dauphin bleu et blanc est présent dans les eaux tempérées aux tropicales (Perrin *et al.*, 1994a). Cette espèce est bien documentée dans le Pacifique occidental et oriental au large du Japon et de l'Amérique du nord. À travers le Pacifique nord, la plupart des observations sont situées au-dessous de 43°N environ et ne descendent pas au sud de 10°S dans quart sud-est du Pacifique. L'espèce suit une distribution continue du Japon à la Nouvelle-Zélande dans la moitié occidentale du Pacifique. On ne sait pas si elle est continue par ailleurs (*cf.* Figure 10 ci-dessous).



**Figure 10. Répartition géographique connue (sombre) et supposée (clair).
Compilation d'après de nombreuses sources.**

Dans l'Océan Atlantique, on observe cette espèce le long des côtes d'Amérique du nord, Golfe du Mexique et Mer Caraïbe ainsi que le long des côtes brésiliennes et nord-argentine. En Europe, et Afrique sa distribution est continue du sud de la Mer du nord et de la Mer d'Irlande au Cap Frio (Angola) et à plusieurs milliers de milles des côtes (Iles Atlantiques en particulier). En outre, la difficulté d'investigation le long des côtes namibiennes laisse présager une continuité avec les observations méridionales et de l'Océan Indien. Le Dauphin bleu et blanc comporte une population très importante en Méditerranée surtout dans le bassin occidental. Il semblerait que sa faible densité en Méditerranée méridionale soit due à une pression d'observation moindre (Forcada & Hammond, 1998; Marchessaux, 1980). Les

observations les plus nordiques semblent suivre l'influence du *Gulf Stream* et les observations enregistrées au Canada, Groënland et Islande seraient selon toute vraisemblance « extra-limitales » (**Bloch et al., 1996**).

Dans l'Océan Indien, le Dauphin bleu et blanc est observé de l'embouchure du fleuve Orange (Afrique du Sud) à Sumatra, en passant par les Mascareignes, Seychelles, le Golfe d'Oman, Sri Lanka et les Maldives mais semblant éviter le nord du Golfe du Bengale et la Mer de Timor. Sa distribution reprend alors au large des côtes d'Australie occidentale. Une fois encore, la distribution du Dauphin bleu et blanc semble interrompue le long de côtes peu explorées scientifiquement.

En conclusion, le Dauphin bleu et blanc est une espèce cosmopolite dont la distribution est probablement continue (*cf.* **Figure 10** ci-dessus) du 50 parallèle nord au 50^e parallèle sud à l'exception notable du quart sud-est du Pacifique, siège de la radiation évolutive du genre *Cephalorhynchus* d'avec qui la compétition a dû contribuer à limiter l'extension de *Stenella caeruleoalba*.

a(ii) En Méditerranée

En Méditerranée, le Dauphin bleu et blanc est l'espèce de Cétacé la plus fréquemment rencontrée (**Beaubrun, 1995; Gannier, 2005; Notarbartolo di Sciara, 2002a**). Sa densité semble la plus élevée dans le bassin liguro-provençal (0,24 individus.km⁻²) et la Mer d'Alborán (0,20 individus.km⁻²), les deux zones réputées avoir la plus forte productivité primaire (Tableau 6). En effet, ces zones, en raison des forts vents d'ouest (*Poniente* en Andalousie) ou de nord-ouest (*Tramuntana* dans le Golfe du Lion) sont des zones d'*upwelling* : la couche d'eau superficielle chaude est repoussée par les vents en provenance du continent et remplacée par une couche d'eau profonde plus froide par un mouvement circulaire. Cette circulation est responsable d'un brassage de nutriments et de micro-organismes favorisant la productivité (**Ekman, 1906**).

Une autre explication à ces différences de densités pourrait être le fait qu'un épisode viral majeur (*cf.* V ci-après) ayant décimé les Dauphins bleu et blanc en 1990-92 a justement commencé dans les eaux actuellement à plus faible densité (Mer Baléare) et que cette dernière pourrait être un stigmate de la grande mortalité observée au début de l'épizootie (**Aguilar & Raga, 1993**).

		Estimateur ponctuel	I.C. 95%
	Nd	5 826	2 193-15 476
« Mer » baléare (1)	Dd	0.09	0.03-0.23
	Nd	30 774	17 433-54 323
Bassin Provençal (2)	Dd	0.23	0.13-0.41
	Nd	14 003	6 305-31 101
Mer Ligure (3)	Dd	0.30	0.14-0.67
	Nd	42 604	24 962-72 716
Bassin Liguro-Provençal (2+3)	Dd	0.24	0.14-0.40
	Nd	48 098	29 388-78 721
Méditerranée NO (1+2+3)	Dd	0.20	0.12-0.32
	Nd	17 728	9 507-33 059
Mer d'Alborán (4)	Dd	0.20	0.10-0.36
	Nd	18 810	8 825-35 940
Secteur Sud-Baléare (5)	Dd	0.08	0.04-0.16
	Nd	39 693	18 206-87 721
Méditerranée SO (4+5)	Dd	0.12	0.05-0.25

Tableau 6. Nombre (Nd) et densités (Dd/km²) de Dauphins b. et b. dans diverses zones de la Méditerranée (d'après (Forcada & Hammond, 1998) ; cf. figure ci-dessous pour la localisation ; I.C. = Intervalle de Confiance).

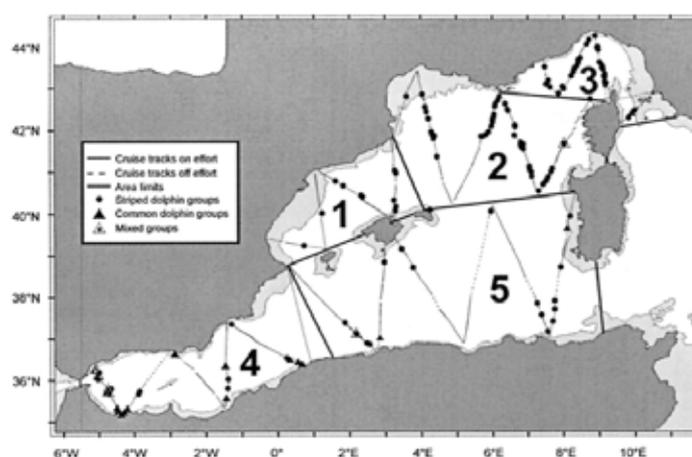


Figure 11. Localisation des zones et effort d'observation de l'étude d'abondance des Dauphins b.et b. (Forcada & Hammond, 1998). (la numérotation des zones suit celle du tableau 4)

3(b) Migrations, répartition saisonnière.

Le Dauphin bleu et blanc est généralement considéré comme une espèce non-migratoire en particulier au large des côtes européennes et nord-américaines. Cependant, il est fréquemment associé à des fronts tièdes de courants océaniques et fait donc des intrusions jusqu'au 46^e parallèle nord au large de l'archipel nippon en été en raison de la présence du courant chaud Kuroshivo alors qu'en hiver, il ne s'aventure pas au nord du 33^e parallèle (Perrin *et al.*, 1994a). Dans les mêmes eaux,

ils ont un habitus plus côtier en septembre-octobre et se dispersent plus au sud en Mer de Chine en hiver. Ils se dispersent au printemps-été dans un habitat beaucoup plus pélagique dans le Pacifique Nord (Miyazaki *et al.*, 1974).

Des mouvements saisonniers pourraient aussi se produire en Méditerranée : les troupes semblent se déplacer vers le nord du bassin occidental l'été, lorsque la température des eaux de surface augmente dans le sud du bassin (Gannier, 1999; Perrin *et al.*, 1994a). Des mouvements circadiens se produisent aussi : la nuit, les Dauphins ont un pic d'activité de prédation le long du talus continental de la Mer Ligure alors que s'opèrent des mouvements vers le continent le soir et vers le large le matin dont les motivations sont inconnues (Gannier, 1999).

II. C. Conservation du dauphin bleu et blanc.

Quelques populations de dauphin bleu et blanc rencontrent des difficultés, notamment dans le Pacifique nord-ouest et en Méditerranée.

Des dauphins bleu et blanc sont capturés volontairement ou incidemment dans de nombreuses autres parties du Monde, mais les stocks ne semblent pas beaucoup s'y réduire. Cependant, rappelons que l'estimation de l'abondance de cétacés est un exercice délicat. Dans la liste rouge de l'IUCN, le dauphin bleu et blanc a été classé en 1996 comme présentant un faible risque d'extinction (« *lower risk* ») mais dépendant de mesures de conservation (« conservation dépendant ») (IUCN, <http://www.redlist.org>).

C.1) Pacifique Nord-ouest (IUCN, 1996).

Les captures de dauphins bleu et blanc au Japon se sont réduites de façon spectaculaire depuis les années 1950, probablement suite à une chasse excessive (ou un changement des méthodes halieutiques, car les captures accidentelles étaient également nombreuses). Les estimations d'abondance dans les eaux japonaises, dans les années 1980, ont donné des valeurs supérieures au demi-million d'individus, mais plusieurs populations continuent d'être pêchées, et les dauphins bleu et blanc ont été complètement ou presque éliminés de certaines zones autrefois fréquentées.

La forte demande de viande de dauphin sur les marchés japonais rend l'accomplissement des mesures de conservation problématique.

C.2) Méditerranée.

Une épizootie de morbillivirose (infection au Dolphin Morbillivirus, cf. V ci-après) a probablement provoqué la mort de milliers d'individus (**Aguilar & Raga, 1993; Bortolotto et al., 1992**), en Méditerranée centrale et occidentale, de 1990 à 1992. La population de Méditerranée occidentale, en 1991, a été estimée à 117 880 individus (intervalle de confiance à 95 % : 68 379 - 214 800 ; Forcada *et al.*, 1994). La pollution et la réduction de la disponibilité des proies sont supposées avoir aggravé les effets de l'épizootie, et persistent aujourd'hui.

III. La Zone D'étude : La mer Méditerranée et le proche Atlantique

III. A. Données physiques (Divers, 1999)

A.1) Géologie

La géologie de la Méditerranée est complexe, impliquant des cassures et des collisions entre les plaques africaines et eurasiennes (*cf.* **Figure 12** ci-après). Bien qu'intracontinentale, on y trouve cependant un fond basaltique au large des côtes françaises, ce qui en fait un « mini-océan du Languedoc ».

La mer Méditerranée se divise en deux bassins bien individualisés, séparés par des hauts fonds situés entre la Sicile et la Tunisie : le bassin - ou Méditerranée - occidental(e) et le bassin - ou Méditerranée - oriental(e), eux-mêmes géologiquement très compartimentés. Le premier recouvre une superficie d'environ 0,85 millions de km² tandis que la seconde recouvre environ 1,65 millions de km². En Méditerranée occidentale, la dorsale corso-sarde isole la mer Tyrrhénienne du reste du bassin. Le bassin oriental est constitué de la mer Ionienne, qui atteint les plus grandes profondeurs, et du bassin levantin. Ils sont séparés par un goulet entre la plate-forme de la mer Égée limitée au sud par la Crète et les hauts fonds au large de la Cyrénaïque.

La mer méditerranéenne a une profondeur moyenne de 1 500 m et le point enregistré le plus profond est de 5 267 mètres en mer Ionienne. Le littoral se déroule sur 46 000 kilomètres. « Mer de montagnes », la Méditerranée possède un littoral en moyenne assez escarpé : le plateau continental y est relativement réduit, sauf dans trois zones : la mer Adriatique, la mer Égée, et le Golfe de Libye. A plus petite échelle cependant, certaines zones côtières possèdent des petites formations sédimentaires comme le Golfe du Lion ou le Delta du Nil. Leur talus peut être le siège d'*upwelling*.

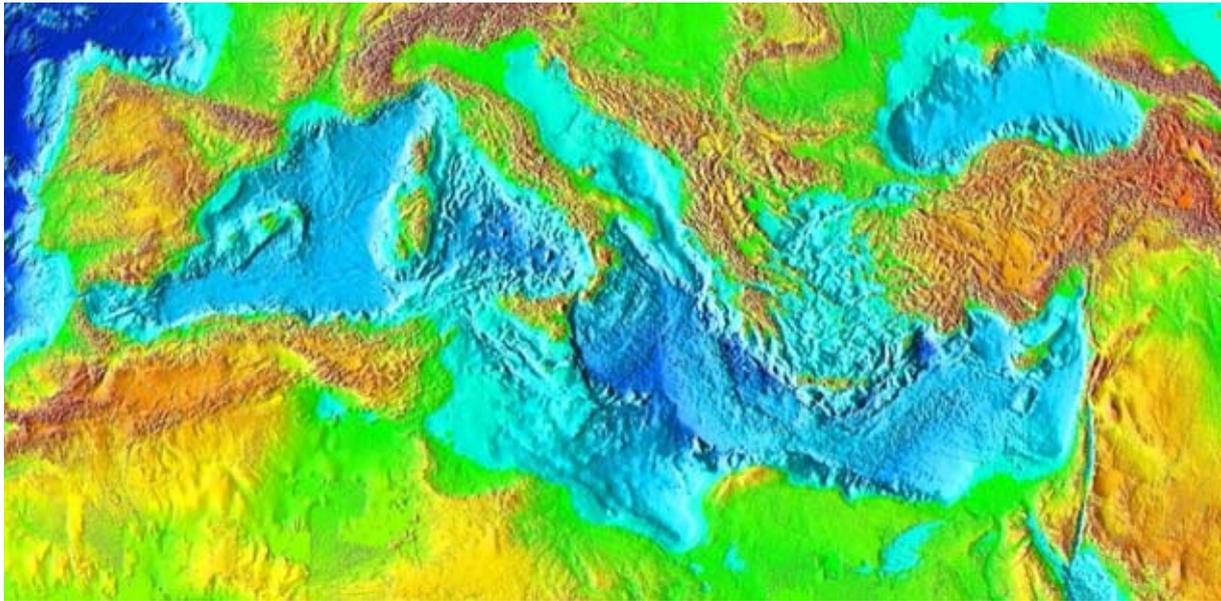


Figure 12. Géographie physique du bassin méditerranéen (cliché météosat).

A.2) Hydrographie

2(a) Les Courants Superficiels Méditerranéens (C.S.M.)

Ces courants sont provoqués par l'afflux des eaux atlantiques et suivent un mouvement cyclonique (anti-horaire). Ces courants superficiels en raison de la plus faible salinité des eaux de l'Atlantique suit les côtes du Maghreb en raison de la force de Coriolis et une petite partie remonte vers les Baléares en raison de la collision avec le courant anticyclonique de la mer d'Alborán. Ils se séparent de nouveau en un courant remontant le long des côtes italiennes et un courant « levantin » qui donne à son tour des courants mineurs cycloniques (Adriatique, Ionien et Égéen) et anticycloniques (Libyen)

2(b) Les Courants Intermédiaires

Ces courants circulent dans la strate entre 200 et 600 mètres de profondeur dans le sens opposé aux C.S.M. Son origine se situe dans le Bassin du Levant où la Méditerranée est la plus salée ($0,39 \text{ g.l}^{-1}$) et lors de l'hiver, l'augmentation de la densité de la couche superficielle « comprime » la strate inférieure, donnant naissance à ce courant. Il se divise en deux rameaux secondaires, dans le Golfe de Syrte et dans la Mer Adriatique, et un rameau principal qui parcourt la Méditerranée, traverse le Détroit de Sicile par deux chenaux de plus grande profondeur dont le plus

septentrional suit le trajet du Courant Liguro-Provençal jusqu'à Gibraltar.

On estime qu'il faut environ 150 ans à l'eau en provenance de l'Atlantique pour effectuer tout le parcours et se trouver profondément modifiée en composition et l'on considère que la totalité de l'eau se renouvelle en un siècle.

2(c) Courants profonds

Les Courants Profonds se situent dans les bassins profonds (Liguro-Provençal et Ionien) où le refroidissement hivernal causé par les vents du nord (*Tramuntana/Mistral & Bora*) provoque une augmentation de la densité des eaux pélagiques (centre du Golfe du Lion et Adriatique). Ces eaux circulent alors à plusieurs milliers de mètres de profondeur et contribuent au lent renouvellement des eaux profondes.

III. B. Ecologie

(Divers, 1999; Quignard & Tomasini, 2000)

B.1) Le talus continental est une zone trophique importante alors que les fonds sont désertiques

La diversité par strate latitudinale baisse très rapidement : on considère que 25% des métazoaires sont localisés dans les 25 premiers mètres, 75% dans les 300 premiers, qu'à 2000 m ne subsistent plus que 7 espèces et à 2904 m seul persiste le grenadier *Chalinura mediterranea*. Cependant la découverte d'écosystèmes particuliers, comme les sources hydrothermales, pourraient laisser penser que les grands fonds méditerranéens, même s'ils sont les plus étudiés de la planète, pourraient réserver quelques surprises malgré l'effet stérilisant de la Crise de Salinité du Messinien (C.S.M. ; cf. 2(a) ci-après) sur ces mêmes grands fonds. Cette pauvreté s'exprime avec acuité au niveau des « groupes écologiques », des guildes : les poissons herbivores sont très rares (Saupé *Boops salpa*, Sars *Diplodus sargus sargus*, Bogue *Boops boops* et quelques Blennies - *Blenniidae*). Cette relative pauvreté a plusieurs origines :

- les eaux atlantiques entrant sont pauvres, « usées » par la production

océanique en amont ;

- les fleuves méditerranéens ne sont pas riches en matières organiques : seuls les fleuves dirigés vers la Mer Noire et la Caspienne le sont ;

- le seuil de Gibraltar et les crises anoxiques que la Méditerranée a connues ont limité l'occupation des grands fonds ;

- la Méditerranée est chaude en profondeur : de 300 à 5 000 m, on constate une homothermie entre 11 et 13°C ce qui limite l'occupation des grands fonds par la plupart des animaux abyssaux et bathyaux que l'on trouve dans l'Atlantique.

B.2) La biodiversité a dû s'adapter à des conditions changeantes

La position charnière entre espace boréal et tropical explique en partie cette diversité mais après la Crise de Salinité du Messinien, ce sont les événements transgressifs/régressifs liés aux variations climatiques du Plio-Pléistocène qui ont construit cette mosaïque de diversité : elle a pu retenir par son orientation géographique est-ouest, ses caractéristiques hydro-climatiques plus constantes et l'effet « cul-de-sac » une faune « chassée » des latitudes boréales par les maxima glaciaires ainsi qu'une faune repoussée vers le nord issue des latitudes tropicales par les chaleurs inter-glaciaires (**Francour et al., 1994**). Il en résulte une faune et une flore très riches : 8 750 métazoaires, 12 000-15 000 dont 25-28% d'endémismes (**Quignard & Tomasini, 2000**). On considère qu'elle abrite entre 7 et 11% de la faune marine mondiale. Cette grande diversité et parfois l'exubérance l'a fait qualifier par certains de mer subtropicale. Mais elle est aussi caractérisée par une relative pauvreté puisque 87% des espèces sont concentrées dans le bassin occidental contre 45% dans le bassin oriental.

En raison de l'assèchement de la Méditerranée pendant la Crise de Salinité du Messinien, le biotope marin de la mer Méditerranée est principalement issu de l'océan Atlantique (**Francour et al., 1994**). L'Atlantique nord-est beaucoup plus froid et plus riche en nutriments que la Méditerranée, et la vie marine a dû s'adapter à des conditions changeantes au cours des 5 millions d'années qui ont suivi son remplissage.

2(a) La Crise de Salinité du Messinien

a(i) Description

La Crise de Salinité du Messinien, vers la fin de Miocène (environ 5,96-5,33 M.a.), marque l'assèchement d'une grande partie de la Méditerranée. Elle est considérée comme l'événement géologique le plus important survenu dans les 20 derniers millions d'années (**Krijgsman et al., 1999**). On en vint à la suspecter lorsqu'on s'aperçut de l'existence de nombreux endémismes et de la relative pauvreté de certains *taxa* bathypélagiques (pour une synthèse cf. **Emig & Geistdoerfer, 2004**). On lui attribue ainsi l'existence de « ruptures phylogéniques » au sein des *Sparidae* (**Bargelloni et al., 2003**) ou l'existence d'espèces cryptiques comme chez les *Gobiidae* du genre *Pomatoschistus* (**Gysels et al., 2004**), entre autres.

L'événement tectonique supposé être à l'origine de la C.S.M. est le soulèvement des communications bético-rifaines (Andalousie-Atlas marocain). Les vents d'ouest, les marées atlantiques ainsi qu'un bilan hydrique négatif ont pu empêcher la sortie des eaux hyper-salines de la couche profonde. La C.S.M. a alors commencé lorsque l'entrée d'eau est devenue inférieure au déficit hydrique. Cela a conduit à l'isolement complet de la Méditerranée, les communications d'avec l'Indo-Pacifique ayant été interrompues dès le Miocène ancien. La communication « rifaine » (méridionale) a été comblée dès la fin du Tortonien (**Weijermars, 1988**) alors que le chenal « bétique » a dû persister laisser place à des formations corallines alternant avec une série de bassins néritiques (**Müller & Hsü, 1987**).

En modélisant la C.S.M., **Blanc (2000)** est arrivé à la conclusion que seule une séparation des bassins occidental et oriental peuvent expliquer une différence dans l'épaisseur des dépôts d'évaporites salins entre les deux bassins. Dans ce modèle, deux « poches » méditerranéennes ont persisté, près du sud de l'actuelle péninsule ibérique et près de la Sicile, à environ 1000 mètres en dessous du niveau actuel (cf. **Figure 13** ci-après). La persistance de telles formations de « lacs marins » (*lago mare*) est en accord avec la disparition des Brachiopodes et de la faune tropicale de l'Éocène-Oligocène des couches superficielles et leur persistance dans les eaux profondes (**Emig, 1989; Emig & Geistdoerfer, 2004**). De plus, les données moléculaires de **Zane et al. (2000)** suggèrent que le Krill de l'espèce *Meganyctyphanes*

norvegica, dont s'alimente le Rorqual commun en Méditerranée, aurait disparu à cette époque puis aurait recolonisé le bassin méditerranéen lors de son remplissage.

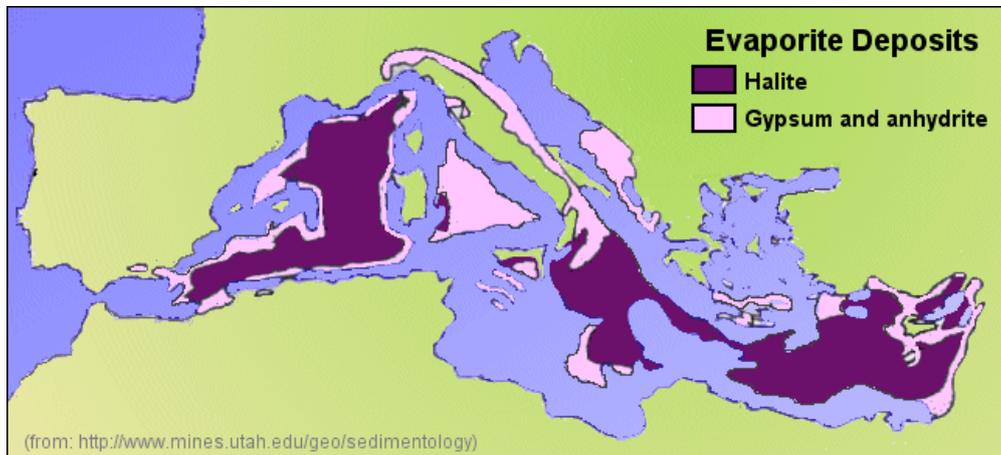


Figure 13. Dépôts d'évaporites dans le bassin méditerranéen.

a(ii) Corrélat paléontologiques et génétiques

La divergence des *Cyprinidæ* de la Méditerranée Nord-Occidentale est datée de cette époque et de nombreuses espèces de ce groupe s'hybrident entre elles (Berrebi, 1995). Par ailleurs, les Gobies du genre *Pomatoschistus* (« Sand gobies ») de l'Adriatique, dont la divergence est contemporaine, auraient pu persister grâce à l'existence de refuges lacustres/fluviaux (Gysels *et al.*, 2004). Cela n'est pas sans rappeler la persistance de lignées paraphylétiques de ce que nous connaissons aujourd'hui comme Dauphins de rivière (*Platanistoidea*, *Inioidea* and *Lipotidæ*), qui s'est effectuée de façon indépendante pour chacun des *taxa* (Cassens *et al.*, 2000).

Par ailleurs, l'isolement complet de la Méditerranée de l'Atlantique lors de la C.S.M. (la fermeture des communications avec l'Indo-Pacifique ayant eu lieu bien avant au Miocène ancien) aurait aussi eu des conséquences globales comme l'événement glaciaire du Kapitéen-Optien (Weijermars, 1988), qui a eu des répercutions sur les écosystèmes terrestres.

2(b) Phénomènes invasifs (Zenetos *et al.*, 2005)

Le percement du canal de Suez en 1869 a créé le premier passage d'eau de mer entre la mer méditerranéenne et la mer rouge. Cette dernière est plus haute que la partie

orientale de la Méditerranée. Aussi, le canal forma un fleuve d'eau salée de la mer rouge qui se déversait dans l'eau de mer de la Méditerranée. Les grands lacs Amer, qui sont des lacs normaux très salés, forment le canal Suez et ont bloqué la migration des espèces de la mer rouge vers la Méditerranée pendant plusieurs décennies, mais comme la salinité des lacs s'est égalisée graduellement avec celle de la mer rouge, la barrière migratoire s'est levée, et les plantes et les animaux de la mer rouge ont commencé à coloniser la Méditerranée orientale.

La mer rouge est plus salée et plus pauvre en nutriments que l'océan atlantique car le détroit du Gibraltar bloque les grands courants de l'Atlantique. Les espèces animales et végétales de la Mer Rouge prennent l'avantage sur les espèces de l'Océan Atlantique dans l'environnement méditerranéen oriental salé et pauvre en nutriments. En effet, la construction du barrage d'Assouan sur le Nil dans les années 60 a réduit l'apport d'eau douce riche en nutriments dans la Méditerranée orientale, ce qui rend l'environnement de la Méditerranée proche de celui de la mer rouge. Cet échange d'espèces est connu sous le nom de migrations lessepsiennes (Ferdinand de Lesseps fut l'ingénieur qui a supervisé la construction du canal de Suez).

B.3) La Méditerranée héberge quelques Vertébrés remarquables

Parmi ces Vertébrés, on peut citer :

- la Caouanne ou Tortue marine commune *Caretta caretta* dont la survie est sérieusement menacée en raison des dégâts sur ses sites de nidification ;

- le Phoque moine *Monachus monachus* qui a une population réduite à 150-250 exemplaires concentrés surtout en Mer Egée avec quelques individus résiduels en Mer Ionienne, côte nord-africaine et des individus erratiques en Mer Tyrrénienne jusqu'en Sardaigne et Corse ;

- la Tortue verte *Chelonia mydas* qui est pratiquement éteinte, ne se trouvant que rarement sur les côtes de Chypre ;

Les Cétacés présents en Méditerranée sont tous considérés comme « à risque » sauf le Dauphin bleu et blanc.

III. C. Menaces pesant sur l'environnement méditerranéen.

C.1) Une menace non ou marginalement anthropique : le phénomène de tropicalisation

(Golani *et al.*, 2002)

On entend par tropicalisation le processus d'installation en Méditerranée d'espèces en provenance de régions tropicales et sub-tropicales. Cela concerne les migrants lessepsiens (migrants en provenance de la Mer Rouge par le Canal de Suez, donc migration artificielle) mais aussi des migrants atlantiques (« herculéens ») originaires des côtes africaines de l'Océan atlantique ainsi que des espèces transportées par les navires de commerce (décharges incontrôlées en mer ou activités portuaires) ou de loisir voire de vidange d'aquariums (institutionnels - *Caulerpa* à Monaco - ou privés - Piranhas dans les lagunes côtières). Une contribution substantielle à ces derniers migrants est le réchauffement climatique par le déplacement vers le nord de la limite des eaux tropicales.

Certaines de ces espèces se sont acclimatées et se reproduisent très bien au point d'arriver à supplanter des espèces autochtones voire d'être pêchées et commercialisées comme :

- le Térodon ou Poisson-balle *Sphoeroides cutaneus*, la Sériole babiane *Seriola fasciata* et d'autres Sérioles d'origine africaine prédateurs très voraces de Poissons et Crustacés (**Bombace, 2001**) ;

- le venimeux Poisson-scorpion *Pteroides miles*, le Rouget-souris bande d'or *Upeneus moluccensis* stabilisé en Mer de Lybie, la Rascasse de Madère *Scorpaena maderensis*, le Barracuda de la Mer Rouge *Sphyraena viridensis* comme d'autres espèces de Barracudas parmi plusieurs centaines de migrants.

Ces changements pourraient avoir des conséquences sur les populations de Cétacés en raison des changements alimentaires que cela pourrait comporter (diminution/disparition de proies par compétition, reports sur d'autres proies pas forcément optimales).

De même, en conséquence de l'augmentation de la température, on assiste à un changement significatif de la distribution de la faune ichthyenne autochtone. Par exemple des espèces d'habitus thermiques plus chauds s'étendent vers le nord

comme le Baliste *Balistes carolinensis* ou le Poisson-perroquet *Sparisoma cretense* (Dulčić & Grbec, 2000)). Enfin, le phénomène d'importation d'espèces allochtones ne concerne pas seulement des Poissons mais aussi des algues : trois algues japonaises sont identifiées depuis la fin des années 60 (*Laminaria japonica*, *Undaria pinnatifida* et *Sargassum muticum*) ainsi que, plus récemment, l'algue tropicale *Caulerpa taxifolia*. Cela est particulièrement préoccupant car elle provoque de graves perturbations dans les herbiers à Posidonies, hôtes et nurseries de nombreuses espèces de Poissons proies de Cétacés côtiers comme le Grand dauphin *Tursiops truncatus*. Des invertébrés ont aussi colonisé la Méditerranée comme le Nudibranche *Melibe fimbriata*, migrant lessepsien.

C.2) Menaces d'origine anthropique

2(a) Les activités halieutiques

La pression de pêche est très intense en tous points de la Méditerranée et elle utilise des techniques très diverses que ce soit par la nature des filets, les techniques de mise en place et les espèces cibles. Elle vulnérabilise de nombreuses espèces tant cibles que non-cibles (**Tudela & Short, 2005**). La capacité en termes de surface de filets de l'Union Européenne est très nettement au-dessus des niveaux soutenables par la ressource (**Lleonart & Maynou, 2003**).

D'une manière générale, la diminution des ressources alimentaires des Cétacés par la pression halieutique et baleinière est considérée comme responsable de la diminution en cascade des populations de Mammifères marins ainsi que la déplétion des population ichthyennes (*cf. infra* et **Estes et al., 1998; Springer et al., 2003**). Elle pourrait intéresser de près les espèces de Cétacés vivant en Méditerranée (**Bearzi et al., 2006**).

Par ailleurs, la pêche est directement responsable de captures accidentelles de Cétacés qui peuvent être extrêmement importantes et touchent particulièrement le Dauphin bleu et blanc (**Aguilar, 2000a; Bearzi, 2002; Silvani et al., 1999**).

La mise en place d'un réseau de plus en plus important d'espaces protégés (« Marine Protected Areas » - MPAs *cf.* <http://www.oceanconservancy.org>) permet non seulement de protéger directement des Cétacés mais aussi d'améliorer très nettement le recrutement de jeunes dans des populations fragiles ce qui a pour effet à moyen et long terme de stabiliser les populations (**Planes et al., 2000**).

2(b) La contamination agricole et industrielle

La Méditerranée est une mer quasiment fermée entourée de contrées très lourdement industrialisées au nord et partout pratiquant une agriculture industrielle très fortement dépendante de l'emploi de pesticides. La concentration de toute une variété de contaminants chimiques est très élevée à tous les niveaux trophiques de l'écosystème (**Ramade et al., 1993**) et on enregistre de très hauts niveaux de métaux lourds et d'organochlorés chez toutes les espèces de petits Odontocètes ((**Aguilar, 2000a; Augier et al., 1993**) et nombreuses références incluses). Il semblerait que les

niveaux observés dans le bassin occidental soient supérieurs à ceux observés dans le bassin oriental (**Troisi et al., 1998**).

La pathogénicité des contaminants sur les mammifères marins est peu connue. Des charges élevées en organochlorés ont été associées avec de nombreux dysfonctionnements physiologiques comme la croissance, la reproduction et le système immunitaire (**Fossi & Marsili, 2003**). Le taux de matières grasses très élevé dans l'organisme des Cétacés favorise d'autant plus la bioaccumulation des contaminants de même que la transmission verticale de la mère au petit par la lactation rend encore plus pérenne la charge au sein des populations (**Karuppiah et al., 2005**). Il est par ailleurs maintenant acquis que la population de Phoques communs de la Mer de Wadden aux Pays-Bas a brutalement chuté dans la période 1960-1975 en raison de l'impact négatif en particulier des Polychlorobiphényles (PCBs) sur l'implantation embryonnaire (**O'Shea et al., 1998**).

Chez le Dauphin bleu et blanc, aucune preuve irréfutable n'existe de l'existence d'un impact négatif des contaminants sur les populations mais certains auteurs émettent l'hypothèse que les PCBs pourraient être à l'origine d'immunodépressions chez les Dauphins morts lors de l'épisode à Morbillivirus de 1990-1992 (**Aguilar & Borrell, 1994**).

2(c) Pression anthropique d'ordre physique

Ces derniers siècles, l'humanité a beaucoup modifié la géologie de la Méditerranée (**Briand & Maldonado, 1996**). Des constructions ont été établies tous le long du littoral, aggravant et déplaçant les zones d'érosion et déséquilibrant la biochimie de cette région. Des plages ont été mal gérées, et l'abus de l'utilisation des ressources naturelles et marines continue d'être un problème. Cet abus se fait au détriment des processus naturels. La géographie a également été modifiée par la construction de barrages, de canaux et le « bétonnage » largement au-delà du raisonnable des côtes.

IV. Le Complexe Majeur d'Histocompatibilité (C.M.H.)

On distingue traditionnellement le Système Immunitaire Inné (*Innate Immune System - IIS*) du Système Immunitaire Adaptatif (*Adaptative Immune System - AIS*). Les Vertébrés ont subi une évolution fondamentale en ce qui concerne leur système immunitaire, passant d'une immunité non spécifique, dominée par l'*IIS*, à une immunité spécifique, dominée par l'*AIS*, importante dans la réussite des Mammifères et Oiseaux.

Le Complexe Majeur d'Histocompatibilité (CMH) est composé d'un ensemble de gènes codant des glycoprotéines transmembranaires de classe I et de classe II dont la fonction principale dans le système mature est de présenter les peptides (endogènes ou exogènes) respectivement aux lymphocytes T CD8+ et CD4+. Chez l'homme, les gènes du CMH de classe II, situés sur le bras court du chromosome 6 (6p21.3 s'étendant sur environ 4 millions de paires de bases - Mb - (**Lawrence *et al.*, 1986; Spring *et al.*, 1985**), constituent la région la plus polymorphe du génome. Elle est aussi la plus dense en exons (**Newell *et al.*, 1993; Trowsdale & Campbell, 1992**). Seuls 40% des gènes de cette région sont impliqués de près ou de loin dans l'immunité (**Danchin *et al.*, 2004**).

La région génomique du C.M.H. est classiquement divisée en trois régions (dans l'ordre centromère vers télomère) : classe II (0,9 Méga-paires de Bases - Mb), classe III (1,1 Mb) et classe I (> 2 Mb). Cette cartographie est schématisée sur la **Figure 14** ci-après.

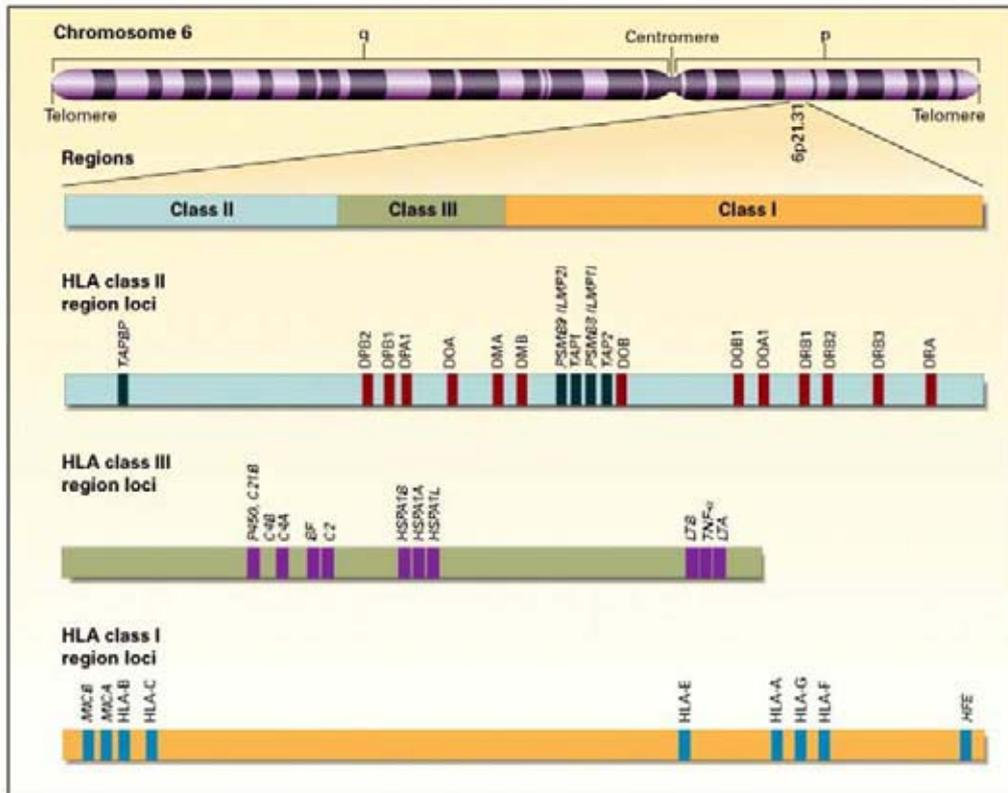


Figure 14. Localisation et organisation du C.M.H. chez l'Homme (Human Leukocyte Antigen - HLA) sur le Chromosome 6 (Klein & Sato, 2000).

IV. A. Mise en évidence du rôle de reconnaissance du « soi modifié » (Doherty, 1996)

Les premières expérimentations sur le virus de la chorio-méningite lymphocytaire des Rongeurs effectuées par Peter Doherty ont montré qu'il s'agissait en fait d'une immunopathologie et ont permis la formulation de l'hypothèse du *single altered self* (le soi modifié). Cette hypothèse consistait en une modification de ce que l'on appellerait plus tard le C.M.H. par le virus. Le « Soi Modifié » est alors reconnu par le Récepteur de la Cellule T (TCR). Cette hypothèse assignait un rôle aux antigènes de transplantation dans la réponse immunitaire aux agents pathogènes. Il a été plus tard démontré que l'antigène viral était issu d'un processus de dégradation *via* une voie de l'endolysosome. Les deux principaux événements survenant lors de l'infection par un agent pathogène sont (exemple de l'infection respiratoire par le virus Sendaï chez la Souris (Doherty, 1996) :

- Le virus envahit l'épithélium respiratoire où il est pris en charge par l'immunité non spécifique (macrophages...) : c'est la phase aiguë de l'inflammation ;

- Les cellules non spécifiques présentent alors l'antigène par l'intermédiaire des molécules du C.M.H. aux lymphocytes qui se multiplient dans le nœud lymphatique. Ces cellules reviennent sur le lieu de l'inflammation par la circulation générale pour potentialiser et achever la destruction de l'agent pathogène (phase effectrice) par l'intermédiaire soit de l'immunité cellulaire (T), soit de l'immunité humorale (B).

IV. B. Histoire Evolutive du Complexe Majeur d'Histocompatibilité

B.1) Emergence du système immunitaire adaptatif

Le Système Immunitaire Adaptatif (AIS) a émergé par le « bricolage » (Jacob, 1977) de protéines recrutées dans des systèmes existants dans le but de les coopter dans une nouvelle fonction (Ganfornina & Sanchez, 1999; Gould & Vrba, 1982; True & Carroll, 2002). Dans le cas de gènes dupliqués, l'un peut être coopté alors que l'autre garde la fonction ancestrale. Des molécules ayant la même activité peuvent alors être présentes dans des compartiments cellulaires différents voire des cellules différentes. Ainsi, le C.M.H. de classe I est exprimé dans la plupart des cellules des Vertébrés alors que le C.M.H. de classe II est principalement exprimé à la surface des cellules présentatrices de l'antigène. Différents niveaux de pression de sélection peuvent ainsi être définis : physiologique, tissulaire, organisme et population. La cooptation peut avoir lieu sans changements de son comportement moléculaire (protéasome impliqué dans la présentation de l'antigène) ou avec changements (M.H.C. de classe I et II avec peut-être des transferts de gènes horizontaux) comme des mutations ou des redistributions d'exons (*exon shuffling*) pouvant survenir par sélection positive. L'AIS en général et la présentation de l'antigène en particulier sont probablement apparus entre l'émergence des Gnathostomes et la divergence Gnathostomes-Agnathes (Klein *et al.*, 2000) lors d'un événement majeur de duplication chez un ancêtre commun aux Gnathostomes (Kasahara *et al.*, 1997) certainement polyploïde (Abi-Rached *et al.*, 2002) entre la divergence Cephalocordata-Craniata (*ca.* -766 Ma) et avant la radiation des Gnathostoma (*ca.*-528 Ma) selon Danchin *et al.* (2004) Un proto-C.M.H. a ainsi pu exister chez l'ancêtre commun des Protostomiens et Deutérostomiens (Danchin *et al.*, 2003). L'AIS a ainsi

pu commencer à apparaître vers la limite Protérozoïque-Phanérozoïque (cf. **Figure 15** ci-après).

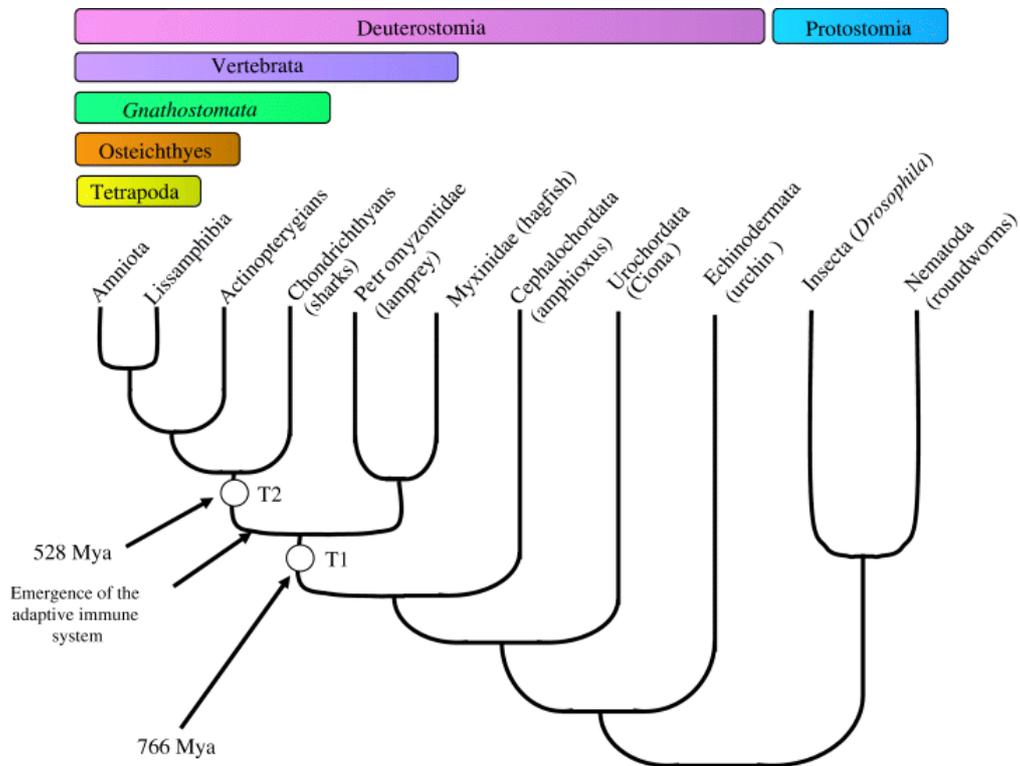


Figure 15. Phylogénie des Métazoaires montrant les étapes de l'émergence du Système Immunitaire Adaptatif (d'après Danchin *et al.* (2004)).
 Les flèches indiquent des duplication *en-bloc* (T1 and T2) dans le proto-CMH. Ainsi, l'ancêtre commun des Céphalochordés et des Chordés ne possédait pas d'AIS.

Après ces duplications à grande échelle, des séries de cis-duplications et cis-redistributions d'exons ont vraisemblablement été à l'origine des configurations actuelles du C.M.H. chez chacun des Vertébrés (Abi Rached *et al.*, 1999). La **Figure 16** est une illustration de ces duplications ayant amené la formation de *clusters* (= grappes) de gènes (en l'occurrence le DRB).

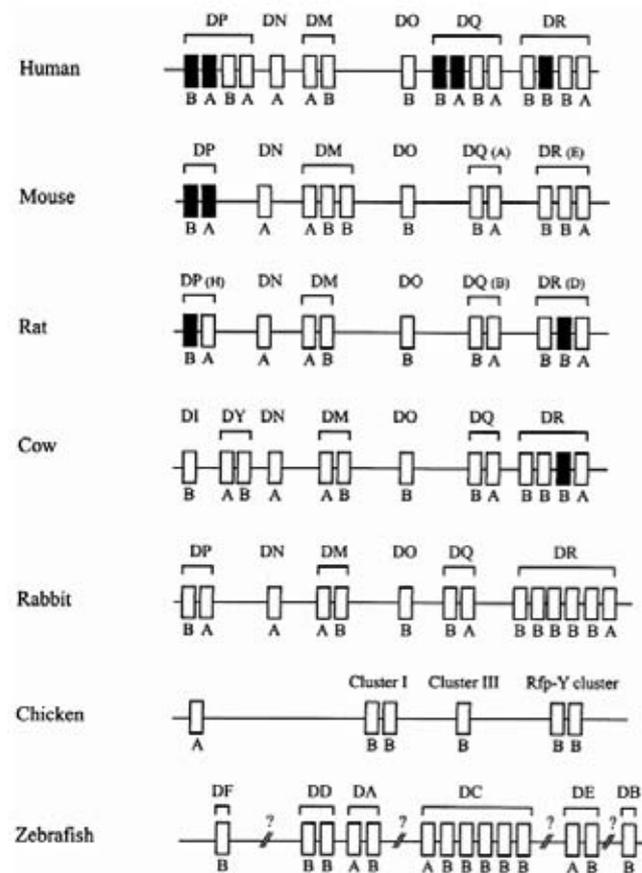


Figure 16. Carte génomique simplifiée des gènes de classe II (Takahashi *et al.*, 2000).

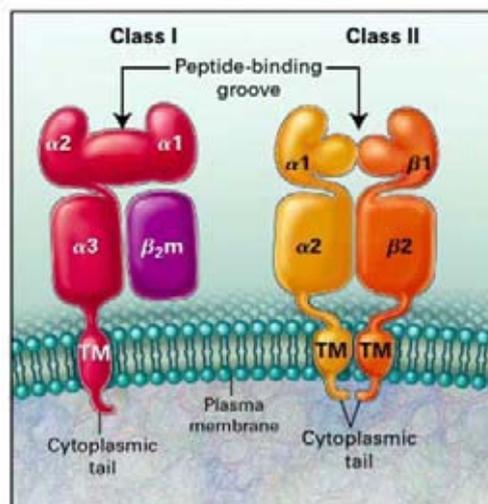


Figure 17. Représentation schématique comparée des molécules du C.M.H. de classe I et II (Klein & Sato, 2000).

B.2) Histoire naturelle des gènes impliqués dans la présentation de l'antigène par le C.M.H. de classe I et II

La région centrale du C.M.H. (classe III) correspond à des gènes codant le Complément et certains membres de la famille du *Tumor Necrosis Factor* (TNF). Toutes les molécules impliquées dans la présentation ne sont pas présentes dans la région génomique du C.M.H. mais on y trouve la plupart. Ce regroupement pourrait résulter d'une pression de sélection positive, ayant conduit au regroupement des gènes entre eux, suivie d'une sélection conservative qui a pu les maintenir ensemble (**Danchin et al., 2004**) mais cette hypothèse reste à tester.

Diverses molécules « de ménage » sont impliquées dans l'assemblage du CMH : calréticuline, HSPA5, calnexine. Les calnexine et calréticuline sont paralogues et toutes ont des orthologues dans l'ensemble des Eucaryotes. Les molécules dédiées impliquées interviennent dans l'assemblage par hétérodimérisation avec les molécules du C.M.H. : TAP et tapasine. La tapasine, retrouvée chez tous les Gnathostomes, possède un domaine IgC paralogue des chaînes β de classe I et II (**Du Pasquier, 2000**). Enfin, les molécules de classe I et II appartiennent à la même famille, ne sont trouvées qu'au sein des Vertébrés mais elles sont composées de trois modules d'origines différentes et non clairement élucidées : le *Peptide Binding Domain* (PBD ou PBR), le module Immunoglobuline (IgC) et le module d'ancrage trans-membranaire (**Klein & O'HUigin, 1993**). L'IgC est une synapomorphie de la famille des butyrophillines, tapasine et C.M.H. et les butyrophillines semblent liées à l'ancêtre des Amniotes. La tapasine est génomiquement proche du C.M.H. dans l'ensemble des *Euteleostomi* (**Sultmann et al., 2000**) ce qui tendrait à prouver son lien à l' AIS précocement dans l'évolution des Vertébrés.

Le système immunitaire des Vertébrés semble avoir recruté les protéasomes qui sont retrouvés chez les *Archea*, *Bacteria* et *Eukaria*. La taille des peptides générés dans les cellules de non-Vertébrés sont de taille et de composition correcte pour le chargement sur les molécules du C.M.H. (**Niedermann et al., 1997**). Les sous-unités immunitaires du protéasome remplacent d'ailleurs les sous-unités « plésiomorphes » après activation par l'interféron γ . Par ailleurs, d'autres protéases extra-lysosomales

sont capables de générer des peptides pour le PBD (Seifert *et al.*, 2003). L'ensemble de ces molécules intra- et extra-lysosomales a subi une co-option sans changement de comportement (*behavior shift*).

Comme vu précédemment, les molécules de classe II ont dû évoluer sans portage de l'antigène au sein du réticulum endoplasmique. Trois hétérodimères $\alpha\beta$ s'associent à un CD74 présent seulement chez les *Euteleostomi*. Seul ce complexe de neuf chaînes est compétent pour le portage de l'antigène.

La nécessité de prendre en charge les pathogènes a dirigé l'évolution du système immunitaire des Vertébrés. Ainsi, non seulement le passage de ceux-ci laisse-t-il une trace dans le génome mais en outre, la direction prise par l'évolution peut influencer la qualité de la réponse immunitaire aiguë en positif ou négatif selon le polymorphisme du C.M.H. Cela confère au C.M.H. un rôle central dans le maintien de l'homéostasie.

IV. C. La Région de Classe II

C.1) Description

Cette région est très riche en gènes : DR, DP & DQ (Campbell & Trowsdale, 1993; Trowsdale, 1995) pour les sous-régions majeures mais aussi bien d'autres comme DM (Radley *et al.*, 1994) ou DO (Lampson & Levy, 1980). Leur produits sont des hétérodimères ((Stern & Wiley, 1994a)) d'une chaîne lourde α (30-34 kDa) et une chaîne légère β (26-29 kDa). Elles sont exprimées de façon constitutive à la surface des cellules dites présentatrices de l'antigène, *i.e.* les monocytes, macrophages, lymphocytes B et cellules dendritiques, ou sur d'autres types cellulaires après activation (lymphocytes T, cellules non immunitaires). Les hétérodimères sont uniques pour DP et DQ, parfois groupés par deux hétérodimères liés pour DR (Bouillon & Mourad, 2003). Chaque chaîne est composée de deux domaines extracellulaires, d'un domaine transmembranaire et d'une courte queue cytoplasmique (12-18 acides aminés). Le polymorphisme de ces molécules est principalement situé dans le premier domaine de la molécule, α_1 et β_1 , et regroupé en 2 à 4 régions hypervariables (McDevitt, 1995). Cette variation structurale

représente la base fondamentale permettant la spécificité antigénique.

C.2) Présentation de l'antigène

Le peptide antigénique de 13 à 18 acides aminés, fruit de l'endocytose et de la dégradation d'un agent pathogène exogène, se fixe dans la niche peptidique située à l'extrémité des domaines α_1 et β_1 où réside la majorité des acides aminés hautement polymorphes (**Stoltze et al., 2000**). Cette voie métabolique est la cible des mécanismes de l'évasion virale (**Vossen et al., 2002**) par perturbation en particulier du transit de l'antigène dans le réticulum endoplasmique (**Lybarger et al., 2005**).

Expérimentalement, la molécule DRB a été associée à la présentation d'un antigène de Morbillivirus (**Ovsyannikova et al., 2004**). Le traitement de l'antigène est schématisé par la **Figure 18** ci-après.

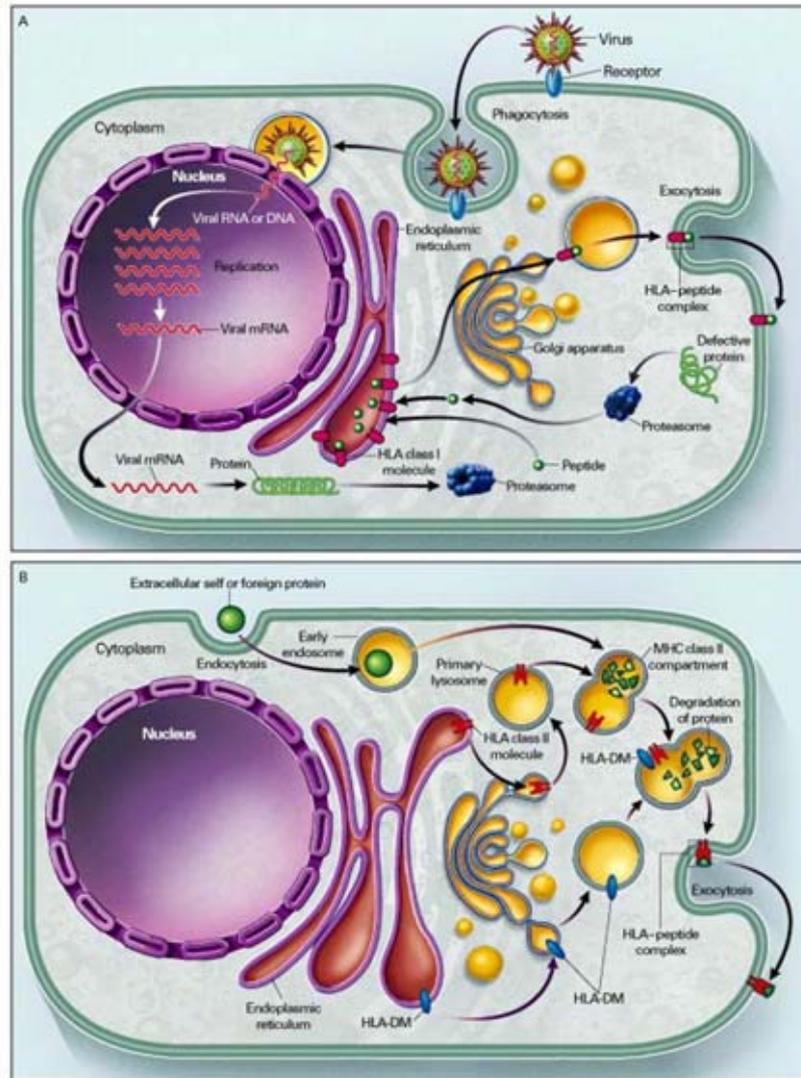


Figure 18. Traitement de l'antigène (Klein & Sato, 2000)

Le cadre A représente la voie principale pour le chargement du peptide sur le CMH de classe I. Les protéines sont dégradées en peptides dans le protéasome. Ceux-ci sont sélectionnés puis chargés dans le RE sur les molécules de classe I et le complexe est exporté *via* l'appareil de Golgi. Dans le cas de virus, certaines protéines virales suivent la même voie de dégradation.

Le cadre B montre le traitement de protéines extra-cellulaires (protéines du Soi ou non-Soi). L'endocytose/phagocytose n'est pas suivie de dégradation dans l'endosome. Les molécules de classe II sont présentes dans le lysosome lorsque celui-ci fusionne avec l'endosome pour dégrader les protéines. Les molécules du CMH-DM aident au chargement du peptide sur les autres molécules de classe II.

C.3) Autres fonctions

Outre leur fonction de présentation antigénique, les molécules du CMH de classe II jouent un rôle important dans la susceptibilité aux maladies auto-immunes telles que la polyarthrite rhumatoïde, le diabète de type I et la sclérose en plaques (**Winchester, 1994**). Il semble que les régions polymorphiques de certains allèles de prédisposition présentent des auto-antigènes qui activent une population de lymphocytes T potentiellement autoréactifs. Mis à part leur rôle dans l'interaction avec le CD4 et le LAG-3 (gène d'activation du lymphocyte-3), les molécules du CMH de classe II agissent comme récepteurs des superantigènes (SAg) (**Lavoie et al., 1999**).

C.4) Conséquences de la présentation de l'antigène par les molécules de Classe II

La reconnaissance du complexe peptide antigénique/CMH de classe II par le TCR (*T-Cell Receptor*)/CD4 induit l'activation des lymphocytes T et des CPA (**Corley et al., 1985**). L'engagement des molécules du CMH de classe II induit l'adhérence, la prolifération et la différenciation des lymphocytes B, la production de cytokines pro-inflammatoires et de médiateurs de l'inflammation, l'expression de molécules co-stimulatrices et, dans certaines circonstances, l'apoptose (**Mehindate et al., 1995**).

La liaison des molécules CMH de classe II entraîne l'activation de deux voies de signalisation distinctes : l'une nécessitant les résidus du domaine cytoplasmique de la chaîne β , et l'autre dépendante du domaine transmembranaire de la chaîne β (**Andre et al., 1994; Harton et al., 1995**). Dans les cellules B quiescentes de souris, l'engagement des molécules du CMH de classe II entraîne l'activation d'une voie de l'AMPcyclique (AMPC) (**Newell et al., 1993**). De plus, seules les cellules B humaines activées sont susceptibles à l'apoptose (**Truman et al., 1994**), suggérant que l'engagement des molécules du CMH de classe II permet de maintenir l'homéostasie de la réponse immunitaire.

Le couplage des molécules du CMH de classe II à ces diverses voies de signalisation

demeure néanmoins un mystère. En effet, le motif du domaine cytoplasmique de la chaîne β ne correspond pas à celui retrouvé habituellement dans les récepteurs couplés aux protéines G. De plus, l'activation des protéines-kinases implique des interactions avec d'autres protéines transmembranaires pour faire le lien avec les protéines-kinases cytoplasmiques.

Bien que leur rôle dans la signalisation relayée par le récepteur des lymphocytes T (TCR) et le récepteur des lymphocyte B (BCR) soit bien établi (**Cheng *et al.*, 1999; Montixi *et al.*, 1998**), leur implication dans la signalisation induite *via* les molécules du CMH de classe II est imparfaitement connue. Par exemple, 50% des molécules de classe II sont associés à des radeaux lipidiques de lignées cellulaires B et de cellules dendritiques (**Anderson *et al.*, 2000**). Dans les cellules B humaines, il a été démontré que les molécules du CMH de classe II, particulièrement les HLA-DR, s'associent avec plusieurs récepteurs de la surface cellulaire tels que le CD19/CD21, CD20, CD40, CD9, CD37, CD53, CD81 & CD82 (**Bradbury *et al.*, 1992; Léveillé *et al.*, 1999a; Léveillé *et al.*, 1999b; Schick & Levy, 1993**).

C.5) La molécule DRB

5(a) Description (chez l'Homme et les Primates)

Le DRB est la région incluant plusieurs gènes homologues dénomés DRBX (X étant un nombre). Il s'agit de la région de classe II la plus polymorphe. Dans les populations humaines, l'organisation génomique et le nombre de gènes du DRB diffère entre les haplotypes contrairement aux région paralogues DQ et DP.

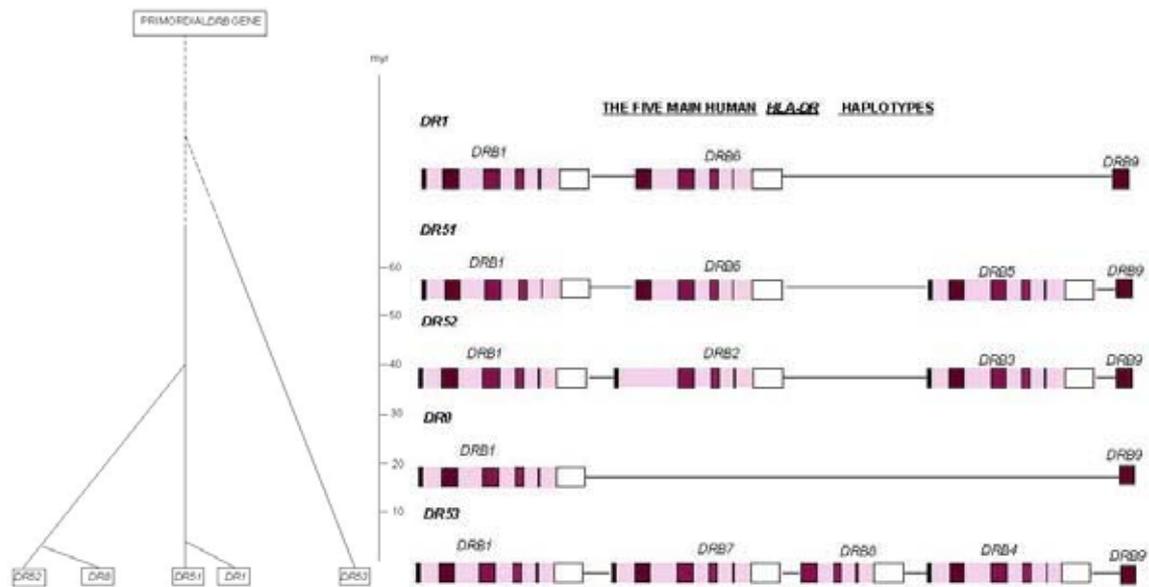


Figure 19. Les 5 principaux haplotypes humains de la portion DRB du C.M.H. ainsi que leur dates de divergence (Andersson, 1998).

A l'intérieur de l'haplotype DR53, les lignages alléliques des gènes DRB1 ont des diversités alléliques variables selon les lignages (*ca.* 135 allèles pour DRB1*04 contre quelques uns pour DRB1*09). La raison de cette contradiction pourrait être le résultat de *bottlenecks* passés (Andersson, 1998) ou d'une néo-cooptation de DRB1*04 pour un éventail de pathogènes différents par glissement de l'affinité antigénique. L'événement de duplication initial qui est survenu entre DR51 et DR52 (60 ± 6 Ma) est donc vraisemblablement propre à la lignée des Primates.

5(b) Le DRB chez les espèces autres que les Primates

Le DR53 ayant divergé beaucoup plus précocement (avant 85 Ma), il pourrait être retrouvé en sein des *Cetartiodactyla*. Il convient par ailleurs de noter que la forte proportion de rétro-éléments dans la région DR (24%) a fortement contribué à l'évolution de cette région (Andersson *et al.*, 1998).

Le DRB a été largement étudié chez le Bovin (Aida *et al.*, 1995; Ellegren *et al.*, 1993; Sigurdardottir *et al.*, 1992) dans des optiques médicales et zootechniques (tests de paternité). Sa diversité a des implications importantes dans la diversité des antigènes portés et l'analyse de la diversité allélique chez des Bovins a montré une différence dans l'efficacité d'une épreuve vaccinale contre l'agent de l'East Coast Fever, le

Sporozoaire *Theileria parva*, parasite intracellulaire des lymphocytes (**Ballingall et al., 2004**). De même, une corrélation a été démontrée entre le polymorphisme de DRB3 dans cette espèce et la résistance à l'infection par le *Bovine Leukaemia Virus* (**Xu et al., 1993**). Enfin, des associations avec des traits de production ont été démontrés (**Crawford et al., 2000; Giphath et al., 1990**).

Il a été caractérisé assez tôt par Gel d'Electrophorèse à Gradient de Dénaturation (DGGE ; (**Aldridge et al., 1998**).

5(c) L'hypothèse de conversion génique

On distingue deux formes de l'hypothèse de conversion génique appelée forte et faible sur la base de la significativité adaptative attribuée au processus de conversion génique dans les gènes du C.M.H. (**Martinson et al., 1999**). L'hypothèse faible se base sur les événements singuliers sans accorder de place majeure à la conversion génique dans l'évolution du C.M.H. En revanche, l'hypothèse forte suppose que la conversion génique est une force puissante qui optimise la variabilité des séquences. Cette forme forte correspond à la l'expression spécifique de l'hypothèse de Haldane d'une réponse des systèmes immunitaires des hôtes (animaux ou végétaux) à la grande variabilité des pathogènes par une grande variabilité des systèmes chargés de l'immunité (**Haldane, 1949**). Cela donne lieu à une « course aux armements » dont les victoires ne sont que des victoires à la Pyrrhus car temporaires. Dans les molécules de classe II, les résidus du PBD/PBR sont au nombre de 16 contre 220 pour toute la chaîne. Les ratios de substitutions synonymes (dN) sur les substitutions synonymes (dS) calculés séparément pour ces deux portions montrent qu'elles évoluent différemment. Au niveau moléculaire, des modulations de la structure chromatinienne pourraient rendre plus accessible la zone à des protéines de recombinaison ou encore cette région pourrait être positionnée dans un optimum local de recombinaison.

5(d) Polymorphisme trans-spécifique

L'évolution concertée est souvent invoquée pour expliquer la diversité et l'évolution des familles multigéniques telles que le C.M.H. et les Immunoglobulines (Ig). Cependant, le polymorphisme trans-spécifique rend ces explications sujettes à

controverse. Ainsi, certains travaux (Nei *et al.*, 1997) ont montré que ces familles semblent évoluer selon un processus de renouvellement (*birth-and-death process*) par lequel des nouveaux gènes sont créés par des duplications géniques répétées. Les molécules de classe I et II semblent évoluer de la même manière selon ce même processus. Le taux de renouvellement est plus lent chez les derniers. Les chaînes α et β sont partagés par les *Eutheria* et *Metatheria* mais il existe des paralogues chez les Vertébrés non-mammaliens. Chez les Placentaires, DO-B est le *cluster* qui a divergé le plus précocément (210-260 Ma) des autres chaînes β , puis les autres ont subi une radiation autour de 170-200 Ma (Takahashi *et al.*, 2000). Par ailleurs, les chaînes α et β ont suivi des évolutions légèrement différentes, même si leurs domaines extracellulaires sont étroitement associés.

En conclusion, le C.M.H a une fonction centrale dans l'immunité. Il est issu du recrutement de nombreuses molécules impliquées dans l'homéostasie cellulaire. Au sein de la région de classe II, la région DR est la plus polymorphe et est impliquée dans la présentation de l'antigène, en particulier ceux issus de pathogènes. Elle est composée de deux chaînes peptidiques (α et β). La particularité de cette région polygénique est que chacune de ces chaînes est représentée par plusieurs gènes homologues. Ceux-ci ont la possibilité de «trans»-recombinaison (conversion génique) et la sélection balancée a certainement permis l'émergence d'un polymorphisme trans-spécifique.

V. L'Épizootie à *Dolphin Morbillivirus* (*D.M.V.*, *Paramyxoviridæ*)

V. A. Épidémiologie descriptive (Dhermain *et al.*, 1994)

En juillet 1990, une mortalité en masse de Dauphins bleu-et-blanc est constatée sur les côtes la Communauté Valencienne et des Baléares (Levant Espagnol) et s'étend en moins de 6 mois à tout le littoral du bassin occidental (**Domingo *et al.*, 1990**). Elle s'étend l'année suivante aux Mer Thyrrénienne et Ionienne. Enfin, en 1992, elle gagne la Mer Égée et la Mer noire (*cf. Figure 20* ci-après).

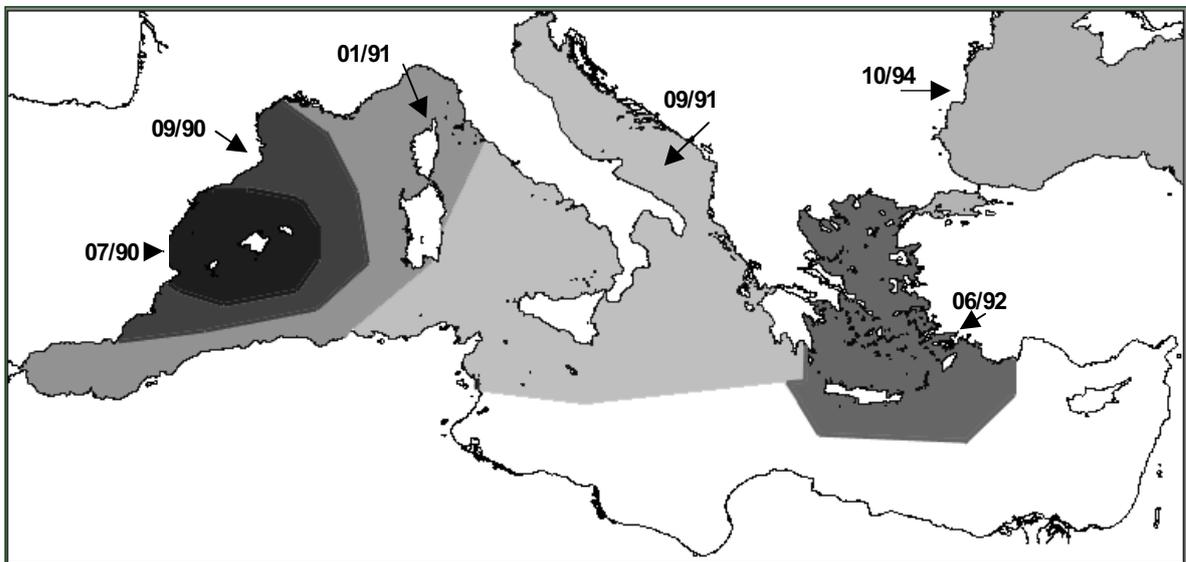


Figure 20. Progression de l'épidémie à Morbillivirus en Méditerranée, 1990-1992 (Dhermain, 1995-2006).

Le nombre d'échouage atteint son maximum en septembre, dans les régions de Valence, de Catalogne et des Iles Baléares ; en octobre, les régions de Murcia et d'Andalousie furent concernées à leur tour, et au nord, le Languedoc, où le maximum de mortalité fut noté à la mi-octobre. En novembre 1990, des cas sont signalés en mer d'Alborán, sur les côtes nord-africaines, en Regione Liguria ainsi qu'en Corse. En janvier 1991, l'épidémie franchit le détroit de Gibraltar, (quelques Dauphins s'échouent sur la côte du Parc de Doñana) mais ne progressa pas en Atlantique.

Durant cette épidémie, au moins 450 dauphins ont été collectés en Espagne, 150 en

France, et 80 sur les côtes nord-ouest italiennes (cf. **Figure 21** ci-après). Ces chiffres ne reflètent bien entendu qu'une partie non quantifiable des décès (manque de prospection, carcasses perdues en mer, coulées ou dévorées par les requins).

Après une complète accalmie (moins de 5 échouages par mois en Catalunya), l'épidémie reprit de manière identique, en juin 1991, dans le sud de l'Italie, atteignant son apogée en août dans le sud de l'Adriatique et en septembre en Mer Ionienne, amenant à la côte 200 cadavres de Dauphins (**Bortolotto et al., 1992**). La Grèce fut à son tour touchée par l'épidémie entre juillet 1991 et août 1992, avec une progression constante vers le nord-est, (**Cebrian, 1995**).

Durant l'été 1994, les échouages de Dauphins communs *Delphinus delphis ponticus* ont augmenté de manière brutale sur les côtes du nord de la mer Noire (**Birkun et al., 1999**). Enfin, des cas non vérifiés ont été rapportés sur les côtes Israéliennes en 1995 (**Tsur et al., 1997**). Ainsi, la vitesse moyenne de progression de l'épidémie peut être calculée à 5 à 16 km/j, très en dessous de la vitesse moyenne de déplacement des dauphins (5km/h). Par ailleurs, il a été démontré qu'il s'agissait de la même souche virale a balayé la Méditerranée (**Di Guardo et al., 2005**).

Actuellement, les échouages imputables au D.M.V. sont devenus peu fréquents en Méditerranée. Il reste un fond de morbidité, avec une fréquence d'échouages de dauphins moribonds un peu plus grande qu'auparavant et les études anatomopathologiques et sérologiques réalisées au cours d'autopsies détaillées sur les côtes catalanes ont montré que l'épizootie de 1990 a fait place à une infection chronique, entraînant des lésions infectieuses subaiguës du système nerveux central (**Domingo et al., 1995**).

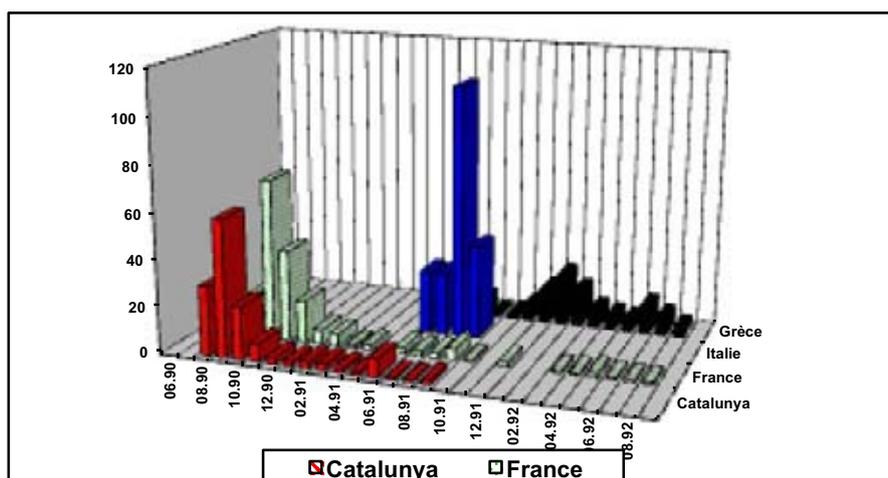


Figure 21. Nombre de *Stenella coeruleoalba* échoués sur les côtes catalanes (Domingo *et al.*, 1992), françaises (Dhermain *et al.*, 1994), italiennes (Bortolotto *et al.*, 1992) et grecques, de juin 1990 à septembre 1992 (Dhermain, 1995-2006).

V. B. Tableau anatomo-clinique

L'épizootie à Morbillivirus est cliniquement caractérisée en premier lieu par un grand nombre de Dauphins bleu-et-blanc s'échouant **vivants** : plus de 25 % en France contre moins de 5 % habituellement, 18 % en Catalogne (Van-Bressem *et al.*, 2001). Les animaux examinés présentaient le plus fréquemment des troubles neurologiques (frissons, attitudes « suicidaires » sur des obstacles, convulsions, voire apathie). Quelques cas de diarrhées furent rapportés en Espagne et sur un Dauphin « post-épizootique » que nous avons examiné en 1998, ainsi qu'un cas de vomissement. Les divers soins vétérinaires entrepris furent toujours inefficaces. Aucune immunothérapie par sérum contre la maladie de Carré n'a été tentée. Les Dauphins meurent en quelques heures à quelques jours (40 heures pour un Dauphin que nous avons pu examiner). Les quelques analyses sanguines effectuées à la Faculté vétérinaire de Barcelone montrèrent le plus souvent des leucopénies et lymphopénies, caractéristiques des infections virales

A l'autopsie, les changements morphologiques les plus importants étaient un collapsus partiel des poumons, avec une atélectasie inégale, un œdème sous-cutané, de l'ictère et une stomatite. D'un point de vue histologique, les lésions les plus fréquentes étaient celles d'une pneumonie bronchiolo-interstitielle (cf. Figure 22 Ci-après) à et d'encéphalite (plus de 70 % des cas en Espagne, chiffre similaire en

France). Le virus a été retrouvé par immuno-histochimie principalement dans le cerveau (77%) et le poumon (70%), et le système lymphatique (60%) (Domingo *et al.*, 1992). Des lésions de congestion hépatique, congestion pulmonaire, gastro-entérite sont également rapportées. Une nécrose de la rate associée à une sévère déplétion lymphoïde a également été soulignée par (Duignan *et al.*, 1992). Domingo *et al.* (1990) ont montré l'existence de lésions secondaires associées : aspergillose, toxoplasmose, actinomycose, favorisées par l'immunodépression consécutive à l'infestation par le Morbillivirus et/ou la contamination par des PCB. De semblables infestations mycologiques secondaires ont par ailleurs été retrouvées de la même façon sur des Grands Dauphins contaminés par les Morbillivirus aux USA (Lipscomb *et al.*, 1994).

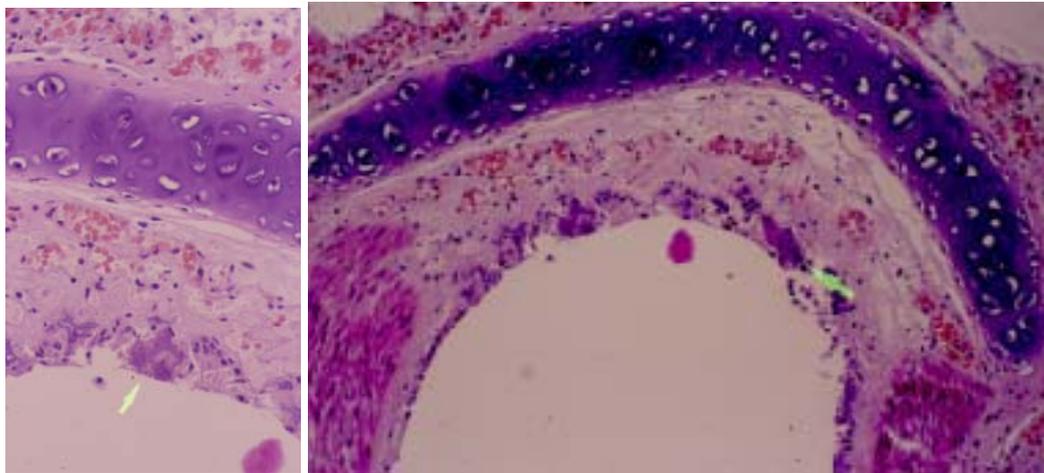


Figure 22. Lésions microscopiques dues au Dolphin Morbillivirus : syncytia dans l'épithélium bronchique (d'après Macé *et al.* 1999).

V. C. Epidémiologie Analytique

Mâles et femelles sont indifféremment touchés, le sex-ratio des Dauphins malades est proche de 1:1. L'âge-ratio (*cf.* **Figure 23** ci-après) montre, pour les deux sexes, une surreprésentation de la tranche d'âge des 11-20 ans (= adultes matures), ainsi que des nouveau-nés, victimes de la mort de leur mère allaitante et/ou facilement infectés à leur côté. En revanche, les individus sub-adultes apparaissent sous-représentés.

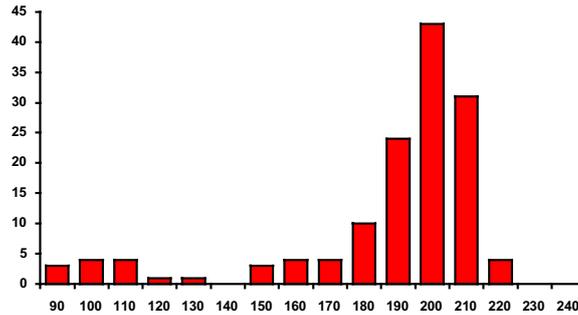


Figure 23. Distribution des dauphins par classe de taille (Calzada *et al.*, 1994).

V. D. Etio-Pathogénie : Emergence des Morbillivirus de Mammifères marins

L'agent pathogène à l'origine de l'épizootie a rapidement été isolé en raison de la mise en relation rapide avec l'épizootie survenue sur les Phoques veau-marin *Phoca vitulina* de Mer du Nord en 1988. Le virus mis en évidence appartient au genre Morbillivirus, virus à ARN monocaténaire de la famille des *Paramyxoviridae* (NCBI TaxId: 11158 ; <http://www.ncbi.nlm.nih.gov/genomes/GenomesGroup.cgi?taxid=11158>).

Un colloque sur les Morbillivirus a d'ailleurs été organisé en 1994 à la Faculté Vétérinaire de Hanovre en 1994 et a fait l'objet d'une publication d'un numéro entier de la revue *Veterinary Microbiology* (Edwards *et al.*, 1995a) détaillant des aspects cliniques, moléculaires, épidémiologiques comparés. Cette famille comprend les virus de la Rougeole humaine (MV, Measles Virus), la Maladie de Carré des Carnivores (CDV, Canine Distemper Virus), la Peste Bovine (RPV RinderPest Virus), la Peste des Petits Ruminants (PPRV) et le Virus de la « Maladie de Carré » des Phoques (PDV-1 proche du CDV) isolé en 1988. Il est à noter que ces virus touchent des *taxa* de Mammifères d'origine phylogénique très différentes (*Cetartiodactyla*, *Carnivora*, *Primata*,...)

Le Dolphin Morbillivirus (DMV) est classiquement divisé en deux sérotypes le DMV-1 affectant les *Delphinidae* et le DMV-2 affectant les *Phocænidæ* (n'entraînant cependant pas d'épizootie). Le Morbillivirus est ensuite rendu responsable *a posteriori* d'une mortalité en masse de Grand dauphins *Tursiops truncatus* ayant eu lieu en 1987-88 sur

la côte Est des Etats-Unis (**Lipscomb et al., 1994**). De très nombreux cétacés sont par ailleurs porteurs d'anticorps anti-DMV, démontrant l'existence sous forme de portage sain ou d'enzootie chez ces espèces (**Duignan et al., 1995**).

Des épizooties à Morbillivirus ont par la suite été observées chez les Lions du Serengeti (Tanzania), Ours blancs (**Follmann et al., 1996**) et des Chevaux dans le Queensland (Australia) donnant lieu à deux mortalités humaines. Le virus isolé (Equine Distemper Virus - EDV) s'est avéré posséder un réservoir sauvage : des Megachiroptères du genre *Pteropus* (**Young et al., 1996**).

Enfin, durant l'été 1997, la plus grande colonie, et la seule viable, de Phoques moines de Méditerranée *Monachus monachus* a été décimée (71% de mortalité des adultes du Cap Blanc en Mauritanie). Les virus isolés sur les populations grecques et mauritaniennes seraient assez divergents entre eux (respectivement proches du DMV-2 et du DMV-1) mais surtout plus proches du DMV que du PDV (**Osterhaus et al., 1997**). En outre, comme pour les Dauphins bleu et blanc (**Aguilar & Borrell, 1994; Borrell et al., 1996; Kannan et al., 1993a**). **Hernandez et al. (1998)** impute la plus grande responsabilité de la contamination par les PCBs et les Dinoflagellés (« marées rouges »).

V. E. Evolution des Morbillivirus

Toutes ces découvertes récentes pourraient laisser penser que les Morbillivirus sont des virus émergents mais l'ancienneté des différents virus adaptés aux différentes espèces plaide plutôt pour une existence au moins aussi ancienne que leurs espèces cibles (**Westover & Hughes, 2001**).

Chapitre II : Recherche d'un marqueur de séquence, potentiellement équivalent à la Région de Contrôle de l'ADN mitochondrial, sur le chromosome Y chez les Cétartiodactyles

I. Introduction

I. A. Utilisation de marqueurs moléculaires en écologie, génétique des populations et évolution

La biologie moléculaire est utilisée depuis plusieurs décennies dans les domaines de la biologie évolutive (phylogénétique, génétique des populations, l'écologie moléculaire, métagénomique), l'étude des maladies complexes, le diagnostic moléculaire, la médecine légale, la zootechnie, etc.

A.1) L'ADN mitochondrial

L'ADN mitochondrial domine le panorama de la biologie évolutive depuis près de trente ans (**Richards & Macaulay, 2001**). En effet, ses propriétés (hérédité matrilineaire chez de nombreuses espèces, non-recombinaison liée à l'haploïdie, grande variabilité de la région de contrôle) ainsi que ses différents niveaux de résolution (peu polymorphe dans ses régions géniques, à l'inverse de la région de contrôle) en ont fait un marqueur de choix à la fois pour les études évolutives interspécifiques et la génétique des populations.

A.2) *Des marqueurs prometteurs portés par le chromosome Y*

D'autres marqueurs sont prometteurs, en particulier ceux portés par le chromosome Y, pour discerner des paramètres populationnels liés à la lignée paternelle. En effet, étant donné que le sexe mâle est le sexe hétérogamétique chez les Mammifères, l'utilisation conjointe du polymorphisme de l'ADN mitochondrial, avec en outre des marqueurs autosomiques, ouvre des opportunités d'exploration des mécanismes de dispersion et ségrégation sexuelle ainsi que des stratégies reproductrices (**Petit et al., 2002**).

2(a) Détermination du sexe

Le genre peut avoir des déterminants environnementaux (*e.g.* Nématodes,

Chéloniens, etc.) (Chandler *et al.*, 2009) ou être déterminé par des régions ou chromosomes sexuels (cas de nombreux Vertébrés). Dans ce dernier cas, la situation est plus ou moins complexe, d'une paire de chromosomes comme chez les Placentaires ou les Oiseaux, à une répartition des *loci* impliqués sur de nombreux chromosomes comme chez l'Ornithorhynque (Veyrunes *et al.*, 2008). La Figure 24 montre l'évolution du système chez les Vertébrés.

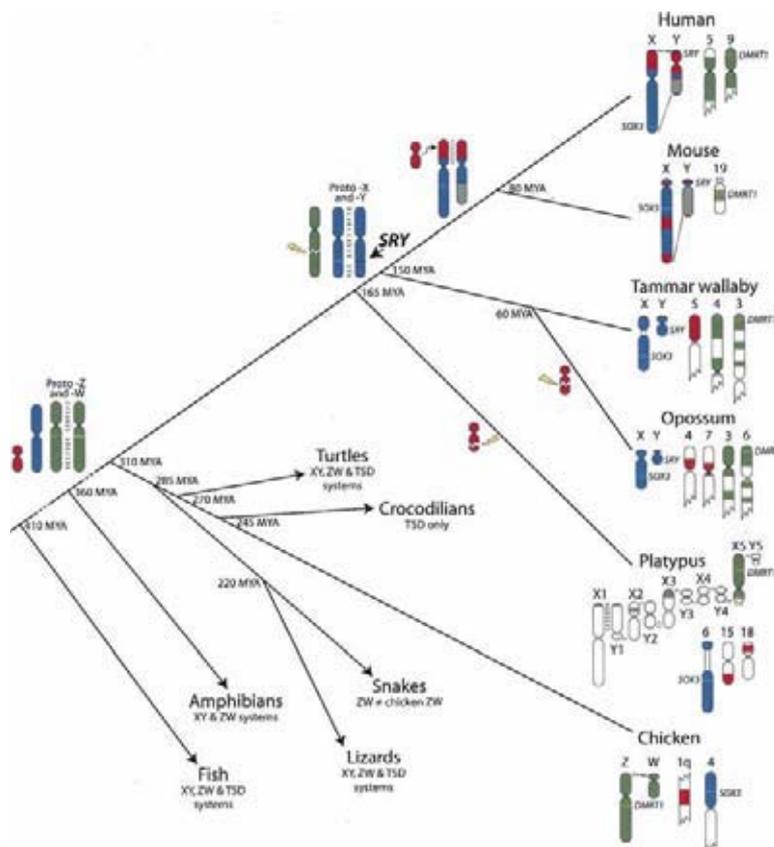


Figure 24. Phylogénie des différents systèmes de chromosomes sexuels des principaux groupes de Vertébrés (Veyrunes *et al.*, 2008).

2(b) Evolution des chromosome sexuels

Les chromosomes sexuels humains ont évolué à partir des autosomes. En effet, 19 gènes autosomaux persistent en tant qu'homologues différenciés sur les chromosomes X et Y (Lahn & Page, 1999) et l'on a découvert une corrélation entre les âges de divergence des différentes paires (X et Y) de gènes et leur localisation chromosomique : l'âge décroît par paliers depuis l'extrémité du bras long. Il en résulte au moins 4 strates évolutives montrant que l'évolution des chromosomes

sexuels a été probablement ponctuée d'autant de macro-événements évolutifs supprimant successivement la recombinaison entre les deux chromosomes dans chacune des régions. Chacun de ces événements semble lié à une étape majeure de l'évolution des Amniotes, le premier étant lié à la divergence entre les lignages Synapside (Reptiles mammaliens) et Sauropside (Anapsides - Chéloniens -, Diapsides - Squamates et Sphénodontes - et Archosauriens - Crocodiliens et Oiseaux) (Matsubara *et al.*, 2006; Organ & Janes, 2008; Smith & Voss, 2007; Veyrunes *et al.*, 2008). Les gènes de ces différentes régions ont alors été recrutés lors de la différenciation X-Y dans le même ordre. La perte de recombinaison sur le chromosome Y a alors permis des inversions sur ce chromosome, et non sur l'X, bousculant l'ordre des strates. La **Figure 25** montre la rétraction progressive du chromosome Y et l'évolution par « blocs » de la région non recombinante (les grands événements sont datés et remis en perspective de l'évolution des Amniotes).

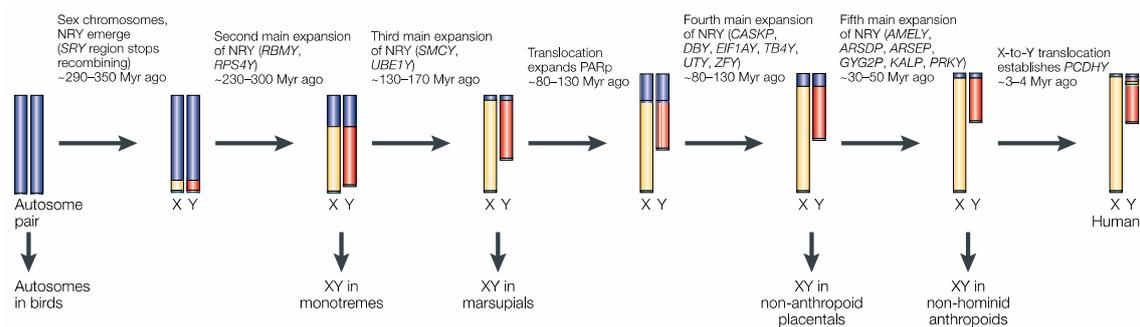


Figure 25. Evolution des chromosomes sexuels (Lahn *et al.*, 2001).

2(c) La région pseudo-autosomique

Sur les chromosomes sexuels humains, dans la limite proximale de cette région s'étend un gène intact sur le chromosome X mais interrompu sur le chromosome Y (Weller *et al.*, 1995) en raison d'une inversion.

La théorie (Barton & Charlesworth, 1998; Charlesworth, 1996; Graves, 1996; Rice, 1996) prédit que dès que la recombinaison cesse dans une région, ses gènes commencent à dégénérer à moins que la pression évolutive ne soit trop importante et les maintienne. Par ailleurs, les gènes du chromosome X sont sujets à l'inactivation (phénomène d'empreinte), processus qui a dû s'étendre gène par gène ou par groupe de gènes (Brown *et al.*, 1997). Ainsi, les strates devenues non-recombinantes (*cf.*

Figure 26 ci-après) le plus récemment sont aussi celles qui ont la plus grande densité en gènes possédant un homologue X-Y et échappant à l'inactivation (Jegalian & Page, 1998). Par ailleurs, le cours du temps a tendance à favoriser les substitutions non-synonymes, altérant la fonction des protéines mais ce phénomène est encadré par les contraintes sélectives exercées sur les protéines codées.

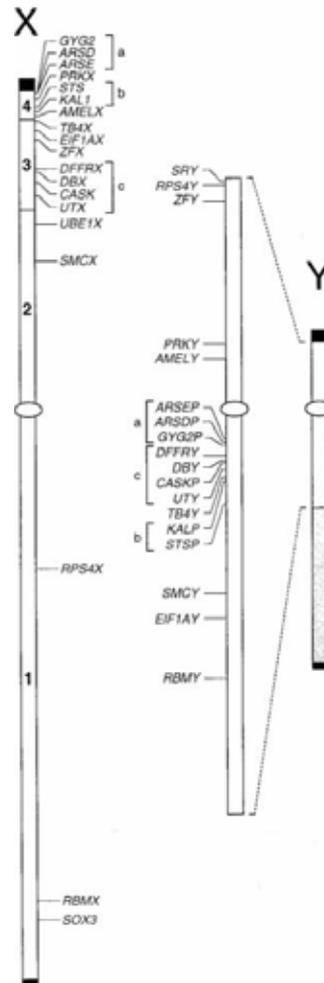


Figure 26. Carte des gènes homologues dans la région non-recombinante des chromosomes X et Y humains.

Les régions pseudo-autosomales de l’X et du Y sont en noir. Les 4 strates décrites dans le texte sont indiquées par des chiffres. Les groupes de gènes (a, b, c) délimités par des crochets sont présents à la fois sur le chromosome X et le chromosome Y.

I. B. Médecine légale et biotechnologies

Le sexe ou détermination du genre à l’aide d’ADN génomique extrait à partir de viande, sang, poil ou biopsies d’embryons est un outil analytique important en médecine légale (*forensic*) et en génotypage de routine (pedigree, études de génétique

médicale). Il a même été utilisé en archéologie en particulier sur le site de Pompéi (**Di Bernardo et al., 2009**). Chez beaucoup de Mammifères, le mâle est identifié par amplification du gène SRY (*sex-determining region Y*) qui est une séquence spécifique du chromosome Y. Cependant, lors de résultat négatif (non-amplification), on ne peut conclure avec certitude qu'il s'agit d'une femelle sans écarter les risques d'erreurs technique (faux négatif). Afin de remédier à ce type de problème, deux choix s'offrent au médecin et/ou biologiste :

- soit amplifier une autre région génomique comme « témoin positif » (comme chez les Ursidés (**Kohn et al., 1995; Taberlet et al., 1993**) mais il est possible d'observer quand même de rares faux positifs ;

- soit d'amplifier une région présentant une homologie sur les deux chromosomes sexuels (avec le risque très négligeable d'observer des mutations), présentant un dimorphisme sexuel important (de taille, séquence ou profil de restriction). Des méthodes ont par exemple été mises au point sur les marqueurs ZFX/ZFY par RFLP (**Aasen & Medrano, 1990**) et par PCR de faible stringence - moins spécifiques -(**Dias Neto et al., 1993**).

I. C. Amélogénine

Le gène de l'amélogénine est une gène simple copie comportant un homologue sur chacun des 2 chromosomes sexuels (dénommées *AMELX* et *AMELY*).

C.1) Aspects physiologiques

L'amélogénin est sécrétée par les améloblastes des bourgeons dentaires. Il s'agit de la matrice protéique qui se combine avec les cristaux d'hydroxyapatite formant ainsi l'émail dentaire. Sa distribution et sa structure sont fortement conservées à l'échelle évolutive (**Satchell et al., 2002**). L'amélogénine se dépose avant que ne survienne la minéralisation et semble organiser les cristaux d'émail dentaire en évitant la prolifération anarchique et en régulant la cinétique de croissance et la taille des cristaux (**Aoba, 1996**).

C.2) Evolution

Le gène de l'amélogénin est un gène considéré comm 'orphelin' c'est-à-dire sans homologues/paralogues (**Delgado et al., 2005**).

Ce gène a été cloné chez plusieurs espèces mammaliennes et non-mammaliennes (**Bonass et al., 1994; Gibson et al., 1991; Girondot & Sire, 1998; Hu et al., 1996; Salido et al., 1992; Snead et al., 1985**). Les séquences complètes des exons de chacune des copies se regroupent, après analyse phylogénique, par espèce, indépendamment de leur origine chromosomique (**Girondot & Sire, 1998**). L'intron 3 de 4 espèces de Primates montre un regroupement en deux lignages, X et Y, dont la composition taxonomique est équivalente. Cela pourrait signifier que deux duplications indépendantes chez les Primates et chez les (Cétar)artiodactyles ont donné naissance aux gènes de l'amélogénine.

Les régions terminales connues sont très similaires (**Sire et al., 2005**) et lors d'analyse phylogénique, le groupement se fait toujours selon l'origine phylogénique et jamais selon l'origine chromosomique (**Girondot & Sire, 1998**). Le gène n'a été détecté dans aucun des trois *taxa* d'Amniotes dépourvus de dents étudiés jusqu'à présent (*Gallus domesticus*, *Dermochelys coriacea*, *Emys orbicularis*) : après être apparues il y a 450 M.a., les dents ont été perdues par les Chéloniens il y a environ 200 M.a. et par les Oiseaux il y a 70-80M.a. (Millions d'années) environ (**Davit-Beal et al., 2009**).

I. D. Le modèle biologique étudié : les *Cetartiodactyla*

Le clade *Cetartiodactyla* (du grec το κῆτος 'monstre marin', ἄρτιος 'bien proportionné (=pair)' et ὁ δάκτυλος 'doigt') est un groupe non rangé (au sein de la classification « classique ») appartenant au Phylum des chordés, Sous-phylum des Vertébrés, Classe des Mammifères, Sous-classe des *Theria*, Infra-classe des *Eutheria*, Super-Ordre des *Laurasiatheria*.

L'analyse phylogénique basée sur des caractères morphologiques considérait traditionnellement les Cétacés et les Artiodactyles en deux ordres monophylétiques. Les Cétacés étaient alors considérés comme dérivant d'un ordre éteint de Mammifères carnivores : les Mésonychiens. A la suite des nombreux travaux de

phylogénie moléculaire (Douzery & Catzeflis, 1995; Gatesy *et al.*, 1996; Gatesy, 1997; Gatesy *et al.*, 1999; Graur & Higgins, 1994; Millinkovitch & Thewissen, 1997; Montgelard *et al.*, 1997; Nikaido *et al.*, 1999; Philippe & Douzery, 1994; Shimamura *et al.*, 1997), les paléontologues ont revisité les caractères morphologiques et se sont accordés à la vision d'un polyphylétisme des Artiodactyles s'ils n'incluaient pas les Cétacés (Boisserie *et al.*, 2005; de Muizon, 2001; Gingerich *et al.*, 2001; Thewissen *et al.*, 2001).

Comprenant l'ordre des Cétacés et l'ordre des Artiodactyles, ce clade est considéré selon les auteurs comme un ordre, un super-ordre, ou non classé. Il montre dans tous les cas que Cétacés et Artiodactyles sont deux groupes-frères. Comme l'indique l'étymologie, une de leurs principales synapomorphies est de posséder un nombre de doigts pair (2 ou 4) ce qui implique que l'axe médian des membres passe entre les doigts III et IV (paraxonie). Les synapomorphies incluent en particulier (Millinkovitch & Thewissen, 1997) :

- un système digestif partiellement (Suina) ou entièrement (autres) multi-loculaire ;
- un *talus* (astragale, os du tarse) à double poulie (*cf. Erreur ! Source du renvoi introuvable.ci-après*)
- un petit processus (*Incus crus brevis* portant l'articulation d'avec le *Malleus*) de l'*Incus* (enclume de l'oreille interne) allongé ;
- un pénis en « S » fibro-élastique avec muscle rétracteur du pénis.

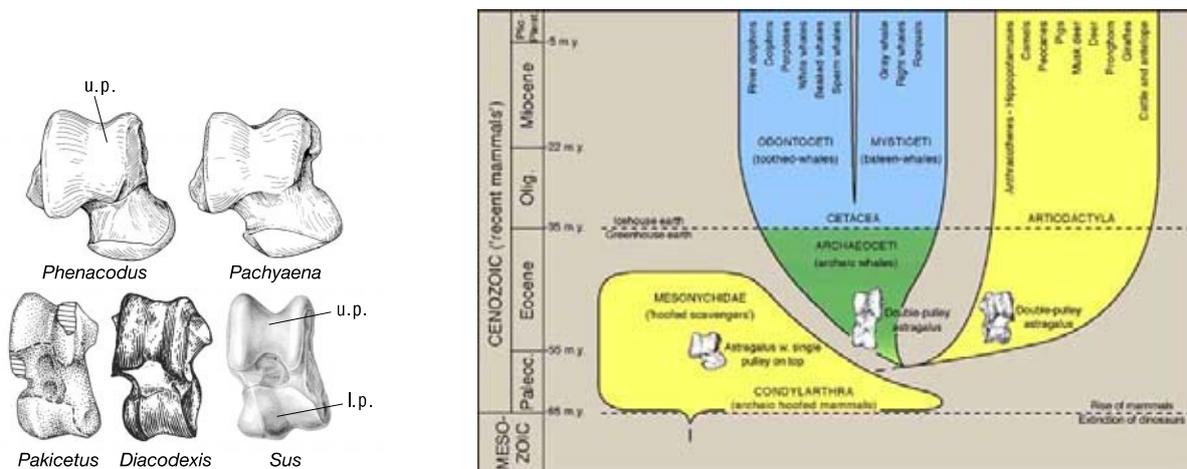


Figure 27. Le Talus affublé d'une double-poulie : une synapomorphie des Cetartiodactyla

- a. Représentation de quelques *Tali* de Cétartiodactyles, 4 fossiles et un actuel, le Porc (de Muizon, 2001) ;
- b. Représentation 'phylético-stratigraphique' des principaux groupes de Cétartiodactyles montrant l'acquisition synapomorphique de la double-poulie au Paléocène-Eocène (courtoisie de Phillip Gingerich - <http://www-personal.umich.edu/~gingeric/PDGwhales/Whales.htm>).

Même si un consensus a émergé quant à la parenté des Hippopotames et des Cétacés, (Geisler & Uhen, 2003) la phylogénie 'basale' des Cétartiodactyles demeure cependant une question ouverte (Luo, 2000) si l'on considère les divergences de vues tant au sein des paléontologues (Boisserie *et al.*, 2005; Thewissen *et al.*, 2001) qu'au sein des phylogénéticiens (*cf.* Figure 28 ci-dessous).

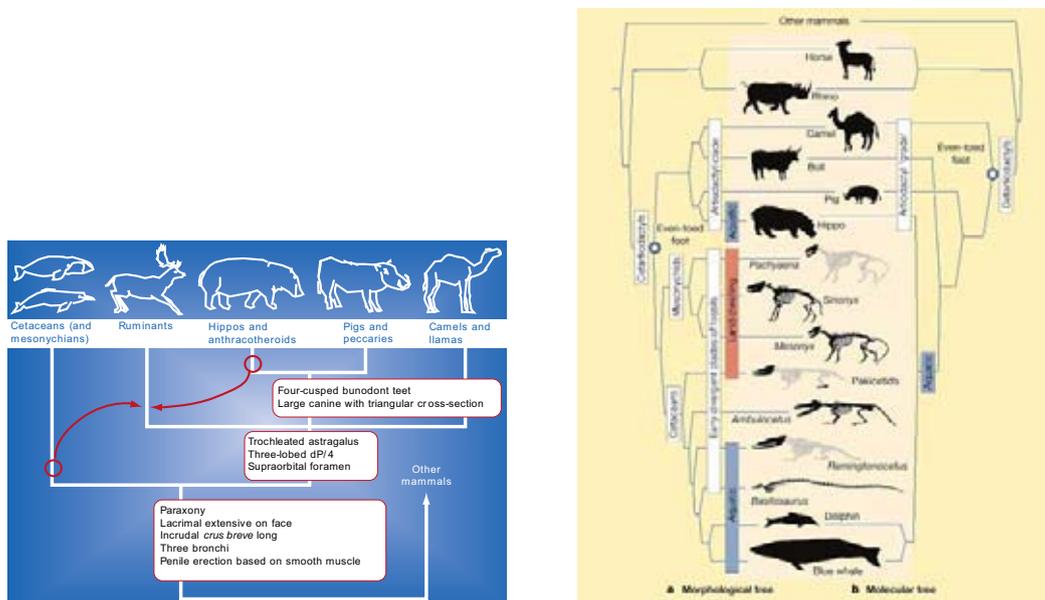


Figure 28. Différentes approches de la phylogénie des Cétartiodactyles.
 a : d'après Millinkovitch & Thewissen (1997)
 b : d'après Gatesy & O' Leary (2001)

II. Méthodologie

II. A. La cladistique

Le mot cladistique provient du grec ancien κλάδος - branche. Le terme de phylogénétique est souvent abusivement utilisé comme synonyme de cladistique. En effet, si la phylogénétique est une discipline, la cladistique n'est qu'une méthode (majoritairement) utilisée dans cette discipline.

Les principes de la cladistique sont apparus dans les années 60 (Cain & Harrison,

1960; Huxley, 1957) et proposés par Willi Hennig (Hennig, 1965). Les disciples de ce dernier ont été taxés de *cladists* par Ernst Mayr (Mayr, 1965). Cependant, la cladistique ne s'est pleinement imposée qu'à partir des années 80 (Dupuis, 1984) par l'utilisation de plus en plus répandue de ce que l'on appelait alors les micro-ordinateurs - lorsqu'ils étaient rares. La cladistique est devenue une méthode majeure en biologie évolutive et dans le même temps, l'explosion de la biologie moléculaire a permis l'application de cette méthode aux caractères biochimiques et moléculaires.

Un des principes fondamentaux de la cladistique est que l'arbre évolutif est si complexe qu'on ne peut l'appréhender en fixant un nombre fixé par avance de niveaux systématiques, ce qui est le cas en Systématique Linnéenne (SL). Les tenants de la Systématique Cladistique (SC) la considèrent supérieure à la SL par les aspects suivants : la nécessité par la SL de créer de nouveaux niveaux taxonomiques dès qu'interviennent de nouvelles découvertes (infra-ordre, super-famille, parvo-ordre etc.) ; l'interdiction de nommer des *taxa* polyphylétiques en SC ; la possibilité ouverte d'évolution permanente de l'arbre en SC *vs* la rigidité de l'arbre établi en SL. Ainsi, la SC se prétend plus reflet de l'évolution biologique que la SL, reflet de l'état phénotypique actuel.

En cladistique, contrairement à l'utilisation qui en est faite en taxonomie, le terme de **taxon** - du grec ancien τᾰξις arrangement (plur. *taxa* ou taxons) - peut s'appliquer à un individu échantillonné au sein d'un groupe dont on veut étudier les parentés. Cela peut être un ensemble d'individus d'une même espèce ou un ensemble d'espèces au sein d'un groupe de rang supérieur.

La **monophylie** - du grec ancien μόνος/όν seul τὸ φύλον feuille - est le lien entre des *taxa* partageant un ancêtre commun au sein du groupe considéré par l'expérimentateur. En effet, si l'on ne tient pas compte de ce dernier point, l'ensemble des *taxa* est monophylétique puisque l'ensemble des êtres vivants ont un ancêtre commun. Un **clade** est un groupe monophylétique de *taxa* partageant un ensemble de caractères dérivés communs ou **synapomorphies**, du grec ancien συν - avec -, ἀπό - indépendamment - et de ἡ μορφή - figure, forme. Un groupe polyphylétique est un groupe de *taxa* échantillonnés à partir de plusieurs groupes monophylétiques

distincts. Un **groupe externe** plus généralement appelé *outgroup* est un taxon ne formant pas partie du group considéré mais suffisamment proche pour que les caractères étudiés ne subissent pas trop de réversions. Un caractère survenant seulement chez des descendants est dit **apomorphe** alors qu'un caractère plésiomorphe, du grec ancien πλεσιός/όν - ancien - est present chez l'*outgroup* et les ancêtres. Les adjectifs 'apomorphe' et 'plésiomorphe' sont utilisés à la place de 'dérivé' et 'primitif' car ce dernier est trop connoté.

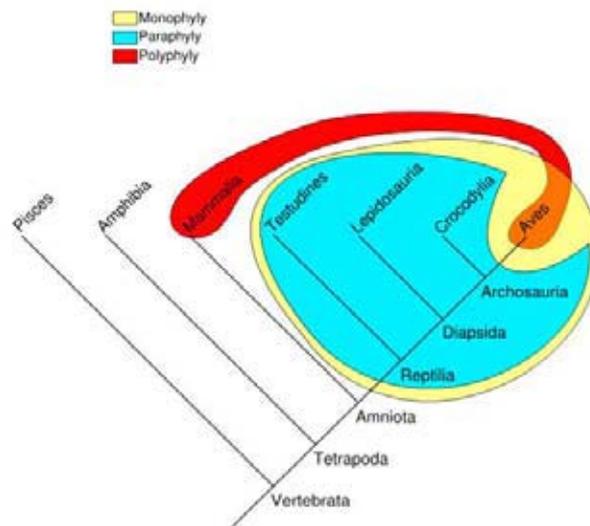


Figure 29. Diagramme de groupes phylogéniques de Vertébrés montrant les taxa mono/para/polyphylétiques (image sous licence GNU, Wikimedia Commons).

Une analyse cladistique vise à obtenir un **cladogramme** (cf. **Figure 29** ci-dessus), représentation graphique d'un processus évolutif sur un jeu de caractères donné. Il est donc lui-même sujet à évolutions lors d'acquisition de nouvelles données. Le nombre de cladogrammes possibles varie factoriellement selon le nombre de *taxa*. Un cladogramme a un axe temporel implicite depuis la base jusqu'aux 'feuilles'.

Le processus consiste tout d'abord à générer une matrice (*taxa vs caractères*) de caractères codés de telle sorte qu'ils puissent être analysés de façon discrète : cela est aisé pour les séquences moléculaires mais plus difficile pour des classes de traits quantitatifs traditionnels (e.g. morphologiques) voire délicat lorsqu'il s'agit de traits où l'on doit définir de façon plus ou moins arbitraire les limites des **états du caractère** (e.g. couleurs).

L'analyse consiste à choisir les états de chacun des caractères qui sont plésiomorphes

et ceux qui sont apomorphes. Un point déterminant est alors d'éviter les **homoplasies** (ex : présence d'ailes), états du caractère partagés entre *taxa* issus de groupes monophylétiques différents (e.g. lors d'évolution convergente). Un des moyens est de préciser le caractère ou de combiner plusieurs caractères (en ce qui nous concerne, séquence nucléotidique par exemple).

Etant donné le grand nombre de cladogrammes possible, seules des méthodes heuristiques s'avèrent d'utilisation pratique par la répétition de boucles de procédures (algorithmes), facilement programmables. Un schéma général de boucle peut être défini par les pas suivants :

- 1) sélection d'un certain nombre de cladogrammes candidats par un score, par exemple la vraisemblance du cladogramme (sachant les données) ;

- 2) génération de candidats alternatifs par modification des cladogrammes (voir inférence phylogénique ci-après) ;

- 3) évaluation des nouveaux candidats, comparaison des scores ;

- 4) arrêt de la procédure si pas d'évolution dans les scores ;

- 5) sélection de nouveaux candidats (et retour à l'étape 1).

La principale divergence entre les différentes méthodes d'inférence est la manière d'identifier le 'meilleur' cladogramme. On peut *grosso-modo* classer les différents algorithmes en méthodes de distance (moindres-carrés, neighbor-joining, parcimonie), maximum de vraisemblance et méthodes bayésiennes.

II. B. Le codage des indels dans les alignements de séquence

La principale difficulté de l'analyse de la région que nous prenons en considération est la présence de nombreux indels, c'est à dire de mutations de type insertion ou délétion lorsqu'on compare les séquences deux à deux.

B.1) Les algorithmes d'alignement

1(a) Cadre théorique : les méthodes de partitionnement des données

Le **partitionnement de données** (*data clustering* en anglais) est une méthode

statistique d'analyse de données qui a pour but de regrouper un ensemble de données en différents paquets homogènes, en ce sens que les données de chaque sous-ensemble partagent des caractéristiques communes, qui correspondent le plus souvent à des critères de proximité que l'on définit en introduisant des mesures de distance (**Wikipedia, 2008**). Pour obtenir un bon partitionnement, il convient de :

- minimiser l'inertie intra-classe pour obtenir des grappes (*cluster* en anglais) les plus homogènes possibles ;

- maximiser l'inertie inter-classe afin d'obtenir des sous-ensembles bien différenciés.

La communauté scientifique francophone utilise différents termes pour désigner cette méthode. Le mot anglais *clustering* est communément employé. On parle également souvent des méthodes de *regroupement*. Le partitionnement de données est une méthode de *classification non supervisée*, à la différence de la classification supervisée où les données d'apprentissage sont déjà étiquetées, et donc parfois dénommées comme telles.

Il existe de multiples méthodes de partitionnement des données, parmi lesquelles la méthode des nuées dynamiques, le regroupement hiérarchique, l'algorithme EM (Expectation Maximisation utilisé par exemple pour le calcul du déséquilibre de liaison en génétique), l'analyse en composantes principales (ACP).

1(b) L'algorithme d'alignement ClustalW

L'algorithme de Clustal W (**Thompson et al., 1994**) est un algorithme procédant par partitionnement de données, alignement partiels et pondération des portions de séquences en fonction de la divergence (les plus divergentes sont pondérées plus fortement). Par ailleurs, il permet d'attribuer des pénalités variables à l'ouverture de gaps (portions de séquences non alignables) ainsi qu'à leur élongation.

B.2) Codage des indels : la méthode MCIC (Müller, 2006)

Les indels apparaissant lors de l'alignement des séquences ne peuvent pas être directement traités par des algorithmes d'inférence phylogénique. Il convient donc de les regrouper et de les coder à la manière de ce qui est pratiqué dans les méthodes

d'association. L'algorithme ou procédure MCIC (*Modified Complex Indel Coding*) permet de coder les indels.

Cette méthode suit 3 étapes (*cf.* Figure 30) :

- 1) choix des indels devant être regroupés (*pooled*) dans un seul caractère ;
- 2) définition des différents états du caractère ;
- 3) définition des coûts de passage d'un état de caractère à un autre.

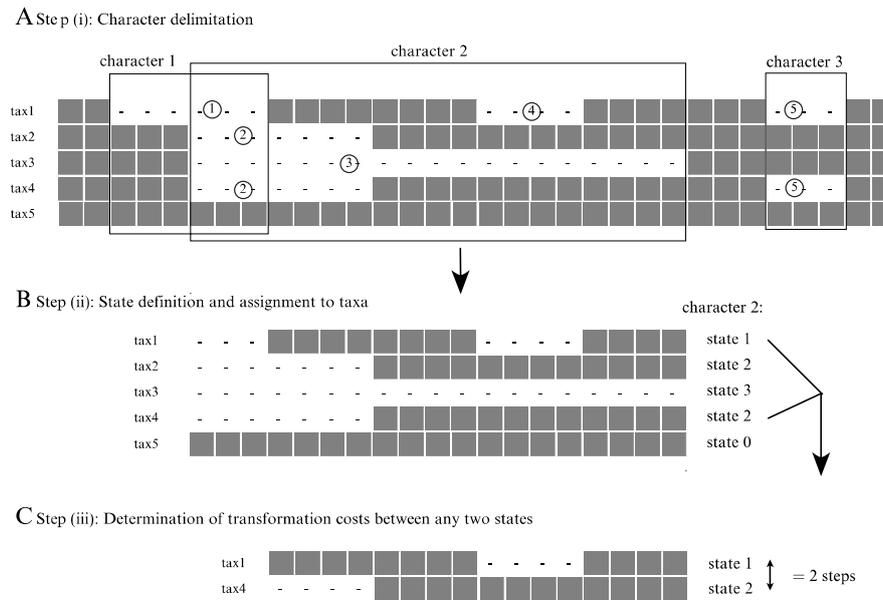


Figure 30. Les trois étapes de la procédure MCIC (Müller, 2006).

- (A) Délimitation du caractère ;
- (B) Définition et attribution des états du caractère ;
- (C) Détermination des coûts de mutation d'un état à un autre.

2(a) Délimitation de l'indel, définition de l'état du caractère

Un indel peut être défini par une paire de nombres : le premier représentant la position de l'extrémité 5' (amont de la séquence) et le second représentant l'extrémité 3' (aval). Si l'on a deux indels $I=(i_1, i_2)$ et $K=(k_1, k_2)$, on définit $I < K \Leftrightarrow (i_1 < k_1 \cdot OU \cdot (i_1 = k_1 \cdot ET \cdot i_2 < k_2))$.

L'algorithme de délimitation du caractère se déplace alors le long d'un vecteur contenant tous les indels, classés dans un ordre croissant tel que défini précédemment. En mettant à jour continuellement l'extension 3' maximale de tous les indels du caractère considéré, l'algorithme est à même de décider si l'indel 'en cours' appartient au caractère considéré ou au prochain.

2(b) Définition et attribution des états du caractère

Un état du caractère est défini par une combinaison différente d'indels tous contenus dans la séquence d'un même taxon. L'état correspondant à une série vide est par convention appelé état 0.

2(c) Estimation des coûts de mutation : évaluation du nombre de pas nécessaires entre chaque état

Sachant deux combinaisons d'alternances de gaps et blocs consécutifs (états), le passage de l'une à l'autre peut être traduit en une paire de séquences de données binaires de longueur L (cf. **Figure 31**). On réduit dans un premier temps l'information redondante (série de paires $\frac{1}{0}$ ou $\frac{0}{1}$, étape A-B) puis on supprime les synapomorphies 'nulles' (coût de transformation nul) communes aux deux séquences ($\frac{0}{0}$, étape B). Après assemblage des paires $\frac{1}{0}$ ou $\frac{0}{1}$ adjacentes (étape C), on ignore à cette étape toutes les paires d'indels ($\frac{1}{1}$, étape D). Le coût équivaut alors au nombre de colonnes. Mathématiquement, si l'on considère les séquences binaires résultantes comme deux séries U_k et U_j (avec $U_x = (c_{m,1}, c_{m,2}, \dots, c_{m,L})$), le nombre de changements de la séquence j à k peut être estimé par :

$$C_{jk} = C_{kj} = \text{nombre}\{ic_j \neq c_{k,i}, 1 \leq i \leq L\}$$

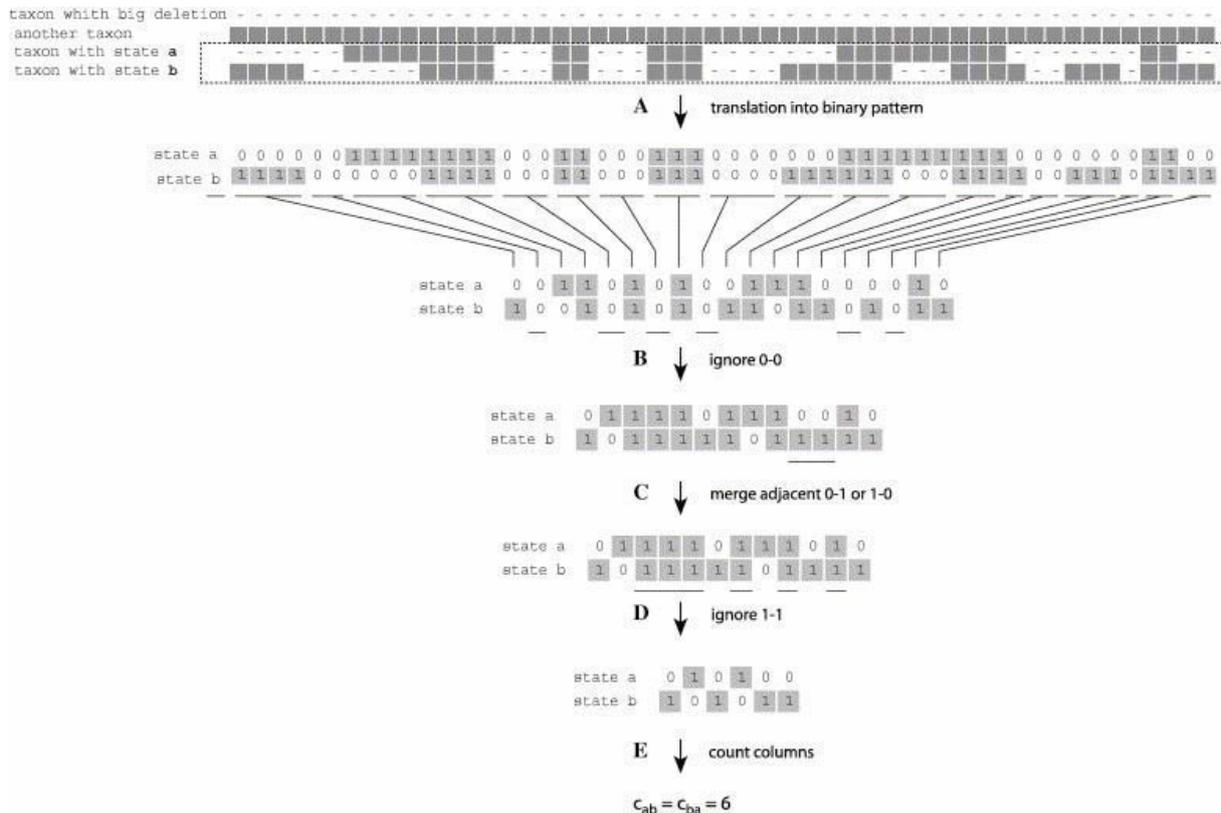


Figure 31. Etapes de la détermination des coûts de transformation entre deux états d'un caractère complexe (Müller, 2006).

Cet algorithme (MCIC) est implémenté dans le logiciel IndelCoder, disponible en JAVA, intégré au logiciel SeqState à l'adresse : <http://www.nees.uni-bonn.de/downloads/>

II. C. Modèles de mutation

C.1) Qu'est-ce qu'un modèle de mutation ?

Un modèle de mutation/substitution est un modèle décrivant les processus par lesquels une séquence de caractères, d'un certain alphabet comportant un nombre défini de caractères, change pour un autre assemblage de caractères (Felsenstein, 2004b). D'un point de vue cladistique, chaque position dans la séquence peut être vue comme propriété apomorphe d'une espèce, prenant alors une valeur selon sa présence ou son absence (par convention 0 ou 1).

En phylogénétique, les séquences sont tout d'abord alignées, et les caractères sont représentés par chacune des positions de l'alignement.

Ces modèles sont utilisés dans deux cas de figure :

- analytique : construction d'arbres évolutifs ;
- simulation : afin de tester des hypothèses et algorithmes

C.2) Caractéristiques, horloge moléculaire

2(a) Caractéristiques

La plupart des modèles de substitution utilisés actuellement sont neutres (la sélection n'opère pas sur eux), indépendants (les changements sur l'un d'entre eux n'affectent pas la probabilité de changement sur un autre site) et « *finite-site* » (entre deux états identiques à des temps t et $t+x$, il a pu y avoir des multiples changements).

2(b) L'horloge moléculaire et les unités de temps

Les différents modèles prennent en compte le temps de manière diverse :

- le plus fréquent est de compter le temps en substitutions. L'intérêt reside dans le fait que l'on n'a pas besoin de se préoccuper de la variabilité du taux de substitution en fonction du temps ;

- la **présomption d'horloge moléculaire** est de même largement répandue. Dans ce cas, on considère que le rapport entre taux de substitution et unité de temps est constant et égal à μ (nombre de substitutions par unité de temps). Cela nécessite de connaître μ à partir d'une seule branche mais cela ouvre la voie à la contestation des résultats. Ce genre de présomption est souvent irréaliste, surtout à travers de longues périodes évolutives, en particulier lorsqu'entre les groupes étudiés les intervalles entre générations ou les structures de populations sont très différents (par exemple, les Rongeurs et Primates sont deux groupes phylogénétiquement proches mais les Rongeurs ont subi un bien plus grand nombre de substitutions par site en moyenne) ;

- enfin, certains travaux ont permis l'émergence de modèles permettant la prise en compte de taux d'évolution variables (**Bromham et al., 1998; Kishino et al., 2001; Rambaut & Bromham, 1998**).

C.3) Aspect mathématique des modèles de substitution

Les modèles neutres, indépendants et *finite-site* (présumant d'un taux d'évolution constant) ont deux paramètres :

- Π , un vecteur de fréquence des caractères (bases, acides aminés) à $t=0$. Pour un modèle temps-réversible, ce vecteur est généralement confondu avec les fréquences des bases à l'équilibre et valable pour tout t ;

- Q , une matrice des taux d'évolution qui décrit le taux de passage d'un état du caractère i au caractère j . Par convention, les diagonales de la matrice sont choisies de telle sorte que les lignes totalisent zéro ($Q_{ii} = -\sum_{i \neq j} Q_{ij}$). La fonction de la matrice de transition est fonction des longueurs de branche (en unité de temps, e.g. substitutions) vers une matrice des probabilités conditionnelles $P(t)$. Ainsi, l'entrée $P_{ij}(t)$ est la probabilité qu'au temps t il y ait une base j à une position donnée sachant qu'il y a eu un i à cette position à $t=0$. Lorsque le modèle est temps-réversible, cette opération peut être effectuée si l'on connaît la longueur de branche totale entre deux séquences, même si l'une n'est pas l'ancêtre de l'autre.

3(a) Propriétés asymptotiques :

Cette propriété s'exprime comme suit : $\lim_{t \rightarrow 0} P_{ij}(t) = \Pi_i$ & $\lim_{t \rightarrow \infty} P_{ij}(t) = \Pi_j$, ce qui signifie d'une part qu'il n'y a pas de changement de composition des bases entre une séquence et elle-même et que la probabilité de trouver une base j à une position lorsque t tend vers l'infini devient indépendante de l'état initial (i.e. la base originelle).

La matrice de transition peut être calculée à partir de la matrice des taux et des fréquences des bases à l'équilibre par $P(t) = e^{Qt}$, une matrice qui peut être approchée par une la série de Taylor : $P(t) = \sum_{n=0}^{\infty} Q^n \frac{t^n}{n!}$. La contrainte de réversibilité/état statique est équivalente à $\Pi Q = 0$ (somme des lignes = 0, fréquences des bases ne devant pas dévier globalement de Π) soit $\Pi P(t) = \Pi, \forall t$.

3(b) Matrice de transition

Différents modèles de mutation ont été implémentés en raison de la grande variation de processus de mutation entre les génomes et à l'intérieur d'un même génome (par exemple entre des régions codantes et non codantes). Ces modèles diffèrent principalement dans la paramétrisation des matrices des taux de substitution et la modélisation des taux d'évolution. On peut considérer l'évolution de l'ADN

considéré comme une Chaîne de Markov “temps-continu” (*Continuous-time Markov Chains, CTMC*). Ces CTMC sont caractérisées par les matrices de transition classiques qui tiennent en outre compte du temps, t . Si E_1, \dots, E_n sont les différents états possibles ($n = 4$ dans la plupart des modèles, soit les quatre nucléotides).

On a alors : $\Pr(t) = \Sigma \Pr_{ij}(t)$ (probabilité du passage de l'état E_i à l'état E_j dans l'espace de temps t). Soit la matrice suivante :

$$\Pr(t) = \begin{pmatrix} \rho_{AA}(t) & \rho_{GA}(t) & \rho_{CA}(t) & \rho_{TA}(t) \\ \rho_{AG}(t) & \rho_{GG}(t) & \rho_{CG}(t) & \rho_{TG}(t) \\ \rho_{AC}(t) & \rho_{GC}(t) & \rho_{CC}(t) & \rho_{TC}(t) \\ \rho_{AT}(t) & \rho_{GT}(t) & \rho_{CT}(t) & \rho_{TT}(t) \end{pmatrix}$$

3(c) Réversibilité temporelle

Un processus Markovien est temps-réversible (*time reversible*), dans son état stationnaire, lorsque la quantité de changements de l'état x à l'état y est égale à celle du passage de l'état y à l'état x , et ce même si les deux états surviennent à des fréquences différentes. Autrement dit : $\pi_x \mu_{xy} = \pi_y \mu_{yx}$.

Tous les processus stationnaires ne sont pas réversibles mais la plupart des modèles d'évolution de l'ADN supposent la réversibilité temporelle comme une présomption raisonnable. Ainsi, $S_{xy} = \frac{\mu_{xy}}{\pi_y} = S_{yx}$, terme symétrique, est appelé ‘échangeabilité’ (*exchangeability*) entre les états x et y . Autrement dit, S_{xy} est la fraction de la fréquence de l'état x qui résulte de la transition de l'état y à l'état x .

Le corollaire est que les 12 entrées sous la diagonale de la matrice Q peuvent être complètement déterminées par 9 variables, 6 étant les termes d'échangeabilité et 3 les fréquences stationnaires π_x (attendu que les fréquences totalisent 1). En d'autres termes, cela signifie que les fréquences relatives de chaque caractère ne changent pas au cours du temps et qu'il n'y a pas de direction préférentielle dans le changement (e.g. T>G>A est équivalent à A>G>T).

La raison pour laquelle il a été nécessaire d'avoir recours à des modèles temps-réversible est l'impossibilité ou pour le moins la difficulté d'accès aux séquences ancestrales. Grâce à ces modèles, il est possible d'enraciner l'arbre sur n'importe quelle branche menant à une espèce actuelle puis de le ré-enraciner sur une autre

branche, ou de le laisser non enraciné, en fonction de nouvelles connaissances.

C.4) Modèles les plus courants d'évolution de l'ADN

4(a) JC69 (Jukes & Cantor, 1969)

Le modèle de Jukes et Cantor est le plus simple des modèles de substitution. Il présume :

- des fréquences de bases égales ($\pi_A = \pi_T = \pi_G = \pi_C = 1/4$) ;
- des taux de mutation égaux.

Ainsi, le seul paramètre autorisé à varier est μ , le taux de substitution global :

$$Q = \begin{pmatrix} * & \frac{\mu}{4} & \frac{\mu}{4} & \frac{\mu}{4} \\ \frac{\mu}{4} & * & \frac{\mu}{4} & \frac{\mu}{4} \\ \frac{\mu}{4} & \frac{\mu}{4} & * & \frac{\mu}{4} \\ \frac{\mu}{4} & \frac{\mu}{4} & \frac{\mu}{4} & * \end{pmatrix}, P = \begin{pmatrix} \frac{1}{4} + \frac{3}{4} e^{-\mu} & \frac{1}{4} - \frac{1}{4} e^{-\mu} & \frac{1}{4} - \frac{1}{4} e^{-\mu} & \frac{1}{4} - \frac{1}{4} e^{-\mu} \\ \frac{1}{4} - \frac{1}{4} e^{-\mu} & \frac{1}{4} + \frac{3}{4} e^{-\mu} & \frac{1}{4} - \frac{1}{4} e^{-\mu} & \frac{1}{4} - \frac{1}{4} e^{-\mu} \\ \frac{1}{4} - \frac{1}{4} e^{-\mu} & \frac{1}{4} - \frac{1}{4} e^{-\mu} & \frac{1}{4} + \frac{3}{4} e^{-\mu} & \frac{1}{4} - \frac{1}{4} e^{-\mu} \\ \frac{1}{4} - \frac{1}{4} e^{-\mu} & \frac{1}{4} - \frac{1}{4} e^{-\mu} & \frac{1}{4} - \frac{1}{4} e^{-\mu} & \frac{1}{4} + \frac{3}{4} e^{-\mu} \end{pmatrix}$$

La distance entre deux séquences est alors donnée par : $d = -\frac{3}{4} \ln(1 - \frac{4}{3} p)$, avec p proportion de sites différant entre les deux séquences.

4(b) K80 (Kimura, 1980)

Ce modèle distingue les transitions des transversions (cf. **Figure 32**) et présume des fréquences nucléotidiques égales $\pi_A = \pi_T = \pi_G = \pi_C = 1/4$. Il est caractérisé par :

- la matrice des taux de substitution $Q = \begin{pmatrix} * & \kappa & 1 & 1 \\ \kappa & * & 1 & 1 \\ 1 & 1 & * & \kappa \\ 1 & 1 & \kappa & * \end{pmatrix};$

- la distance 'Kimura-deux paramètres' $d = -\frac{1}{2} \ln(1 - 2p - q) - \frac{1}{4} \ln(1 - 2q)$, avec p ,

proportion de sites présentant des transitions et q , proportion de sites présentant des transversions.

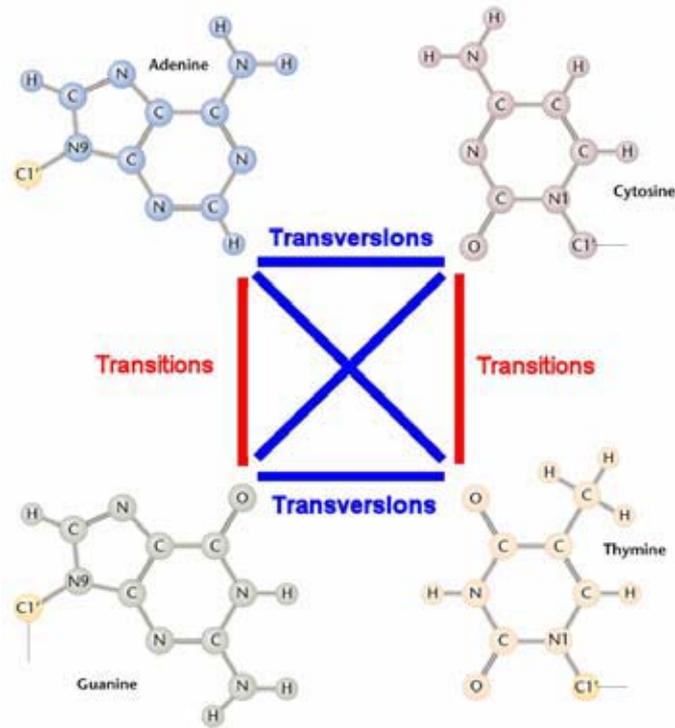


Figure 32. Illustration des transitions (purine \leftrightarrow purine) et transversions (purine \leftrightarrow pyrimidine) ; Courtoisie du Dr. Steven M. Carr, Memorial Univ. Newfoundland.

4(c) F81 (Felsenstein, 1981)

Fréquences nucléotidiques inégales $\pi_A \neq \pi_T \neq \pi_G \neq \pi_C$

Matrice des taux de substitution
$$Q = \begin{pmatrix} * & \pi_T & \pi_T & \pi_T \\ \pi_C & * & \pi_C & \pi_C \\ \pi_A & \pi_A & * & \pi_A \\ \pi_G & \pi_G & \pi_G & * \end{pmatrix}$$

4(d) HKY85 (Hasegawa et al., 1985)

Il s'agit d'un modèle du même type que le modèle de Felsenstein mais qui distingue les transitions des transversions.

Sa matrice des taux de substitution est
$$Q = \begin{pmatrix} * & k\pi_T & \pi_T & \pi_T \\ k\pi_C & * & \pi_C & \pi_C \\ \pi_A & \pi_A & * & k\pi_A \\ \pi_G & \pi_G & k\pi_G & * \end{pmatrix}$$

4(e) T92 (Tamura, 1992)

Il est caractérisé par une seule fréquence : π_{GC} d'où l'on déduit :

$$\pi_G = \pi_C = \frac{\pi_{GC}}{2}$$

$$\pi_A = \pi_T = \frac{(1 - \pi_{GC})}{2}$$

$$Q = \begin{pmatrix} * & \frac{\kappa(1 - \pi_{GC})}{2} & \frac{(1 - \pi_{GC})}{2} & \frac{(1 - \pi_{GC})}{2} \\ \frac{\kappa\pi_{GC}}{2} & * & \frac{\pi_{GC}}{2} & \frac{\pi_{GC}}{2} \\ \frac{(1 - \pi_{GC})}{2} & \frac{(1 - \pi_{GC})}{2} & * & \frac{\kappa(1 - \pi_{GC})}{2} \\ \frac{\pi_{GC}}{2} & \frac{\pi_{GC}}{2} & \frac{\kappa\pi_{GC}}{2} & * \end{pmatrix}$$

La distance entre deux séquences non codantes est ici :

$d = -h \ln \left(1 - \frac{\rho}{h} - Q \right) - \frac{1}{2} (1 - h) \ln (1 - 2Q)$ où $h = 2\vartheta(1 - \vartheta)$ avec $\vartheta \in (0, 1)$ étant la proportion de GC.

4(f) TN93 (Tamura & Nei, 1993)

Le modèle Tamura-Nei distingue les deux types de transition (taux A \leftrightarrow G différent de C \leftrightarrow T).

Ce modèle présume des fréquences inégales : $\pi_A \neq \pi_T \neq \pi_G \neq \pi_C$.

$$Q = \begin{pmatrix} * & \kappa_1 \pi_T & \pi_T & \pi_T \\ \kappa_1 \pi_C & * & \pi_C & \pi_C \\ \pi_A & \pi_A & * & \kappa_2 \pi_A \\ \pi_G & \pi_G & \kappa_2 \pi_G & * \end{pmatrix}$$

Cette matrice montre clairement la parenté avec les modèles F81 et HKY95.

4(g) GTR (Generalised time reversible (Tavaré, 1986))

Il s'agit du modèle neutre, indépendant, temps-réversible le plus général possible. Il est caractérisé par les paramètres suivants :

- un vecteur des fréquences : $\Pi = (\pi_1, \pi_2, \pi_3, \pi_4)$;
- une matrice des probabilités de mutation

$$Q = \begin{pmatrix} -\left(\frac{\pi_1 X_1}{\pi_2} + \frac{\pi_1 X_2}{\pi_3} + \frac{\pi_1 X_3}{\pi_4}\right) & \frac{\pi_1 X_1}{\pi_2} & \frac{\pi_1 X_2}{\pi_3} & \frac{\pi_1 X_3}{\pi_4} \\ X_1 & -\left(X_1 + \frac{\pi_2 X_4}{\pi_3} + \frac{\pi_2 X_5}{\pi_4}\right) & \frac{\pi_2 X_4}{\pi_3} & \frac{\pi_2 X_5}{\pi_4} \\ X_2 & X_4 & -\left(X_2 + X_4 + \frac{\pi_3 X_6}{\pi_4}\right) & \frac{\pi_3 X_6}{\pi_4} \\ X_3 & X_5 & X_6 & -(X_3 + X_5 + X_6) \end{pmatrix}.$$

Dans ce cas (4 caractères, comme généralement le cas en phylogénie moléculaire), le modèle nécessite 6 paramètres des taux de substitution ainsi que 4 paramètres de fréquences nucléotidiques à l'équilibre. Le modèle est cependant simplifié à 9 paramètres additionnés de μ , le nombre total de substitutions par unité de temps. Ainsi, lorsqu'on mesure le temps en nombre de substitution (*i.e.* $\mu=1$), le modèle est complètement caractérisé par 9 paramètres.

En dehors du cas des séquences nucléotidiques ($n > 4$ valeurs/traits par site), on obtient le nombre de paramètres en comptant le nombre d'entrées au-dessus de la diagonale :

$$\underbrace{\frac{n^2 - n}{2}}_{\text{taux_de_substitution}} + \underbrace{n}_{\text{fréquences}} - \underbrace{1}_{\mu} = \frac{1}{2}n^2 + \frac{1}{2}n - 1$$

Ainsi, par exemple dans le cas de séquence d'acides amines (20 états) :

Codons ($4^3=64$ codons) = $\frac{20 \times 19}{2} \times 3 + 64 - 1 = 633$ paramètres (on considère les taux de transition entre codons qui diffèrent de plus d'une base comme nuls).

II. D. Inférence phylogénétique

Des travaux de Charles Darwin (**Darwin, 1859**) ont émergé deux principes qui ont eux-même donné naissance à deux disciplines, à la manière d'un arbre phylogénétique :

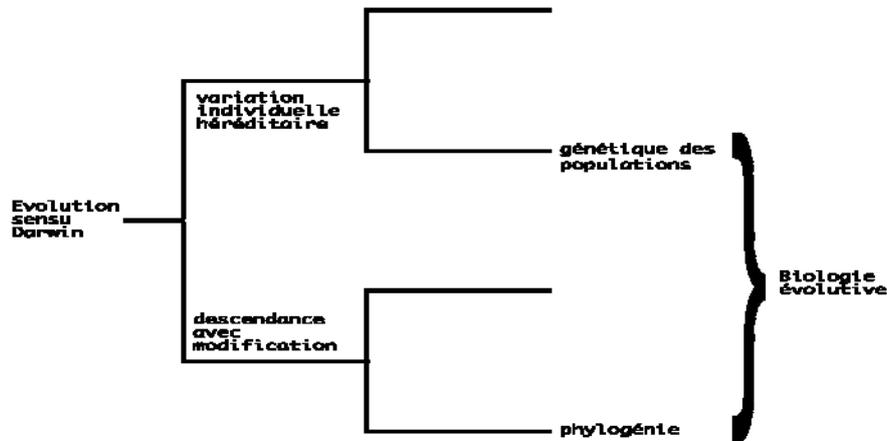


Figure 33. La 'phylogénie des phylogénies'.

D.1) Principe

La reconstruction phylogénétique vise à déterminer les arbres de *taxa* les plus proches de la réalité évolutive. Or le nombre d'arbres possibles selon le nombre de *taxa* est une fonction dont la croissance est factorielle (plus de 10^{21} pour 20 *taxa*, voir tableau ci-après Lowe *et al.*, 2004).

Nombre de <i>taxa</i>	2	3	4	5	6	7	8	9	10	N
Nombre de cladogrammes	1	3	15	105	945	10395	135135	2027025	34459425	$1 \times 3 \times 5 \times 7 \times \dots \times (2N-3)$

Tableau 7. Correspondance entre le nombre de cladogrammes possibles en fonction du nombre de *taxa* considérés.

La reconstruction phylogénétique se base sur une analyse cladistique (*cf.* II. A) de la variation de caractères observés (dans notre cas, l'alignement de séquences des *taxa* étudiés).

D.2) Les trois grandes familles de méthodes

2(a) Méthodes de distance

Le principe des méthodes de distance est de calculer des degrés d'homologie des séquences deux à deux, soit dans le cas de notre illustration le nombre de barres liant

les caractères dans le même état des deux séquences, sur le nombre total de caractères :

```
Taxon 1  ATGCATGCATGC
          |||  ||  |||
Taxon 2  ATGGAAACCTGC
          ||  ||  |||  |||
Taxon 3  ATTGACCCCTGC
```

On exprime en général cette distance en pourcentage et l'on représente les liens phylogéniques dans une matrice de distance :

	Taxon 1	Taxon 2	Taxon 3
Taxon 1			
Taxon 2	4		
Taxon 3	5	3	

L'algorithme de partitionnement repère dans la matrice la distance la plus faible (ici 2-3) et opère une sommation des *taxa* :

	Taxon (2+3)	Taxon 1
Taxon (2+3)		
Taxon 1	4,5	

Ce genre d'opération se poursuit sur chaque matrice obtenue successivement jusqu'à obtention d'une matrice 2x2.

La représentation se fait sous forme d'un arbre. On mesure aisément la faiblesse de ces méthodes : elle ne prennent en compte qu'une partie de l'information (homologie ou non) et ne tiennent pas compte de mutations indépendantes au même site (homoplasies). Pour ces raisons ont été développées des méthodes soit par agglomération de *taxa* (*Neighbor-Joining*, *Weighted/Unweighted Pair Group Method with Arithmetic mean*) soit des méthodes paramétriques reconstruisant des dendrogrammes selon des critères *a priori* (Moindres carrés).

2(b) Cladistique et Parcimonie

Dans ces méthodes, l'information de chaque site est prise en compte de manière cladistique. Les groupes sont donc définis par des synapomorphies. Le problème qui se pose alors est la multiplicité des arbres possibles. Une approche consiste en un *a priori* évolutif : la parcimonie retient l'arbre consistué du moins grand nombre de pas possibles (arbre le plus « parcimonieux »).

2(c) Approches probabilistes

Afin de résoudre les problèmes de manière non paramétrique et d'éviter des biais comme l'attraction des longues branches observée avec les méthodes de parcimonie (**Felsenstein, 1981**), on utilise des approches probabilistes basées sur la probabilité à chaque site de passer d'un état à un autre. Deux approches voisines se présentent alors : une approche par optimisation (maximisation de la vraisemblance, *cf.* Chapitre III Chapitre I.IV. B.1(b)) et une approche par densité de probabilité

(inférence Bayésienne, cf. Chapitre III Chapitre I.IV.B.3).

III. Matériel et méthodes

III. A. Biologie moléculaire

Le matériel biologique est issu de tissus mous d'animaux échoués et l'ADN a été extrait selon le protocole classique 'phénol-chloroforme' (Sambrook & Russell, 2001). Nous avons d'abord amplifié le gène de l'amélogénine à l'aide d'amorces hétérologues utilisées au sein de l'équipe sur du matériel issu de Primates. De nouvelles amorces (SC1 et SC2) spécifiques des *Cetartiodactyla* ont par la suite été dessinées après clonage, séquençage et alignement des fragments obtenus (Cétacés) avec des références issues de la base de données nucléotidiques GenBank (*Artiodactyla*, cf. [article n°1](#) pour le détail). Nous les avons placées de telle sorte qu'elles permettent l'amplification complète de l'homologue du 4^e intron. Les conditions optimales d'amplification ont été déterminées sur des appareils à PCR à gradient de température.

La procédure a été reproduite sur 8 espèces de Cétacés : le Dauphin bleu et blanc (*Stenella æruleoalba*), le Rorqual commun (*Balaenoptera phsalus*), le Grand dauphin (*Tursiops truncatus*), le Dauphin commun (*Delphinus dlphis*), la Baleine grise (*Eschrichtius robustus*), le Cachalot (*Physeter acrocephalus*), le Rorqual à museau pointu (*Balaenoptera autorostrata*) et le Dauphin de Risso (*Grampus giseus*). Les fragments ont par la suite été séquencés, après clonage ou en séquence directe.

III. B. Analyses bioinformatiques

Les séquences obtenues ont été alignées à l'aide de l'algorithme basé sur une méthode de partitionnement de données Clustalw avec les séquences suivantes issues de la base de données de GenBank :

- le Grand dauphin : [AY787743S2](#)

- Bovin : [AB091789](#) & [AB091790](#)

- Porc : [AF328419](#) & [AB091792](#)

- Ovin : [AY604731](#)

- Humain : [NT_011757](#) (de la position 9098117 à la position 9098612) &

[NC_000024](#) (de la position 6796200 à la position 6796719)

Les séquences générées ont été stockées dans la base de données de l'EMBL, interrogée sur son site miroir GenBank (accession AM744958 à AM744971).

La difficulté de l'analyse de cette région a résidé dans la grande quantité d'indels. Etant donné cette particularité, nous avons divisé les données en deux partitions : une matrice de séquence nucléotidique de 1120 caractères et une matrice de caractères binaires (présence/absence d'indels). Cette dernière a été codée à l'aide de l'algorithme MCIC (Modified Complex Indel Coding) en suivant les caractères parcimonieusement informatifs (n=47, figure 5b de l'article) issus de l'alignement non ambigu. Le modèle de mutation le plus approprié déterminé par une démarche de sélection hiérarchique par *Likelihood Ratio Test* (ModelTest ; cf. Chapitre III Chapitre I.IV. B.1(c) était le modèle GTR (Tavaré, 1986) pour les données nucléotidiques, F81 (Felsenstein, 1981) avec un biais de sélection variable (*variable ascertainment bias*) pour les données binaires (0 pour absence d'indel - probabilité π_0 - , 1 pour la présence - probabilité π_1 , avec π_0/π_1 le taux de transition de l'état 0 vers l'état 1).

L'inférence phylogénique a été menée selon une approche bayésienne par échantillonnage de chaînes de Markov (MrBayes ; cf. Chapitre III Chapitre I.IV.B.3). L'algorithme utilisé permet en particulier de déterminer la probabilité postérieure de chaque nœud de l'arbre phylogénique.

IV. Objectifs et principaux résultats

IV. A. Objectifs

Les *Cetartiodactyla* (Artiodactyles - ou Ongulés à nombre de doigts pairs + Cétacés - Baleines et Dauphins, cf. I. D et Chapitre I II.A.2) ont subi une radiation évolutive initiale il y a environ 70 à 80 millions d'années. Leur rapports évolutifs à la base de cette radiation demeurent difficiles à appréhender.

Les questions auxquelles répondent potentiellement les marqueurs portés par le chromosome Y sont importantes d'une part pour la compréhension fondamentale de l'évolution de la diversité ainsi que pour la mise en place de politiques de conservation des espèces et plus généralement des communautés.

C'est dans cette seconde voie que nous avons choisi de rechercher une région candidate sachant que la RFLP couramment utilisée augmente le risque de contamination ou de faux-diagnostic. Notre attention s'est donc naturellement portée sur le gène de l'amélogénine (AMEL), qui est présent sur les deux chromosomes sexuels, dont une portion est déjà utilisée chez l'Homme comme test officiel de sexage en médecine légale (**Sullivan et al., 1993**) ou encore chez les bovins (**Ennis & Gallagher, 1994**). En revanche, les polymorphismes de taille sont (67 pb chez les bovins, 6 pb chez l'Homme) et surtout, il s'agit de méthodes très taxon-spécifiques. Nous avons donc « screené » l'ensemble des séquences disponibles sur GenBank et nous avons trouvé une région présentant un grand polymorphisme de taille correspondant, sur le génome humain, à l'intron situé entre l'exon 5 et l'exon 6. Vu la présence de ce polymorphisme chez d'autres Artiodactyles, nous avons décidé de tester la présence de ce polymorphisme chez les Cétacés, groupe-frère des Artiodactyles au sein du groupe des *Cetartiodactyla*.

IV. B. Principaux résultats

On observe le même polymorphisme d'environ 500 pb entre le chromosome X et le chromosome Y chez les Cétacés et Artiodactyles disponibles à l'exception notable des Suidés. En outre, cette insertion de 500 pb sur le chromosome Y, dont l'apparition date vraisemblablement du Paléocène (-65 à -55 millions d'années), est caractérisée chez le Dauphin bleu et blanc par un polymorphisme notable.

Un article conclut cette partie méthodologique : une zone très polymorphe du gène de l'amélogénine du chromosome Y est trouvée, qui permettra de développer l'étude génétique des Cétartiodactyles.

Article n°1 : Macé, M. and B. Crouau-Roy (2008). A highly polymorphic insertion in the Y-chromosome amelogenin gene can be used for evolutionary biology, population genetics and sexing in Cetacea and Artiodactyla. *BMC Genetics* 9: 64. <http://dx.doi.org/0.1186/1471-2156-9-64>

Methodology article

Open Access

A highly polymorphic insertion in the Y-chromosome amelogenin gene can be used for evolutionary biology, population genetics and sexing in Cetacea and Artiodactyla

Matthias Macé*^{1,2} and Brigitte Crouau-Roy¹

Address: ¹UMR 5174 UPS/CNRS EDB "Evolution et Diversité biologique", Bât 4R3b2, Université Paul Sabatier, 118 route de Narbonne, 31062 TOULOUSE cedex 9, France and ²Centre de Physiopathologie de Toulouse Purpan, INSERM U563, CHU Purpan, F-31300 Toulouse, France

Email: Matthias Macé* - matthias.mace@orange.fr; Brigitte Crouau-Roy - bcrouau@cict.fr

* Corresponding author

Published: 16 October 2008

Received: 28 September 2007

BMC Genetics 2008, 9:64 doi:10.1186/1471-2156-9-64

Accepted: 16 October 2008

This article is available from: <http://www.biomedcentral.com/1471-2156/9/64>

© 2008 Macé and Crouau-Roy; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: The early radiation of the *Cetartiodactyla* is complex, and unambiguous molecular characters are needed to clarify the positions of hippotamuses, camels and pigs relative to the remaining taxa (*Cetacea* and *Ruminantia*). There is also a need for informative genealogic markers for Y-chromosome population genetics as well as a sexing method applicable to all species from this group. We therefore studied the sequence variation of a partial sequence of the evolutionary conserved amelogenin gene to assess its potential use in each of these fields.

Results and discussion: We report a large interstitial insertion in the Y amelogenin locus in most of the *Cetartiodactyla* lineages (cetaceans and ruminants). This sex-linked size polymorphism is the result of a 460–465 bp inserted element in intron 4 of the amelogenin gene of Ruminants and Cetaceans. Therefore, this polymorphism can easily be used in a sexing assay for these species.

When taking into account this shared character in addition to nucleotide sequence, gene genealogy follows sex-chromosome divergence in *Cetartiodactyla* whereas it is more congruent with zoological history when ignoring these characters. This could be related to a loss of homology between chromosomal copies given the old age of the insertion.

The 1 kbp *Amel-Y* amplified fragment is also characterized by high nucleotide diversity (64 polymorphic sites spanning over 1 kbp in seven haplotypes) which is greater than for other Y-chromosome sequence markers studied so far but less than the mitochondrial control region.

Conclusion: The gender-dependent polymorphism we have identified is relevant not only for phylogenetic inference within the *Cetartiodactyla* but also for Y-chromosome based population genetics and gender determination in cetaceans and ruminants. One single protocol can therefore be used for studies in population and evolutionary genetics, reproductive biotechnologies, and forensic science.

Background

About 240 to 320 million years ago, shortly after the divergence of mammalian and avian lineages, progressive X-Y differentiation began, following chromosomal interstitial rearrangements [1]. This resulted in a partial loss of homology between both chromosomes which reached its maximal extent in primates [2]. Amelogenin is the enamel matrix protein that combines with hydroxyapatite crystals to form enamel prisms in teeth [3]. The gene encoding the amelogenin protein (*Amel*) is among the few genes expressed from both X and Y chromosomes in placental mammals (*Eutheria*) [4].

Evolutionary uncertainties about the basal diversification of Cetartiodactyla

The *Cetartiodactyla* (even-toed ungulates, whales and dolphins) radiated approximately 70–80 Myrs ago. The relative positions of the camelid, suiform (pigs), hippopotamus, ruminant and cetacean (whales and dolphins) groups remain unclear, whether morphological or molecular characters are used for attribution [5-9]. Moreover, polytomies (unresolved tree nodes) within some *Cetartiodactyla* taxa [8] highlight areas for further data collection (both species and markers) and phylogenetic research. This is a particularly delicate problem within cetaceans, probably due to adaptative radiations within a short period of time [10,11].

Y-chromosome sequence markers are needed for population genetics

Males are the heterogametic sex in mammals, and usually, unequal numbers of males and females transmit genes from one generation to the next. Y-specific polymorphisms should allow the inference of sex-specific population parameters and decryption of breeding system patterns and dispersal strategies. Overall, the use of Y-specific markers has been restricted to evolutionary studies of human history and some scarce studies in population genetics, perhaps because of the low diversity of these markers [12]. Within *Cetartiodactyla*, genetic structure or admixture, e.g. in sheep [13] or cattle [14,15], has made use of a few Y-specific markers including microsatellites, SNPs and indels.

The matrilineally transmitted mitochondrial control region is commonly used as an informative sequence for population genetics. An equivalent had not been found to date on the Y chromosome. We considered the well-known amelogenin gene to be of particular interest because parts of it do not recombine between X and Y chromosomes.

Molecular sexing

Sex-chromosome recombination discrepancies have been exploited to develop many molecular sexing techniques.

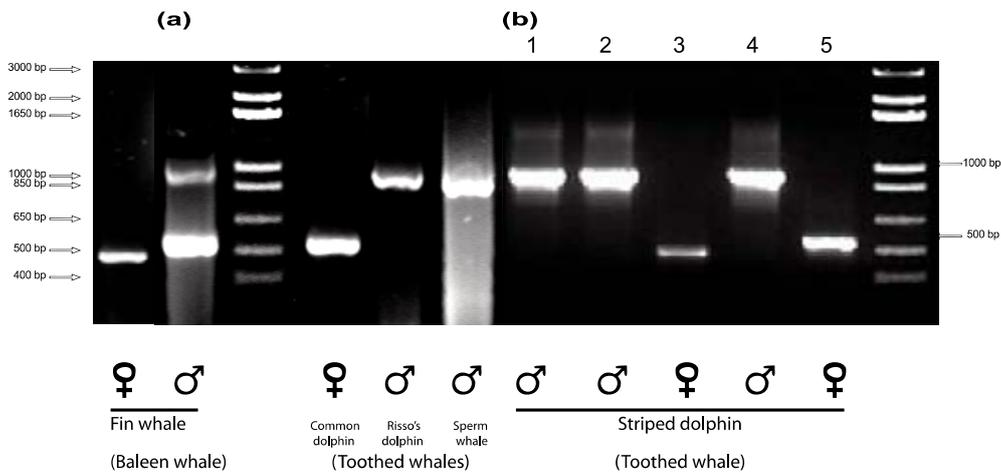
Although it can be ambiguous in some small populations, the amelogenin locus is the most commonly used for gender determination in humans [16]. Accurate gender determination in mammals is crucial to various applications in reproductive technologies, forensic investigations and population management. Some techniques rely on specific amplification of loci localized on the Y chromosome (such as Sry [17]) while others are based on amplification of homologous fragments from both X and Y chromosomes (e.g. ZFX/ZFY [18]) or use both markers [19]. Each of these has limitations, such as the need for multiplexing with other markers or additional steps such as digestion, labelling or sequencing. Several amelogenin-based techniques have expanded the taxonomic coverage of molecular sexing for *Artiodactyla* [20-22] but they have not yet been extended to *Cetacea*. Therefore, there is a need both for new methods that apply to a greater number of species and to increase the number of cross-checking sexing methods, especially in conservation biology [23].

In this study, we found out that sequence variations in the amelogenin locus can be applied in evolutionary and population genetics as well as for molecular sexing in the highly diversified *Cetartiodactyla* group. We therefore carried out an evolutionary study of orthologous *Amel-Y* and *Amel-X* sequences (exons 4 to 5) in *Cetartiodactyla*. We studied four *Cetacea* (the striped dolphin, the bottlenosed dolphin, Risso's dolphin and the fin whale) and three *Artiodactyla* (cow, pig and sheep) species.

Results and Discussion

Amelogenin can be used for molecular sexing and evolutionary genetics in Cetartiodactyla

Amplification of the studied segment of the amelogenin locus using the species-specific SC1-SC2 primers resulted in an obvious sex-related size polymorphism in all *Cetacea* (Fig. 1) with a unique 521 bp band for females (two *Amel-X* copies) and an additional 980 bp band for the *Amel-Y* in males. This pattern was obvious in male Baleen whales (Mysticetes) but there was no corresponding *Amel-X* amplification in male dolphins unless by using the primers X5-X6 derived from the human amelogenin sequence. Previous studies showed that amelogenin amplification was prone to allelic drop-out or at least to preferential amplification [24]. These phenomena may be explained by several factors. Usually, amplification of the lesser sized allele is favoured when the amount of polymerase is a limiting factor or in case of template DNA degradation [25]. Small amounts of DNA may also increase stochasticity of the annealing [26]. However, our results are not consistent with these situations since the allele favoured (*Amel-Y*) is always the greatest one. On the other hand, differences in GC content and mismatches in the annealing sequences may account for differential amplification. The amelogenin fragments that we studied are character-

**Figure 1**

Sex-related size polymorphism of amelogenin fragment in Cetacean. (Molecular weight markers is Biolabs' 1 kb + ladder): a) Agarose gel showing differences between male amplification in a Baleen (toothless) whale (left of the ladder) and Toothed whales (on the right). b) Agarose gel showing differences between males and females in Striped dolphin. 1,000 bp band for *Amel-Y*, 500 bp band for *Amel-X*. Each lane represents a single sample (#1 to 5). Symbols ♂ and ♀ are for male and female samples respectively.

ized by a higher GC content when amplified from the X chromosome (56%) than from Y-chromosome (47%). This difference may result from a non-insertion in the *Amel-X* fragment. This feature as well as a 2 bp-long mismatch between dolphin's *Amel-X* and the 5' end of the reverse primer SC2 (Fig. 2) may favour preferential amplification of the Y copy in dolphins (Fig. 1b). Indeed, amplifying male dolphin samples SC3 (primer without mismatch, see Fig. 2) instead of the SC2, restores the two bands, seen in baleen whales. The presence of this large insertion in the *Amel-Y* copy can be used for sex determination in probably all cetacean species.

In order to define the breakpoints of the Y insertion location and investigate its evolutionary history, we sequenced various cetaceans (listed in Methods; sequences deposited under the following accessions: EMBL:AM744958 to AM744971). After alignment with available sequences from *Artiodactyla* (see list in Methods), we detected the same polymorphism in all other *Cetartiodactyla* except the Pig (Fig. 3): a 460–465 bp insertion (size is a function of indels within different individuals or species) located between the 4th and 5th exons (188th

to 651st position of Y sequences e.g. EMBL:AM744958). Haplotype names and their corresponding accessions are given in Table 1. Sequence similarity was checked by running BLAST (Basic Local Alignment Search Tool) over GenBank nr/nt nucleotide collection sequences with megablast algorithm (intended for high similarity sequences). In addition to the bovine and ovine *Amel-Y*, the only two relevant (78 and 83% homology, E-values 4.10^{-68} and 3.10^{-53}) hits matched a fragment on the seventh chromosome in pigs (ca. 250 bp), suggesting the insertion might be a transposable element.

We interpret the presence of this insertion as a synapomorphy (shared character) of the *Cetartiodactyla* excluding pigs and probably other early derived groups (camels, hippopotamuses; [27], see Fig. 3). In addition to this long insertion, 46 other indels were detected by sequence alignment (positions and sizes detailed in Figure 5). Indels are particularly useful for testing phylogenetic hypotheses, as they can provide information about ancient divergences rather than population information. We therefore assessed whether phylogenetic topologies differed if we took into account or not the information

<pre> --CAAGCATGCATTTCAATTCCC----- ATCAAGCATGCATTTCAATTCCCTTTTA ATTAAGCATGCATTTCAATTCCCTTTTA ATCAAGCATGCATTTCAATTCCCTTTTA ATCAAGCATGCATTTCAATTCCCTTTTA GTGAAGCATGCATTTCAATTCCCTTTTA ATTAAGCATGCATTTCAATTCCCTTTTA ATTAGGCATGCATTTAAAATTCCCATATT </pre>	<p>Forward Primer "SC1" (5'-3')</p> <p>Dolphin AMELY Dolphin AMELX Whale AMELY Whale AMELX Cattle AMELY Cattle AMELX Man AMELX</p>
<pre> CT-----CCGATGTTCCCC--ATGCAG GCCT-----CCGATGTTCCCC--ATGCAGAATC GCCT-----CCGATGTTCCCCGCATGCAGCCCT GCCT-----CCGATGTTCCCC--ATGCAGAATC GCCT-----CCGATGTTCCCC--ATGCAGAATC GCCT-----CCGATGTTCCCC--ATGCAGAATC GCCT-----CCGATGTTCCCC--ATGCAGAATC CAGCCCCAGTCACCCATGCACCCC--ATCCAGCCCT CAGCCCCAGCCACCTGTGCACCCC--ATGCAGCCCC CT-----CCGATGTTCCCCGCATGC </pre>	<p>Reverse Primer "SC2" (3'-5')</p> <p>Dolphin AMELY Dolphin AMELX Whale AMELY Whale AMELX Cattle AMELY Cattle AMELX Man AMELX</p> <p>Reverse Primer "SC3" (3'-5')</p>

Figure 2

Sequence alignment of the oligonucleotide primers with target sequences in Cetacea, Cattle and Man. Species and chromosomal location are given on the right side. Shaded columns represents the nucleotide mutated in Dolphins. Accession numbers of sequences follow: Dolphins (EMBL:AM744958-AM744964, EMBL:AM744970-AM744971, EMBL:AM744968, EMBL:AY787743S2 - Y and EMBL:AM744965 - X) and Whales (EMBL:AM744967, EMBL:AM744969 -X- and EMBL:AM744966 - Y), Cattle (GenBank:AB091789 -X- and GenBank:AB091790 - Y) and Man (GenBank:NT_011757 -X- from 9098117 to 9098612 and GenBank:NC_000024 -Y- from 6796200 to 6796719).

contained in these indels. Thus, the cetacean sequences summarized in Table 1 as well as *Artiodactyla* sequences were analyzed first classically, with gaps coded as missing characters, and secondly, with gaps coded as supplementary binary characters (see Fig. 5). For each analysis, two independent Bayesian searches were performed. The phylogenetic trees presented in Figure 4 result from a consensus of 20,000 trees sampled after standard deviation between the two runs dropped below 0.01. They show highly supported nodes. The phylogenetic analysis performed on the complete segment (Fig. 4a) confirmed the clustering by sex-chromosome copy in *Cetartiodactyla* (*Stenella caeruleoalba*, *Balaenoptera physalus*, *Grampus griseus*, *Tursiops truncatus*, *Bos taurus* and *Ovis aries*) whereas *Amel-X* and *Amel-Y* clustered together in other mammals (*Homo sapiens*, *Sus scrofa*) together with *Amel-X* from *Cetartiodactyla*. On the other hand, phylogeny inferred without taking into account the insertion gave a different result (Fig. 4b): whereas haplotypes also clustered by chromosome in

cetaceans, no signal related either to species history or to chromosome bearing could be seen in the other *Cetartiodactyla*. Hence, the phylogenetic signal related to species history seems to strengthen as we follow the tree from *Cetartiodactyla* towards primates. This partial, homoplastic, persistence of the phylogenetic signal may be explained by the influence of the region surrounding the insertion. This could be the result of the old age of the insertion (74-87 myrs [27]). The subsequent loss of homology may have given rise to a more divergent evolution between chromosomes in some taxa (*Cetacea*) than in others (*Artiodactyla*).

It would be interesting to study this region at the whole clade level by combining sequence and indel characters in the same analysis. This could give clues to test the many hypotheses about basal radiation of *Cetartiodactyla* (e.g. [5,6,8]). Given the presumably basal position of the *Suoidae* and *Tylopoda* in the *Cetartiodactyla* phylogeny ([7] and

Table 1: List of Amel-X and Amel-Y haplotype names in Cetaceans and their EMBL accession numbers

Haplotype Name	EMBL Accession
<i>Stenella caeruleoalba</i> YA1	AM744963
<i>Stenella caeruleoalba</i> YA2	AM744964
<i>Stenella caeruleoalba</i> YB1	AM744958
<i>Stenella caeruleoalba</i> YB2	AM744959
<i>Stenella caeruleoalba</i> YB3	AM744960
<i>Stenella caeruleoalba</i> YB4	AM744961
<i>Stenella caeruleoalba</i> YB5	AM744962
<i>Delphinus delphis</i> Y1	AM744970
<i>Delphinus delphis</i> Y2	AM744971
<i>Stenella caeruleoalba</i> X	AM744965
<i>Grampus griseus</i> Y	AM744968
<i>Balaenoptera physalus</i> Y	AM744966
<i>Balaenoptera physalus</i> X	AM744967
<i>Eschrichtius robustus</i> X	AM744969
<i>Tursiops truncatus</i> X	AY787743S2

Fig. 3), we hypothesize that the major evolutionary event represented by the insertion (illustrated by an arrow Figure 4a) occurred once in the *Cetacea-Ruminantia* clade and not in the remaining *Cetartiodactyla*.

The presence of this large insertion in the *Amel-Y* copy can be useful for sex determination.

Evolutionary history also indicates that our sexing technique is applicable, in addition to cetaceans, to over a wide range of *Cetartiodactyla* species including domestic and wild species, in particular the widespread *Ruminantia* (*Bovidae*, *Capridae* and most likely *Cervidae*). It is however not suitable to Suiformes and further studies are required to confirm that the technique is also not applicable to Camelidae, given their even more basal position in the *Cetartiodactyla* phylogeny.

Use in pedigree assessment and population genetics

In dolphins, the *Amel-Y* fragments amplified with the SC1-SC2 primer pair were easily sequenced without the

need of cloning since amplification was Y chromosome-specific. From the ten Striped dolphin samples sequenced, nine were males, and we could deduce seven distinct Y-haplotypes (one haplotype represented by three individuals and four individual haplotypes) bearing 64 polymorphic sites (nucleotide diversity $\pi = 0.004 \pm 0.0007$). Half of these were in the ~460 bp insertion. An alignment of polymorphic sites is presented in Figure 5 (a). Strikingly, these sequences showed two highly divergent haplogroups, diverging by a mean of 49 substitutions. This concurs with our results that support the probable existence of two subspecies within the Mediterranean sea (unpublished data). Moreover, one of these haplogroups displayed a high degree of polymorphism, with 24 informative sites, whereas the others showed only eight. These values are sufficient for use in pedigree analysis and population genetics, as the Y chromosome counterpart of the mitochondrial d-loop in this species. Indeed, in striped dolphin the intra-specific (inter-group) divergence is greater than inter-specific divergence with a mean of 45 nucleotide substitutions between the striped dolphin and fin whale sequences. There is an average of 0.048 ± 0.01 substitutions per site when comparing the two striped dolphin populations. This is comparable to the divergence observed between each population and the Common dolphin (0.058 ± 0.03) and confirms that nucleotide diversity is one order of magnitude higher than the range observed (10^{-4}) for Y chromosome markers in mammals [12]. As for the mitochondrial d-loop, the size of the amplified fragment slightly limits the use of the technique. Some degraded samples do not amplify; even so, a particularly degraded sperm whale sample was still amplifiable (data not shown).

Since the Y chromosome population is expected to have a small effective size, it is more likely to be affected by genetic drift. Thus, it reflects more recent demographic events such as bottlenecks, expansions or founder effects [28]. To study this sort of event, one needs a marker whose diversity is high enough to allow the reconstruction of gene genealogies with the least ambiguities and in regions where recombination does not interfere with the uniqueness of the trees. For this purpose, highly variable microsatellites represent valuable markers but they require intensive computing methods to take into account uncertainties in the trees arising from alleles that are identical by state and not by descent (homoplasies) [29,30]. Adding a new sequence marker is therefore of interest for Y-chromosome population genetics in *Cetartiodactyla*. Moreover, the Bayesian estimate of mutation rate on each edge of both trees in Fig. 4, jointly computed with phylogenetic inference, shows high values for a marker of nuclear DNA: between 10^{-8} and 10^{-10} substitutions per site and per year in all *Cetartiodactyla* branches. This value is

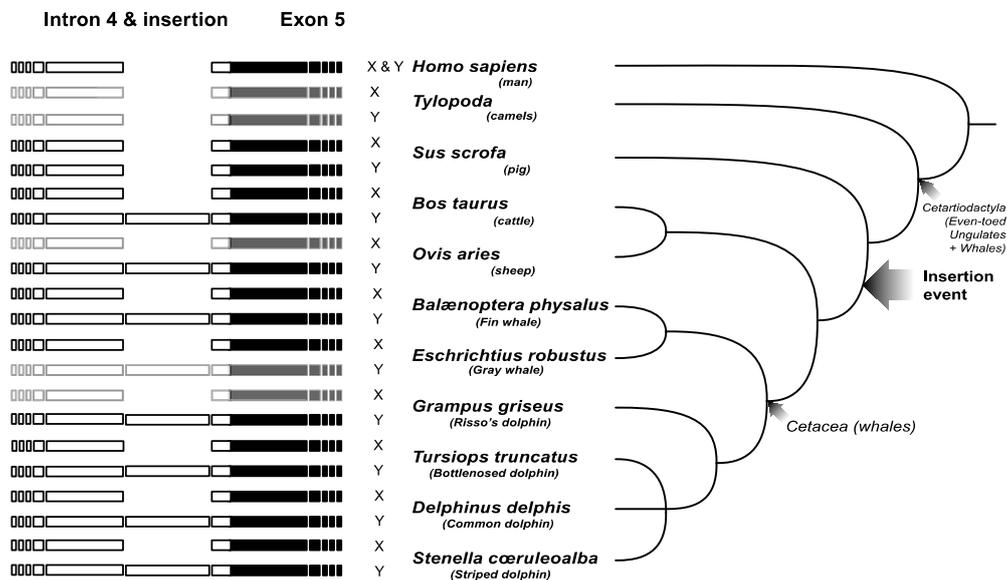


Figure 3
Schematic representation of the sex-related polymorphism of the amelogenin locus in an evolutionary perspective. Insertion and intron 4 are represented by a white bar, whereas exon 5 is in black. Shaded bars stands for absent data, deduced from evolutionary relationships. The vertical order links to the "tree of life" view (according [27] among others) provided on the right.

intermediate between those of mitochondrial d-loop and nuclear DNA in mammals [31,32].

Functional perspectives in amelogenin evolution

We found two stop codons at amino acid positions 98 and 99 of exon 5 in all Y chromosome copies of amelogenin in the four studied cetacean species (positions 988–993 of sequence EMBL:AM744959). The *Amel-Y* gene product may therefore be truncated in these species or represent a pseudogene as already observed in species from most of the other eutherian clades [33].

Conclusion

The 460 bp insertion studied represents a single-event synapomorphy among most *Cetartiodactyla*. Together with the presence of other numerous indels informative at the order-level, it could help resolve the phylogenetic discrepancies between hippopotamuses, pigs, camels and other *Cetartiodactyla* observed by many authors [7–9]. In addition, we demonstrate higher diversity within a single sequence than has yet been observed in multi-sequence assays [34]. This high diversity should allow the use of this sequence as the male counterpart of the mitochondrial

control region. The applications would include inference of male-driven evolution in population genetics as applied to breeding management; domestication studies in archaeogenetics [35]; conservation biology (population history, sex-biased dispersal, admixture); or for testing sex-biased selection [28]. Amelogenin intron 4 amplification will also be an efficient tool for sexing *Ruminantia* and *Cetacea*. This will be useful for many fields of veterinary and forensic science (embryo technologies, *in vitro* fertilization, meat products). Finally, amelogenin amplification could also be a helpful tool for conservation biology through sampling of dead animals, faecal remains and biopsies of free-ranging animals like whales and dolphins. Amelogenin amplification in *Cetartiodactyla* therefore is a simple, single-step procedure with a wide range of applications.

Methods

Laboratory procedures

Biological material was isolated from soft tissues sampled from dead stranded cetaceans and extracted using a classical phenol-chloroform protocol [36]. We used heterologous primers, X5 (5'-GTGCTTACCCCTTGAAGTG-3')

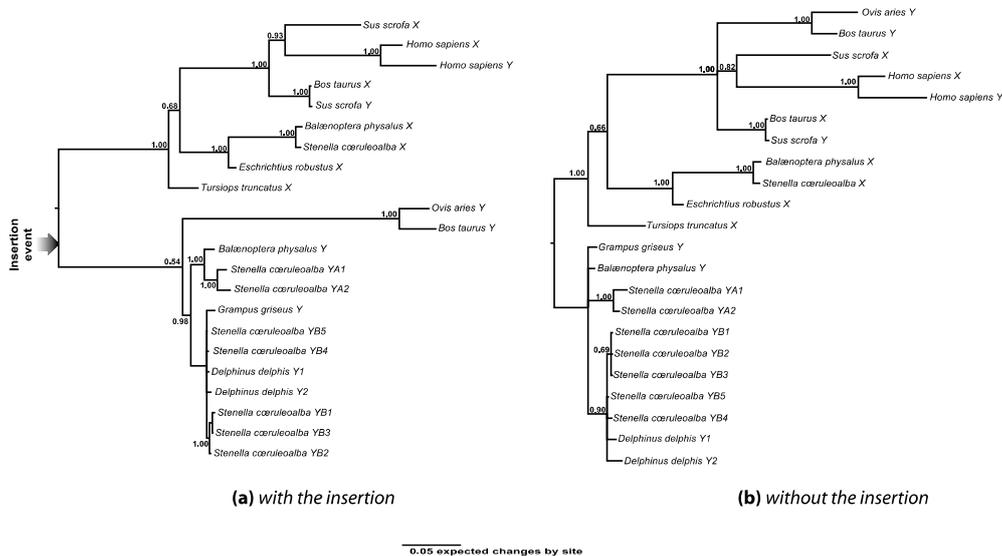


Figure 4
Comparison of phylogenetic trees of the Amel-X and Amel-Y fragments inferred (a) with the insertion (b) without the insertion. (a) The phylogenetic tree of the complete fragment shows trans-specific clustering by sex chromosome in *Cetartiodactyla*. Tip labels are haplotypes as deposited in the EMBL database; Y and X are for Amel-Y and Amel-X haplotypes respectively. *Stenella caeruleoalba* haplotypes were named according to population origin (YA/Group 1, YB/Group 2, see Methods). (b) The inferred phylogeny after removing the insertion gives a slightly different picture: trans-specific clustering by sex-chromosome is lost except in Cetaceans.

and X6 (5'-CITCCTCCCGCTTGGTCITG-3'), designed from the amelogenin intron 4 and exon 5 of *Homo sapiens* X chromosome (GenBank:NC_000023, chrX:11221454-11228802) reference assembly Build 36.3, to amplify the homologous region in cetacean amelogenin. This region of Amel-X is 92% identical to Amel-Y.

We subsequently cloned and sequenced these PCR fragments and designed oligonucleotide primers specific to the *Artiodactyla*. They are anchored in exons 4 and 5 of Amel-X and Amel-Y, which allows complete amplification of the 4th intron (SC1: 5'-CAAGCATGCATTCAATCC-3' and SC2: 5'-CTGCATGGGGAACATCGGAG-3'). Optimal PCR conditions were adjusted by using temperature (49-62°C) and MgCl₂ (1-2.5 mM) gradients. Following this optimization step, the PCR amplifications were conducted in a reaction mix consisting of 477 mM KCl, 1 mM Tris/HCl pH 8.3, 1.5 mM MgCl₂, 250 μM each dNTP, 2 pmol/μl of each primer, 3 units of Taq polymerase in a final volume of 25 μl. Cycling was conducted as follows: 95°C for 2 min, followed by 30 cycles of denaturation at 94°C for 1 min, annealing at 55°C for 45 s, extension at

72°C for 1 min, and a final extension at 70°C for 10 min. PCR products were run on 1.2% agarose gel ethidium bromide stained, alongside a 1 kbp ladder (New England Biolabs, County Road, MA). In order to validate the assay, the gender, if identified during examination of the stranded carcasses, was recorded (22 out of 38 samples). These amplifications were conducted in eight Cetacean species: 20 Striped dolphin (*Stenella caeruleoalba*), five Fin whales (*Balaenoptera physalus*), four Bottlenosed dolphins (*Tursiops truncatus*), three Common dolphins (*Delphinus delphis*), three Gray whales (*Eschrichtius robustus*), two Sperm whales (*Physeter macrocephalus*), one Minke whale (*Balaenoptera acutorostrata*) and one Risso's dolphin (*Grampus griseus*).

Of these, we sequenced striped dolphins (9 males and one female), fin whales (two males and two females), common dolphin (2 males), gray whale (one female) and Risso's dolphin (one female). Sequencing was performed on an ABI Prism sequencer (Applied Biosystems, Foster City, CA) with the dye terminator protocol directly for fragments amplified using the same (SC1-SC2) primer

BC was responsible for funding, supervision of the research project and manuscript writing.

Acknowledgements

Many thanks go to H. Etchevers, N. Lambert and P. Monget for their much valuable comments on the manuscript as well as to P. Sudour and J. Butterworth.

We also wish to thank B. Jacobsen from the UAF Museum and F. Dhermain from the GECEM who kindly provided a significant part of the samples as well as A.-M. Gasc for technical help.

We are indebted to three anonymous reviewers for valuable comments that significantly helped to improve manuscript.

References

- Lahn BT, Page DC: **Four evolutionary strata on the human X chromosome.** *Science* 1999, **286**(5441):964-967.
- Lahn BT, Pearson NM, Jeggalian K: **The human Y chromosome, in the light of evolution.** *Nat Rev Genet* 2001, **2**(3):207-216.
- Satchell PG, Anderton X, Ryu OH, Luan XH, Ortega AJ, Opamen R, Berman BJ, Witherspoon DE, Gutmann JL, Yamane A, et al.: **Conservation and variation in enamel protein distribution during vertebrate tooth development.** *J Exp Zool (Mol Dev Evol)* 2002, **294**(2):91-106.
- Delgado S, Giron-dot M, Sire JY: **Molecular evolution of amelogenin in mammals.** *J Mol Evol* 2005, **60**(1):12-30.
- Boisserie JR, Lihoreau F, Brunet M: **The position of hippopotamidae within cetartiodactyla.** *Proc Natl Acad Sci USA* 2005, **102**(5):1537-1541.
- Gatesy J, O'Leary M: **Deciphering whale origins with molecules and fossils.** *Trends Ecol Evol* 2001, **16**(10):562-570.
- Nikaido M, Rooney AP, Okada N: **Phylogenetic relationships among cetartiodactyls based on insertions of short and long interspersed elements: hippopotamuses are the closest extant relatives of whales.** *Proc Natl Acad Sci USA* 1999, **96**(18):10261-10266.
- Price SA, Bininda Emonds ORP, Gittleman AL: **A complete phylogeny of the whales, dolphins and even-toed hoofed mammals (Cetartiodactyla).** *Biol Rev* 2005, **80**(3):445-473.
- Thewissen JG, Williams EM, Roe LJ, Hussain ST: **Skeletons of terrestrial cetaceans and the relationship of whales to artiodactyls.** *Nature* 2001, **413**(6853):277-281.
- Leduc RG, Perrin WF, Dizon AE: **Phylogenetic relationships among the delphinid cetaceans based on full cytochrome B sequences.** *Mar Mamm Sci* 1999, **15**(3):619-648.
- Milinkovitch MC: **Molecular phylogeny of cetaceans prompts revision of morphological transformations.** *Trends Ecol Evol* 1995, **10**(8):328-334.
- Hellborg L, Ellegren H: **Low levels of nucleotide diversity in mammalian Y chromosomes.** *Mol Biol Evol* 2004, **21**(1):158-163.
- Meadows JR, Hanotte O, Drogemuller C, Calvo J, Godfrey R, Colman D, Maddox JF, Marzanov N, Kantanen J, Kijas JW: **Globally dispersed Y chromosomal haplotypes in wild and domestic sheep.** *Anim Genet* 2006, **37**(5):444-453.
- Anderung C, Hellborg L, Seddon J, Hanotte O, Gotherstrom A: **Investigation of X- and Y-specific single nucleotide polymorphisms in taurine (*Bos taurus*) and indicine (*Bos indicus*) cattle.** *Anim Genet* 2007, **38**(6):595-600.
- Edwards CJ, Gaillard C, Bradley DG, MacHugh DE: **Y-specific microsatellite polymorphisms in a range of bovid species.** *Anim Genet* 2000, **31**(2):127-130.
- Brinkmann B: **Is the amelogenin sex test valid?** *Int J Legal Med* 2002, **116**(2):63.
- Bryja J, Konečný A: **Fast sex identification in wild mammals using PCR amplification of the Sry gene.** *Folia Zool* 2003, **52**(3):269-274 [http://www.ivb.cz/folia/52/3/269-274.pdf].
- Shaw CN, Wilson PJ, White BN: **A reliable molecular method of gender determination for mammals.** *J Mammal* 2003, **84**(1):123-128.
- Pomp D, Good BA, Geisert RD, Corbin CJ, Conley AJ: **Sex identification in mammals with polymerase chain reaction and its use to examine sex effects on diameter of day-10 or -11 pig embryos.** *J Anim Sci* 1995, **73**(5):1408-1415.
- Ennis S, Gallagher TF: **A PCR-based sex-determination assay in cattle based on the bovine amelogenin locus.** *Anim Genet* 1994, **25**(6):425-427.
- Pfeiffer I, Brenig B: **X- and Y-chromosome specific variants of the amelogenin gene allow sex determination in sheep (*Ovis aries*) and European red deer (*Cervus elaphus*).** *BMC Genet* 2005, **6**(1):16.
- Weikard R, Pitra C, Kuhn C: **Amelogenin cross-amplification in the family Bovidae and its application for sex determination.** *Mol Reprod Dev* 2006, **73**(10):1333-1337.
- Robertson BC, Gemmill NJ: **PCR-based sexing in conservation biology: Wrong answers from an accurate methodology?** *Conserv Genet* 2006, **7**(2):267-271.
- Findlay I, Quirke P: **Fluorescent polymerase chain reaction: Part I. A new method allowing genetic diagnosis and DNA fingerprinting of single cells.** *Hum Reprod Update* 1996, **2**(2):137-152.
- Walsh PS, Erlich HA, Higuchi R: **Preferential PCR amplification of alleles: mechanisms and solutions.** *PCR Methods Appl* 1992, **1**(4):241-250.
- Findlay I, Ray P, Quirke P, Rutherford A, Lilford R: **Allelic drop-out and preferential amplification in single cells and human blastomeres: implications for preimplantation diagnosis of sex and cystic fibrosis.** *Hum Reprod* 1995, **10**(6):1609-1618.
- Bininda-Emonds OR, Cardillo M, Jones KE, MacPhee RD, Beck RM, Grenyer R, Price SA, Vos RA, Gittleman JL, Purvis A: **The delayed rise of present-day mammals.** *Nature* 2007, **446**(7135):507-512.
- Petit E, Balloux F, Excoffier L: **Mammalian population genetics: why not Y?** *Trends Ecol Evol* 2002, **17**(1):28-33.
- Beerli P, Felsenstein J: **Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach.** *Proc Natl Acad Sci USA* 2001, **98**(8):4563-4568.
- Wilson IJ, Balding DJ: **Genealogical inference from microsatellite data.** *Genetics* 1998, **150**(1):499-510.
- Ayala FJ: **Vagaries of the molecular clock.** *Proc Natl Acad Sci USA* 1997, **94**(15):7776-7783.
- Moritz C, Dowling TE, Brown WM: **Evolution of Animal Mitochondrial DNA: Relevance for Population Biology and Systematics.** *Annu Rev Ecol Syst* 1987, **18**:269-292.
- Sire JY, Delgado S, Fromentin D, Giron-dot M: **Amelogenin: lessons from evolution.** *Arch Oral Biol* 2005, **50**(2):205-212.
- Hatch LT, Dopman EB, Harrison RG: **Phylogenetic relationships among the baleen whales based on maternally and paternally inherited characters.** *Mol Phylogenet Evol* 2006, **41**(1):12-27.
- Zeder MA, Emshwiller E, Smith BD, Bradley DG: **Documenting domestication: the intersection of genetics and archaeology.** *Trends Genet* 2006, **22**(3):139-155.
- Sambrook J, Russell DW: **Molecular Cloning: A Laboratory Manual.** 3rd edition. Cold Spring Harbor: Cold Spring Harbor Laboratory Press; 2001.
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG: **The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools.** *Nucleic Acids Res* 1997, **25**(24):4876-4882.
- Ronquist F, Huelsenbeck JP: **MrBayes 3: Bayesian phylogenetic inference under mixed models.** *Bioinformatics* 2003, **19**(12):1572-1574.
- Müller K: **Incorporating information from length-mutational events into phylogenetic analysis.** *Mol Phylogenet Evol* 2006, **38**(3):667-676.
- Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R: **DnaSP, DNA polymorphism analyses by the coalescent and other methods.** *Bioinformatics* 2003, **19**(18):2496-2497.
- Nei M: **Molecular evolutionary genetics.** United States: Columbia University Press, New York, NY; 1987.

V. *Perspectives*

Outre l'aspect méthodologique du sexage et l'éventuelle utilisation comme marqueur de séquence en génétique des populations, la présence de cette insertion polymorphe est intéressante du strict point de vue évolutif pour deux raisons :

- d'une part elle va permettre de préciser les nœuds de divergence entre les diverses branches « basales » des *Cetartiodactyla* comme celles des *Suoidae* et *Tylopoda*. En effet, leur absence chez le Porc argue en faveur d'une divergence de ce groupe des Ruminants étonnamment plus précoce que celle des Cétacés ;

- la persistance de ce polymorphisme permet de donner une borne « supérieure » à la date de la perte d'homologie entre les deux chromosomes sexuels de cette zone située à la frontière d'une région pseudo-autosomale (**Graves, 1996; Iwase et al., 2003**). Cette perte d'homologie pourrait alors expliquer le degré de polymorphisme de l'insertion à l'instar des nombreux réarrangements de séquence observés dans les copies Y au moins chez les Primates, Artiodactyles et Equidés.

Un petit « bémol » à cette technique est qu'elle semble être sensible à la qualité de l'ADN puisque nous n'avons été en mesure d'amplifier qu'à partir des échantillons les plus « frais », ce qui explique le faible nombre de séquences (13 séquences - 10/50 Dauphins bleu et blanc, 2/8 Rorqual commun, 1/1 Dauphin de Risso - l'échantillon le plus frais puisque prélevé immédiatement après la mort).

***Chapitre III : Étude de la Variabilité Génétique du
Dauphin Bleu et Blanc en Méditerranée***

I. Introduction

I. A. Quelques observations naturalistes

Des observations anatomiques effectuées sur des animaux échoués entre 1994 et 2000 nous ont amené à émettre des doutes quant à l'unicité de la population de Dauphin bleu et blanc de Méditerranée. En particulier, nous avons remarqué, sur un petit nombre d'individus toutefois, que le nombre de vertèbres cervicales soudées qui est classiquement de 2 chez cette espèce, semblait suivre une distribution biphasique, un certain nombre d'individus adultes possédant deux vertèbres soudées (Atlas+Axis - C1 et C2) alors que d'autres possédaient au moins 5 vertèbres soudées (C1 à C5). Par ailleurs, nous avons aussi observé des différences de coloration des animaux, les uns présentant un fort contraste entre zones colorées et zones claires, les autres présentant un patron de coloration moins contrasté, la zone abdominale présentant en outre des reflets rosés. Ces observations semblaient convergentes avec celles de naturalistes et scientifiques (P. Beaubrun, F. Dhermain,...) travaillant sur le terrain (animaux vivants, en pleine mer). Aucune étude n'a cependant été publiée sur ce sujet.

I. B. Travaux précédents concernant la morphologie et le développement

Des différences significatives de la taille des Dauphins bleu et blanc adultes ont été mises en évidence entre la Méditerranée et l'Atlantique (Di Méglia *et al.*, 1996). De même, des variations géographiques ont été observées au sein de la Méditerranée (Calzada & Aguilar, 1995). Enfin, une asymétrie fluctuante a été observée chez des animaux morts lors de l'épizootie à Morbillivirus (Pertoldi *et al.*, 2000). Cependant, dans sa monographie sur l'espèce dans *Mammalian Species*, l'auteur indique : (Archer II & Perrin, 1999b). Des doutes subsistent donc quant à l'homogénéité morphologique de cette espèce.

I. C. Travaux précédents de génétique des populations

C.1) Inbreeding et épizootie

Certains auteurs (Valsecchi *et al.*, 2004a) ont tenté de mettre l'*inbreeding* en relation avec l'épisode à D.M.V. utilisent un indice de parenté « interne » (*Internal Relatedness*, IR) calculé par la formule suivante :

$$IR = \frac{(2H - \sum f_i)}{(2N - \sum f_i)}$$

où H est le nombre de *loci* homozygotes, N le nombre de *loci* et f_i la fréquence du i^e allèle contenu dans le génotype. Ils ont trouvé que l'IR mesuré sur un échantillon d'animaux échoués était significativement supérieur à ce que l'on pouvait attendre d'un échantillon aléatoire. Leur explication était soit que les animaux d'IR élevé étaient plus susceptibles à des affections, soit que leur comportement en « phase terminale » était modifié.

Localisation des échantillons	Assigné à :						Total
	GV	LS	NT	NI	SI	SA	
GV, Golfe de València	10	6	5		2	5	28
LS, Mer Ligure	3	16	3	2	5	1	30
NT, Nord Thyrrénienne	2	4	9			5	20
NI, Nord Ionienne		3	1		1		5
SI, Sud Ionienne	2	3	3				8
SA, Sud Adriatique		1	4			2	7

Tableau 8. Niveaux d'assignation des génotypes individuels sur un échantillon aléatoire en Méditerranée (Valsecchi *et al.*, 2004a).

Comme on peut le remarquer sur le tableau ci-dessus, non seulement on ne peut assigner des individus issus d'un échantillon aléatoire méditerranéen qu'à concurrence de 50% pour une zone (LS), mais en sus, l'assignation semble moins efficace dans le bassin oriental (0 à 28%) et certainement due au hasard. Ils attribuent le faible niveau d'assignation à l'insuffisance d'échantillonnage. Par ailleurs (Valsecchi *et al.*, 2004a) attribuent à la différence d'IR entre bassin occidental et bassin oriental à la progression géographique de l'épisode à DMV qu'à des différences géographiques intrinsèques. Or les individus collectés étaient groupés en fonction de la période (« pendant » l'épisode et « après » l'épisode) sans diagnostic de certitude d'infection à DMV.

Ils n'observent pas de subdivision de l'échantillon méditerranéen en fonction de l'origine géographique.

C.2) Structure génétique inter-océanique

Une étude des profils de restriction d'ADN mitochondrial (combinaison de bandes issues du découpage par des enzymes) d'échantillons de Dauphins bleu et blanc a révélé en Méditerranée l'existence exclusive d'haplotypes « privés », suggérant l'existence d'une population endémique à la Méditerranée, avec un flux de gènes très limité avec l'Atlantique (**García-Martínez *et al.*, 1999**). Auparavant, les mêmes auteurs avaient observé une grande diversité haplotypique (0,789) ainsi que des divergences de séquences comprises entre 0,148% et 0,919% sur un échantillon en provenance des eaux ibériques (**García-Martínez *et al.*, 1995**). Ces valeurs sont faibles mais à ramener à la totalité du génome mitochondrial.

I. D. Etudes préliminaires

D.1) Polymorphisme trans-spécifique chez les petits Delphinidæ

Nous avons encadré une étudiante en Master de l'Ecole Nationale Vétérinaire de Toulouse (Myriam Guillochon) sur un mini-projet concernant l'inférence des relations phylogénétiques d'espèces de petits *Delphinidæ*.

Ce travail a consisté en la collecte dans GenBank de séquences homologues de région de contrôle de l'ADN mitochondrial chez les genres *Tursiops*, *Stenella*, *Delphinus* et *Lagenorhynchus*. L'analyse a ensuite été effectuée par reconstruction phylogénique au moyen d'un algorithme de maximum de vraisemblance implémenté dans une version parallélisée de *DNAML* (**Felsenstein, 1981; Felsenstein, 1989**). Le résultat de cette analyse montre l'existence d'un polymorphisme trans-spécifique : les *taxa* étudiés ne se regroupent pas par genre ou espèce mais plutôt par grappes où l'on retrouve des individus de différents genres, particulier le genre *Tursiops* qui se retrouve systématiquement associé à d'autres genres (*cf.* **Figure 34**).

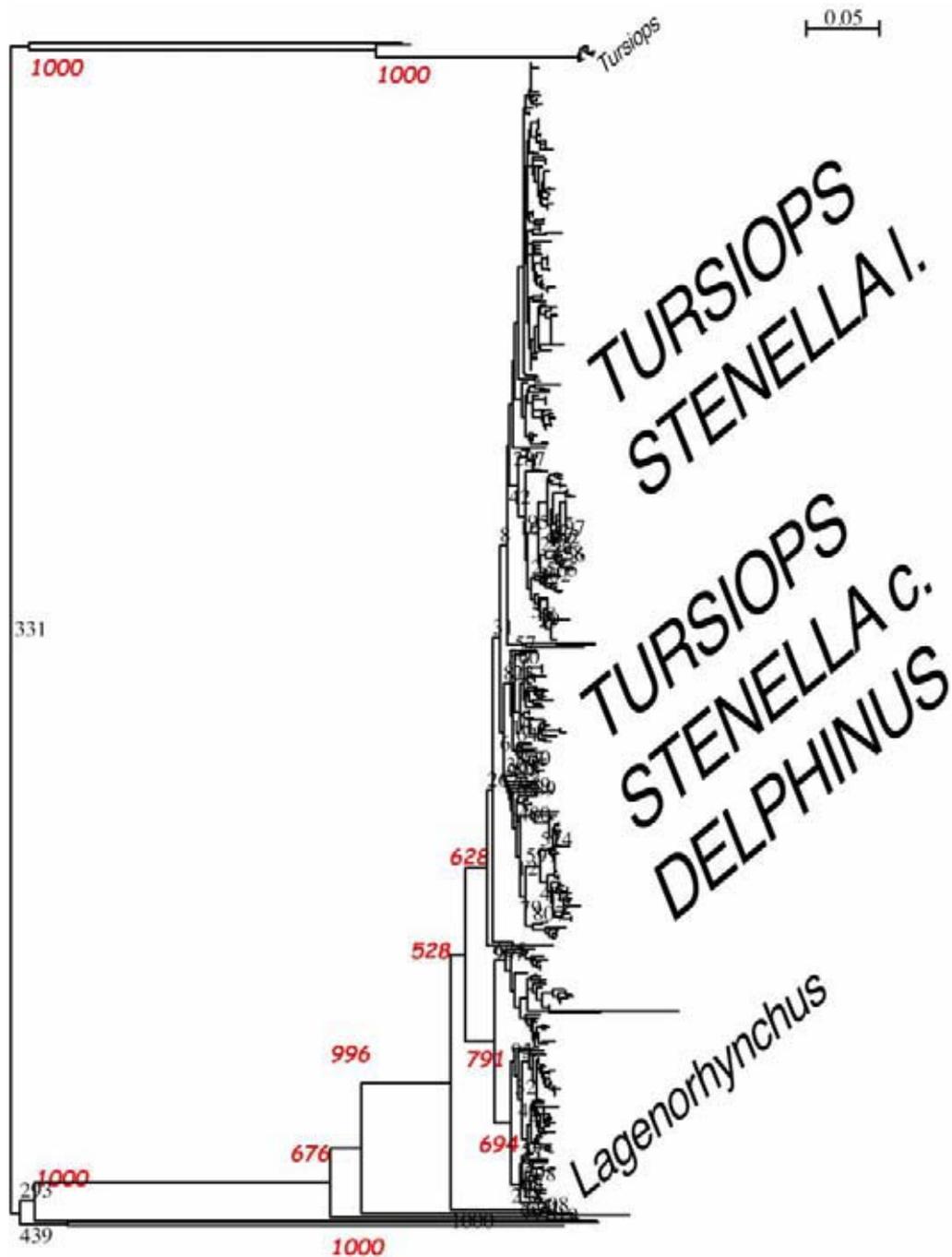


Figure 34. Analyse phylogénétique des séquences d'ADN mitochondrial montrant un polymorphisme trans-spécifique chez les petits *Delphinidae*.

D.2) Recherche de microsatellites utilisables chez les Cétacés et le Dauphin bleu et blanc en particulier

Nous avons aussi encadré un étudiant (Vincent Bourret) en thèse de doctorat

vétérinaire soutenue en 2006 sur la caractérisation de microsatellites chez le Dauphin bleu et blanc.

Un travail préliminaire a consisté en une synthèse bibliographique de l'ensemble des marqueurs microsatellites amplifiables chez les Cétacés ainsi que la caractérisation moléculaire (séquence) de certains d'entre eux destinés à être utilisés chez le Dauphin bleu et blanc.

Ce travail a fait l'objet d'une publication dans *The Open Marine Biology Journal*.

Article n°2 : Bourret V, Macé M, Bonhomme M, Crouau-Roy B (2008) Microsatellites in Cetaceans: An Overview. *The Open Marine Biology Journal* 2, 38-42. <http://dx.doi.org/10.2174/1874450800802010038>. La liste des microsatellites, des espèces cibles et des conditions de PCR est disponible en tant que *Supplementary Online Material* à l'adresse : <http://www.bentham-open.org/pages/content.php?TOMBJ/2008/00000002/00000001/38TOMBJ.SGM>

Microsatellites in Cetaceans: An Overview

Vincent Bourret, Matthias Macé, Maxime Bonhomme and Brigitte Crouau-Roy*

Laboratoire Evolution & Diversité Biologique Bâtiment 4R3 Université Paul Sabatier 118, Route de Narbonne 31062 Toulouse Cedex 9, France

Abstract: This paper presents a comprehensive overview of the published literature on microsatellites studied in cetaceans from 1989 to 2007. We inventoried 246 loci isolated from 18 and amplified in 51 cetacean species representing 11 families. The majority of loci (68%) were dinucleotide (CA) repeats, which were also shown to be on average more variable than tetranucleotide repeats. For each of these loci and the 1610 locus/species combinations, we present the species in which the markers were isolated and tested, specific diversity parameters (number of alleles, expected and observed heterozygosities), together with the primer sequences, the size range of the PCR products and the GenBank accession number. Simple analyses were performed on the assembled data and the widespread use of cross-species amplification, an important source of microsatellites in cetaceans, is discussed using laboratory data from *Stenella caeruleoalba*. No significant ascertainment bias was detected when considering all dinucleotide or tetranucleotide loci. This comprehensive database should help to inform those working on population and conservation genetic studies in most cetacean species.

Keywords: Cetacea, microsatellite, genetic diversity, conservation genetics, cross-species amplification, ascertainment bias, *Stenella caeruleoalba*.

INTRODUCTION

Several cetacean populations and species have been severely depleted during the last century and some are today considered at risk of extinction [1]. Until the mid-late 20th century, intensive commercial whaling was a major factor in the collapse of a number of stocks worldwide. At present, the whaling effort has been markedly reduced, but anthropogenic degradation of cetacean habitats is increasing, causing effects that are difficult to quantify. Such changes include the over-exploitation of marine resources, chemical and acoustic pollution, and heavy maritime traffic (see [2] for a review). The conservation of species or stocks living at the edge of extinction has raised international concern for several decades; however, consideration should also be given to the management of threatened populations not yet critically endangered [3].

Evaluating the structure and diversity of animal populations in conservation biology today involves the use of accurate genetic tools that support the planning of management schemes. Microsatellites are short, tandemly repeated DNA sequences (e.g. CACACACA) that are widely dispersed throughout the eukaryotic nuclear genome. Due to their high polymorphism and codominance, they are considered cornerstone molecular markers to address wildlife conservation issues [4]. They are among the most variable genetic markers available today and hence are particularly suitable to detect low levels of genetic diversity that are often found in vulnerable, declining or isolated populations [3]. They also allow investigations into the demographic history of populations and potentially correlated anthropogenic factors [5].

Furthermore, these markers are essential for forensic investigations, allowing the reliable identification of individuals [6-8] and estimation of the numbers of individuals entering illegal or unreported trades [9]. In this context, an increasing number of population structure inference tools allow the assignment of individuals to populations [10-12]. Finally, microsatellites are helpful for elucidating breeding strategies and social relationships which are characterized by complex behaviors in cetaceans [13], and for the fine-scale description of kinship that is of great interest in conservation biology. From a practical standpoint, microsatellite genotyping can be readily achieved from tiny quantities of even poor quality DNA. This allows conservative sampling of living as well as dead individuals, such as stranded, by-caught, harvested, or museum cetaceans [14, 15].

Specific microsatellites can be isolated by screening a species' genomic library with any repetitive probe (e.g. (CA)_n); however, this method is both costly and time-consuming. An alternative approach is to test loci that have been previously characterized in related species. Generally, the more closely related the species, the more successful the cross-amplification [16, 17]. However, Rico *et al.* [18] reported successful cross-amplifications between different fish Superclasses believed to have diverged some 470 million years (My) ago, suggesting variable conservation patterns among different taxonomic groups.

This paper presents the microsatellite markers isolated from cetacean genomic libraries from 1989 to 2007, and their polymorphism across a wide range of cetacean species. An extensive search of the published literature was carried out to create a synthetic database that can be used for population or conservation genetic studies in cetaceans.

*Address correspondence to this author at the Laboratoire Evolution & Diversité Biologique Bâtiment 4R3 Université Paul Sabatier 118, Route de Narbonne 31062 Toulouse Cedex 9, France; Tel: +33 (0) 561 557 328; Fax: +33 (0) 561 556 259; Email: bcrouau@cict.fr

METHODS

Database assembly. Ninety papers published between 1989 and 2007 were assembled, reporting either the characterization of new microsatellite loci or their use in cetacean species through specific or cross-species amplifications. From this material, a database was constructed including 246 microsatellites with the following information for each locus: i) name, repeat type, GenBank accession number, and sequence of each primer; ii) species in which the locus was tested; iii) specific genetic diversity parameters (number of individuals tested, number of alleles, expected (H_e) and observed (H_o) heterozygosities) and size range of the PCR products, whenever available. This database consists of four tables, one for each type of repeat: di-, tri-, tetranucleotide and complex repeats (Tables A-D, Supplementary Material).

Statistical analyses. We investigated which type of locus (di-, tri-, tetranucleotide, complex) is the most variable, and whether the variability is greater in species where the loci were isolated as opposed to species where the loci were cross-amplified (ascertainment bias). For this purpose, permutation tests were performed, which are suitable for analyzing the assembled database since they make no particular assumption of the distribution of the data or the characteristics of the datasets to be compared. The aim of these permutation tests is to test for a significant difference between means. The permutation procedure first involves calculating the difference between the observed means of two vectors of the H_e values (d_{obs}). These two vectors are then concatenated to create a new vector. Two vectors of the same respective lengths as the original two vectors to be compared are randomly sampled into this new vector. The difference between the estimated means of the two new vectors of H_e values is recorded (d_{est}). The procedure is repeated 10,000 times. The p-value of the permutation test is estimated as the proportion d_{est} superior to d_{obs} for positive d_{obs} , and the proportion d_{est} inferior to d_{obs} for negative d_{obs} . This permutation procedure was performed using R statistical package (version 2.0.1).

RESULTS & DISCUSSION

Microsatellites in cetacean species. Microsatellite data are reported from 51 cetacean species encompassing 11 families. The 244 species-specific loci have been isolated from genomic libraries in four mysticete and 14 odontocete species from 1989 to 2007 (Table I and Tables A-D). These 18 species represent six families (Balænidæ, Balænopteriðæ, Delphinidæ, Monodontidæ, Phocœnidæ and Physeteridæ) among the 14 recognized by Berta *et al.* [19]. In particular, the family Delphinidæ was largely studied with 91 specific microsatellites isolated from nine delphinid species. This was not unexpected as this family is comprised of 34 species and is largely represented worldwide [19], with some species such as *Tursiops truncatus* (Montagu) being among the most studied and best-known cetacean models. Five families (Eschrichtiidæ, Iniidæ, Kogiidæ, Pontoporiidæ and Ziphiidæ) were studied only through cross-species amplifications. No microsatellite was isolated nor amplified in three cetacean families, namely the poorly known Neobalænidæ, endangered Platanistidæ, and critically endangered Lipotidæ [1]. Two additional microsatellites used in cetaceans were primarily isolated from the cattle *Bos taurus* (Linnæus), the total number of loci reported therefore being 246.

Table 1. Cetacean Species for which Microsatellites Could be Amplified. Bold Characters Indicate the Taxa from which Specific Microsatellites were Isolated. Other Taxa Reported here were Studied Only Through Cross-Species Amplifications

Suborder	Family	Genus	Species	
Mysticeti	Balænidæ	<i>Balæna</i>	<i>B. mysticetus</i>	
			<i>E. australis</i>	
		<i>Eubalæna</i>	<i>E. glacialis</i>	
			<i>E. japonica</i>	
			<i>E. spp.</i>	
	Balænopteriðæ	<i>Balænoptera</i>		<i>B. acutirostrata</i>
				<i>B. borealis</i>
				<i>B. brydei</i>
				<i>B. musculus</i>
				<i>B. physalus</i>
	<i>Megaptera</i>	<i>M. novæangliæ</i>		
Eschrichtiidæ	<i>Eschrichtius</i>	<i>E. robustus</i>		
Odontoceti	Cephalorhynchus	<i>Cephalorhynchus</i>	<i>C. commersonii</i>	
			<i>D. delphis</i>	
			<i>D. spp.</i>	
		<i>Globicephala</i>		<i>G. macrorhynchus</i>
				<i>G. melas</i>
		<i>Grampus</i>	<i>G. griseus</i>	
		<i>Lagenorhynchus</i>		<i>L. acutus</i>
				<i>L. albirostris</i>
				<i>L. obliquidens</i>
				<i>L. obscurus</i>
			<i>Orcælla</i>	<i>O. brevirostris</i>
	Delphinidæ	<i>Orcinus</i>	<i>O. orca</i>	
		<i>Pseudorca</i>	<i>P. crassidens</i>	
			<i>S. fluviatilis</i>	
		<i>Sotalia</i>	<i>S. guianensis</i>	
		<i>Sousa</i>		<i>S. chinensis</i>
				<i>S. attenuata</i>
			<i>S. clymene</i>	
		<i>Stenella</i>		<i>S. cœruleoalba</i>
				<i>S. frontalis</i>
				<i>S. longirostris</i>
	<i>Steno</i>	<i>S. bredanensis</i>		
	<i>Tursiops</i>		<i>T. aduncus</i>	
			<i>T. spp.</i>	
			<i>T. truncatus</i>	
	Monodontidæ	<i>Delphinapterus</i>	<i>D. leucas</i>	
		<i>Monodon</i>	<i>M. monoceros</i>	
	Kogiidæ	<i>Kogia</i>	<i>K. simus</i>	
	Phocœnidæ	<i>Neophocæna</i>		<i>N. phocœnoides</i>
				<i>P. dioptrica</i>
<i>Phocæna</i>			<i>P. phocæna</i>	
			<i>P. sinus</i>	
<i>Phocœnoides</i>		<i>P. spinipinnis</i>		
	<i>P. dalli</i>			
Physeteridæ	<i>Physeter</i>	<i>P. macrocephalus</i>		
Iniidæ	<i>Inia</i>	<i>I. geoffrensis</i>		
Pontoporiidæ	<i>Pontoporia</i>	<i>P. blainvillei</i>		
Ziphiidæ	<i>Berardius</i>	<i>B. bairdii</i>		
	<i>Hyperoodon</i>	<i>H. ampullatus</i>		
		<i>M. bidens</i>		
	<i>Mesoplodon</i>	<i>M. stejnegeri</i>		
	<i>Ziphius</i>	<i>Z. cavirostris</i>		

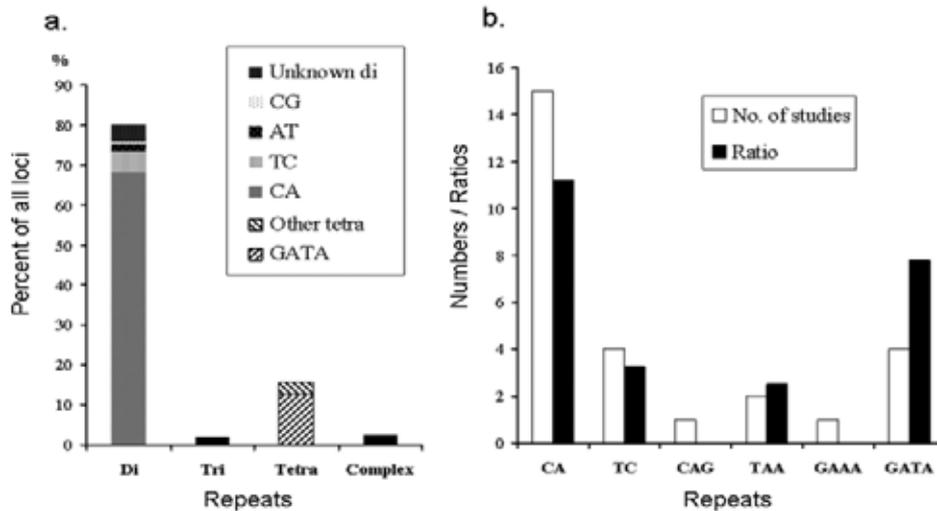


Fig. (2). Numbers and proportions of the different microsatellite motifs isolated in cetaceans from 1989 to 2007: (a) relative frequency of each type of repeat (percent), (b) numbers of published studies in which a cetacean genomic library was screened for a specific microsatellite motif (white bars), and ratios between the number of loci found and the corresponding number of studies (black bars) (see text for details).

(10 values ranging from 0.12 to 0.85), $0.48 (\pm 0.32)$ for tetranucleotide (68 values ranging from 0.00 to 0.90), and $0.72 (\pm 0.13)$ for complex repeats (5 values ranging from 0.58 to 0.87). High SD values show that the genetic diversity levels are heterogeneous among cetacean populations and species due notably to differences in effective population size and population history.

We investigated which type of locus (di-, tri-, tetranucleotide, complex) is the most variable, and whether the variability is greater in species where the loci were isolated as opposed to species where the loci were cross-amplified (ascertainment bias). Dinucleotide repeats were found to be significantly more polymorphic than tetranucleotide repeats (p -value < 0.0001). The CA and GATA repeats were the most common repeat types and were responsible for this difference in polymorphism (mean H_e 0.63 ± 0.23 and 0.49 ± 0.32 , respectively, $p = 0.0002$). This is consistent with studies in humans showing that dinucleotide loci evolve at a rate 1.5-2 times faster than tetranucleotide loci [30]. Similar results are found in a range of organisms, and can be expected when considering the mutation mechanism in microsatellite regions [31]. These observations suggest that, where possible, variability comparisons between different populations should be made using similar sets of markers.

No significant ascertainment bias was detected when considering all dinucleotide ($p = 0.24$) or tetranucleotide ($p = 0.16$) loci. This result indicates that the use of cross-amplified loci does not *a priori* bias the estimation of genetic diversity in cetaceans, suggesting that comparisons are possible between studies based on cross-amplification and studies using specifically isolated markers.

CONCLUSION

The legacy of past intensive whaling, combined with an increasing anthropogenic pressure on the marine environment where human and animal communities have diverging needs and interests, make cetacean conservation problematic. In this context, a thoughtful use of molecular genetic markers will help in defining conservation priorities. This database presents a comprehensive overview of the microsatellites studied in cetaceans from 1989 to 2007 with their technical characteristics and genetic diversity, and includes data from 246 loci studied among 51 cetacean species (1610 locus/species combinations). Dinucleotide repeats, and especially (CA) repeats, are the most frequent and the most variable loci, as observed in other species. The widespread use of cross-species amplification appears to be a convenient approach due to the high degree of conservation of the repeats and flanking sequences and the absence of significant ascertainment bias in cetacean species. Such genetic information should fruitfully be integrated with more traditional approaches for the design of realistic and efficient cetacean management policies.

ACKNOWLEDGEMENTS

We are grateful to A. Rozzi and F. Magné for technical laboratory assistance, and L. Chikhi for statistical advice. We also thank K. Karolemeas, S. Hatch, P. Mirleau, V. Vervoort, A. Ribéron, E. Lecompte and two Journal referees for comments on the manuscript, and W. Dabin, A. Dewez, F. Dhermain, J.-L. Fabre, B. Jakobsen, B. Lafitte, A.-S. Lemaire and E. Valsecchi who provided striped dolphin samples. Finally, we are grateful to J. Galián Albaladejo for advice at various stages of the project.

II. Matériel

Divers tissus (mucles, foie, rein, hypoderme et/ou peau) ont été prélevés sur des animaux échoués en Méditerranée par les organismes ou personnes suivantes :

- membres du Groupe d'Etude des Cétacés En Méditerranée

Dr Frank Dhermain (13)

Dr Jean-Michel Bompar (83)

Dr Nicolas Keck (Services Vétérinaires Hérault)

Dr Didier Sitte (34)

Dr Sylvie Heinzmeier (83)

Laetitia Cornil (34)

M. Dupuy de la Grand Rive (34)

- Anne-Sophie Lemaire (stagiaire chez le vétérinaire du Passage – opérant à la FERSO-BIO, Le Passage (Agen)) ;

- Matthias Macé.

D'autres animaux ont été prélevés sur le littoral Atlantique par des membres du Réseau National d'Echouages dont Olivier Van Canneyt.

Enfin, des prélèvements issus d'animaux échoués sur le littoral Pacifique américain ou pris accidentellement dans le Pacifique ont été fournis par M. Brandy Jacobsen de l'*University of Alaska Museum of the North* (UAF) et 35 ADN de Dauphins bleu et blanc de Méditerranée ont par ailleurs été fournis par le Dr Elena Valsecchi de l'*Università degli Studi di Milano*.

III. Variabilité inter-océanique

Ce travail confié à Vincent Bourret dans le cadre de sa thèse de Doctorat Vétérinaire a consisté à étudier la variabilité de 5 marqueurs microsatellites sur 138 Dauphins bleu et blanc issus de 3 bassins océaniques (Méditerranée, Atlantique Nord-Est et Pacifique).

Les principales conclusions sont :

- une différenciation génétique significative entre l'échantillon méditerranéen et l'échantillon atlantique ;

- l'existence d'un écart à l'équilibre de Hardy-Weinberg et un déséquilibre de liaison entre paires de microsatellites observés au sein de l'échantillon méditerranéen. Nous avons discuté des origines possibles de ce déséquilibre dont l'effet Wahlund (mélange de populations).

Ce travail a fait l'objet d'une publication dans le *Journal of the Marine Biological Association of the United Kingdom*.

Article n°3: Bourret V, Macé M, Crouau Roy B (2007) Genetic variation and population structure of western Mediterranean and northern Atlantic *Stenella coeruleoalba* populations inferred from microsatellite data. *Journal of the Marine Biological Association of the United Kingdom* **87 (1) Special Issue**, 265-269. (<http://dx.doi.org/10.1017/S0025315407054859>)

Genetic variation and population structure of western Mediterranean and northern Atlantic *Stenella caeruleoalba* populations inferred from microsatellite data

Vincent J.R. Bourret*, Matthias R.J.M. Macé and Brigitte Crouau-Roy†

UMR 5174 UPS/CNRS 'Evolution et Diversité Biologique' EDB, Bât 4R3b2 2e étage, Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse cedex 9—France. *Corresponding author, e-mail: vincent@um.es
†E-mail: bcrouau@cict.fr

The patterns of genetic differentiation and levels of genetic diversity among striped dolphin (*Stenella caeruleoalba*) populations from the North Atlantic Ocean (N=45 individuals) and the central and western Mediterranean Sea (N=78) were investigated using five polymorphic microsatellite loci. A North Pacific sample (N=14) was added as an out-group. Two of the markers were tetranucleotide repeats tested for the first time in this species. The Mediterranean, Atlantic and Pacific samples displayed a mean number of alleles per locus of 11.2, 13.4, and 9.6 respectively, suggesting a high but variable polymorphism across loci. The Mediterranean sample displayed particular characteristics: (i) the lowest allelic richness and expected heterozygosity ($H_{e, \text{Mediterranean}}=0.76$, while $H_{e, \text{Atlantic}}=0.83$ and $H_{e, \text{Pacific}}=0.85$); (ii) a significant departure from Hardy–Weinberg equilibrium ($P<0.001$; $F_{is}=0.050$); and (iii) a significant linkage disequilibrium between two pairs of loci. These last two features, present neither in the Atlantic sample nor in the Pacific one, suggest that the western Mediterranean population might possibly be further subdivided. Significant genetic differentiation was detected between the Mediterranean and Pacific populations, and between the Mediterranean and Atlantic populations. However, pairwise Wright's F_{st} was not significantly different from zero between the two geographically isolated Atlantic and Pacific populations.

INTRODUCTION

The striped dolphin *Stenella caeruleoalba* (Meyen, 1833) is cosmopolitan in most of the world's warm and temperate seas, and is generally considered as a relatively abundant species (Archer II & Perrin, 1999; Reeves et al., 2003). According to the International Union for the Conservation of Nature and Natural Resources (IUCN), the species is at 'lower risk' but 'conservation dependent' (Cetacean Specialist Group, 1996). In the western Mediterranean, the striped dolphin is the most abundant cetacean species according to transect survey estimations (95% confidence interval: 68,379 to 214,800 individuals after Forcada et al. (1994), although this estimate is not based on the most robust current prospecting methods). However, the Mediterranean population has recently been facing various threats such as morbillivirus epizootics, pollution, or fishing interactions, the effects of which are difficult to quantify precisely (Aguilar, 2000).

In the humanized areas of south-western Europe, assessing the striped dolphin's population genetic structure and genetic diversity within populations could be an early step towards the design of proper conservation policies (Beebee & Rowe, 2004). In a study of mitochondrial DNA (mtDNA) restriction polymorphism, García Martínez et al. (1999) found that none of the 27 restriction haplotypes was shared between a Mediterranean sample (N=76) and an Atlantic sample (N=22). However, due to the maternal mode of inheritance of the mtDNA, studying nuclear autosomal markers is also useful, especially in species displaying a complex social

behaviour. Previous studies on cetaceans (e.g. Bérubé et al., 1998) have suggested that a nuclear gene flow may occur, whereas mtDNA reveals a strong geographical structure among populations. Using polymorphic nuclear markers, Valsecchi et al. (2004) found a significant differentiation between striped dolphin samples from the North Sea (N=6) and the Mediterranean Sea (N=98). The present study compares the patterns of genetic differentiation and levels of genetic diversity, using selectively neutral markers, between striped dolphin samples from the Mediterranean Sea (N=78), the Atlantic Ocean (N=45) and the North Pacific Ocean as an outgroup (N=14).

MATERIALS AND METHODS

Sample collection

Samples, consisting of muscle, skin and blubber, liver, or kidney from 123 dead stranded individuals, were collected in the western Mediterranean and the North Atlantic between 1989 and 2004. Samples from 14 dead by-caught individuals from the North Pacific were added as an out-group (Figure 1). Samples were stored at -20°C .

Laboratory protocols

Deoxyribonucleic acid (DNA) extraction was performed using standard phenol:chloroform methods (Sambrook & Russell, 2000). Individuals were genotyped at five microsatellite loci. Three loci were dinucleotide repeats:



Figure 1. (A) Sampling locations of the 78 Mediterranean and 45 Atlantic stranded individuals. The Mediterranean samples came from the Gulf of Valencia (GV), the Gulf of Lions (GL), the Ligurian Sea (LS), the northern Tyrrhenian Sea (NT), southern Ionian Sea (SI), the northern Ionian Sea (NI) and the southern Adriatic Sea (SA). The 'Atlantic' samples came from the southern Bay of Biscay (SB), the northern Bay of Biscay (NB), the English Channel (EC) and the North Sea (NS); N, number of individuals sampled at each location; (B) sampling locations of the 14 dead by-caught individuals from the North Pacific; each black spot indicates a sampling location.

MK6 and MK9 (isolated from *Tursiops aduncus*, Krützen et al., 2001), EV92Mn (isolated from *Megaptera novaeangliae*, Valsecchi & Amos, 1996). Two loci were tetranucleotide repeats: GATA053 and GATA098 (isolated from *Megaptera novaeangliae*, Palsbøll et al., 1997). Polymerase chain reactions (PCR; Mullis & Faloona, 1987) were conducted in 20 μ l volumes containing 10 μ l of Qiagen Multiplex Mix (Qiagen GmbH, Hilden, Germany), primers to a final concentration of 0.2 to 0.5 μ M, and purified water to a final volume of 20 μ l. One primer of each pair (provided by Proligo, France) was fluorescently labelled. Reactions were cycled using the Eppendorf Gradient Mastercycler 5331 (Eppendorf AG, Hamburg, Germany) as follows: 15 min at 95 °C, then 34 cycles of 30 s at 94 °C, 90 s at 51 °C and 90 s at 72 °C. A final extension step of 10 min at 72 °C was added. Migrations were performed on a ABIprism automated sequencer (Applied Biosystems, Foster City, CA, USA) with the Genescan 500 TAMRA size marker, and gels were analysed using ABI GENESCAN and GENOTYPER softwares (Applied Biosystems, Foster City, CA, USA).

Statistical analyses

The allelic richness (R) in each sample was computed as described in the FSTAT v. 2.9.3.2 package (Goudet, 1995). The unbiased expected heterozygosities (He) and genetic distances were estimated after Nei (1978) and computed using the GENETIX (Belkhir et al., 1996–2004) package. Significance of the difference between the mean variability indices among the basins was tested using Student's paired *t*-test. The Hardy–Weinberg (HW) exact test was performed after Guo & Thompson (1992) using the GENEPOP (Raymond & Rousset, 1995) package. Frequencies of null alleles were estimated using the MICROCHECKER (Van Oosterhout et al., 2004) package. Population structure was analysed computing Wright's (1969) F-statistics according to Weir & Cockerham (1984), using the GENETIX package. Population differentiation was estimated using F_{ST} , whereas

departure from HW equilibrium (HWE) within the samples was measured using F_{IS} . Values of R_{ST} , accounting for allele sizes, were computed after Michalakis & Excoffier (1996) using the GENEPOP package. Significance of the deviation from zero of Nei's (1978) genetic distances (D) and Wright's (1969) F-statistics were assessed by permutation tests.

RESULTS

Amplifications were successful for the five loci, among which were GATA098 and GATA053, two tetranucleotide markers which had not been previously amplified in this species.

Linkage disequilibrium

No significant linkage disequilibrium was detected when all samples were pooled. However, when pooling the samples by basin, significant linkage disequilibrium was observed only within the Mediterranean, between two pairs of loci: GATA098/GATA053 ($P=0.038$) and MK9/GATA053 ($P=0.028$).

Genetic diversity

The number of alleles per locus and population ranged from six (GATA053 in the Mediterranean and the Pacific) to 21 (MK6 in the Atlantic) (Table 1). The allelic richness (R) per locus and population ranged from 3.8 (GATA053 in the Mediterranean) to 13.0 (MK6 in the Pacific). The unbiased expected heterozygosity (He) per locus and per population ranged from 0.497 (GATA053 in the Mediterranean) to 0.931 (MK6 in the Pacific), suggesting a high but variable level of polymorphism across loci. The observed heterozygosity (Ho) per locus and per population ranged from 0.323 (GATA053 in the Mediterranean) to 1.000 (MK6 in the Pacific).

The mean number of alleles per locus in each population ranged from 9.6 in the Pacific to 13.4 in the Atlantic (Table 1). The mean allelic richness (R) was significantly higher

Table 1. Genetic variation at five microsatellite loci in *Stenella coeruleoalba*.

Locus/ parameters	Mediterranean	Atlantic	Pacific	Total
<i>GATA098</i>				
n	73	42	14	129
N _A	11	11	10	11
R	6.980	8.014	9.152	7.923
He	0.839	0.871	0.897	0.865
Ho	0.808	0.833	0.857	0.821
<i>MK9</i>				
n	65	42	14	121
N _A	11	13	9	16
R	6.395	8.865	8.132	7.865
He	0.754	0.882	0.868	0.827
Ho	0.754	0.786	0.857	0.777
<i>GATA053</i>				
n	62	41	14	117
N _A	6	9	6	10
R	3.760	5.861	5.142	4.822
He	0.497	0.621	0.669	0.564
Ho	0.323	0.610	0.714	0.470
<i>MK6</i>				
n	65	40	12	117
N _A	19	21	15	25
R	9.141	11.086	12.998	10.536
He	0.850	0.914	0.931	0.895
Ho	0.908	0.925	1.000	0.923
<i>EV92Mn</i>				
n	56	38	10	104
N _A	9	13	8	15
R	6.853	8.060	8.000	8.093
He	0.840	0.848	0.863	0.865
Ho	0.625	0.526	0.500	0.577
<i>Mean ±SD</i>				
n	64.2 ±6.14	40.6 ±1.67	12.8 ±1.79	
N _A	11.2 ±4.82	13.4 ±4.56	9.6 ±3.36	
R	6.63 ±1.92	8.38 ±1.88	8.68 ±2.83	
He	0.756 ±0.150	0.827 ±0.118	0.846 ±0.102	
Ho	0.684 ±0.226	0.736 ±0.164	0.786 ±0.189	

n, number of genotyped individuals; N_A, number of different alleles detected; R, allelic richness based on a minimum sample size of ten individuals; He, Nei's (1978) unbiased expected heterozygosity; Ho, observed heterozygosity; SD, standard deviation.

Table 3. Multilocus indices of genetic differentiation between three basins. Values below diagonals were computed using data from the five microsatellite loci, while values above diagonals were computed excluding the data from locus EV92Mn. For D and F_{ST}, asterisks indicate significant deviations from zero (*, P<0.05; **, P<0.01; ***, P<0.001).

D	Med	Atl	Pac	F _{ST}	Med	Atl	Pac	R _{ST}	Med	Atl	Pac
Med		0.066***	0.106***	Med		0.023***	0.036***	Med		0.003	0.032
Atl	0.085***		0.045***	Atl	0.024***		0.009	Atl	0.013		-0.009
Pac	0.166***	0.092***		Pac	0.042***	0.015*		Pac	0.026	-0.010	
(A)				(B)				(C)			

(A), pairwise values of Nei's (1978) genetic distance; (B) pairwise values of Wright's F_{ST}; (C) pairwise values of R_{ST}; Med, Mediterranean; Atl, Atlantic; Pac, Pacific.

Table 2. Values of F_{IS} and results of the Hardy-Weinberg equilibrium exact test. The 'HWE' rows indicate the P-value of HWE exact test.

Locus/parameters	Mediterranean	Atlantic	Pacific
<i>GATA098</i>			
F _{IS}	0.0363	0.0437	0.0459
HWE	0.2927	0.4984	0.8816
<i>MK9</i>			
F _{IS}	-0.0002	0.1105*	0.0127
HWE	0.4243	0.0230	0.9177
<i>GATA053</i>			
F _{IS}	0.3531*	0.0182	-0.0700
HWE	0.0000	0.6532	0.1915
<i>MK6</i>			
F _{IS}	-0.0691	-0.0119	-0.0776
HWE	0.4546	0.8143	0.8428
<i>EV92Mn</i>			
F _{IS}	0.2579*	0.3823*	0.4340*
HWE	0.0000	0.0000	0.0213
<i>All loci</i>			
F _{IS}	0.0965*	0.1115*	0.0746
HWE	< 0.001	< 0.001	0.3007

*, indicates that the F_{IS} value is significantly different from zero (P<0.05).

both in the Atlantic (R=8.38) and the Pacific (R=8.68) than in the Mediterranean (R=6.63) (P<0.05). The mean multilocus He was also significantly higher in the Atlantic (He=0.827) and the Pacific (He=0.846) than in the Mediterranean (He=0.756) (P<0.05) with no significant difference between the Atlantic and the Pacific. The mean multilocus Ho was lower in the Mediterranean (Ho=0.684) than in the Atlantic (Ho=0.736) and the Pacific (Ho=0.786), but the pairwise differences in Ho were not significant (P>0.05).

Hardy-Weinberg equilibrium (HWE)

Most of the loci were in HWE. However, the heterozygote deficiency was significant for five locus/basin combinations among 15 (indicated by asterisks in Table 2). The multilocus computation of the F_{IS} and HWE exact test showed a significant heterozygote deficiency in the Mediterranean and the Atlantic. One of the markers (EV92Mn) was

suspected of bearing null allele(s) in the three basins, with an estimated frequency of 0.12 in the Mediterranean, 0.19 in the Atlantic and 0.19 in the Pacific. However, results for genetic variability and HWE remained similar when excluding EV92Mn from the analyses (data not shown, available upon request), except that the departure from HWE in the Atlantic was no longer significant ($P=0.142$, $F_{IS}=0.0413$).

Genetic differentiation between the basins

The overall multilocus F_{ST} value (accounting for allele identity) was 0.025 considering all loci, and 0.023 excluding EV92Mn. The overall multilocus R_{ST} value (accounting for allele size) was lower with values of 0.011 (all loci) and 0.005 (without EV92Mn).

Pairwise Nei's (1978) genetic distance (D), F_{ST} and R_{ST} values indicated that the highest nuclear genetic differentiation occurred between the Mediterranean and Pacific populations (Table 3). Computation of multilocus pairwise Nei's (1978) genetic distance (D) and Wright's (1969) F_{ST} indicated a slight but highly significant differentiation between the Mediterranean and Atlantic populations ($P<0.001$; Table 3A,B). Most analyses indicated a higher or more significant differentiation between the Mediterranean and the Atlantic than between the Atlantic and the Pacific.

DISCUSSION

A higher genetic diversity within the Atlantic

In order to compare the patterns of genetic differentiation and the levels of genetic diversity among striped dolphin populations, 45 Atlantic samples mainly from the Bay of Biscay and 78 Mediterranean samples were analysed. The values computed for R and He within each basin showed a higher nuclear genetic diversity within the Atlantic than within the Mediterranean; this trend was also observed for the mean number of alleles (N_A) per locus largely dependent on the number of individuals sampled. This conclusion is consistent with Valsecchi et al.'s observations (2004) suggesting a higher level of allelic diversity in a North Sea sample (N=6) than in a Mediterranean sample (N=98), although this difference was not significant. Regarding mitochondrial DNA, García Martínez et al. (1999) found that mtDNA nucleotide diversity was higher in an Atlantic sample (N=22) than in a Mediterranean sample (N=76). Thus, the Atlantic population would be significantly more polymorphic than the Mediterranean population, regarding both nuclear and mitochondrial DNA.

Various hypotheses could explain the Atlantic population's higher genetic diversity. One explanation would be that the older a population is, the more polymorphic it is, as suggested by García Martínez et al. (1999). A second explanation would be a higher effective population size of the Atlantic population in relation to its large geographical range. The genetic diversity in the Pacific sample, despite its low number of individuals (N=14), is approximately similar to, or even slightly higher than, the diversity in the Atlantic sample.

Genetic isolation between the basins

Overall, based on autosomal microsatellites allelic frequencies information (F_{ST}), the differentiation between

Stenella ceruleoalba populations is low. The fact that the greatest genetic differentiation occurs between the Mediterranean and Pacific samples is consistent with the permanent isolation between those two basins since the worldwide radiation of *S. ceruleoalba* (LeDuc et al., 1999).

The computation of Nei's (1978) genetic distance and Wright's (1969) F_{ST} indicates a slight but significant genetic differentiation between the Atlantic and the Mediterranean samples, suggesting a partial genetic isolation. When assuming an island model (Wright, 1969), the estimated number of migrants per generation (Nm) can be computed from F_{ST} , where N is the effective population size and m the fraction of the population replaced by migrants at each generation. The computation of Nm gives a rough estimate of 10.30 with all loci, and 10.60 without EV92Mn. Due to sampling characteristics and model assumptions, this estimation may not reflect an actual number of migrants. However, it seems to differ notably from the mtDNA Nm estimation of 0.98 by García Martínez et al. (1999), where N is the effective number of females (Slatkin, 1989). The magnitude difference between these two rough estimates arising from different markers might suggest that the male-mediated gene flow could be higher than the female-mediated gene flow, but a precise quantitative assessment remains to be undertaken.

Surprisingly, most indices of genetic diversity indicate that the nuclear genetic differentiation between the Mediterranean and the Atlantic populations is greater than the differentiation between the Atlantic and the Pacific populations. The striped dolphin is essentially a tropical and temperate water species, and it is usually found neither in the waters off southern Argentina or Chile, nor in the Canadian Archipelago (Archer II & Perrin, 1999). Moreover, the closure of the Panama seaway dates back to the late Pliocene (3–4 million years ago; Coates et al., 2004). Consequently, the low genetic differentiation between the Atlantic and Pacific samples is hardly accounted for by a significant current gene flow between these two oceans.

A significant heterozygote deficiency within the Mediterranean sample only

Though the Atlantic and Pacific samples are in Hardy–Weinberg equilibrium, the Mediterranean sample displays significant heterozygote deficiency, in particular regarding GATA053, one of the newly-amplified markers. Several hypotheses could explain this observation (Crouau-Roy, 1988), such as a bias due to the sampling of stranded animals, a significant inbreeding within the population, or a 'Wahlund effect' due to the sampling of two or more reproductively distinct populations. The hypothesis that stranded individuals might be, in some cases, more homozygous than the population they represent is discussed by Valsecchi et al. (2004). In the present study, though the Atlantic sample consists solely of stranded individuals, it does not display such a significant heterozygote deficiency. Furthermore, significant linkage disequilibrium (involving the two newly-tested tetranucleotides) is observed only within the Mediterranean sample, which can also be the consequence of the sampling of two or more reproductively distinct Mediterranean populations. So the Mediterranean

striped dolphin population would itself be subdivided (this would be consistent with the bimodal distribution of the allelic frequencies observed within this basin for the tetranucleotide locus GATA098). Performing assignment tests, Valsecchi et al. (2004) found some slight indices of genetic differentiation according to geographical distance between sampling regions within the Mediterranean (e.g. Gulf of Valencia, Ligurian Sea). Gaspari (2004) also found significant genetic differentiation within the Mediterranean, between 'inshore' and 'offshore' striped dolphin groups. Further studies, such as analyses based on haplotypic data, could be carried out to investigate the exact relationship between possible Mediterranean distinct lineages.

Conservation of the Mediterranean striped dolphins

Though it does not seem to be in imminent danger of extinction, the Mediterranean striped dolphin population lives in a highly humanized area, and faces a range of threats whose effects are difficult to quantify. This study shows that, regarding nuclear markers, the Mediterranean population as a whole is genetically less variable than the Atlantic or Pacific populations, and is partially reproductively isolated. Given its genetic characteristics, the Mediterranean striped dolphin population should therefore benefit from an adequate conservation effort.

We are grateful to W. Dabin, A. Dewez, F. Dhermain, J.-L. Fabre, B. Jakobsen, B. Lafitte, A.-S. Lemaire, and E. Valsecchi who provided samples, and to M. Bonhomme, F. Magné and A. Rozzi for technical laboratory assistance. We are also grateful to P. De la Rúa, J. Galían, A. Ortiz and J. Serrano for advice at various stages of the project and helpful comments on the manuscript, and to M. Oñate and B. Séverac for reviewing the English.

REFERENCES

Aguilar, A., 2000. Population biology, conservation threats and status of Mediterranean striped dolphins (*Stenella ceruleoalba*). *Journal of Cetacean Research and Management*, **2**, 17–26.

Archer II, F.I. & Perrin, W.F., 1999. *Stenella ceruleoalba*. *Mammalian Species*, **603**, 1–9.

Beebee, T. & Rowe, G., 2004. *An introduction to molecular ecology*. New York: Oxford University Press.

Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N. & Bonhomme, F., 1996–2004. *GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations*. Laboratoire Génome, Populations, Interactions, CNRS UMR 5171, Université de Montpellier II, Montpellier, France.

Bérubé, M. et al., 1998. Population genetic structure of North Atlantic, Mediterranean Sea and Sea of Cortez fin whales, *Balaenoptera physalus* (Linnaeus 1758): analysis of mitochondrial and nuclear loci. *Molecular Ecology*, **7**, 585–599.

Cetacean Specialist Group, 1996. *Stenella ceruleoalba*. In *2004 IUCN Red list of threatened species*. www.redlist.org downloaded on 24 June 2005.

Coates, A.G., Collins, L.S., Aubry, M.P. & Berggren, W.A., 2004. The geology of the Darien, Panama, and the late Miocene–Pliocene collision of the Panama arc with north-western South America. *Geological Society of America Bulletin*, **116**, 1327–1344.

Crouau-Roy, B., 1988. Genetic structure of cave-dwelling beetles populations: significant deficiencies of heterozygotes. *Heredity*, **60**, 321–327.

Forcada, J., Aguilar, A., Hammond, P.S., Pastor, X. & Aguilar, R., 1994. Distribution and numbers of striped dolphins in the Western Mediterranean Sea after the 1990 epizootic outbreak. *Marine Mammal Science*, **10**, 137–150.

García Martínez, J., Moya, A., Raga, J.A. & Latorre, A., 1999. Genetic differentiation in the striped dolphin *Stenella ceruleoalba* from European waters according to mitochondrial DNA (mtDNA) restriction analysis. *Molecular Ecology*, **8**, 1069–1073.

Gaspari, S., 2004. *Social and population structure of striped and Risso's dolphins in the Mediterranean Sea*. PhD thesis, University of Durham, Durham, UK.

Goudet, J., 1995. FSTAT (version 1.2): a computer program to calculate F-statistics. *Journal of Heredity*, **86**, 485–486.

Guo, S.W. & Thompson, E.A., 1992. Performing the exact test of Hardy–Weinberg proportion for multiple alleles. *Biometrics*, **48**, 361–372.

Krützen, M., Valsecchi, E., Connor, R.C. & Sherwin, W.B., 2001. Characterization of microsatellite loci in *Tursiops aduncus*. *Molecular Ecology Notes*, **1**, 170–172.

LeDuc, R.G., Perrin, W.F. & Dizon, A.E., 1999. Phylogenetic relationships among the delphinid cetaceans based on full cytochrome b sequences. *Marine Mammal Science*, **15**, 619–648.

Michalakis, Y. & Excoffier, L., 1996. A generic estimation of population subdivision using distances between alleles with special interest to microsatellite loci. *Genetics*, **142**, 1061–1064.

Mullis, K.B. & Faloona, F.A., 1987. Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. *Methods in Enzymology*, **155**, 335–350.

Nei, M., 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, **89**, 583–590.

Palsboll, P.J., Bérubé, M., Larsen, A.H. & Jørgensen, H., 1997. Primers for the amplification of tri- and tetramer microsatellite loci in baleen whales. *Molecular Ecology*, **6**, 893–895.

Raymond, M. & Rousset, F., 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Reeves, R.R., Smith, B.D., Crespo, E.A. & Di Sciara, G.N., 2003. *Dolphins, whales and porpoises: 2002–2010 conservation action plan for the World's cetaceans*. Gland and Cambridge: IUCN.

Sambrook, J. & Russell, D.W., 2000. Preparation of genomic DNA from mouse tails and other small samples. In *Molecular cloning—a laboratory manual on the web*. Chapter 6, Protocol 5. http://www.molecularcloning.com/members/protocol_print.jsp?protocol=4038 downloaded on 10 November 2005.

Slatkin, M., 1989. Population structure and evolutionary progress. *Genome*, **31**, 196–202.

Valsecchi, E. & Amos, W., 1996. Microsatellite markers for the study of cetacean populations. *Molecular Ecology*, **5**, 151–156.

Valsecchi, E., Amos, W., Raga, J.A., Podestà, M. & Sherwin, W., 2004. The effects of inbreeding on mortality during a morbillivirus outbreak in the Mediterranean striped dolphin (*Stenella ceruleoalba*). *Animal Conservation*, **7**, 139–146.

Van Oosterhout, C., Hutchinson, W.F., Wills, D.P.M. & Shipley, P., 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**, 535–538.

Weir, B.S. & Cockerham, C.C., 1984. Estimating F-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

Wright, S., 1969. *Evolution and the genetics of populations*. Vol. 2. *The theory of gene frequencies*. Chicago: University of Chicago Press.

Submitted 15 June 2006. Accepted 17 October 2006.

IV. Polyphyllie et Crise de Salinité du Messinien

IV. A. Méthodes

A.1) Extraction

Les échantillons, préservés dans l'alcool (UAF) ont dû être élués entre une et cinq fois dans de l'eau désionisée avant extraction.

A.2) Amplification et séquence

Les amorces utilisées pour l'amplification de la région de contrôle ("boucle-D") de l'ADN mitochondrial ont leurs points d'ancrage dans les ARN de transfert (ARNt) situés de part et d'autre de ladite région :

- ARNT-Thr pour l'amorce "sens" (forward) 5'-GAATCCCCGGTCTTGTAAC-3'

- ARNt-Phe pour l'amorce "anti-sens" (reverse) 5'-TCTCGAGATTTTCAGTGTCTTGCTTT-3'

Les produits d'amplification ont été contrôlés sur gel d'agarose à 1,3%.

Les quantités d'ADN amplifié ont été mesurées par spectrophotométrie et par mesure de la fluorescence sur le gel. Certains amplicons ont été purifiés à l'aide du MinElute PCR Purification Kit (Qiagen GmbH, Hilden, Germany, #28004) et ceux dont la concentration était la plus faible ont été resuspendus dans 10 µl au lieu de 30 µl.

Le programme GENALYS (Masazumi Takahashi, <http://software.cng.fr/docs/genalys.html>) a été utilisé pour la correction en parallèle des séquences mais un « bogue » dans ce programme empêche d'ouvrir une partie des électrophorégrammes qui ont dû être corrigés individuellement.

IV. B. Méthodologie

B.1) Inférence statistique par Vraisemblance (L) et Critère d'Information d'Akaiké (AIC)

1(a) Concept et principe de vraisemblance

Le principe de vraisemblance est un principe de statistique inférentielle qui affirme que toute l'information d'un échantillon est contenue dans la fonction de vraisemblance. L'estimation du maximum de vraisemblance est une méthode statistique utilisée couramment pour inférer les paramètres de la distribution de probabilité d'un échantillon donné. Après avoir tenté de trouver un critère absolu dans les problèmes de régression et présenté la procédure numérique (**Fisher, 1912**), Ronald Fisher a exploré les concepts de vraisemblance, suffisance, efficacité et information qui lui ont permis de formaliser plus tard la méthode de maximum de vraisemblance (**Fisher, 1922**).

La Fonction de Vraisemblance (*likelihood function*), notée $L(x_1, \dots, x_n \mid \theta_1, \dots, \theta_k)$ est une fonction de probabilités conditionnelles qui décrit les paramètres θ_i d'une loi statistique en fonction des valeurs x_j supposées connues. Elle s'exprime à partir de la

fonction de densité $f(x \mid \theta)$ par $L(x_1, \dots, x_n \mid \theta) = \prod_{i=1}^n f(x_i; \theta)$.

C'est donc le contraire d'une densité de probabilité qui donne, en fonction de paramètres θ_i connus, les valeurs prises par les x_i . D'une certaine manière, la vraisemblance fonctionne à l'inverse de la probabilité : sachant B , nous avons la probabilité $\Pr(A \mid B)$ afin de raisonner sur A et sachant A , on utilise la fonction de vraisemblance $L(B \mid A)$ afin de raisonner sur B . Ce formalisme est donné par le théorème de Bayes (**Bayes & Price, 1763**) :

$$\Pr(B \mid A) = \frac{\Pr(A \mid B) \cdot \Pr(B)}{\Pr(A)}$$

Cette dernière relation se lit comme suit : la probabilité de l'évènement B sachant la survenue de A équivaut au produit des probabilités de la survenue de A sachant la survenue de B par la probabilité de B ramenée à la probabilité de A .

En statistiques, une fonction de vraisemblance est une fonction de probabilité conditionnelles considérée comme une fonction de son second argument, son

premier argument étant maintenu fixe : $b \mapsto \Pr(A|B=b)$. La fonction de vraisemblance de B est la classe de fonctions équivalentes telle que : $L(b|A) = \alpha \Pr(A|B=b)$ pour $\alpha > 0$. La valeur numérique obtenue étant immatérielle, on utilise les ratios de la forme $\frac{L(b_2|A)}{L(b_1|A)}$.

1(b) Statistique par maximum de vraisemblance

On cherche à trouver le maximum de cette vraisemblance pour que les probabilités des réalisations observées soient aussi maximales. On a donc recours à l'optimisation. Ainsi, si L est dérivable et si elle admet un maximum global pour une valeur $\theta = \hat{\theta}$, sa dérivée première s'annule et sa dérivée seconde est négative en ce point. Ces deux conditions permettent donc d'affirmer que $\theta = \hat{\theta}$ est un maximum local. La première permet de trouver la valeur $\theta = \hat{\theta}$ alors que la deuxième permet d'affirmer que $\hat{\theta}$ est un maximum local (cf. Figure 35).

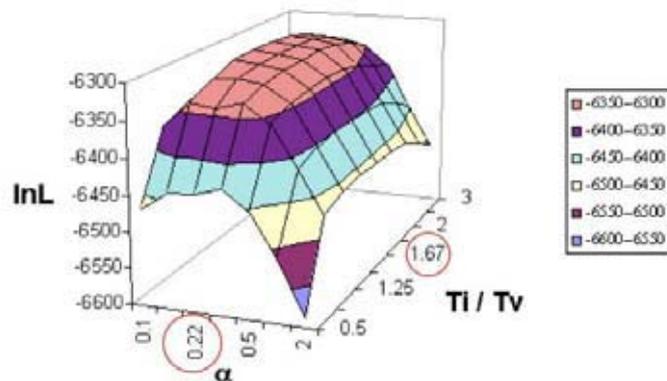


Figure 35. Exemple de représentation d'un point de maximum de vraisemblance (ici son logarithme) sur un espace à deux paramètres.

1(c) Test de quotient de vraisemblance LRT

Le test de quotient de vraisemblance *Likelihood Ratio Test*, LRT, est une méthode de test d'hypothèses qui consiste en un quotient des probabilités maximales d'une observation sous deux hypothèses différentes :

$$\Lambda(x) = \frac{\sup\{L(\theta|x) : \theta \in \Theta_0\}}{\sup\{L(\theta|x) : \theta \in \Theta\}}$$

(Une hypothèse nulle est énoncée en définissant θ comme une partie d'un sous-ensemble Θ_0 de l'espace des paramètres Θ).

Cette fonction de x est une statistique. Le *likelihood-ratio test* rejette l'hypothèse nulle si cette valeur est trop petite. De la petitesse de cette valeur dépend la significativité de ce test, c'est à dire quelle probabilité d'erreur de Type 1 est considérée comme tolérable. On considère une valeur critique c en-deçà de laquelle ($\Lambda \leq c$) on accepte H_0 .

Dans un certain nombre de cas, on ne peut dissocier le processus décisionnel en petites unités, utiles pour faciliter le calcul, mais l'on peut aussi approcher la distribution exacte de Λ par une distribution approchée. En effet, lorsque la taille de l'échantillon $n \rightarrow \infty$, $-2\log(\Lambda)$ est asymptotiquement distribué à la manière d'un χ^2 de nombre de degrés de liberté égal à la différence de dimensions entre Θ_0 et Θ .

1(d) Le Critère d'Information d'Akaiké, AIC

Hirotsugu Akaike a énoncé « un critère d'information » comme une mesure de l'adéquation (*goodness of fit*) d'un modèle statistique (Akaike, 1974). Cette notion se base sur le principe d'entropie en mesurant la perte d'information lors de l'utilisation de ce modèle pour décrire la réalité. Autrement dit, cette mesure décrit le compromis entre le biais et la variance dans la construction du modèle. Ce n'est donc pas un test statistique mais un outil de sélection de modèle. Sa formulation générale est la suivante :

$$AIC = 2k - 2\ln(L) = 2k + n\ln(RSS/n)$$

où k est le nombre de paramètres du modèle, L la valeur maximale de la fonction de vraisemblance, n le nombre d'observations et RSS la somme des carrés résiduelle. Ainsi, l'AIC récompense-t-il la complexité du modèle (nombre de paramètres libres k) mais inclut aussi une pénalité qui est une fonction croissante du nombre de paramètres estimés, décourageant par là-même le « surapprentissage » (*overfitting*).

Pour de petits échantillons, on utilise le AIC_C (qui converge vers le AIC lorsque n devient grand) :

$$AIC_C = AIC + \frac{2k(k+1)}{n-k-1}$$

B.2) Test hiérarchique de modèles de substitution (ModelTest)

Dans le cas de modèles de substitution, qui nous intéresse ici, l'équation donnant le LRT peut s'énoncer comme suit :

$$\Lambda = \frac{\max[L_0(\text{Modèle_nul} \mid \text{Données})]}{\max[L_1(\text{Modèle_alternatif} \mid \text{Données})]}$$

où L_0 est la vraisemblance sous l'hypothèse nulle (modèle le plus simple) et L_1 est la vraisemblance sous l'hypothèse alternative (modèle plus complexe, plus riche en paramètres). On a vu que lorsque les modèles comparés sont 'nichés' (*nested*), c'est à dire que l'hypothèse nulle est un cas particulier de l'hypothèse alternative et qu'elle est correcte, la statistique δ est distribuée asymptotiquement à la manière d'un χ^2 à q degrés de liberté (q étant la différence du nombre de paramètres libres entre les deux modèles). Le logiciel MODELTEST (Posada & Crandall, 1998) propose d'utiliser l'AIC pour tester les modèles de substitution deux-à-deux dans un sens de complexité croissante. Dans son cadre classique (non bayésien), MODELTEST calcule la statistique δ du LRT ainsi que la *p-value* associée en utilisant χ^2 à q degrés de liberté. La procédure suit une démarche hiérarchique telle que présentée dans la Figure 36 ci-après :

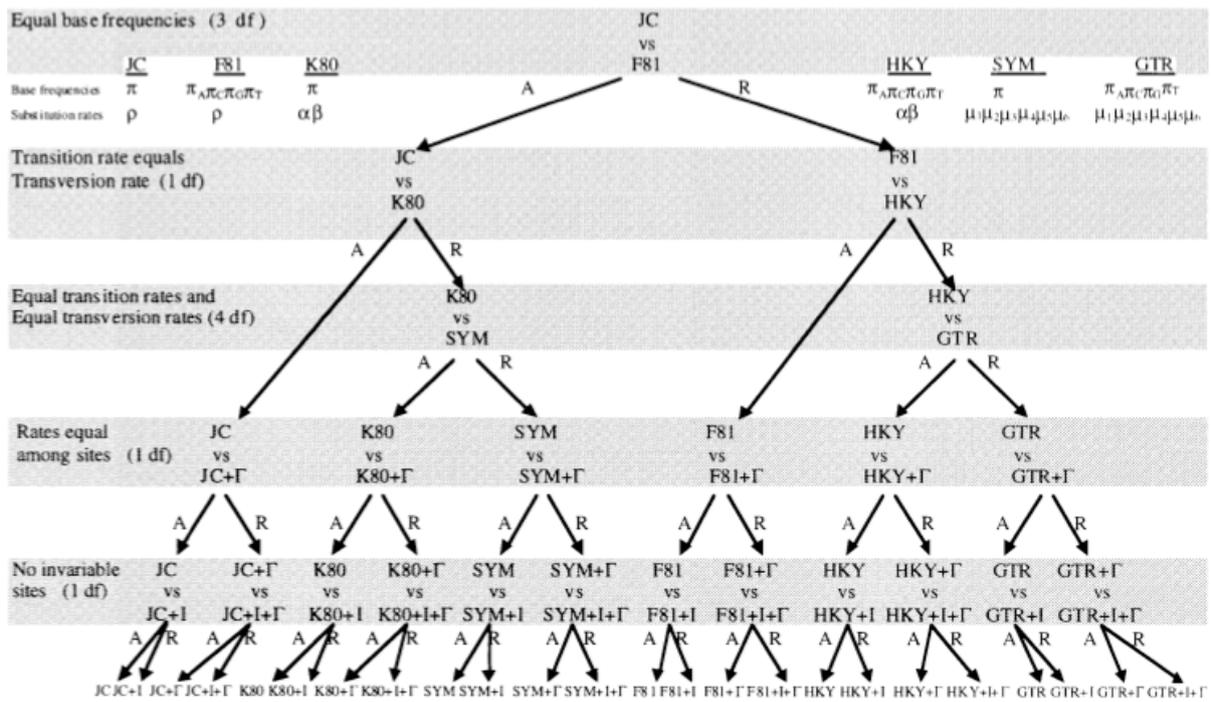


Figure 36. Test d'hypothèses hiérarchique utilisé par MODELTEST.

A chaque niveau, l'hypothèse est acceptée (A) ou rejetée (R).

B.3) Inférence bayésienne

L'inférence bayésienne est une inférence statistique dans laquelle des observations ou des faits démontrés sont utilisés pour procéder à une nouvelle inférence de la probabilité qu'une hypothèse soit vraie. Le principe ontologique de cette méthode est basé sur le fait qu'à mesure que des faits s'accumulent, le degré de vraisemblance d'une hypothèse évolue jusqu'à devenir très haut ou très bas. Elle peut donc en théorie se convertir en une base logique capable de discriminer entre des hypothèses conflictuelles. En pratique, elle nécessite d'assigner *a priori* aux hypothèses des probabilités qui seront revues tout au long du processus d'inférence. Ainsi, cette méthode utilise une estimation numérique du degré de probabilité d'une hypothèse avant l'observation et calcule une autre estimation après l'observation. Elle utilise la relation de Bayes que nous avons déjà vue :

$$\Pr(H_0|E) = \frac{\Pr(E|H_0) \cdot \Pr(H_0)}{\Pr(E)} \quad \text{où :}$$

- H_0 représente l'hypothèse nulle qui a été inférée avant que la nouvelle

observation E ne soit disponible ;

- $\Pr(H_0)$ appelée probabilité a priori - *prior probability* - ou tout simplement *prior* ;

- $\Pr(E | H_0)$ est appelée probabilité conditionnelle d'observer E étant donnée l'hypothèse H_0 considérée comme vraie (on dit aussi « sachant H_0 »). Il s'agit de la fonction de vraisemblance (*cf. supra*) de son second argument (H_0) lorsque le premier est maintenu constant ;

- $\Pr(E)$ est appelée la probabilité marginale de E : la probabilité d'observer une nouvelle occurrence de E sous l'ensemble des hypothèses mutuellement exclusives. Elle peut être calculée comme la somme du produit de l'ensemble des probabilités conditionnelles : $\sum_i \Pr(E|H_i)\Pr(H_i)$;

- $\Pr(H_0 | E)$ est appelée la probabilité postérieure de H_0 sachant E .

Ainsi, le facteur $\Pr(E | H_0) / \Pr(E)$ est une mesure de l'**impact** que l'observation a sur la probabilité de l'hypothèse.

Les détracteurs de cette méthode soutiennent que ce procédé a des limitations inhérentes à l'existence de *priors*. En effet, même si les « bayésiens » soutiennent que même lorsque des personnes attribuent des probabilités *a priori* diverses et subjectives, de nouvelles observations répétées auront tendance à faire converger les probabilités postérieures vers des valeurs voisines, on peut légitimement se demander si ces inférences ne peuvent pas être perturbées par des différences d'appréciation suffisamment radicales.

Ainsi, le ratio de vraisemblance $\Lambda = \frac{L(H_0|E)}{L(notH_0|E)} = \frac{\Pr(E|H_0)}{\Pr(E|notH_0)}$

peut être réécrit : $\Pr(H_0|E) = \frac{\Pr(E|H_0)\Pr(H_0)}{\Pr(E|H_0)\Pr(H_0) + \Pr(E|notH_0)\Pr(notH_0)} = \frac{\Lambda \Pr(H_0)}{\Lambda \Pr(H_0) + \Pr(H_0)}$

La règle de Bayes peut être utilisée itérativement lorsqu'il y a plus d'une observation :

$\Pr(H_0|E_1, E_2) = \frac{\Pr(E_1|H_0) \cdot \Pr(E_2|H_0) \cdot \Pr(H_0)}{\Pr(E_1) \cdot \Pr(E_2)}$ (exemple pour 2 évènements).

Le raisonnement Bayésien peut être illustré comme suit (Veyrunes *et al.*, 2008) :

For billions of years, the sun has risen after it has set.

The sun has set tonight.

With very high probability (or I strongly believe that or it is true that) the sun will rise tomorrow.

With very low probability (or I do not at all believe that or it is false that) the sun will not rise tomorrow.

B.4) Théorie de la Coalescence

La théorie de la coalescence (Kingman, 1982a) est un modèle rétrospectif de génétique des populations. Il s'agit d'un processus stochastique qui est une bonne approximation des distributions des histoires généalogiques issues des modèles *forward-time* comme le modèle Fisher-Wright (Fisher, 1930; Kingman, 1982b; Wright, 1931). L'avantage des estimations basées sur les généalogies provient du fait qu'elles tiennent compte de la non-indépendance des séquences échantillonnées.

Un arbre est une collection de branches, nœuds et âges de nœuds qui ensemble caractérisent complètement une histoire évolutive. Alors que de nombreuses méthodes coalescentes ne se réfèrent qu'à une seule phylogénie (Fu, 1994; Nee *et al.*, 1995; Pybus *et al.*, 2000) obtenue par des méthodes phylogéniques traditionnelles - et souvent entâchées d'incertitudes considérables -, les méthodes probabilistes (Maximum de Vraisemblance et Bayésienne) autorisent l'incertitude sur la généalogie et le calcul de la vraisemblance moyenne des paramètres populationnels. Le calcul fait appel à l'intégration sur les généalogies distribuées selon la coalescence (*i.e.* les plus probables sont échantillonnées plus souvent) et fait appel à des algorithmes d'échantillonnage variable selon la probabilité (*e.g.* Metropolis-Hastings). On estime ainsi le paramètre populationnel $\Theta = 2N_e\mu$, le taux d'accroissement exponentiel k tel que $\frac{\Theta_0}{\Theta(t)} = e^{-kt}$ (attention : $t \leq 0$!), les taux de migrations et de recombinaison. Des méthodes basées sur des Chaines de Markov avec technique de Monte Carlo (MCMC) plus ou moins couplées avec des algorithmes capables par exemple de générer des marches aléatoires en utilisant des densités et une méthode de rejet des mouvements proposés.

Par exemple l'algorithme de Metropolis-Hastings (**Hastings, 1970; Metropolis et al., 1953**) calcule le quotient de vraisemblance $a_1 = \frac{\Pr(x')}{\Pr(x)}$ entre l'échantillon proposé

(potentiel nouvel état de la MCMC) et l'état précédent. Il le multiplie au quotient de

densité de proposition $a_2 = \frac{Q(x';x)}{Q(x;x')}$. Le processus de décision est alors le suivant :

- Si $a = a_1 \cdot a_2 \geq 1 \Rightarrow x^{t+1} = x'$
- Si $a = a_1 \cdot a_2 < 1 \Rightarrow x^{t+1} = \begin{cases} x'; \Pr(x') = a \\ x; \Pr(x) = 1 - a \end{cases}$

Ainsi, les mouvements « descendants » sont pénalisés alors que ceux « montants » sont favorisés comme le montre la 0suivante.

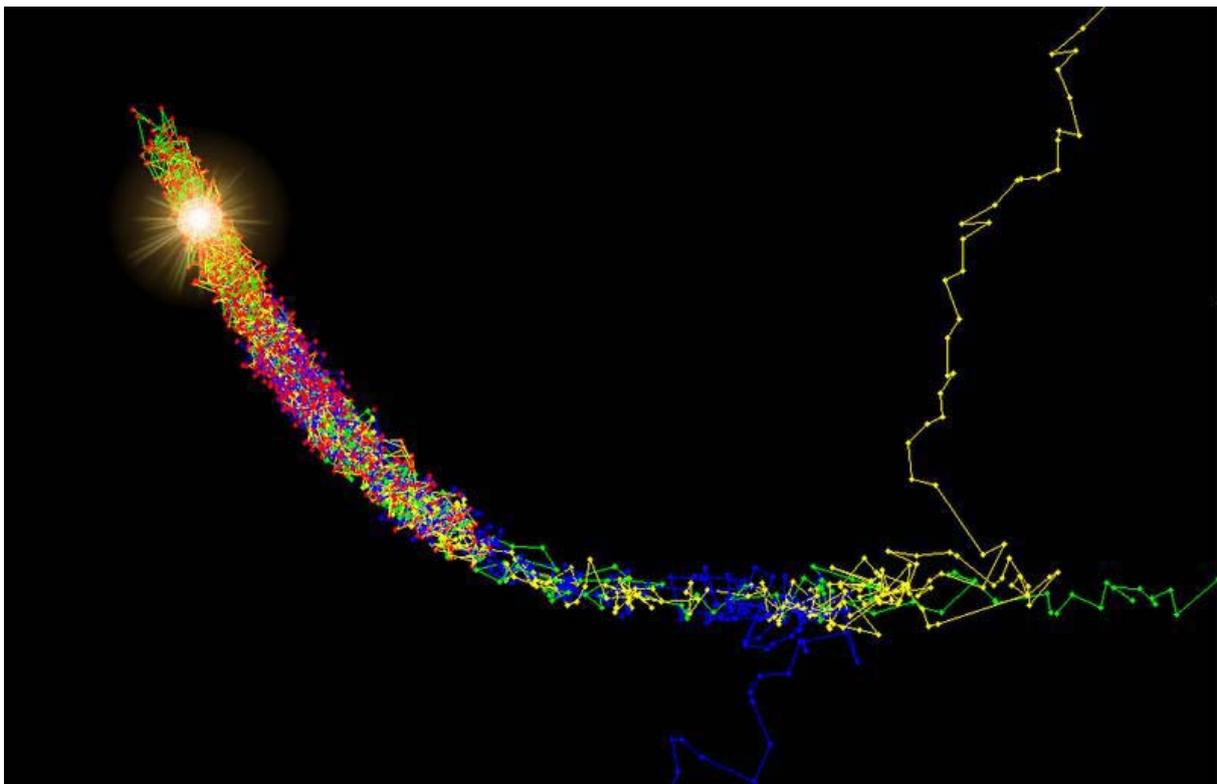


Figure 37. Exemple de représentation graphique du parcours de 3 chaînes de Markov indépendantes.

L'algorithme de Metropolis-Hastings, convergeant sur la fonction de Rosenbrock, échantillonne préférentiellement dans les régions où la probabilité postérieure est élevée et où les chaînes commencent à converger. Le point lumineux est la position approximative du maximum de probabilité postérieure.

Ainsi, un point de l'espace des probabilités est plus fréquemment visité lorsque sa probabilité postérieure est plus grande.

B.5) Une méthode d'inférence démographique non-paramétrique : le Bayesian Skyline Plot

En sus de l'inférence en génétique des populations basées sur la coalescence sur des échantillons actuels, on peut disposer maintenant d'échantillons non synchrones. Ces séquences espacées dans le temps peuvent être issues de virus évoluant rapidement (*e.g.* VIH ; **Rodrigo, 1999**) ou de vestiges archéo-paléontologiques (*e.g.* Manchots de la Terre Adélie ; **Lambert et al., 2002**). Ces séquences espacées dans le temps offrent le potentiel d'observer les mutations survenues au cours du temps et donc d'estimer les taux de mutation et leur variation temporelle (**Drummond et al., 2001**).

5(a) Prise en compte de l'horloge moléculaire et de séquences échantillonnées de façon asynchrone

Cette méthode se situe dans le cadre d'un arbre issu d'un processus de coalescence dont les extrémités sont déportées dans le temps (**Drummond et al., 2002**).

On cherche dans un premier temps à évaluer la densité de la Coalescence pour un modèle de Fisher-Wright de taille constante. Le temps est compté positivement vers le passé. On estime $\theta = N_e \rho$ (ρ est le nombre d'unités calendaires de temps, N_e la taille efficace) qui est un facteur d'échelle permettant de passer du « temps de coalescence » au temps mesuré calendaire. L'arbre g est un arbre enraciné binaire composé de n nœuds terminaux (*leaf nodes*) et $n-1$ nœuds ancestraux. Le nœud i est caractérisé par un âge t_i .

Si l'on considère que :

- E_g représente l'ensemble des arêtes de l'arbre entre les nœuds i et j tels que $i > j$,
- t_Y représente le temps de divergence des nœuds ancestraux,

alors, la densité de probabilité d'un arbre $g = (E_g, t_Y)$ de n nœuds externes et $n-1$ nœuds ancestraux s'écrit :

$$f_G(g|\theta) = \frac{1}{\theta^{n-1}} \cdot \prod_{i=2}^{2n-1} e^{(-k_i(k_i-1)/2\theta)(t_i - t_{i-1})}$$

où k_i représente le nombre de lignages présents dans l'intervalle de temps compris entre le nœud $i-1$ et le nœud i .

On peut interpréter cette valeur comme suit : en **fixant un temps t** et en supposant que **k lignages** sont présents à ce moment, un événement de coalescence entre chacune des $k(k-1)/2$ paires de lignages a lieu à un taux instantané de $1/\theta$ (comme prédit par la théorie de Kingman). Ainsi, dans l'intervalle de temps t_i-t_{i-1} précédent un nœud externe i \square I , « rien ne se passe » avec la probabilité $e^{-k_i(k_i-1)/2\theta(t_i-t_{i-1})}$. L'intervalle de temps $t-t_{i-1}$ précédent le nœud de coalescence i \square Y est une variable aléatoire de densité $k(k-1)/2\theta \cdot e^{-k(k-1)/2\theta(t-t_{i-1})}$. Pour plus de détails sur l'extension de la coalescence de Kingman à des arbres dont les extrémités sont non contemporaines, cf. **Rodrigo & Felsenstein (1999)**.

5(b) Principe du Bayesian Skyline Plot

Le problème de la plupart des méthodes basées sur la coalescence pour déterminer la dynamique des populations est que celles-ci reposent sur un modèle démographique prédéfini. Le *Bayesian Skyline Plot (BSP)* est une méthode non paramétrique visant à estimer la dynamique passée d'une population au cours du temps sans avoir recours à un modèle d'histoire démographique spécifié au préalable (**Drummond et al., 2005**). Il consiste en une procédure d'échantillonnage MCMC tirant ses échantillons d'un *skyline plot* généralisé (**Strimmer & Pybus, 2001**), étant donné les séquences. Il combine alors ces relevés graphiques afin de générer une distribution postérieure de la taille efficace de la population au cours du temps (cf. **Figure 38**). L'avantage par rapport aux méthodes de skyline plot précédemment mises au point (**Pybus et al., 2000; Strimmer & Pybus, 2001**) est qu'il intègre les incertitudes de la phylogénie dans des intervalles de crédibilité.

La formulation de sa *prior* sur Θ est la suivante :

$$f_{\Theta}(\Theta) \propto \frac{1}{\theta_1} \prod_{j=2}^m \theta_{j-1} e^{-\theta_j/\theta_{j-1}}$$

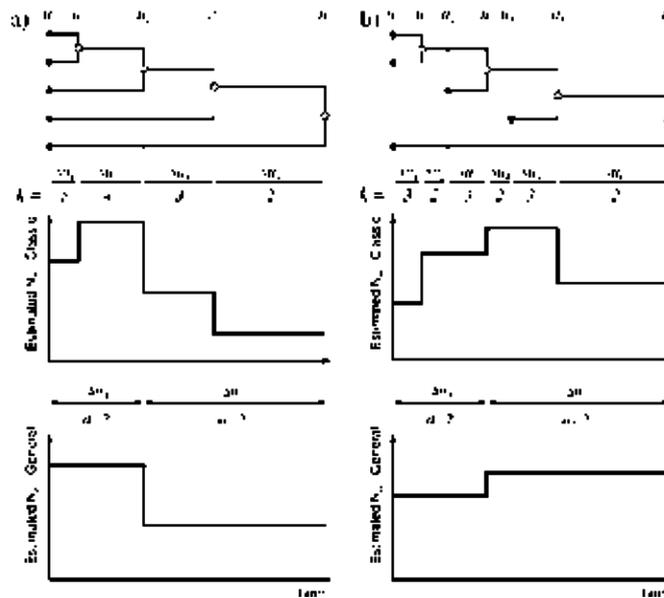


Figure 38. Illustration théorique de 5 individus échantillonnés de façon contemporaine (a) ou asynchrone (b) avec les skyline plots classiques (milieu) et généralisés (bas) associés (abscisse=temps, ordonnée=taille efficace estimée).

5(c) Implémentation MCMC et logicielle : BEAST

L'implémentation non paramétrique reprend la méthode paramétrique précédemment publiée par les mêmes auteurs (**Drummond et al., 2002**) afin d'échantillonner les paramètres Θ et A (groupe ordonné des *taxa* considérés) du *BSP*. La méthode s'applique aussi bien à un jeu de séquences contemporaines qu'à un jeu de séquences datées (échantillonnées à des dates variables, ce qui permet de s'affranchir d'une *prior* sur la profondeur de l'arbre et de relâcher l'hypothèse d'horloge moléculaire).

Le résultat de l'analyse bayésienne du processus de MCMC est alors un échantillon de la distribution postérieure décrit comme suit :

$$f_{het}(\Theta, A, \Omega, g, \mu | D) = \frac{1}{Z} \Pr\{D | \mu, g\} f_G(g | \Theta, A) f_\Theta(\Theta) f_A(A) f_\Omega(\Omega) f_\mu(\mu)$$

Le logiciel BEAST (*Bayesian Evolutionary Analysis Sampling Trees*) est un outil disponible pour tous les systèmes d'exploitation (Mac OS X, Unix/Linux, Windows) de par sa programmation en JAVA. Il est entièrement orienté vers la reconstruction phylogénique enracinée et temps-dépendante par la prise en compte d'horloge moléculaire stricte ou relâchée. BEAST utilise des MCMC pour explorer l'espace des

arbres possibles par un échantillonnage pondéré en fonction de leurs probabilités postérieures.

Cette exploration se fait par des mécanismes permettant des propositions de modification de la topologie des arbres à mesure qu'avance l'analyse markovienne (cf. **Figure 39**).

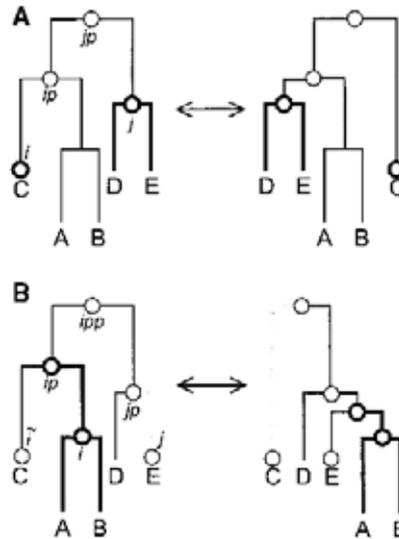


Figure 39. Diagramme de deux mécanismes de propositions utilisés pour modifier la topologie lors d'une analyse par Monte Carlo Markov Chain (Drummond et al., 2002).

Le changement (A) est appelé un *narrow exchange* et est similaire à un échange du plus proche voisin. (B) Ce mouvement, initialement proposé pour les généalogies issues de microsatellites (Wilson & Balding, 1998), consiste en la suppression d'un sous-arbre et son rattachement à une nouvelle branche parente.

La nouveauté de l'approche réside non pas dans le cadre bayésien mais dans le type de données (séquences espacées dans le temps), le nombre important de paramètres estimés et les procédures d'échantillonnage de type MCMC. La Coalescence donnant la fréquence attendue de chaque généalogie dans un modèle de Fisher-Wright, elle peut donc être traitée comme un partie du processus d'observation - en définissant la Vraisemblance des paramètres populationnels -, ou comme *prior* de la généalogie « vraie », par définition inconnue. Les logiciels, qu'ils utilisent le cadre bayésien ou l'approche purement vraisemblance, sont ainsi les mêmes jusqu'au point de

l'utilisation de *prior*.

A la fin de l'analyse, il convient de vérifier la validité des distributions postérieures par examen entre autres de la convergence afin d'éviter l'altération des distributions postérieures par des *priors* non-informatives ou contre-informatives. En effet, cette altération peut mener à des interprétations erronées. Cette vérification peut s'effectuer facilement en examinant graphiquement l'estimation des divers paramètres au cours de la course des chaînes de Markov.

Cette méthode a donné des résultats spectaculaires dans divers domaines : reconstitution de la démographie des Bisons des steppes américains et découverte d'un *bottleneck* sévère datant de l'arrivée des premiers eurasiatiques en Amérique du Nord (**Drummond et al., 2005**), ou encore évaluation du taux d'évolution de l'ADN mitochondrial sur des Manchots de Terre Adélie (**Lambert et al., 2002**). Elle a permis en outre la datation des souches du Virus de l'Immunodéficience Humaine et de l'Hépatite C ayant infecté 111 enfants Lybiens et ainsi de disculper un médecin Libanais et cinq infirmières bulgares (**de Oliveira et al., 2006**).

IV. C. Etude de la polyphylie du Dauphin bleu et blanc en Méditerranée

C.1) Introduction et hypothèses testées

Nous avons vu précédemment que l'échantillon méditerranéen de Dauphins bleu et blanc considéré est caractérisé par des déséquilibres tant au niveau des fréquences alléliques (Hardy-Weinberg) que de la non indépendance apparente des marqueurs entre eux (Déséquilibre de Liaison).

Afin de mieux appréhender l'histoire évolutive du Dauphin bleu et blanc en Méditerranée, nous avons examiné la structure génétique en combinaison avec une inférence phylogénique de la région de contrôle de l'ADN mitochondrial.

Pour ce faire, nous avons tout d'abord testé trois séries d'hypothèses :

- Notre première hypothèse était que les coefficients de consanguinité individuels F_{IS} et coefficients de parentalité (*relatedness coefficients*) de chacun des

sous-échantillons méditerranéens n'étaient pas différents les uns des autres et étaient significativement différents de zéro. En relation avec cette hypothèse, nous avons testé la significativité d'une structure de population suivant les limites de ces mêmes zones d'échantillonnage. Le non-rejet de ces hypothèses sous-tend l'unicité de la population méditerranéenne.

- Notre seconde hypothèse était qu'il existe des preuves d'Isolation par la Distance (*Isolation-by-Distance, IBD*) en tant que force structurant la population méditerranéenne du Dauphin bleu et blanc. Nous avons également testé l'hypothèse liée de l'existence de barrières physiques au flux de gènes.

- Notre troisième hypothèse, alternative aux précédentes, était qu'un certain niveau de sympatrie avec un niveau d'hybridation limité était l'explication la plus plausible des déséquilibres observés.

Pour plus de détails sur les méthodes utilisées, cf. paragraphes précédents et méthodes de l'article.

C.2) Principaux résultats

Pour plus de détails, cf. paragraphe *Results* de l'article n°4.

2(a) Rejet des hypothèses sous-tendues par l'unicité de la population méditerranéenne

La réjection de ces hypothèses implique qu'il n'existe pas une population unique méditerranéenne du Dauphin bleu et blanc. L'existence de plusieurs populations est l'origine vraisemblable de l'effet Wahlund que nous avons précédemment observé.

2(b) Rejet de l'hypothèse d'une Isolation par la distance

Cette hypothèse a été testée à l'aide test de Mantel (**Legendre & Legendre, 1998**) sur les matrices de distances génétiques et géographiques. Le détail de la méthodologie est résumé dans le paragraphe *Material and Methods* de l'article n°3. La corrélation entre les deux matrices est mesurée par un coefficient de corrélation r dont la significativité est évaluée par un test unidirectionnel (*one-tailed Z-test*) de l'hypothèse nulle de corrélation positive ($r > 0$) ou négative ($r < 0$).

Nous avons corrélativement testé, de façon non paramétrique, l'hypothèse de

l'existence de barrières physiques au sein de la Méditerranée qui auraient pu jouer un rôle dans la structure génétique de la population. Cette hypothèse a été testée d'une part au moyen d'une Analyse Spatiale de la Variance Moléculaire (*Spatial AMOVA, SAMOVA*), d'autre part en comparant la fréquence des haplogroupes mitochondriaux de part et d'autre du Déroit de Sicile. L'AMOVA spatiale utilise une triangulation de Delauney et l'algorithme de Montmonnier pour rechercher de manière itérative la structure regroupant des échantillons adjacents qui maximise la proportion de la variance génétique totale qui résulte des différences entre groupes. La comparaison des fréquences des haplogroupes mitochondriaux entre le bassin occidental (GV+GL+LS+NT) et le bassin oriental (IS+SA) a été effectuée au moyen d'un test du T bi-directionnel.

Les principaux résultats de ces analyses sont qu'il n'existe pas d'évidences d'IBD au sein de l'échantillon méditerranéen (*cf. figure 2 de l'article n°4*), que l'analyse par SAMOVA n'a pas pu déterminer de manière non ambiguë le regroupement optimal des sous-échantillons méditerranéens, mais qu'il existe une différence significative dans les fréquences des haplogroupes mitochondriaux entre le bassin occidental et le bassin oriental. Ces résultats vont dans le sens de l'existence d'une structure cachée non déterminée par la géographie physique.

2(c) Echec de la réjection de l'hypothèse de cohabitation sympatrique au sein de la population méditerranéenne

En raison du fait qu'une structure de population cachée peut biaiser les estimateurs traditionnels basés sur une structure définie *a priori* (Weir, 1996), nous avons basé l'analyse suivante sur des méthodes bayésiennes. D'une part, nous avons utilisé une méthode de partitionnement fondée sur un modèle visant à minimiser les déséquilibres gamétique et génotypique. Le modèle suppose l'existence de K populations chacune caractérisée par une série de fréquences alléliques à chaque locus. Chaque individu est assigné à une population (ou plusieurs si son génome s'avère être d'origine multiple - ou *admixed*) avec la probabilité $\Pr(X|K)$, X étant le génotype observé. Un vecteur Q, appelé Coefficient d'*admixture*, est alors calculé, définissant la proportion du génome de chaque individu dont l'origine putative est située dans chacune des populations inférées. Cette méthode est implémentée dans le

logiciel STRUCTURE (Pritchard *et al.*, 2000).

Par ailleurs, nous avons utilisé une approche basée sur la Coalescence de généalogies dans un contexte bayésien basé sur l'évolution de chaînes de Markov (MCMC) échantillonnées par un algorithme de Metropolis-Hastings pour estimer conjointement les flux géniques (*gene flow*), le taux de croissance - présumant d'une croissance exponentielle, et les tailles effectives de chacun des groupes (*cf.* le paragraphe *Material and Methods* de l'article n°3 pour plus de détails). L'algorithme de Metropolis-Hastings (Hastings, 1970; Metropolis *et al.*, 1953) est utilisé pour créer une chaîne de Markov qui converge sur une séquence d'échantillons tirés d'une Loi de probabilité difficile à échantillonner directement. Cette méthode, implémentée dans le logiciel LAMARC (Beerli & Felsenstein, 2001), permet d'obtenir des intervalles de crédibilité (pendant des intervalles de confiance de l'inférence statistique fréquentiste en statistique bayésienne) à 95% ainsi que des distributions de probabilité postérieure des taux de migration. Par ailleurs, elle permet de découpler les migrations asymétriques. La limitation de la méthode est qu'elle ne permet pas de différencier migration récente de migration ancienne. Nous avons un aperçu de cette « ancienneté » par le coefficient *d'admixture* Q .

2(d) Phylogénie et inférence du temps écoulé depuis le plus récent ancêtre (TMRCA)

L'existence d'un certain degré de polyphylie au sein de l'espèce *Stenella caeruleoalba* à l'intérieur de la Méditerranée a été testée au moyen de la reconstruction phylogénique des séquences de la région de contrôle de l'ADN mitochondrial. Après alignement à l'aide de l'algorithme présenté plus haut CLUSTALW (Thompson *et al.*, 1994), et choix du modèle le plus adapté de mutation (Posada & Crandall, 1998), nous avons inféré la phylogénie par trois méthodes : une approche de distance par Neighbour-Joining (Saitou & Nei, 1987), une approche par maximum de vraisemblance (Felsenstein, 1981) ainsi qu'une approche bayésienne (Ronquist & Huelsenbeck, 2003). De même, les T_{MRCA} (Time to the Most Recent Common Ancestor) des différents lignages d'haplotypes ont été inférés au moyen de deux méthodes : l'une semi-paramétrique utilisant une maximisation de la vraisemblance pénalisée (Sanderson, 2002), l'autre bayésienne estimant conjointement les taux de

mutation (**Drummond *et al.*, 2002**). Pour l'une et l'autre méthode, nous avons calibré les nœuds comme indiqué dans la **Figure 40** ci-dessous.

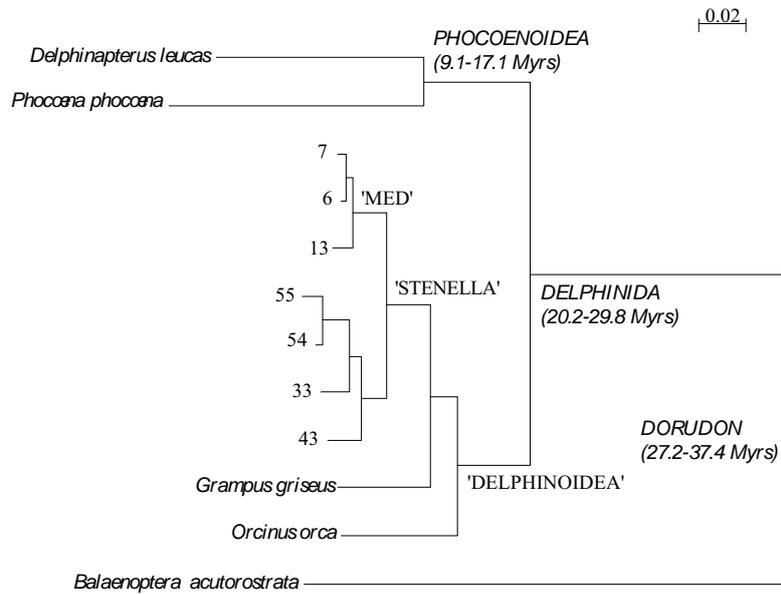


Figure 40. Calibrations des nœuds pour l'estimation des TMRCA.

Les nombres correspondent aux haplotypes du tableau S3 et de la figure 3 de l'article n°4. Les noms en italiques sont les espèces utilisées comme groupe externes (*Grampus griseus* séquencé dans le laboratoire, *Balaenoptera acutorostrata* (GenBank #[AY230267](#)), *Delphinapterus leucas* (GenBank #[U18117](#)), *Phocæna phocæna* (GenBank #[AJ554063](#)) and *Orcinus orca* (GenBank #[M60409](#))). 'MED', 'STENELLA' & 'DELPHINOIDEA' sont les nœuds à dater, PHOCOENOIDEA, PHOCOENOIDEA & DORUDON (fossile proche de l'ancêtre commun des *Mysticeti* et *Odontoceti*) sont les points de calibration dont des intervalles sont donnés entre parenthèses sur la base de calibrations à partir des fossiles (Muizon, 1988; Thewissen & Williams, 2002) et moléculaires (Nikaido *et al.*, 2001).

Article n°4 : Macé M, Bourret V, Crouau-Roy B « Polyphyly in the Mediterranean Striped dolphin (*Stenella coeruleoalba*) and the Messinian Salinity Crisis », en révision dans la revue PLoS ONE

1 Title
2 Polyphyly in the Mediterranean Striped dolphin (*Stenella coeruleoalba*) and the Messinian
3 Salinity Crisis
4

5 Authors and Affiliations
6 Matthias Macé^{1,2§}, Brigitte Crouau-Roy¹

7
8 ¹UMR 5174 UPS/CNRS EDB “Evolution et Diversité biologique”, Bât 4R3b2, Université
9 Paul Sabatier, 118 route de Narbonne, 31062 TOULOUSE cedex 9, France

10 ²Centre de Physiopathologie de Toulouse Purpan, INSERM U563, CHU Purpan, F-31300
11 Toulouse, France

12

13 [§]Corresponding author

14 Telephone: +33 561 556259

15 FAX : +33 561 557327

16 e-mail: matthias.mace@orange.fr

17

18

19 Abstract

20 *Background.* Marine vertebrates are models of major interest to study the impact of global
21 change on biodiversity. A disruptive geological change, the Messinian Salinity Crisis, has
22 been extensively investigated but little is yet known about its palæoecology and the survival
23 of higher marine Vertebrate fauna during this period. Among them, the Mediterranean Striped
24 dolphin is an interesting model because it is both widespread and its systematics
25 (parapatry/sympatry, polyphyly,...) is still unresolved. Moreover, Mediterranean Striped
26 dolphin stock assessment is necessary as this population faces numerous threats such as
27 *Morbillivirus* infection and fisheries. Here, we provide the first assessment of
28 sympatry/parapatry in the widely diversified genus *Stenella*.

29 *Methodology/Principal findings.* We analysed mitochondrial DNA (mtDNA) and nuclear
30 markers of dolphin samples from the Mediterranean Sea, the North-Eastern Atlantic and the
31 Pacific Oceans. We first tested the following groups of hypotheses: i) there are significant
32 differences in inbreeding and genetic structure among Mediterranean sampling areas ii)
33 genetic differentiation is explained by isolation as a function of distance and/or physical
34 barriers. After global rejection of these working hypotheses, we propose that the polyphyletic
35 aspect of Mediterranean Striped dolphin genetic diversity is best explained by an ongoing
36 process of variance. Indeed, the mtDNA revealed two highly divergent haplogroups within
37 the Mediterranean (Groups 1 and 2). Group 1 is characterised by much lower nucleotide
38 diversity as compared to Group 2 and other geographic areas. Genotypic and gametic
39 disequilibria were not observed with microsatellites in either of the two groups. Bayesian
40 estimation of admixture confirmed the distinct nuclear gene pools and a low proportion of
41 admixed individuals. Group divergence was dated to the late Miocene. All these features
42 support the existence of two distinct Mediterranean populations/subspecies. They might be

43 the result of allopatric evolution during the Mio-Pliocene and behavioural and/or
44 immunogenetic factors that would have prevented more admixture as expected by the
45 secondary contact, than observed.
46 *Conclusions/Significance.* Two yet undescribed striped dolphin populations/subspecies are
47 reported in the Mediterranean. In addition, these may be the first Mediterranean higher
48 Vertebrates to have survived the Messinian Salinity Crisis. These results suggest that the
49 Striped dolphin is a good model for multidisciplinary research (molecular and behavioural
50 ecology, palaeoecology and systematics) on global change and conservation.

51

52 Introduction

53 Given the challenges of global changes for biodiversity, model species are needed to assess
54 how they were able to go through past major ecological changes. Biodiversity is facing a
55 major extinction crisis characterized by an accelerating rate (**Mace *et al.*, 2005; Vitousek *et***
56 ***al.*, 1997**). Therefore, emergency conservation strategies based on phylogenetic
57 considerations are necessary (**Isaac *et al.*, 2007**). Local water cycle changes are already taking
58 place in the Mediterranean area and could have local (possible marine life disruption around
59 the Gibraltar salinity valve) and global consequences (**European Commission, 2007**). They
60 will certainly have major impacts on human welfare (**Balvanera *et al.*, 2001; Worm *et al.*,**
61 **2006a**).

62 In this context, biological models of past major environmental crisis are of special interest for
63 understanding the decline in biodiversity: the Mediterranean underwent such a crisis and is
64 still going through great human-induced pressure (urbanisation, maritime traffic, pollution
65 and heavy fisheries activities) that threatens marine fauna (**Silvani *et al.*, 1999**). In addition,
66 its geological characteristics (*e.g.* semi-closed sea, highly segmented water body (**Gili, 2001**))
67 and high biological diversity and endemisms (**Myers *et al.*, 2000**) render Mediterranean
68 species interesting research models for monitoring global changes.

69 One example, the so-called Messinian Salinity Crisis (MSC; (**Hsü *et al.*, 1973**)), induced by
70 the dramatic closure of the communications between the Atlantic and Mediterranean basins
71 during the late Miocene, led to endemisms, and abyssal fauna impoverishment (**Emig &**
72 **Geistdoerfer, 2004**). This major environmental change that faced the Mediterranean
73 accompanied global large-scale changes such as ongoing Northern hemisphere glaciations
74 (**Weijermars, 1988**). To date, a limited number of species are considered to have survived
75 this event by either dispersal or adaptation (**Boury-Esnault *et al.*, 1992; Durand *et al.*, 2003;**
76 **Gilles *et al.*, 1998; Gysels *et al.*, 2004; Jaume & Boxshall, 1996; Maldonado & Uriz, 1995;**
77 **Penzo *et al.*, 1998; Wilke, 2003**). However, uncertainties remain about the extent of *taxa* and
78 therefore the palaeoenvironments during such dramatic periods. It would therefore be
79 compelling to seek whether and how marine species, especially higher Vertebrates, traversed
80 such abrupt ecological changes.

81 Among Mammals, the Striped dolphin *Stenella coeruleoalba* is a good model species for this
82 purpose since it is the most common dolphin species in the Mediterranean (**Notarbartolo di**
83 **Sciara, 2002b**), with a worldwide presence (**Archer II & Perrin, 1999a**) and it is an
84 opportunistic feeder (**Aguilar, 2000b**). Moreover, although this species is not considered
85 endangered, it could be at risk in the Mediterranean Sea because it recently went through
86 severe epizootics (**Domingo *et al.*, 1990**). A sister-species, *Delphinus delphis*, disappeared
87 rather suddenly from the Western Mediterranean in spite of having been prominent until
88 recent decades (**Bearzi *et al.*, 2003**).

89 Morphological variability in dolphins has been revealed within Mediterranean populations
90 (**Calzada & Aguilar, 1995; Pertoldi *et al.*, 2000**). In addition, in a previous study, we
91 observed genotypic and gametic disequilibrium, suggesting, among others hypotheses, the
92 existence of two distinct Mediterranean populations (**Bourret *et al.*, 2007**).

93 To further unravel the Striped dolphin evolutionary and demographic history we here examine
94 the genetic structure at individual and population levels as well as the phylogenetic and
95 coalescence-based population parameter inference on both mitochondrial DNA (mtDNA,
96 control region) and nuclear microsatellite markers. European samples from the Mediterranean

97 and North-Eastern Atlantic were also studied and a small North Pacific sample was added as
98 an outgroup.

99 The following hypotheses were tested. First, the null hypothesis that sampling areas,
100 individual inbreeding (F_{IS}) and relatedness coefficients do not differ from each other but
101 significantly differed from zero. Related to this question, we also tested the significance of the
102 genetic structure based on Mediterranean geographic sampling locations. These hypotheses
103 were rejected, suggesting the significant presence of relatively inbred individuals, or of a
104 mixture or Wahlund effect within sampling locations and in the Mediterranean overall.

105 Our second hypothesis was that there was evidence of Isolation by Distance (IBD) within
106 Mediterranean samples. We also tested the related hypothesis that physical barriers within the
107 Mediterranean could play a role in within-basin genetic structure. This hypothesis was also
108 rejected, but a significant difference in distribution of individuals whose mitochondrial
109 control region haplotype belongs to two divergent haplogroups between eastern and western
110 basins was detected, suggesting some level of female philopatry.

111 Finally, a third hypothesis was tested, alternate to the previous ones, *i.e.* that sympatry with a
112 significant but limited level of admixture was the most likely explanation for the observed
113 genetic structure pattern. This hypothesis received high support by Bayesian admixture
114 analyses as well as from phylogenetic-based analyses (divergence times and demographic
115 inferences). The question about the age of group divergence in relation to the radiation of
116 Delphinidæ and a possible vicariant event is discussed.

117 Our results supporting a deep-rooted polyphyly may result in a new view of MSC
118 palæoecology with respect to survival of macrofauna. They are of relevance to other marine
119 species regarding sympatric/parapatric differentiation (**Andrews, 1911; Borsa & Quignard,**
120 **2001; Nielsen *et al.*, 2003; Ross, 1977b**) and theoretical issues regarding vicariance and
121 dispersal processes. The presented animal model also displays specific characteristics
122 significant for taxonomical-paleontological issues in Delphinidae (**Escorza-Treviño *et al.*,**
123 **2002; Leduc *et al.*, 1999**). From a wider perspective, we discuss how maintaining reciprocal
124 monophyly might be triggered by either biological or behavioural processes in a higher
125 Vertebrate. Finally, comparing the precise present and past population status is meaningful for
126 efficient conservation policies *via* the assessment of Evolutionary Significant Units (ESU) in
127 the framework of the Mediterranean “International Sanctuary for Cetaceans” aiming to
128 regulate human activities in one of the world’s most crowded areas (**Anonymous, 1999**).

129

130 Results

131

132 Panmixia, inbreeding and genetic structure

133 *Single locus statistics*

134 *Overall data.* Microsatellite *loci* polymorphism are summarized in Table S1

135 *Per sample statistics.* For each oceanic basin, we report the *loci* that depart from
136 Hardy-Weinberg expectations (HWE) frequencies (supplementary Table S1) and the pairs of
137 *loci* in potential genotypic disequilibrium (Table 1). The Mediterranean samples displayed the
138 highest level of disequilibrium with 3 out of 5 *loci* whose frequencies deviated from HWE
139 and 7 out of 10 pair-wise *loci* comparisons showing Linkage Disequilibrium (LD).

140 *Inbreeding and relatedness coefficients*

141 We first examined if inbreeding coefficients F_{IS} , reflecting the level of genetic
142 inbreeding within subpopulations, differed among sampling locations. The two-sided test
143 performed after 1000 permutations was significant ($p=0.01$). However, the 99% confidence
144 interval, obtained after 1000 bootstrap replicates over *loci*, for whole Mediterranean F_{IS} was [-
145 0.031;0.326]. There is therefore little or no evidence that global F_{IS} significantly differed from
146 zero. Likewise, the relatedness coefficient for the whole Mediterranean sample did not
147 significantly differed from zero (99%C.I. [-0.037;0.023] after 1000 bootstrap over *loci*) and
148 there were no significant differences between samples (two-sided p-value after 1000
149 permutations $p=0.34$).

150 These results suggest the existence of relatively inbred individuals, non-random
151 mating, and/or the possibility of mixing (*i.e.* Wahlund effect) within each of the sampling
152 locations.

153 *Genetic structure*

154 At the global level, we previously demonstrated (Bourret *et al.*, 2007) that significant
155 differentiation was detected between the Mediterranean and the other basins. At the
156 Mediterranean level, an analysis of F_{ST} revealed no evidence of population genetic structure
157 between samples within the Mediterranean : none of the pairwise test for F_{ST} differentiation
158 was found to be significant (Table 2(a)). Moreover, distribution of θ and f , estimators for F_{ST}
159 and F_{IS} obtained after jackknifing over *loci* showed that F_{ST} was slightly below F_{IS} ($\theta=$
160 0.003 ± 0.007 and $f=0.106\pm 0.077$).

161 When each basin is considered separately, hierarchical AMOVAs indicated that in all
162 cases the variance distribution is skewed towards subpopulation level but the this was a little
163 more evident in the Mediterranean sample. A two-level hierarchical AMOVA, with oceanic
164 basins and sampling locations within basins as the two levels showed that less than 1% of the
165 genetic variance is distributed among populations whereas more than 90% of the genetic
166 variance is attributable to within sampling location level (Table 2(b)).

167
168 All these findings concord to reject the first hypotheses i) of the existence of genetic
169 structure according to sampling locations and ii) of no significant difference between the
170 empiric individual inbreeding coefficient and the population inbreeding coefficient. This
171 suggests the existence of mixing populations on the scale of the whole Mediterranean. This
172 so-called Wahlund effect could result from IBD and/or admixture.

173
174 Isolation by distance, Spatial AMOVA and physical barrier to gene flow

175 *Isolation by distance.* The analysis based on pairwise Slatkin's M estimates suggests
176 that there is no correlation between genetic and geographical distances when the 6 sampling
177 locations are considered separately: one-tailed p-values for null hypothesis of negative or
178 positive correlation were all insignificant (Table S2) and plotting M against various
179 transformations of geographic distance displayed low correlation coefficients r^2 (Figure 2).

180 *Spatial AMOVA.* We applied the SAMOVA algorithm to mitochondrial and
181 microsatellite data first including Atlantic and Pacific populations and then considering
182 Mediterranean sample alone. Computation was performed over 10000 iterations searching for
183 $K \in \{2, \dots, 8\}$ groups from each of 500 random initial conditions. For each K, the
184 configuration with the largest F_{CT} values was retained as the best clustering of sampling
185 populations. The pattern of F_{CT} values as a function of K was therefore examined: we checked
186 the number of groups necessary for F_{CT} to reach a plateau. Considering the global sample, a
187 plateau of F_{CT} values was reached for $K=4$ ($F_{CT}\sim 0.02$). Regardless of these results, SAMOVA
188 performed on only the Mediterranean sample did not unambiguously identify the number (K)
189 of groups of populations that had the highest differentiation among groups F_{CT} . Indeed, no
190 plateau was reached instead the F_{CT} values followed a decreasing curve form $K=1$ (0.13) to
191 $K=5$ (0.009) when running the analysis on microsatellite data and an increasing curve form
192 $K=1$ (0.09) to $K=5$ (0.27) when considering mitochondrial sequences.

193 *Physical barrier to gene flow.* A significant difference (T-value of difference: -2.987;
194 df-t: 26; double-sided p-value=0.0058) in the frequencies of Group 1 individuals (please see
195 below in phylogeny for the definition of mitochondrial haplogroups) was observed between
196 the Eastern basin and other Mediterranean locations with respectively around 90% and 50%
197 (see also pie-charts in Figure 1).

198 These results are consistent to reject the second hypothesis of an effect of isolation by
199 distance but are not able to reject the hypothesis of a hypothetical physical barrier to
200 mitochondrial gene flow, suggesting some level of matrilinearity.

201
202 Evolutionary relationships between mitochondrial haplotypes

203 Mitochondrial control regions were amplified in 74 individuals (Pacific $n=7$, Atlantic
204 $n=19$, Mediterranean $n=48$). Sequences were deposited in the EMBL database under the
170

205 accession numbers AM498667 to AM498740. The alignment of 627-628 bp sequences
206 allowed the identification of a total of 59 distinct haplotypes (Pacific n=6, Atlantic n=19,
207 Mediterranean n=34, see Table S3 for details) and therefore displayed high values of
208 haplotype diversity (overall $h=0.988$, between 0.952 for the Pacific and 1 for the Atlantic).
209 The sequences displayed three indels and 89 variable sites; four were 3-state characters and
210 one was a 4-state character. Among the polymorphic sites, 47 were informative for
211 parsimony. Nucleotide diversity showed medium values for the total sample ($\pi=0.022$), and it
212 varied between 0.017 and 0.024 in the Mediterranean and Atlantic samples respectively. Two
213 groups of haplotypes could be unambiguously identified: the haplogroup 1 (see below and
214 Figure 3), only found in the Mediterranean, and the haplogroup W which was worldwide
215 distributed. Within the haplogroup W, we thereafter identify the group of Mediterranean
216 samples as Group 2.

217 The Table 3 summarises diversity indices at the three hierarchical levels: whole
218 sample, geographic samples and both groups found in the Mediterranean. The mean number
219 of nucleotide differences between Group 1 haplotypes and the others was 17, whereas it was
220 only 11.3 within the whole dataset. This divergence is corroborated by the 12 apomorphic
221 sites characterising the haplogroup 1. Furthermore, the Mediterranean bimodal mismatch
222 distribution of pair-wise differences between haplotypes disappeared when considering Group
223 1 or Group 2 alone. The whole Mediterranean sample showed slightly lower diversity indices
224 than other geographic samples and even lower values were observed in Group 1, while Group
225 2 values were comparable to the Atlantic and Pacific values.

226 The three phylogenetic approaches (NJ, ML and Bayesian) gave the same topology
227 pattern: *i.e.*, two well-supported clusters. We focus hereafter on the Bayesian approach as it
228 was shown to be the most reliable for testing clustering exclusion (**Douady *et al.*, 2003**).
229 MODELTEST showed that Hasegawa-Kishino-Yano (HKY; (**Hasegawa *et al.*, 1985**) model
230 best fits to the data (invariant sites=64%, gamma-parameter=0.409, transition/transversion
231 rate=8.575). The Group 1 (58.8% of the haplotypes and 63.8% of the individuals from the
232 Mediterranean) diverged significantly from all other haplotypes (Bayesian posterior
233 probability $Bp=0.91$, Figure 3). Within the Group 2, no significant geographic structure was
234 observed since haplotypes from diverse origins clustered together (*e.g.* 5 Atlantic and 1
235 Mediterranean, $Bp=0.95$) and Pacific haplotypes did not. The median-joining network among
236 haplotypes furthermore specifically showed that Group 2 haplotypes are mainly located at the
237 periphery of the network, suggesting a recently derived origin (see Figure S1).

238

239 Bayesian estimation of coefficients of admixture

240 Bayesian analysis of coefficients of admixture Q (**Pritchard *et al.*, 2000**) on the whole
241 sample (138 individuals) confirmed the co-existence of two populations within the
242 Mediterranean, with little if any admixture. This structure was more evident (*i.e.* $\text{Prob}(\text{data}|\mathbf{K})$
243 above 0.8 as given by Bayes' rule) using learning samples (mitochondrial Group 1 and Pacific
244 individuals prior probability of belonging to their respective population set at 1) otherwise,
245 probabilities, although significant, were less elevated (above 0.5). The probability
246 $\text{Prob}(\text{data}|\mathbf{K})$ was maximum for a number of populations $K=3$ ($\ln P=-2530$, $\text{prob}(K=3)=0.99$),
247 and considering the admixture model with correlated frequencies. Figure 4 shows the
248 admixture coefficients of the individuals given their respective geographic origin. As a
249 consequence, we no longer observed deviation from HW expectations in these populations
250 and LD was found in only one pair of *loci* in comparison to 7 pairs in the whole
251 Mediterranean sample (Table 1). Among the 68 Mediterranean individuals with complete
252 genotype, 24 individuals were assigned to the Group 1 ($Q(1) > 0.5$), 33 to the Group 2 ($Q(2) >$
253 0.5) and 11 not assigned ($0.1 < Q(n) < 0.4$). This analysis gave the following proportions of
254 each Group: Group 1 (25-47%, 95% C.I.), Group 2 (37-60%) and not assignable or admixed
255 (9-26%).

256

257 Inference of Time to the Most Recent Common Ancestor

258 Assuming the fossil calibrations, both Penalised-Likelihood (PL) and Bayesian
259 Uncorrelated Relaxed-Clock (UCR) methods gave high T_{MRCA} estimations *i.e.* within the

260 Pliocene for the Group 1 haplotypes while the *Stenella* haplotype divergence is placed in late
261 Miocene and the radiation of Delphinidæ/Delphininæ in the middle Miocene (see Figure 5 for
262 UCR and Table 4 for PL and UCR methods). To obtain these results, three independent
263 chains were run for 10 million steps each after running preliminary shorter chains in order to
264 optimise the scale factors to an acceptance rate probability of 0.25 for each parameter as
265 advocated by the authors (**Drummond et al., 2002**). A graphical examination of the three
266 runs combined suggested that the chains sampled from similar probability density
267 distributions, and that mixing during the search was adequate. The quality of the Monte Carlo
268 Markov Chain (MCMC) search was also ascertained by ESS values, all well over the
269 recommended value of 100. Bayesian estimations gave invariably the same ranges of values
270 whatever the priors on node heights and models of relaxed clock (exponential or log-normal).
271 Moreover, while performing Bayesian search without any priors (with strict and relaxed
272 clocks), we roughly obtained the same results, with quite greater support intervals and with
273 the need of much longer chains. Within this analysis, molecular clock values ranged between
274 1.10^{-9} and 1.10^{-8} substitutions/site/year following Cetacean clade branches. These results are
275 in agreement with the molecular clock values below 10^{-8} given for *Cetacea* by robust
276 estimates (**Kitazoe et al., 2007**). Given this deep rooting of the Striped dolphin's haplotypes,
277 we performed a phylogenetic inference including *in silico* mitochondrial DNA haplotypes from
278 other Delphinids (*Delphinus* spp., *Tursiops* spp., other *Stenella* spp. and *Lagenorhynchus*
279 spp.) together with our haplotypes. The Stripe dolphin polyphyly was highlighted by the
280 much higher divergence of haplogroup 1 from other Striped dolphin haplogroups than from
281 other species' haplotypes (Figure 3, bottom-left).

282

283 Demographic inference of Mediterranean Striped dolphin in Bayesian
284 coalescent frameworks

285 Demographic patterns and gene flow within basins confirmed the existence of a strong barrier
286 to gene flow between the studied Striped dolphin populations.

287 LAMARC's Bayesian estimates of gene flow between the four populations are summarised in
288 Table S4. They strongly support an inner Mediterranean structure as mentioned above in
289 Bayesian coefficient of admixture estimates. The level of gene flow was significantly higher
290 between the Atlantic and the Mediterranean Group 2 than between the oceanic basins. In
291 addition, there was a much stronger barrier to gene flow between Group 1 and Group 2 than
292 between the Atlantic and Group 2 and this barrier is comparable to the one observed between
293 the oceanic (Pacific and Atlantic) samples. The gene flow between Mediterranean stocks
294 displays a sex-biased asymmetry: mitochondrial gene flow appears significantly higher in
295 mitochondrial than in nuclear markers in the Group 2 towards Group 1 direction but it is very
296 weak in the opposite direction for both nuclear and mitochondrial markers (Figure S2).
297 Effective population size estimates gave about the same order of magnitude for both
298 Mediterranean populations (0.02 and 0.06) and were not low enough to suggest a long-lapsed
299 bottleneck (**Hauser et al., 2002**). On the Bayesian Skyline plot, the oceanic population
300 revealed a high expansion (~10-fold) whereas Mediterranean populations showed a
301 noteworthy quite constant size (Figure 6). Another non-parametric approach based on reverse-
302 jumping MCMC (**Opgen-Rhein et al., 2005**) gave roughly the same trends but with a more
303 fluctuating shape for the Mediterranean Pleistocene (available upon request).

304

305 Discussion

306 In the present study, we provide evidence that Striped dolphin populations exhibit
307 higher levels of population structure at local (Mediterranean) levels than at worldwide (inter-
308 oceanic) level. This population structure did not appear to follow sampling areas (the genetic
309 variance was mainly assignable to covariance at the individual level) and Mediterranean
310 geographic subsamples display significant levels of genotypic and gametic disequilibrium. In
311 addition, neither isolation by distance nor geographic barriers could explain this Wahlund
312 effect even if a difference in mitochondrial DNA haplotype frequencies exists between the
313 western and eastern basins of the Mediterranean. However, we concede that the limited

314 number of samples and microsatellite *loci* potentially limit the scope the study, *e.g.* the
315 statistical power to detect genetic structure between the Atlantic and the Pacific samples.

316 Since the two hypotheses of no genetic structure or genetic structure arising from
317 geography were rejected, we hereafter propose an alternate explanation for the observed
318 pattern. By investigating individual coefficient of admixture, we revealed that this population
319 subdivision encompassed the whole Mediterranean. By combining it with mitochondrial DNA
320 haplotype phylogenetic reconstruction, greater support was provided, increasing the power to
321 discriminate between the two populations by using them as priors: marginal posterior
322 probabilities $\text{Prob}(\text{data}|\text{K})$ then exceeded 0.8 (instead of 0.5 without). This might be viewed
323 potentially as too influential a role for the priors but as stated by the software's authors
324 (**Rosenberg *et al.*, 2005**), if this influence had been too strong compared to the information
325 given by the markers, it would have lead to the same posterior assignation of all the
326 individuals belonging to the same haplogroup, which was not the case since STRUCTURE
327 would have distributed the membership coefficients of all individuals evenly across clusters
328 rather than assigning each individual a membership coefficient. Moreover, T_{MRCA} inference
329 confirmed an ancient split within the haplotype tree dating back to the Miocene. All these
330 arguments make a sympatric coexistence of two well-differentiated populations conceivable.
331 Similar results, mainly a putative inshore/offshore differentiation, were reported in a study
332 investigating the patterns of social kinship in a sample of live free-ranging Mediterranean
333 Striped dolphin (**Gaspari *et al.*, 2007**). Our study, further analysing spatial and temporal
334 correlates of genetic structure, demonstrates an unexpected polyphyly whose possible origins
335 and factors of persistence throughout time are discussed below.

336
337 Vicariance vs dispersal in Mediterranean Striped dolphin evolution: an
338 echo of the Messinian Salinity Crisis ?

339 Vicariance and dispersal are two important contrasting evolutionary forces (**Sanmartin,**
340 **2003**). Studying them within a Bayesian molecular clock framework provides essential
341 knowledge for studying global changes (**Kaustuv *et al.*, 2001; Lieberman, 2000; Voelker,**
342 **1999**). We tried to give new insights on vicariance by using an original biological model, a
343 highly mobile marine Vertebrate, in relation to one of the major ecological crises to which it
344 was exposed, the Messinian Salinity Crisis.

345 The analysis of the mitochondrial DNA revealed an unexpected deep-rooted polyphyly
346 (Figure 5) and global genetic diversity values ($\pi \sim 0.02$, $h \sim 0.95$) akin to those exhibited by
347 pelagic Bottlenosed dolphin populations (**Natoli *et al.*, 2005**) but greater than those observed
348 in other Cetacean and Tuna populations with low effective size (**Carlsson *et al.*, 2004;**
349 **Rosenbaum *et al.*, 2000**). Individuals bearing haplotypes belonging to one of the two
350 haplogroups, namely the Group 1 ($\sim 40\%$ of the sample), only dwell in a limited area of the
351 Mediterranean basin (Figure 1). Since gametic and genotypic disequilibria observed in the
352 whole Mediterranean sample disappear in each of these two groups, they can be considered as
353 separate genetic stocks coexisting in sympatry or parapatry. This is confirmed by the number
354 of admixed individuals ($\sim 10\%$ of the sample), surprising low for a pelagic wide-ranging
355 species, as well as by the similarly deep divergences observed in the Y chromosome
356 amelogenin gene observed on the same sample (**Macé & Crouau-Roy, 2008**). Of note, since
357 the split predates the radiation of the subfamily Delphininae, a few million years ago (**Xiong**
358 ***et al.*, 2009**), and therefore occurred in a *Stenella/Delphinus/Tursiops* ancestor (T_{MRCA} ranges
359 in the late Miocene – 8 ± 3 Myrs ago), we should not necessarily expect reciprocal
360 monophyly for all of these species, though it is possible for some. Group 1 haplotypes could
361 also be observed in other extant species: this is in accordance with the slow lineage sorting
362 requiring between N and $4N$ generations, where N is the effective size for the largest
363 population (**Neigel & Avise, 1986**).

364 Polyphyly can have various origins such as imperfect taxonomy, inadequate
365 phylogenetic information, interspecific hybridisation, paralogy or incomplete lineage sorting
366 (reviewed in (**Funk & Omland, 2003**)). The available information on the Striped dolphin's
367 natural history (**Aguilar, 2000b; Archer II & Perrin, 1999a; Perrin *et al.*, 1994b**)
368 potentially rules out the first four hypotheses. The most likely explanation, an incomplete

369 lineage sorting, may be the result of at least three different *scenari* following the degree of
370 vicariance and dispersal components: (1) a recent divergence of sympatric populations (**Funk**
371 **& Omland, 2003**) (both weak vicariance and weak dispersal) as observed *e.g.* in the Atlantic
372 Bigeye Tuna (**Alvarado Bremer et al., 1998**), (2) an allopatric evolution followed by range
373 expansion and secondary contact with admixture (**Posada & Crandall, 2001**) (weak
374 vicariance and strong dispersal) and (3) an allopatric evolution followed by a secondary
375 contact with little or no admixture (strong vicariance). The first two hypotheses are poorly
376 supported as they would entail only little divergence between both Mediterranean groups.
377 Indeed, the deep Group 1 divergence, together with the low level of admixture between
378 Mediterranean groups, suggests that the first scenario (sympatric split) is unlikely. Group 2
379 haplotypes are more likely located on the tips of a median-joining network, suggesting the
380 scenario 2 but the limited gene flow between both groups is inconsistent with this second
381 hypothesis.

382 Our best explanation for the observed pattern is therefore an allopatric evolution due
383 to a discontinuity in range followed by a secondary contact with limited admixture, leading to
384 the coexistence of two highly differentiated sympatric/parapatric genetic stocks in the
385 Mediterranean. It is interesting to question the origins of their deep divergence (strong
386 vicariance), and the factors that have maintained it over time. By examining the degree of this
387 intra-specific divergence, we are able to roughly include species from the Atlanto-
388 Mediterranean basin into three categories: i) species with the most recent divergence
389 (Holocene), probably triggered by weak ecological discontinuities (*e.g.* in Lobster (**Stamatis**
390 **et al., 2004**)); ii) species with well defined Atlantic and Mediterranean populations possibly
391 originating in Pleistocene Glacial *refugia* (most of the Fishes and Crustaceans *e.g.* (**Almada**
392 **et al., 2001; Bargelloni et al., 2003; Zane et al., 2000**)) likewise European terrestrial fauna
393 (**Godoy et al., 2004; Hewitt, 2001**); iii) species undergoing cryptic speciation under the
394 pression of strong Vicariance, possibly begotten by the MSC (*e.g.* Sand Goby and Sponges
395 (**Gysels et al., 2004; Sole Cava et al., 1991**)). It seems that Mediterranean Striped dolphin
396 falls within the last category since the T_{MRC} is at a higher value for intra-specific lineages,
397 being comparable to inter-specific divergence in *Mysticetes* (**Sasaki et al., 2005**).
398 Since coalescence time should predate the geological event leading to a range discontinuity
399 (**Avise, 2000**), we postulate that the MSC may be the event founding this divergence.

400 The MSC, one of the most dramatic episodes of oceanic changes, occurred about 5.5
401 Myrs ago (**Krijgsman et al., 1999**). Lithospheric movements closed atlanto-mediterranean
402 communications (**Duggen et al., 2003**), inducing negative water budget that dried up the
403 Mediterranean into shallower hypersaline (**Blanc, 2000**) and oligohaline lakes such as
404 Panonian and Euxinic basins (**Clauzon et al., 2005**) (See Figure 1). Group 1 ancestors could
405 have found refuge in these freshwaters (geographically corresponding to where the Group 1
406 reaches its highest frequencies, Figure 1), as the Adriatic Sand Goby, currently exhibiting
407 cryptic speciation. Likewise, river dolphins seemed to owe the survival of their lineages to a
408 secondary riverine adaptation from marine species earlier in the Miocene (**Cassens et al.,**
409 **2000**). It is also noteworthy that marine Myctophidae fishes, a major prey of dolphins from
410 the genus *Stenella* (**Archer II & Perrin, 1999a; Perrin, 1998**), were found in Messinian
411 layers in the Italian peninsula (**Carnevale et al., 2006**). They might have allowed the
412 persistence of the Group 1 lineage whose low diversity may be explained by the isolation
413 during MSC and possible genetic drift (see Bayesian estimation of allelic distributions in each
414 group and putative ancestral distribution in Figure S3). Interestingly, the Bottlenosed dolphin,
415 species showing subspecies differentiation, does not show genetic structure in the
416 Mediterranean in spite of physical barriers to dispersal (**Natoli et al., 2005**). Hence, because
417 of the deeper divergence in the Striped dolphin, we propose that Group 1 and Group 2 should
418 be considered as subspecies. Indeed, theoretical works showed that speciation can occur
419 despite the presence of moderate gene flow, and that low to moderate gene flow may actually
420 promote adaptive divergence (**Church & Taylor, 2002; Rieseberg et al., 2003**).

421
422 Maintaining two sympatric Striped dolphin subspecies: Behavioral Ecology,
423 Immunogenetics, Palaeoecology and Systematics implications

424 The limited gene flow between these Mediterranean putative subspecies raises the
425 question of the nature of the barriers that limiting admixture over long periods of time: are
426 they of ecological, behavioural and/or genetic nature?

427 In several other Cetacean species, ecological and behavioural features maintain
428 parapatric or sympatric divergent populations (*e.g.* by co-inheritance of cultural and genetic
429 traits favoured by philopatric females (**Rendall & Whitehead, 2001**)). Indeed, Groups 1 and
430 2 could live in parapatry (*cf. inshore* and *offshore* Bottlenosed dolphins subspecies (**Herch &**
431 **Duffield, 1990**)). Alternatively, the two Mediterranean groups could also live in sympatry (*cf.*
432 Killer whale ecotypes or even species displaying different feeding preferences and/or
433 breeding avoidance (**Barett-Lennard & Ellis, 2001; Pitman et al., 2007; Schrope, 2007**)).
434 We might consider similar differentiation for the Striped dolphin given the geographic
435 variation in haplogroup frequencies and sexual asymmetry in gene flow between the two
436 groups (Figure S2). However, to date, no morphological polymorphisms have been reported
437 for this species, although some field naturalists suspect the existence of slight polymorphisms.
438 The present study is limited by sampling on stranded animals, not accounting for school
439 origin of the animals, in contrast to a previous study (**Gaspari et al., 2007**) which lacked
440 mitochondrial information. It is therefore not known whether Group 1 and Group 2
441 individuals co-occur in the same school. It will be interesting to test if the observation of
442 possible genetic differentiation between sampling locations and between putative *inshore* and
443 *offshore* schools is the result of school segregation between the two groups.
444 Inbreeding and loss of fitness (*e.g.* pathogen susceptibility) can be the result of barriers to
445 gene flow such as foeto-maternal incompatibility linked to the Major Histocompatibility
446 Complex (MHC) (**Laitinen et al., 1993a; Tregenza & Wedell, 2000; Zeh & Zeh, 1996**).
447 These deleterious effects of inbreeding can be balanced out by polyandry (**Stockley, 2003;**
448 **Wolff & Macdonald, 2004**), a common reproductive strategy in Delphinidae (**Whitehead &**
449 **Mann, 2000**). Within-species differences in reproductive strategies (polyandry *vs*
450 polygynandry), observed in the two forms of the sister-species *Stenella longirostris* (**Perrin**
451 **& Mesnick, 2003**), may therefore be extended to differences in inbreeding and fitness
452 between *Stenella caeruleoalba* subspecies. In support of this view, Group 1 is less diverse and
453 its MHC DR-B haplotypes display particular combination (data not shown). Moreover, the
454 Striped dolphins that perished during Morbilliviral epizootic episodes were the more inbred
455 (**Valsecchi et al., 2004b**), which could lead to an overrepresentation of Group 1 individuals.
456 We may therefore speculate that outbreeding avoidance, needed for the maintain of sympatry
457 or parapatry, by increasing directional selection and decreasing balancing selection on MHC,
458 would result in fitness cost *i.e.* pathogen susceptibility induced by MHC's more limited
459 repertoire. In the same way, the existence of 'recently admixed' genotypes (maternal
460 mitochondrial markers from one group and paternal from the other) and low admixture could
461 be the result of admixed offspring's lesser fitness.

462 As we stated above, the divergence associated with the MSC may be present in other
463 Delphininae whose radiation could be interesting to investigate, as for other mammalian
464 families, because they help in shedding light on ecological and intrinsic patterns leading to
465 evolutionary success (**Hare et al., 2002; Michaux et al., 2001; Purvis, 1995**). In the
466 Mediterranean, epipelagic Cetacean (Kentriodontidae and Ziphiidae) expansion occurred
467 during the Miocene (**Bianucci & Landini, 2002**). However, it remains unclear whether the
468 Delphinidae went through the Messinian Salinity Crisis. Some authors state that they appeared
469 only after the Mediterranean replenishment at the end of the MSC (**Bianucci, 2001; Muizon,**
470 **1988**), while other researchers suspect a possible previous presence of Delphinidae in the
471 Mediterranean during the Miocene (**Cigala-Fulgosi & Pilleri, 1985; Pilleri, 1985**). Since our
472 results agree with the latter hypothesis, they call into question the completeness of the
473 Mediterranean cetacean fossil record and/or the phylogenetic place of Kentriodontids
474 (Delphinids sister-group or ancestors?). In both cases, they emphasise the need of a
475 taxonomic update of the family (see (**Leduc et al., 1999**)).

476 In addition, geographical variations in morphology (**Calzada & Aguilar, 1995; Di**
477 **Méglio et al., 1996; Pertoldi et al., 2000**) may be the consequence of sampling bias of both

478 “subspecies”. There is hence a need for in field-visual identification in the framework of
479 conservation studies.

480

481 Conclusion and perspectives

482 The opportunistic Mediterranean Striped dolphin, or rather its ancestor, survived the
483 stringent conditions of the MSC. This particular evolutionary history makes it a good model
484 for multidisciplinary investigations in evolutionary biology (paleontology, systematics,
485 behavioural ecology, immunogenetics, palaeoecology...). For example, it may be pertinent to
486 investigate whether natural or human factors affect differently the two “subspecies”. Indeed,
487 investigating whether MHC patterns involve differences in susceptibility to *Morbillivirus*, or
488 if foraging behaviour differences exist and lead to contrasting fisheries by-catch rates, are of
489 great interest in conservation biology.

490 In addition, Striped dolphins now face various aspects of human activities among
491 which the most damaging are over-fishing, pollution and coast overcrowding. An extension of
492 the marine sanctuary, towards the Adriatic and the Balearic islands, may therefore be
493 necessary for the preservation of this remarkable biological model.

494

495 Material and Methods

496 Sample collection and DNA extraction

497 A total of 138 Striped dolphins *Stenella coeruleoalba* were sampled representing the
498 Mediterranean (n=79), the North-Eastern Atlantic (n=45) and the North-Eastern Pacific
499 (n=14). Mediterranean individuals came from various localities of the Western basin and
500 some Eastern localities (Figure 1). Various tissues (muscle, liver, kidney, skin) were sampled
501 on animals, stranded and by-caught between 1989 and 2004 and Genomic DNA was extracted
502 using classical phenol-chloroform protocol (Sambrook & Russell, 2001).

503

504 Laboratory procedures

505 The mtDNA control region (627 bp) was amplified using the primers 5'-
506 GAATCCCCGGTCTTGTAACC-3' and 5'-TCTCGAGATTTTCAGTGTCTTGCTTT-3'
507 as previously described (Hoelzel *et al.*, 1991). Genomic DNA (20-100 ng) was amplified in
508 the following assay conditions (final volume of 50 µl): 477 mM KCl, 1 mM Tris/HCl pH 8.3,
509 1.5 mM MgCl₂, 250 µM each dNTP, 1 µM of each primer, and 2.5 unit of Taq polymerase.
510 Cycling was programmed as follows: 30 cycles (93°C/30 s, 50°C/1 min., 70°C/1 min.) and
511 one final step (72°C/10 min). Both positive (already amplified samples) and negative
512 (distilled water) controls were added to each assay. Amplicons were purified using the
513 Qiaquick PCR Purification Kit (Qiagen GmbH, Hilden, Germany) and rediluted in 10 µl
514 water. Sequencing was carried out on a CEQ™ 2000XL Sequencer (Beckman Coulter,
515 Fullerton, CA) using the Dye Terminator Cycle Sequencing kit (Beckman Coulter) with the
516 following conditions: 60-80 ng of PCR product, 4 µl of the Quick Start mix, 1 µM of each
517 primer (same primers as for PCR reaction) in a final volume of 20 µl.

518 In addition, a hundred and thirty-eight individuals were genotyped at five microsatellite *loci*
519 (3 di- and 2 tetranucleotide repeats) as described in (Bourret *et al.*, 2007).

520

521 Data analysis

522 *F*-statistics and AMOVA testing for the first hypothesis.

523 We provide calculation of linkage disequilibrium (LD) using the expectation
524 maximization (EM) algorithm (Slatkin & Excoffier, 1996) as well as a test analogous to
525 Fisher's exact test for Hardy-Weinberg equilibrium performed locus by locus on alternative
526 contingency tables generated by Markov Chains (Guo & Thompson, 1992).

527 For each of the Mediterranean samples, we computed estimates of Wright's population
528 inbreeding coefficient (F_{IS}) across all *loci* as well as Hamilton's relatedness coefficient

529 $relat = \frac{2 F_{IS}}{1 + F_{IS}}$ (Hamilton, 1963), calculated using an estimator strictly equivalent (Queller &

530 Goodnight, 1989). This measure is the average relatedness of individuals within samples

531 when compared to the whole data set. All these computations were performed using the
532 software FSTAT (Goudet, 2002).

533 We also conducted two-level hierarchical Analysis of Molecular Variance (AMOVA).
534 We used this approach first to estimate the proportion of variance distributed among oceanic
535 basins (Pacific, Atlantic, Mediterranean). We also conducted a separate hierarchical AMOVA
536 on subpopulations within the Mediterranean basin. This analysis was performed thanks to the
537 help of the ARLEQUIN software (Schneider *et al.*, 2000).

538
539 *Isolation-by-distance (IBD) and analyses seeking for barriers to gene flow testing for*
540 *the second hypothesis.*

541 The correlation between geographical and genetic distances among samples was
542 assessed by performing a Mantel test (Legendre & Legendre, 1998) implemented in the
543 program IBD 1.5 (Bohonak, 2002). Each test was based on 10000 randomizations. Slatkin's
544 similarity measure (Slatkin & Voelm, 1991) for each population pair is estimated by

545
$$F_{ST} = (1 - \frac{1}{F_{CT}}) / 4$$
 and the corresponding Z-statistics is computed as $Z = \frac{F_{ST} - E(F_{ST})}{\sqrt{Var(F_{ST})}}$ (matrix B

546 corresponding to the geographic distance matrix). The geographical distance input matrix
547 contained the shortest waterway distances between sampling locations. Correlation between
548 the two matrices is measured by r ($-1 < r < 1$). Significance is assessed by comparing this
549 value to a distribution of r -scores obtained by randomizing rows/columns of the B. One-tailed
550 p-values for null hypothesis of negative ($r < 0$) or positive ($r > 0$) correlation are provided.

551 Alternatively, to test for the existence of significant barriers to gene flow within the
552 Mediterranean, we first assessed their possible existence in a non-parametric way. A spatial
553 analysis of genetic variance based on a simulated annealing procedure was used to define
554 groups of populations that are geographically homogeneous and maximally differentiated
555 from each other. By using Delauney triangulation and Monmonnier algorithm the SAMOVA
556 software (Dupanloup *et al.*, 2002), iteratively seeks the composition of a user-defined
557 number K of groups of geographically adjacent populations that maximizes F_{CT} , *i.e.* the
558 proportion of total genetic variance resulting from differences among groups of population.
559 We also tested if the strait of Sicilia between the Occidental and Oriental basins could be a
560 significant barrier by computing fixation index F_{ST} between the occidental group
561 (GV+GL+LS+NT) and oriental group (IS+SA) and proportions in mitochondrial haplogroups,
562 between both basins (two-sided T-test).

563
564 *Estimation of admixture and gene flow testing for the third hypothesis.*

565 Then, since hidden population structure may confound traditional estimates of *a priori*
566 defined genetic structure (Weir, 1996), we used a model-based clustering method to infer
567 population structure, by minimizing the Hard-Weinberg and gametic phase disequilibria, and
568 assign individuals to populations, using multilocus genotype data. The model assumes an
569 unknown number (K) of populations each characterised by a set of allele frequencies at each
570 locus. Individuals in the sample are assigned (with a probability $\text{prob}(X|K)$, X being the
571 observed genotype) to one population, or more if their genotypes indicate that they are
572 admixed. This method allows the determination of a vector Q denoting the admixture
573 proportions of each individual and it is implemented in a Bayesian framework in STRUCTURE
574 (Pritchard *et al.*, 2000). We used prior assignation of some individuals (learning samples)
575 following information from mtDNA data (see Results section).

576 To further test for population divergence and differences in evolutionary history, we used the
577 MCMC approach using Metropolis-Hastings sampling of random coalescent genealogies
578 implemented in LAMARC (Likelihood Analysis with Metropolis Algorithm using Random
579 Coalescences (Beerli & Felsenstein, 2001)) in its Bayesian framework. This approach led to
580 the assessment of the gene flow level between groups ($M_{ij} = m_{ij}/\mu$, m_{ij} being the migration rate
581 from population i to j), the growth rate (g in an exponential growth model: $N(t) = N(t_0) \cdot e^{-(g(t-t_0))}$)
582 and each group's theta ($\Theta_i = 4N_e\mu$). Using the Metropolis-Hastings sampling scheme, the
583 MCMC concentrates the sampling in areas of the coalescent most contributing to the final

584 likelihood. In its Bayesian framework, it gives 95% Credible Intervals and posterior
585 probability distributions of the migration rates. This approach assumes that population are at
586 drift-migration equilibrium and has the major benefit, unlike F-statistics, to allow for the
587 possibility of non-symmetrical migration and differences in population sizes, scenarios more
588 biologically realistic for highly mobile marine species. However, since migration equilibrium
589 involves that migration rate is averaged over the coalescent, it does not account for
590 differences between recent and former migrants. This difference is therefore taken into
591 account by the individual level of admixture inferred by STRUCTURE described above.
592

593 *Phylogeny-based analyses testing for the third hypothesis*

594 The existence of potential polyphyletic patterns in *Stenella caeruleoalba* within the
595 Mediterranean was tested using control region phylogenetic inference. DNA sequences in
596 both directions were assembled to obtain a consensus sequence for each sample and aligned
597 using CLUSTALX (Thompson *et al.*, 1997). Mitochondria DNA data previously released study
598 (Galov *et al.*, 2008) were included (5 haplotypes identical to ours plus one new). Mutation
599 model was selected by MODELTEST (Posada & Crandall, 1998). Phylogenetic inference was
600 performed using Bayesian Monte Carlo Markov Chain (MCMC) approach (MRBAYES
601 (Ronquist & Huelsenbeck, 2003)), maximum likelihood (DNAML (Felsenstein, 1981)), and
602 bootstrapped neighbour-joining (NJ) distance of Nei (Saitou & Nei, 1987). Nucleotide - π -
603 and haplotype - h - diversities were estimated using DNASP (Rozas *et al.*, 2003).

604 Time to the most recent common ancestor (T_{MRCA}) of haplotype lineages was inferred
605 by two methods: (1) a semi-parametric smoothing method using penalised likelihood
606 (Sanderson, 2002) (R8S software (Sanderson, 2003)) and (2) the Bayesian approach that
607 jointly estimates the evolution rate implemented in BEAST (Drummond *et al.*, 2002). In both
608 analysis, we used the present sequences of Striped dolphin haplotypes as well as orthologous
609 sequences from five species representing the main Cetacean lineages (Monodontidae,
610 Phocoenidae, Balaenopteridae and two Delphinoidea – a super-family including Delphinidae
611 and related families-: one Killer whale *Orcinus orca* and one Risso's dolphin *Grampus*
612 *griseus* sequenced in this study – EMBL accession AM498741 – see Figure S4 for details).
613 Both methods were constrained by fossil calibration of nodes (Phocoenoidea [9.1-17.1 Myrs],
614 Delphinida [20.2-29.8 Myrs], Cetacea [27.2-37.4 Myrs], see Figure S4 for details). In
615 addition, Bayesian estimation was done assuming exponential or log-normal relaxed clock, a
616 coalescent skyline tree prior and HKY (Hasegawa-Kishino-Yano) mutation model with a
617 proportion of invariant sites. The analysis was run simultaneously with the skyline
618 demographic reconstruction described implemented in BEAST (Drummond *et al.*, 2002).
619 Bayesian skyline estimates past population dynamics by assessing the uncertainty in tree
620 topology, dates of divergence and mutation parameters. Using calibrated evolutionary rate,
621 analysis was run for 25 000 000 iterations, with sampling every 1 000 chains and 10% burn-
622 in, which is comparable to studies with same purpose (de Oliveira *et al.*, 2006; Shapiro *et*
623 *al.*, 2004).
624

625 Supporting Information

626 **Figure S1 Median joining network based on the variation of the control region**
627 **sequence from 73 Striped dolphins.**

628 Found at:

629 **Figure S2 Gene flow between the Mediterranean groups and the Atlantic.**

630 Found at:

631 **Figure S3 Bayesian estimated allele frequencies for microsatellite *loci* GATA053 (a) and**
632 **EV92Mn (b).**

633 Found at:

634 **Figure S4 Fossil calibration of the nodes used for T_{MRCA} estimations.**

635 Found at:

636 **Table S1 Microsatellite polymorphism summary by locality.**

637 Found at:

638 **Table S2 Mantel test for matrix correlation between genetic similarity/log (genetic**
639 **similarity) and geographic distance/log(geographic distance).**

640 Found at:

641 **Table S3 Variable sites in the control region of 73 Striped dolphin defining 59**
642 **haplotypes and their geographical distribution.**

643 Found at:

644 **Table S4 Gene flow Support Intervals (95%) among Striped dolphin populations given**
645 **by Bayesian estimation from microsatellite and mtDNA markers.**

646 Found at:

647

648 Acknowledgments

649 We are particularly indebted to Frank Dhermain from the GECEM, for providing the greatest
650 part of samples and to Heather Etchevers for her help improving the manuscript. Special
651 thanks also go to A.S. Lemaire and J. Lafitte from the FERSO-BIO, W. Dabin and O. Van
652 Canneyt from the CRMM, B.K. Jacobsen from the UAF Museum. Great thanks also go to M.-
653 A. Sol, M. Bonhomme, A. Rozzi, A. Ribéron and E. Douzery for both technical assistance
654 and valuable discussions.

655

656 Author contributions

657 Conceived and designed the experiments: MM BC. Performed the experiments: MM VB.

658 Analysed the data: MM. Wrote the paper: MM BC.

659

660 References

- 661 1. Mace GM, Masundire H, Baillie J, Ricketts T, Brooks T, et al. (2005) Biodiversity.
662 Millennium Ecosystem Assessment, 2005 Current State and Trends: Findings of the
663 Condition and Trends Working Group Ecosystems and human well-being (Vol 1).
664 Washington, DC: Island Press.
- 665 2. Vitousek PM, Mooney HA, Lubchenco J, Melillo JM (1997) Human Domination of Earth's
666 Ecosystems. *Science* 277: 494.
- 667 3. Isaac NJB, Turvey ST, Collen B, Waterman C, Baillie JE (2007) Mammals on the EDGE:
668 Conservation Priorities Based on Threat and Phylogeny. *PLoS ONE* 2: e296.
- 669 4. European Commission (2007) Climate Change Impacts on the Water Cycle, Resources and
670 Quality. Brussels: European Commission. 149 p.
- 671 5. Balvanera P, Daily GC, Ehrlich PR, Ricketts TH, Bailey SA, et al. (2001) Conserving
672 biodiversity and ecosystem services. *Science* 291: 2047.
- 673 6. Worm B, Barbier EB, Beaumont N, Duffy JE, Folke C, et al. (2006) Impacts of
674 Biodiversity Loss on Ocean Ecosystem Services. *Science* 314: 787.
- 675 7. Silvani L, Gazo M, Aguilar A (1999) Spanish driftnet fishing and incidental catches in the
676 western Mediterranean. *Biological Conservation* 90: 79-85.
- 677 8. Gili JM (2001) Biodiversitat Marina Mediterrània. *L'Atzavara* 9: 25-35.
- 678 9. Myers N, Mittermeier RA, Mittermeier CG, da-Fonseca GA, Kent J (2000) Biodiversity
679 hotspots for conservation priorities. *Nature* 403: 853-858.
- 680 10. Hsü KJ, Ryan WB, Cita MB (1973) Late miocene desiccation of the mediterranean.
681 *Nature* 242: 240-244.
- 682 11. Emig CC, Geistdoerfer P (2004) Faune profonde en Mer Méditerranée : les échanges
683 historiques, géographiques et bathymétriques. *Carnets de Géologie* 3: 1-10.
- 684 12. Weijermars R (1988) Neogene tectonics in the Western Mediterranean may have caused
685 the Messinian Salinity Crisis and an associated glacial event. *Tectonophysics* 148:
686 211-219.

- 687 13. Boury-Esnault N, Pansini M, Uriz MJ (1992) A new Discorhabdella (Porifera,
688 Demospongiae), a new Tethyan relict of pre-Messinian biota? Journal of Natural
689 History 26: 1-7.
- 690 14. Durand JD, Bianco PG, Laroche J, Gilles A (2003) Insight Into the Origin of Endemic
691 Mediterranean Ichthyofauna: Phylogeography of Chondrostoma Genus (Teleostei,
692 Cyprinidae). Journal of Heredity 94: 315-328.
- 693 15. Gilles A, Chappaz R, Cavalli L, Lortscher M, Faure E (1998) Genetic differentiation and
694 introgression between putative subspecies of *Leuciscus soufia* (Teleostei : Cyprinidae)
695 of the region of the Mediterranean Alps. Canadian Journal of Fisheries and Aquatic
696 Sciences 55: 2341-2354.
- 697 16. Gysels ES, Hellemans B, Patarnello T, Volckaert FAM (2004) Current and historic gene
698 flow of the sand goby *Pomatoschistus minutus* on the European continental shelf and
699 in the Mediterranean Sea. Biological Journal of the Linnean Society 83: 561-576.
- 700 17. Jaume D, Boxshall GA (1996) The persistence of an ancient marine fauna in
701 Mediterranean waters: new evidence from misophrioid copepods living in anchihaline
702 caves. Journal of Natural History 30: 1583-1595.
- 703 18. Maldonado M, Uriz MJ (1995) Biotic Affinities in a Transitional Zone Between the
704 Atlantic and the Mediterranean: A Biogeographical Approach Based on Sponges.
705 Journal of Biogeography 22: 89-110.
- 706 19. Penzo E, Gandolfi G, Bargelloni L, Colombo L, Patarnello T (1998) Messinian salinity
707 crisis and the origin of freshwater lifestyle in western Mediterranean gobies.
708 Molecular Biology and Evolution 15: 1472-1480.
- 709 20. Wilke T (2003) *Salenthydrobia* gen. nov (Rissooidea : Hydrobiidae): a potential relict of
710 the Messinian salinity crisis. Zoological Journal Of The Linnean Society 137: 319-
711 336.
- 712 21. Notarbartolo di Sciara G (2002) Cetacean species occurring in the Mediterranean and
713 Black Seas. In: G. Notarbartolo di Sciara (Ed.), Cetaceans of the Mediterranean and
714 Black Seas: state of knowledge and conservation strategies. Monaco: A report to the
715 ACCOBAMS Secretariat. 3 3. 17 p.
- 716 22. Archer II F, Perrin WF (1999) *Stenella coeruleoalba*. Mammalian Species: 1-9.
- 717 23. Aguilar A (2000) Population biology, conservation threats and status of Mediterranean
718 striped dolphins (*Stenella coeruleoalba*). Journal of Cetacean Research and
719 Management 2: 17-26.
- 720 24. Domingo M, Ferrer L, Pumarola M, Marco A, Plana J, et al. (1990) Morbillivirus in
721 dolphins. Nature 348: 21.
- 722 25. Bearzi G, Reeves RR, Notarbartolo Di Sciara G, Politi E, Canadas A, et al. (2003)
723 Ecology, status and conservation of short-beaked common dolphins *Delphinus delphis*
724 in the Mediterranean Sea. Mammal Review 33: 224-252.
- 725 26. Calzada N, Aguilar A (1995) Geographical variation of body size in Western
726 Mediterranean striped dolphins (*Stenella coeruleoalba*). Zeitschrift fur Säugetierkunde
727 60: 257-264.
- 728 27. Pertoldi C, Podesta M, Loeschcke V, Schandorff S, Marsili L, et al. (2000) Effect of the
729 1990 die-off in the northern Italian seas on the developmental stability of the striped
730 dolphin *Stenella coeruleoalba* (Meyen, 1833). Biological Journal of the Linnean
731 Society 71: 61-70.
- 732 28. Bourret V, Macé M, Crouau-Roy B (2007) Genetic variation and population structure of
733 western Mediterranean and northern Atlantic *Stenella coeruleoalba* populations
734 inferred from microsatellite data. Journal of the Marine Biological Association of the
735 United Kingdom 87: 265-269.
- 736 29. Andrews RC (1911) Description of an apparently new Porpoise of the genus *Tursiops*,
737 with remarks upon a Skull of *Tursiops gilli* Dall. Bulletin of the American Museum of
738 Natural History 30: 233-237.
- 739 30. Borsa P, Quignard JP (2001) Systematics of the Atlantic-Mediterranean soles *Pegusa*
740 *impar*, *P. lascaris*, *Solea aegyptiaca*, *S. senegalensis*, and *S. solea* (Pleuronectiformes:
741 Soleidae). Canadian Journal of Zoology 79: 2297-2302.

- 742 31. Nielsen EE, Hansen MM, Ruzzante DE, Meldrup D, Gronkjaer P (2003) Evidence of a
743 hybrid-zone in Atlantic cod (*Gadus morhua*) in the Baltic and the Danish Belt Sea
744 revealed by individual admixture analysis. *Molecular Ecology* 12: 1497-1508.
- 745 32. Ross GJB (1977) The Taxonomy of Bottlenosed Dolphins *Tursiops* Species in South
746 African Waters, with Notes on Their Biology. *Annals of the Cape Provincial Museum*
747 11: 135-194.
- 748 33. Escorza-Treviño S, Lang A, Dizon AE (2002) Genetic differentiation and intraspecific
749 structure of Eastern Tropical Pacific Spotted dolphins, *Stenella attenuata*, revealed by
750 Mitochondrial And Microsatellite DNA Analyses. La Jolla, CA: South-West Fisheries
751 Science Center. LJ-02-38 LJ-02-38. 20 pp p.
- 752 34. Leduc RG, Perrin WF, Dizon AE (1999) Phylogenetic relationships among the delphinid
753 cetaceans based on full cytochrome B sequences. *Marine Mammal Science* 15: 619-
754 648.
- 755 35. Anonymous (1999) Agreement on the creation of a Mediterranean Sanctuary for marine
756 mammals. Final agreement signed in Roma on 25 November 1999 by the Ministers of
757 Italy, France and Monaco.
- 758 36. Douady CJ, Delsuc F, Boucher Y, Doolittle WF, Douzery EJP (2003) Comparison of
759 Bayesian and maximum likelihood bootstrap measures of phylogenetic reliability.
760 *Molecular Biology and Evolution* 20: 248-254.
- 761 37. Hasegawa M, Kishino H, Yano T (1985) Dating of the human-ape splitting by a molecular
762 clock of mitochondrial DNA. *Journal of Molecular Evolution* 22: 160-174.
- 763 38. Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using
764 multilocus genotype data. *Genetics* 155: 945-959.
- 765 39. Drummond AJ, Nicholls GK, Rodrigo AG, Solomon W (2002) Estimating mutation
766 parameters, population history and genealogy simultaneously from temporally spaced
767 sequence data. *Genetics* 161: 1307-1320.
- 768 40. Kitazoe Y, Kishino H, Waddell PJ, Nakajima N, Okabayashi T, et al. (2007) Robust time
769 estimation reconciles views of the antiquity of placental mammals. *PLoS ONE* 2:
770 e384.
- 771 41. Hauser L, Adcock GJ, Smith PJ, Ramirez JH, Carvalho GR (2002) Loss of microsatellite
772 diversity and low effective population size in an overexploited population of New
773 Zealand snapper (*Pagrus auratus*). *Proceedings of the National Academy of Sciences*
774 of the United States of America 99: 11742-11747.
- 775 42. Opgen-Rhein R, Fahrmeir L, Strimmer K (2005) Inference of demographic history from
776 genealogical trees using reversible jump Markov chain Monte Carlo. *BMC*
777 *Evolutionary Biology* 5: 6.
- 778 43. Rosenberg NA, Mahajan S, Ramachandran S, Zhao C, Pritchard JK, et al. (2005) Clines,
779 clusters, and the effect of study design on the inference of human population structure.
780 *PLoS genetics* 1: e70.
- 781 44. Gaspari S, Azzellino A, Airoidi S, Hoelzel AR (2007) Social kin associations and genetic
782 structuring of striped dolphin populations (*Stenella coeruleoalba*) in the Mediterranean
783 Sea. *Molecular Ecology* 16: 2922-2933.
- 784 45. Sanmartin I (2003) Dispersal vs. vicariance in the Mediterranean: historical biogeography
785 of the Palearctic Pachydeminae (Coleoptera, Scarabaeoidea). *Journal of Biogeography*
786 30: 1883-1897.
- 787 46. Kaustuv R, Jablonski D, Valentine JW (2001) Climate change, species range limits and
788 body size in marine bivalves. *Ecology Letters* 4: 366-370.
- 789 47. Lieberman BS (2000) *Paleobiogeography: Using Fossils to Study Global Change, Plate*
790 *Tectonics, and Evolution*; Jones NHLADS, editor. New York: Kluwer
791 Academic/Plenum Publishers. 208 p.
- 792 48. Voelker G (1999) Dispersal, Vicariance, and Clocks: Historical Biogeography and
793 Speciation in a Cosmopolitan Passerine Genus (*Anthus*: Motacillidae). *Evolution* 53:
794 1536-1552.

- 795 49. Natoli A, Birkun A, Aguilar A, Lopez A, Hoelzel AR (2005) Habitat structure and the
796 dispersal of male and female bottlenose dolphins (*Tursiops truncatus*). Proceedings of
797 the Royal Society of London Series B Biological Sciences 272: 1217-1226.
- 798 50. Rosenbaum HC, Brownell RL, Brown MW, Schaeff C, Portway V, et al. (2000) World-
799 wide genetic differentiation of *Eubalaena*: questioning the number of right whale
800 species. Molecular Ecology 9: 1793-1802.
- 801 51. Carlsson J, McDowell JR, Diaz-Jaimes P, Carlsson JE, Boles SB, et al. (2004)
802 Microsatellite and mitochondrial DNA analyses of Atlantic bluefin tuna (*Thunnus*
803 *thynnus thynnus*) population structure in the Mediterranean Sea. Molecular Ecology
804 13: 3345-3356.
- 805 52. Macé M, Crouau-Roy B (2008) A highly polymorphic insertion in the Y-chromosome
806 amelogenin gene can be used for evolutionary biology, population genetics and sexing
807 in Cetacea and Artiodactyla. BMC Genetics 9: 64.
- 808 53. Xiong Y, Brandley MC, Xu S, Zhou K, Yang G (2009) Seven new dolphin mitochondrial
809 genomes and a time-calibrated phylogeny of whales. BMC evolutionary biology 9: 20.
- 810 54. Neigel JE, Avise JC (1986) Phylogenetic relationships of mitochondrial DNA under
811 various demographic models of speciation. In: Nevo SKaE, editor. Evolutionary
812 Processes and Theory. New York: Academic Press, Inc. pp. 515-534.
- 813 55. Funk DJ, Omland KE (2003) Species-level paraphyly and polyphyly: Frequency, causes,
814 and consequences, with insights from animal mitochondrial DNA. Annual Review of
815 Ecology Evolution and Systematics 34: 397-423.
- 816 56. Perrin WF, Wilson CE, Archer II F (1994) Striped dolphin *Stenella coeruleoalba* (Meyen,
817 1833). In: Ridgway SH, Harrison R, editors. Volume 5: The First Book of Dolphins:
818 Academic Press. pp. 129-159.
- 819 57. Alvarado Bremer JR, Stequert B, Robertson NW, Ely B (1998) Genetic evidence for inter-
820 oceanic subdivision of bigeye tuna (*Thunnus obesus*) populations. Marine Biology
821 132: 547-557.
- 822 58. Posada D, Crandall KA (2001) Intraspecific gene genealogies: trees grafting into
823 networks. Trends in Ecology and Evolution 16: 37-45.
- 824 59. Stamatis C, Triantafyllidis A, Moutou KA, Mamuris Z (2004) Mitochondrial DNA
825 variation in northeast atlantic and mediterranean populations of norway lobster,
826 *Nephrops norvegicus*. Molecular Ecology 13: 1377-1390.
- 827 60. Almada VC, Oliveira RF, Gonçalves EJ, Almeida AJ, Santos RS, et al. (2001) Patterns of
828 diversity of the north-eastern Atlantic blennioid fish fauna (Pisces : Blenniidae). Global
829 Ecology and Biogeography 10: 411-422.
- 830 61. Bargelloni L, Alarcon JA, Alvarez MC, Penzo E, Magoulas A, et al. (2003) Discord in the
831 family Sparidae (Teleostei): divergent phylogeographical patterns across the Atlantic-
832 Mediterranean divide. Journal of Evolutionary Biology 16: 1149-1158.
- 833 62. Zane L, Ostellari L, Maccatrozzo L, Bargelloni L, Cuzin Roudy J, et al. (2000) Genetic
834 differentiation in a pelagic crustacean (*Meganyctiphanes norvegica* : Euphausiacea)
835 from the North East Atlantic and the Mediterranean Sea. Marine Biology 136: 191-
836 199.
- 837 63. Godoy JA, Negro JJ, Hiraldo F, Donazar JA (2004) Phylogeography, genetic structure and
838 diversity in the endangered bearded vulture (*Gypaetus barbatus*, L) as revealed by
839 mitochondrial DNA. Molecular Ecology 13: 371-390.
- 840 64. Hewitt GM (2001) Speciation, hybrid zones and phylogeography - or seeing genes in
841 space and time. Molecular Ecology 10: 537-549.
- 842 65. Sole Cava AM, Klautau M, Boury Esnault N, Borojelic R, Thorpe JP (1991) Genetic
843 evidence for cryptic speciation in allopatric populations of two cosmopolitan species
844 of the calcareous sponge genus *Clathrina*. Marine Biology 111: 381-386.
- 845 66. Sasaki T, Nikaido M, Hamilton H, Goto M, Kato H, et al. (2005) Mitochondrial
846 phylogenetics and evolution of mysticete whales. Systematic Biology 54: 77-90.
- 847 67. Avise J (2000) Phylogeography: the History and Formation of Species. Cambridge, MA:
848 Harvard University Press.

- 849 68. Krijgsman W, Hiigeni FJ, Raffi I, Sierro FJ, Wilson DS (1999) Chronology, causes and
850 progression of the Messinian salinity crisis. *Nature* 400: 652-655.
- 851 69. Duggen S, Hoernle K, van-den-Bogaard P, Rupke L, Morgan JP (2003) Deep roots of the
852 Messinian salinity crisis. *Nature* 422: 602-606.
- 853 70. Blanc PL (2000) Of sills and straits: a quantitative assessment of the Messinian Salinity
854 Crisis. *Deep Sea Research, Part I Oceanographic Research Papers* 47: 1429-1460.
- 855 71. Clauzon G, Suc J-P, Popescu S-M, Marunteanu M, Rubino J-L, et al. (2005) Influence of
856 Mediterranean sea-level changes on the Dacic Basin (Eastern Paratethys) during the
857 late Neogene: the Mediterranean Lago Mare facies deciphered. *Basin Research* 17:
858 437-462.
- 859 72. Cassens I, Vicario S, Waddell VG, Balchowsky H, Van-Belle D, et al. (2000) Independent
860 adaptation to riverine habitats allowed survival of ancient cetacean lineages.
861 *Proceedings of the National Academy of Sciences of the United States of America* 97:
862 11343-11347.
- 863 73. Perrin WF (1998) *Stenella longirostris*. *Mammalian Species*: 1-7.
- 864 74. Carnevale G, Landini W, Sarti G (2006) Mare versus Lago-mare: marine fishes and the
865 Mediterranean environment at the end of the Messinian Salinity Crisis. *Journal of the*
866 *Geological Society* 163: 75-80.
- 867 75. Church SA, Taylor DR (2002) The evolution of reproductive isolation in spatially
868 structured populations. *Evolution* 56: 1859-1862.
- 869 76. Rieseberg LH, Raymond O, Rosenthal DM, Lai Z, Livingstone K, et al. (2003) Major
870 ecological transitions in wild sunflowers facilitated by hybridization. *Science* 301:
871 1211-1216.
- 872 77. Rendall L, Whitehead H (2001) Culture in whales and dolphins. *Behavioral and Brain*
873 *Sciences* 24: 309-382.
- 874 78. Herch SL, Duffield DA (1990) Distinction between northwest Atlantic offshore and
875 coastal bottlenose dolphins based on hemoglobin profile and morphology. In: Reeves
876 SLaR, editor. *The Bottlenose Dolphin*. New York: Academic Press. pp. 129-142.
- 877 79. Barrett-Lennard LG, Ellis GM (2001) Population Structure and Genetic Variability in
878 Northeastern Pacific Killer whale: Towards an Assessment of Population Viability.
879 Ottawa: Fisheries and Oceans Canada. 35 pp p.
- 880 80. Pitman RL, Perryman WL, LeRoi D, Eilers E (2007) A dwarf form of Killer Whale in
881 Antarctica. *Journal of Mammalogy* 88: 43-48.
- 882 81. Schrope M (2007) Food chains: killer in the kelp. *Nature* 445: 703-705.
- 883 82. Laitinen T, Koskimies S, Westman P (1993) Foeto-maternal compatibility in HLA-DR, -
884 DQ, and -DP *loci* in Finnish couples suffering from recurrent spontaneous abortions.
885 *European Journal of Immunogenetics*, official journal of the British Society for
886 Histocompatibility and Immunogenetics 20: 249-258.
- 887 83. Zeh JA, Zeh DW (1996) The Evolution of Polyandry I: Intragenomic Conflict and Genetic
888 Incompatibility. *Proceedings of the Royal Society of London Series B Biological*
889 *Sciences* 263: 1711-1717.
- 890 84. Tregenza T, Wedell N (2000) Genetic compatibility, mate choice and patterns of
891 parentage: invited review. *Molecular Ecology* 9: 1013-1027.
- 892 85. Stockley P (2003) Female multiple mating behaviour, early reproductive failure and litter
893 size variation in mammals. *Proceedings of the Royal Society of London Series B*
894 *Biological Sciences* 270: 271-278.
- 895 86. Wolff JO, Macdonald DW (2004) Promiscuous females protect their offspring. *Trends in*
896 *Ecology & Evolution* 19: 127-134.
- 897 87. Whitehead H, Mann J (2000) Female Reproductive Strategies and Social Bonds. In: J.
898 Mann RCC, P. L. Tyack & H. Whitehead, editor. *Cetacean Societies: Field Studies of*
899 *Dolphins and Whales*. Chicago: University of Chicago Press. pp. 219-246.
- 900 88. Perrin WF, Mesnick SL (2003) Sexual ecology of the Spinner dolphin, *Stenella*
901 *longirostris*: geographic variation in mating system. *Marine Mammal Science* 19: 462-
902 483.

- 903 89. Valsecchi E, Amos W, Raga JA, Podesta M, Sherwin W (2004) The effects of inbreeding
904 on mortality during a morbillivirus outbreak in the Mediterranean striped dolphin
905 (*Stenella coeruleoalba*). *Animal Conservation* 7: 139-146.
- 906 90. Hare MP, Cipriano F, Palumbi SR (2002) Genetic Evidence On The Demography Of
907 Speciation In Allopatric Dolphin Species. *Evolution* 56: 804-816.
- 908 91. Michaux J, Reyes A, Catzeflis F (2001) Evolutionary History of the Most Speciose
909 Mammals: Molecular Phylogeny of Muroid Rodents. *Molecular Biology and*
910 *Evolution* 18: 2017-2031.
- 911 92. Purvis A (1995) A Composite Estimate of Primate Phylogeny. *Philosophical Transactions*
912 *of the Royal Society B: Biological Sciences* 348: 405-421.
- 913 93. Bianucci G, Landini W (2002) Change in diversity, ecological significance and
914 biogeographical relationships of the Mediterranean Miocene toothed whale fauna.
915 *Geobios* 35: 19-28.
- 916 94. Bianucci G (2001) A new genus of Kentriodontid (Cetacea : Odontoceti) from the
917 Miocene of south Italy. *Journal of Vertebrate Paleontology* 21: 573-577.
- 918 95. Muizon Cd (1988) Les relations phylogenetiques des Delphinida (Cetacea, Mammalia).
919 *Annales de Paléontologie* 74: 159-227.
- 920 96. Cigala-Fulgosi F, Pilleri G (1985) The lower Serravalian cetacean fauna of Visiano
921 (Northern Apennines, Parma, Italy). *Investigations on Cetacea* 17: 55-93.
- 922 97. Pilleri G (1985) The Miocene Cetacea of the Belluno Sandstones. *Memorie di Scienze*
923 *Geologiche* 36: 1-250.
- 924 98. Di Méglion N, Romero-Alvarez R, Collet A (1996) Growth comparison in striped dolphins,
925 *Stenella coeruleoalba*, from the Atlantic and Mediterranean coasts of France. *Aquatic*
926 *Mammals* 22: 11-21.
- 927 99. Sambrook J, Russell DW (2001) *Molecular Cloning: A Laboratory Manual*. Cold Spring
928 Harbor: Cold Spring Harbor Laboratory Press.
- 929 100. Hoelzel AR, Hancock JM, Dover GA (1991) Evolution of the cetacean mitochondrial D-
930 loop region. *Molecular Biology and Evolution* 8: 475-493.
- 931 101. Slatkin M, Excoffier L (1996) Testing for linkage disequilibrium in genotypic data using
932 the Expectation-Maximization algorithm. *Heredity* 76: 377-383.
- 933 102. Guo SW, Thompson EA (1992) Performing the exact test of Hardy-Weinberg proportion
934 for multiple alleles. *Biometrics* 48: 361-372.
- 935 103. Hamilton WD (1963) The Evolution of Altruistic Behavior. *The American Naturalist* 97:
936 354-356.
- 937 104. Queller DC, Goodnight KF (1989) Estimating relatedness using genetic markers.
938 *Evolution* 43: 258-275.
- 939 105. Goudet J (2002) FSTAT, a program to estimate and test gene diversities and fixation
940 indices (version 2.9.3.2).
- 941 106. Schneider S, Roessli D, Excoffier L (2000) ARLEQUIN ver 2.0: a software for
942 population genetics data analysis. Genetics and Biometry Laboratory, University of
943 Geneva, Geneva. Geneva.
- 944 107. Legendre P, Legendre L (1998) *Numerical ecology*. Amsterdam: Elsevier.
- 945 108. Bohonak AJ (2002) IBD (Isolation by Distance): A Program for Analyses of Isolation by
946 Distance. *Journal of Heredity* 93: 153-154.
- 947 109. Slatkin M, Voelm L (1991) FST in a hierarchical island model. *Genetics* 127: 627-629.
- 948 110. Dupanloup I, Schneider S, Excoffier L (2002) A simulated annealing approach to define
949 the genetic structure of populations. *Molecular Ecology* 11: 2571-2581.
- 950 111. Weir BS (1996) *Genetic Data Analysis II: Methods for Discrete Population Genetic*
951 *Data*. Sunderland, MA: Sinauer Associates. 376 p.
- 952 112. Beerli P, Felsenstein J (2001) Maximum likelihood estimation of a migration matrix and
953 effective population sizes in n subpopulations by using a coalescent approach.
954 *Proceedings of the National Academy of Sciences of the United States of America* 98:
955 4563-4568.

- 956 113. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The
957 CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment
958 aided by quality analysis tools. *Nucleic Acids Research* 25: 4876-4882.
- 959 114. Galov A, Lauc G, Nikolić N, Šatović Z, Gomerčić T, et al. (2008) Records and genetic
960 diversity of striped dolphins (*Stenella coeruleoalba*) from the Croatian coast of the
961 Adriatic Sea. *JMBA2 - Biodiversity Records* Published online: 1-6.
- 962 115. Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution.
963 *Bioinformatics Oxford, England* 14: 817-818.
- 964 116. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under
965 mixed models. *Bioinformatics Oxford, England* 19: 1572-1574.
- 966 117. Felsenstein J (1981) Evolutionary trees from DNA sequences: a maximum likelihood
967 approach. *Journal of Molecular Evolution* 17: 368-376.
- 968 118. Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing
969 phylogenetic trees. *Molecular Biology and Evolution* 4: 406-425.
- 970 119. Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA
971 polymorphism analyses by the coalescent and other methods. *Bioinformatics Oxford,*
972 *England* 19: 2496-2497.
- 973 120. Sanderson MJ (2002) Estimating absolute rates of molecular evolution and divergence
974 times: a penalized likelihood approach. *Molecular Biology and Evolution* 19: 101-109.
- 975 121. Sanderson MJ (2003) r8s: inferring absolute rates of molecular evolution and divergence
976 times in the absence of a molecular clock. *Bioinformatics Oxford, England* 19: 301-
977 302.
- 978 122. de Oliveira T, Pybus OG, Rambaut A, Salemi M, Cassol S, et al. (2006) Molecular
979 epidemiology: HIV-1 and HCV sequences from Libyan outbreak. *Nature* 444: 836-
980 837.
- 981 123. Shapiro B, Drummond AJ, Rambaut A, Wilson MC, Matheus PE, et al. (2004) Rise and
982 fall of the Beringian steppe bison. *Science* 306: 1561-1565.
- 983
984

985 Figure legends

986

987 **Figure 1. Map showing sampling locations and their relative proportions of Group 1 and**
988 **Group 2 individuals.**

989 Atlantic: NS = North Sea & Channel (4 individuals), NBB = Northern Bay of Biscay (21
990 ind.), SBB = Southern Bay of Biscay (20 ind.). Mediterranean Western basin: GV = Bay of
991 Valencia (6 ind.), GL = Gulf of Lions (23 ind.), LS = Ligurian Sea (26 ind.), NT = Northern
992 Tyrrhenian Sea (4 ind.). Mediterranean Eastern basin: IS = Ionian Sea (12 ind.), SA =
993 Southern Adriatic Sea (5 ind.). Pacific Ocean sample is not shown: all samples were located
994 between 30-40°N and 165-180°W. Pie-charts represent the respective proportions in 4
995 Mediterranean sub-regions of Group 1 individuals (red) and Group 2 individuals (yellow)
996 following their mitochondrial DNA haplotype. Eastern basin sample shows significantly ($p <$
997 0.05 , 0.001 if including sequences from (Galov *et al.*, 2008)) higher Group 1 individuals
998 proportions than other samples.

999 Location of oligohaline lakes during the Messinian Salinity Crisis are indicated in faded
1000 yellow (Panonian and Euxinic basins). Interrogation marks indicate putative locations.

1001

1002 **Figure 2. Genetic isolation by distance of Mediterranean samples of *S. caeruleoalba***
1003 **inferred from multilocus estimates of F_{ST} genetic distances versus geographical distance**
1004 **(Mantel test).**

1005 Plot coordinates are as follows: (a) genetic vs geographic; (b) genetic vs log(geographic); (c)
1006 log(genetic) vs geographic; (d) log(genetic) vs log(geographic). Regression r^2 are given above
1007 each plot.

1008

1009 **Figure 3. Phylogenetic relationships among Striped dolphin *mtDNA* haplotypes.**

1010 The Bayesian phylogenetic tree represents a consensus of 9.000 trees sampled proportionally
1011 to the posterior probability of their branches lengths after a burn-in of 10.000 trees. Posterior
1012 probabilities above 0.5 are represented left to their supported node. Sampling origin (see
1013 Figure 1) is indicated right to the haplotype number. Black-highlighted origins are for Group
1014 2 haplotypes; "PAC" is for Pacific haplotypes.

1015 At the bottom-left is presented a schematic phylogeny among Delphinid haplotypes showing
1016 the polyphyly of the Striped dolphin in the Mediterranean (SC= *Stenella caeruleoalba*, bold
1017 type represents lineages including SC haplotypes underlined). The Bayesian tree is the result
1018 of the consensus of 45.000 trees after a burn-in of 20.000 trees.

1019

1020 **Figure 4. Estimated population structure of 138 Striped dolphin samples.**

1021 A single vertical line divided into 3 colours (number assumed populations), represents each
1022 individual. The length of coloured segment is the Bayesian estimated proportion of admixture
1023 coefficient within each individual's genome *i.e.* the estimated proportion of membership of
1024 the individual in the corresponding population.

1025 Above legend indicates the geographic sampling origin. Below legend indicates oceanic
1026 origin. Red colour represents Group 1 genetic pool in each individual's genome while yellow
1027 represents the Atlantic-Group 2 genetic pool.

1028

1029 **Figure 5. T_{MRCA} of Mediterranean Striped dolphin lineages as estimated by the**
1030 **Uncorrelated Exponential Relaxed Clock.**

1031 Sample names and species names used as calibration are shown as tree's tip labels and font
1032 Colours denote population (red for Group 1, yellow for Group 2). Numbers below nodes are
1033 Bayesian posterior probabilities of the respective nodes.

1034 T_{MRCA} 95% Credible Intervals are represented as horizontal bar-plots on their respective node,
1035 referring to the Bayesian probability densities presented above (colours correspond).

1036 Branch colours denote inferred evolutionary rates (in units of 10^{-9} /site per year) as follows:
1037 black, 0-2; blue, 2-4; purple, 4-8; dark red, 6-8; and light red, 8-11.

1038 A stratigraphic scale is reported at the bottom.

1039

1040 **Figure 6. Reconstructed demographic history within the 3 sampling areas as inferred on**
1041 **mitochondrial DNA sequences.**

1042 Estimates of effective population size are represented against time backwards since present.

1043 Estimates are based on mtDNA control region sequences for Atlantic+Pacific+Mediterranean
1044 population (n=74), Atlantic+Mediterranean populations (n=66) and the whole Mediterranean
1045 sample (n=48). The Bayesian skyline plot shows median estimates as bold lines and 95%
1046 confidence bounds as greyish lines.

1047

1048 Tables

1049

1050 **Table 1. Linkage disequilibrium (LD) between microsatellite *loci* as tested by the EM**

1051 **algorithm (10000 Markov steps).**

Locus	GATA098	MK9	GATA053	MK6	EV92Mn
GATA098		-	-	+(1)	-
MK9	+		+(2)	-	+(A)
GATA053	+	+		-	-
MK6	+	+	-		+(P)
EV92Mn	-	+	-	+	

1052

1053 Positive linkage disequilibrium is represented by a cross at the intersection of the rows and

1054 columns of the respective microsatellites. Below diagonal are shown LD detected within the

1055 whole Mediterranean sample and above diagonal, LD detected within the Atlantic sample (A),

1056 Pacific sample (P), Group 1 (1), Group 2 (2).

1057 **Table 2. F-statistics and AMOVA results.**

1058

	GV	GL	LS	NT	IS	SA
GV		-0.00216	0.00344	0.00340	-0.08402	0.01726
GL	0.41441+/- 0.0433		-0.01278	0.00125	-0.10693	-0.03495
LS	0.40541+/- 0.0434	0.98198+/- 0.0096		-0.02194	-0.09711	-0.04325
NT	0.41441+/- 0.0388	0.30631+/- 0.0492	0.76577+/- 0.0245		0.00400	-0.01000
IS	0.87387+/- 0.0334	0.99099+/- 0.0030	0.99099+/- 0.0030	0.37838+/- 0.0466		0.05904
SA	0.27928+/- 0.0550	0.82883+/- 0.0354	0.94595+/- 0.0205	0.49550+/- 0.0475	0.09009+/- 0.0235	

1059 **(a) Pairwise F_{ST} comparisons between Mediterranean sampling locations.**

1060 Above diagonal: pairwise F_{ST} values, below diagonal, p-values.

1061

1062

	Among subpopulations			Within subpopulations			Overall	
	d.f	S.S.	% variance	d.f.	S.S.	% variance	F_{ST} (p-value)	F_{IS} (p-value)
Pacific <i>N</i> = 14	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.		
Atlantic (NS, NBB, SBB sampling locations) <i>N</i> = 45 (4, 20, 21)	2	3.740	0.44	87	146.026	99.56	0.00438 (0.32356)	
Mediterranean (GL, GL, LS, NT, IS, AS sampling locations) <i>N</i> = 76 (6, 23, 26, 4, 12, 5)	5	0.815	-4.07	144	196.005	104.07	-0.04068 (0.99707)	0.12 (0.98338)

1063 **(b) Hierarchical AMOVA within each basin**

1064 N.A.: not applicable

1065

1066 **Table 3. Estimates of genetic diversity of the mitochondrial control region in the Striped**
 1067 **dolphin samples.**

Sample	Number of polymorphic sites	Number of mutations	Number of haplotypes	Haplotype diversity (SD)	Nucleotide diversity π (SD)	Average number of nucleotide differences (SD)
Whole	89	95	58	0.988 (0.006)	0.022 (0.001)	13.8 (6.25)
Atlantic	52	52	18	1.0	0.024 (0.001)	14.9 (7.02)
Pacific	34	35	6	0.952 (0.096)	0.022 (0.003)	14.0 (7.16)
Mediterranean	65	67	35	0.974 (0.013)	0.017 (0.001)	10.65 (4.93)
Group 1	29	29	20	0.940 (0.032)	0.005 (0.0007)	3.03 (1.41)
Group 2	51	52	14	0.967 (0.030)	0.023 (0.001)	14.47 (6.80)

1068
 1069 The diversity indices are reported stratified by sample origin. The number of polymorphic
 1070 sites is the number of nucleotide positions showing mutations at least in one individual. The
 1071 total number of mutations is the number of mutations at each polymorphic site taking into
 1072 account n -times the sites bearing n mutations. The number of haplotypes is the number of
 1073 unique sequences and the haplotype diversity is the number of haplotypes divided by the
 1074 number of individuals. The nucleotide diversity π is the average number of nucleotide
 1075 differences per site between two sequences. The average number of nucleotide differences is
 1076 the number of difference between same nucleotides from each sequence divided by the total
 1077 number of nucleotides. Standard deviation is shown between brackets (SD).

1078 **Table 4. Time to the Most Recent Common Ancestor estimates (in Myrs).**

	NPRS	Bayesian estimate	
		Uncorrelated Exponential Clock	Uncorrelated Log-normal Clock
Striped dolphin	11.1 (8.8-13.5) *	8.8 (3.8-14.4)	7.5 (5.5-11.7)
Group 1	4.61 (2.3-6.9) *	3.4 (1.1-6.4)	1.7 (0.9-2.5)
Striped dolphin excluding Group 1	N/A	6.5 (3.3-11.5)	6.4 (5.1-10.4)

1079

1080 T_{MRCA} are reported as median estimates in million of years (main number) and 95% C.I.

1081 (Confidence or Credible Intervals; between brackets). NPRS is the Non Parametric Rate

1082 Smoothing method (**Sanderson, 2002**)(N/A=not applicable).

1083

1084

1085

Figure 1

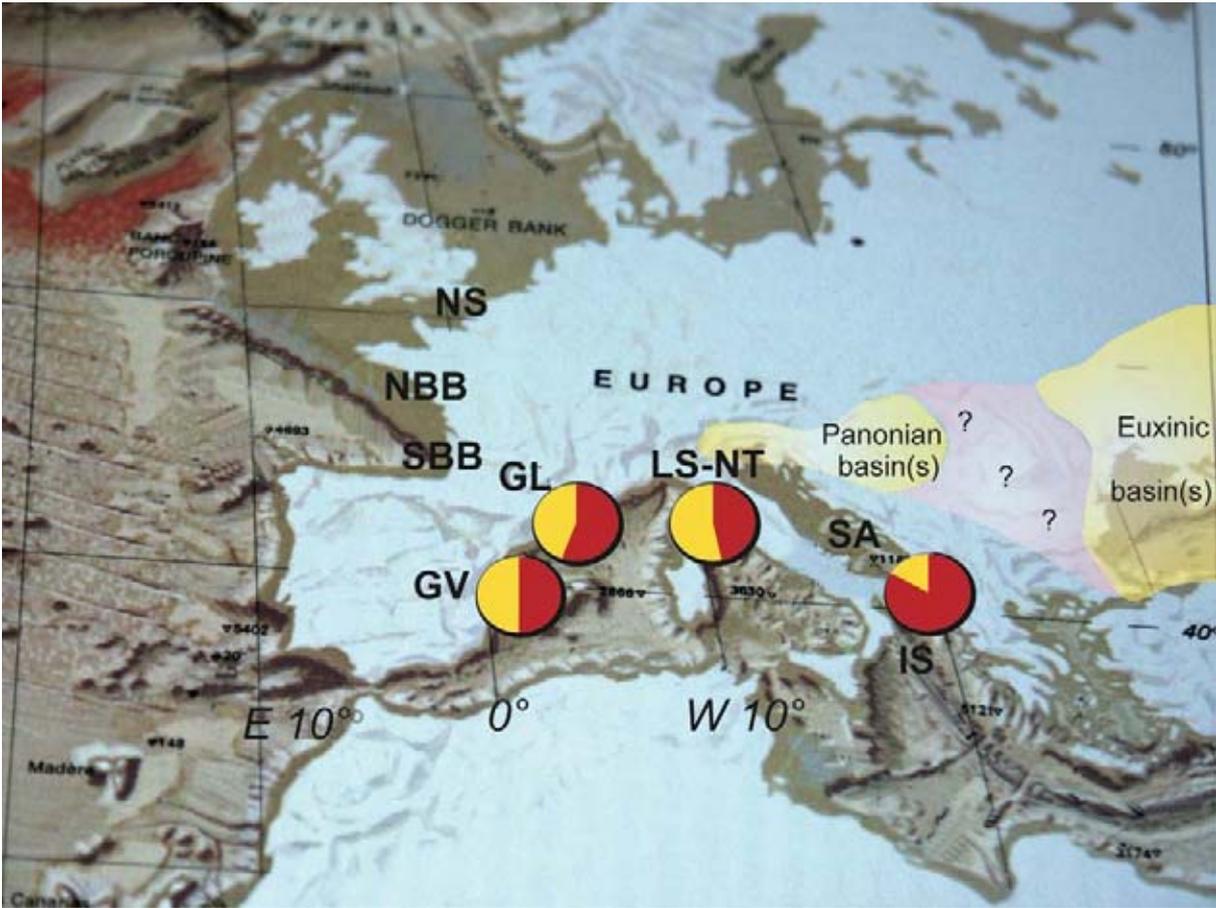


Figure 2

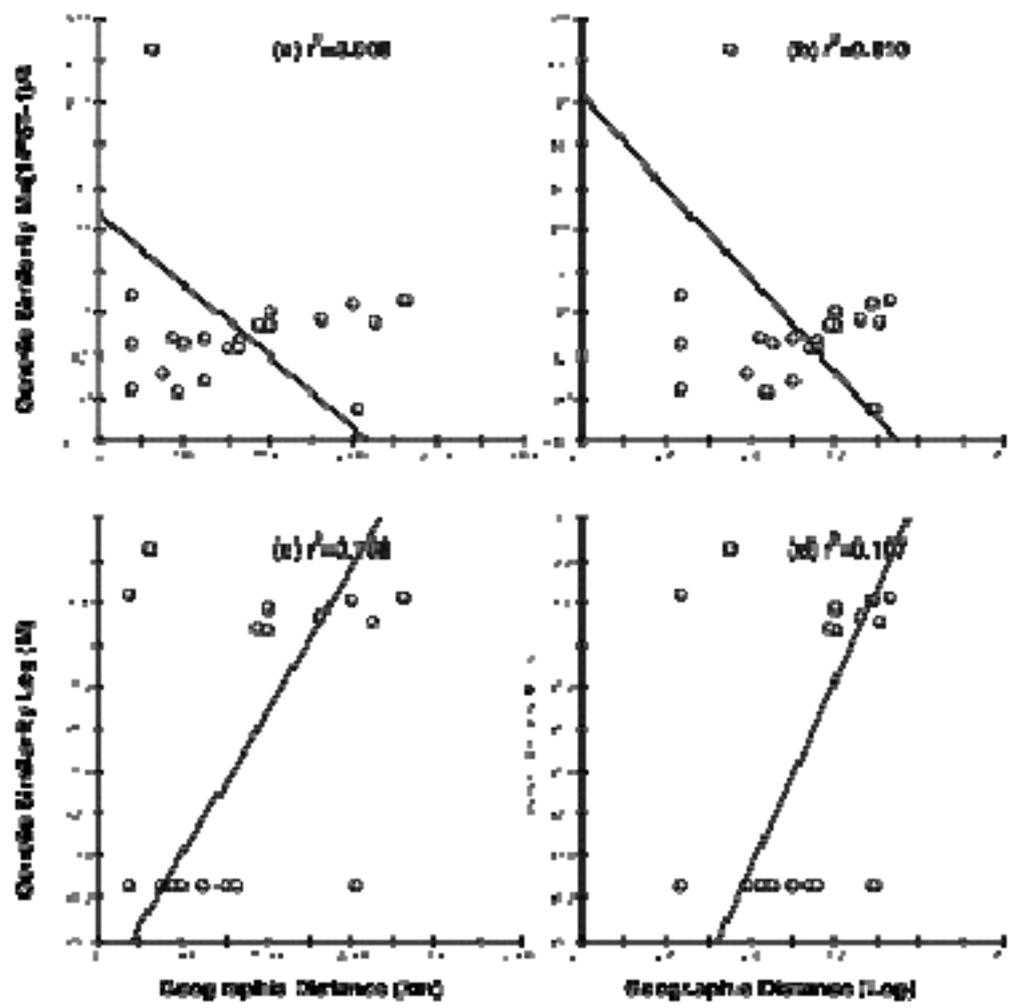


Figure 3

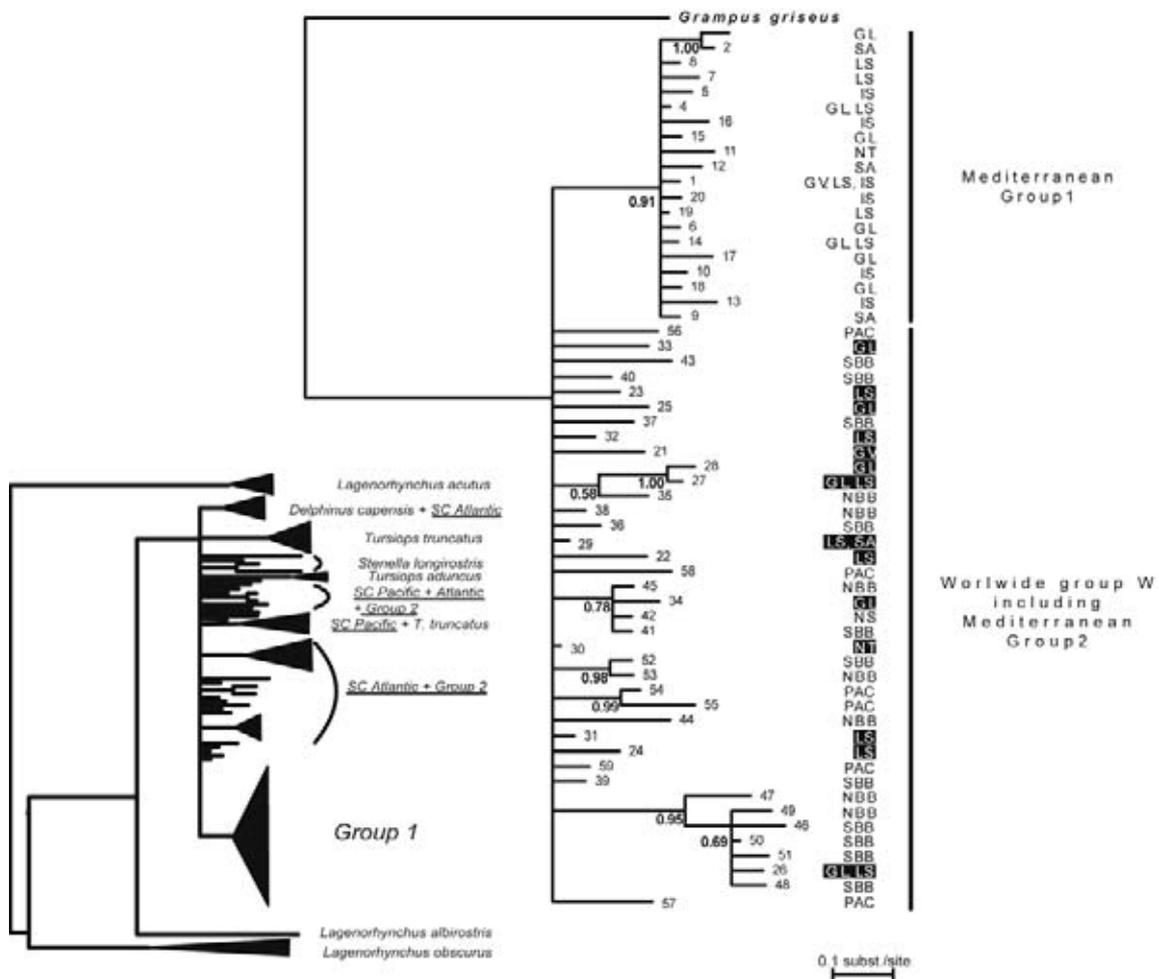


Figure 4

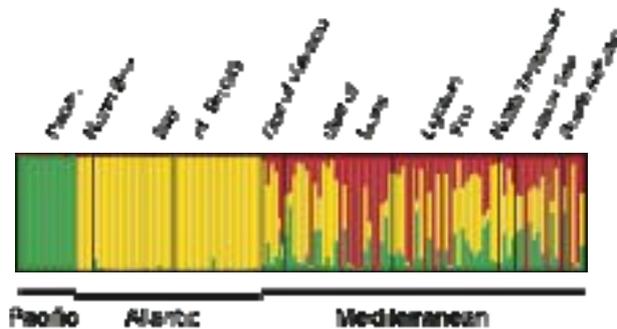


Figure 5

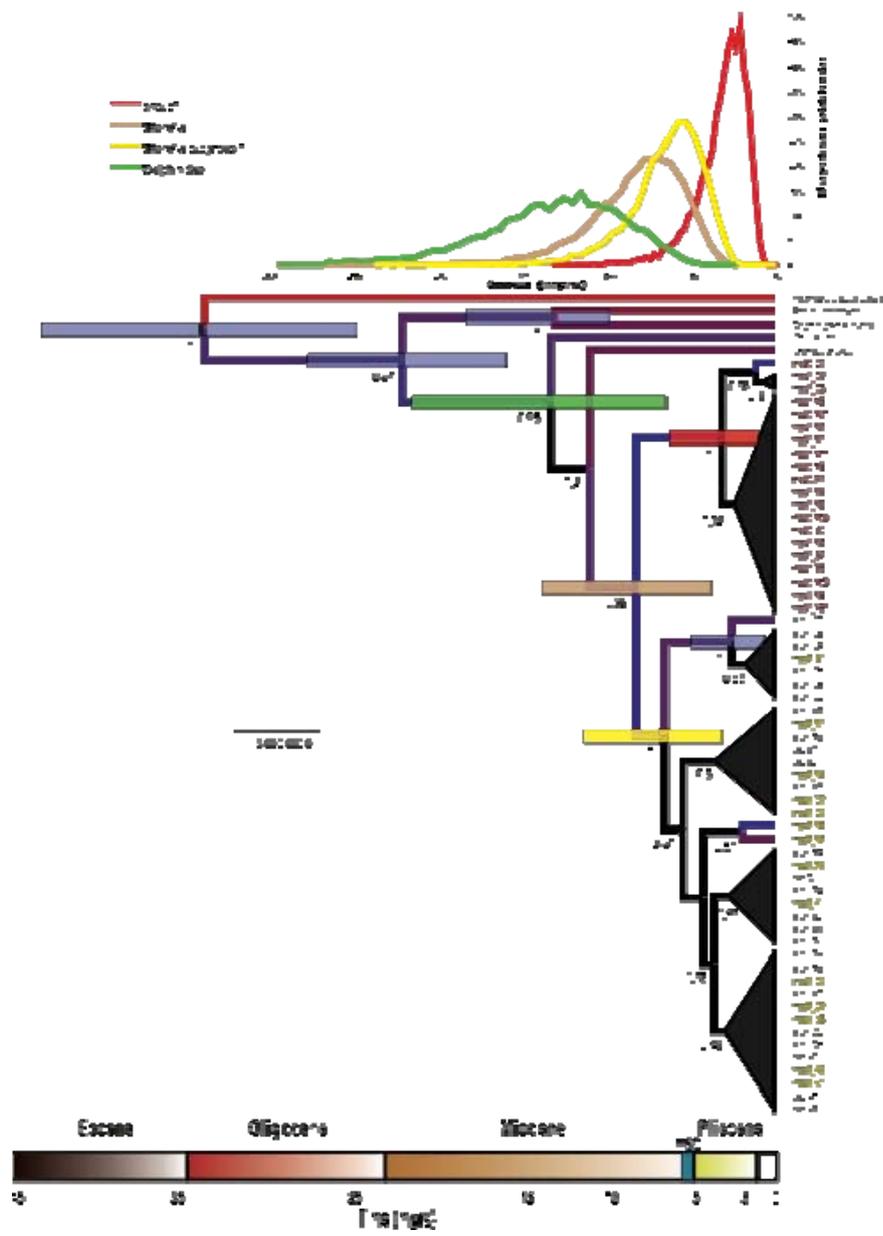


Figure 6

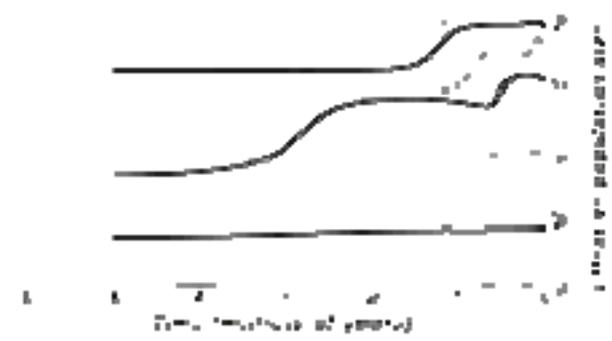
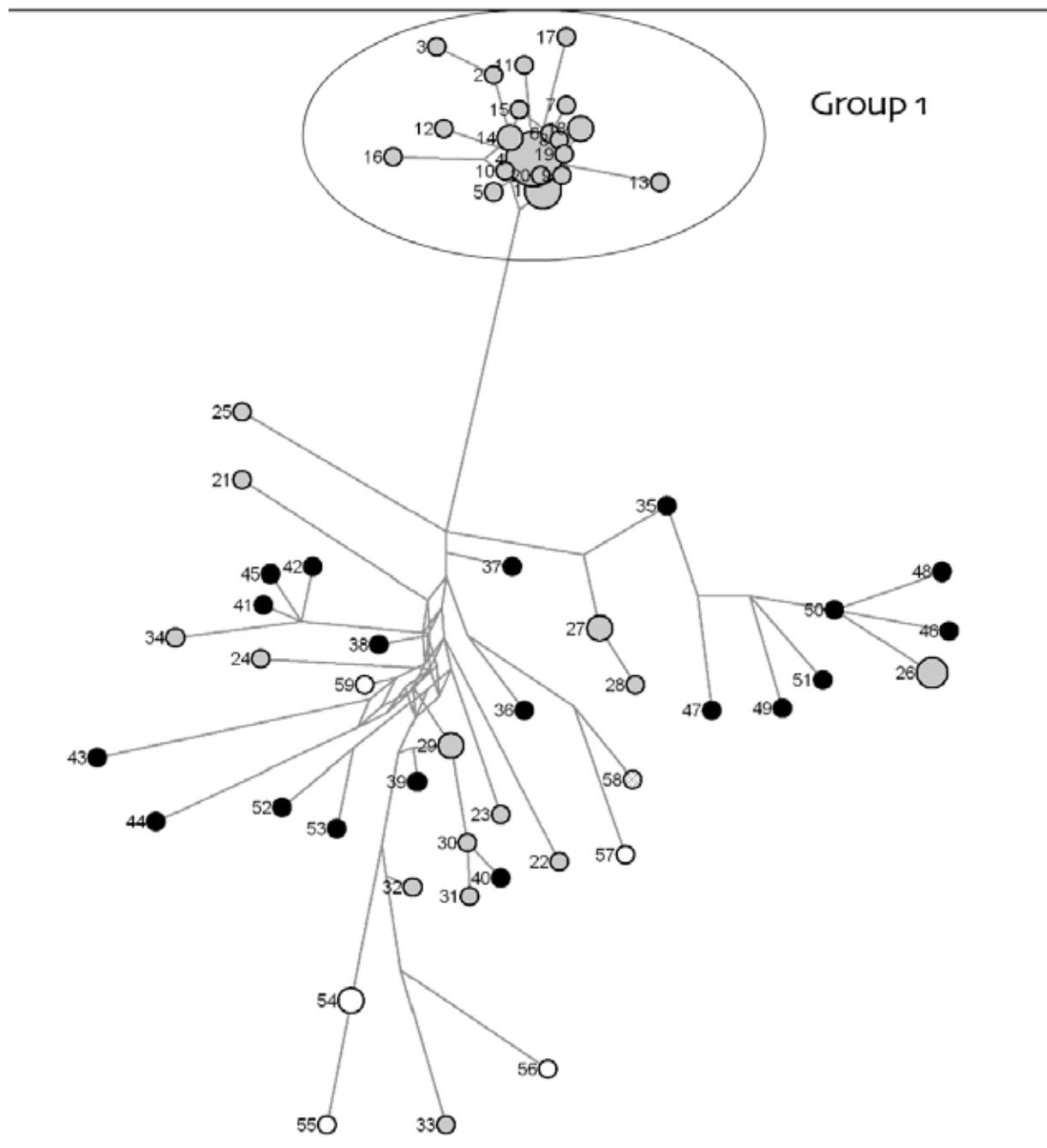


Figure S1: Median joining network based on the variation of the control region sequence from 73 Striped dolphins.



Numbers near each node denotes sample as reported in Figure 2 and Table S1. Node radius is proportional to the haplotype frequency in the whole dataset. Black nodes are Atlantic haplotypes, white are Pacific haplotypes, greyish are Mediterranean haplotypes.

Figure S2

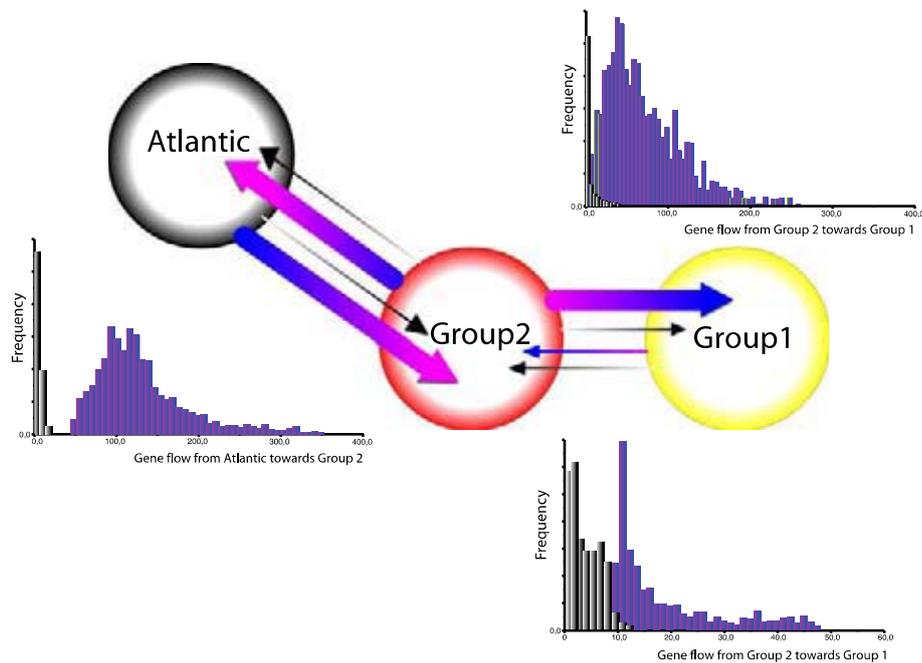
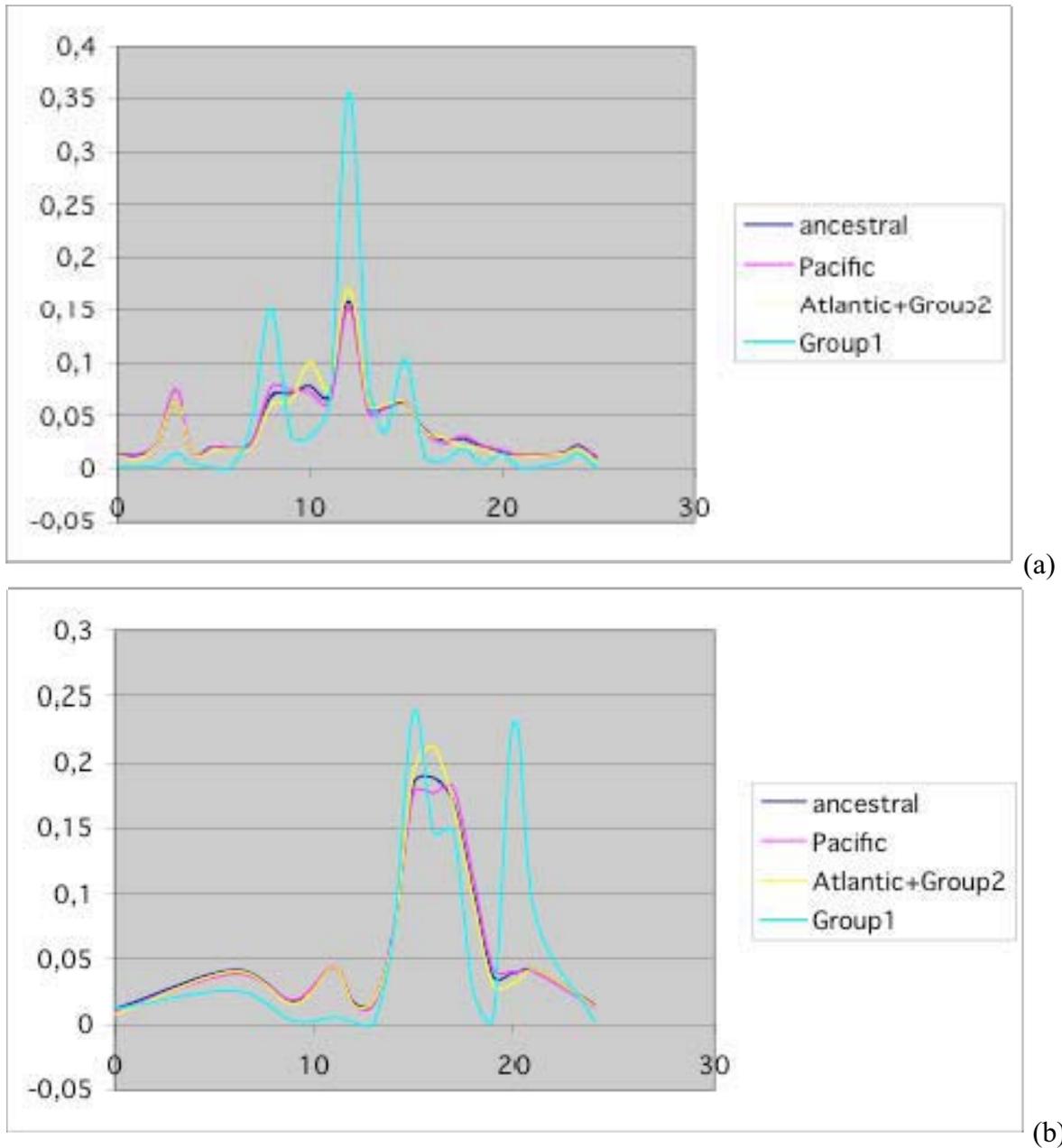


Figure S2. Gene flow between the Mediterranean groups and the Atlantic.

The genetic stocks are identified by circles and gene flows by arrows. Blue-pink arrows represent mitochondrial gene flow (*i.e.* female mediated gene flow) and black-white arrows represent gene flow as inferred by nuclear markers (female+male gene flow).

Graphs represent the probability density of gene flow in both direction for Group 1 vs Group 2 (blue-pink density for mitochondrial, black-white for microsatellite). Note the clear asymmetry denoted by the great difference in the median of the microsatellite and mitochondrial curves. In contrast, the gene flow is only represented in one direction for Atlantic vs Group 2 given its symmetry.

Figure S3: Bayesian estimated allele frequencies for microstallite *loci* GATA053 (a) and EV92Mn (b)

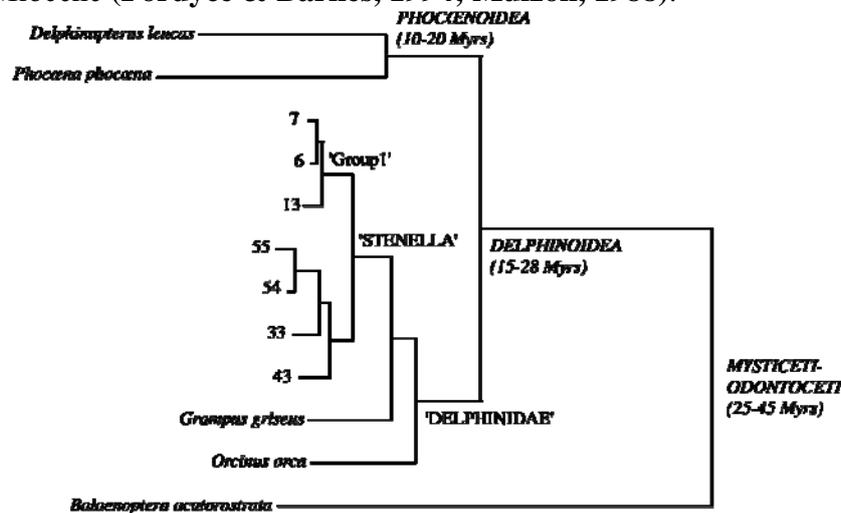


Curves represent allele frequencies plotted against allele size. Different colours are for different populations. Pacific, Atlantic+Group 2 and estimated ancestral profiles follow the same trend while Group 1 trend is quite different.

Figure S4. Fossil calibration of the nodes used for T_{MRCA} estimations.

In the following tree, 'Group1' and 'STENELLA' and 'DELPHINIDAE' are the nodes to be dated, the *DELPHINOIDEA*, *PHOCOENOIDEA* and *MYSTICETI-ODONTOCETI* are the calibration points. The tree is based on control region homologous sequences selected among this study's ones (numbers refer to table 1 allele numbering) as well as other Cetacea species' sequences. Other species are from laboratory-owned samples for *Grampus griseus* (EMBL #AM498741) and *Balaenoptera acutorostrata* (GenBank #AY230267) and database sequences for *Delphinapterus leucas* (EMBL #U18117), *Phocaena phocaena* (EMBL #AJ554063), and *Orcinus orca* (EMBL #M60409).

The *MYSTICETI-ODONTOCETI* age (tree root height) set at 25–45 Myrs is a consensus from (Fordyce, 1992; McLeod *et al.*, 1993; Milinkovitch, 1995). We also used very conservative prior for *DELPHINOIDEA* from both fossil and molecular calibrations: 15–28 Myrs (Fordyce & Barnes, 1994; Nikaido *et al.*, 2001; Waddell *et al.*, 2000). 10–20 Myrs was chosen for the *PHOCOENOIDEA* (Phocoenidae+Monodontidae clade) node following its dating in middle-mate Miocene (Fordyce & Barnes, 1994; Muizon, 1988).



REFERENCES

1. Fordyce RE (1992) Cetacean evolution and Eocene/Oligocene environments. In: Prothero DR, Berggren WA, editors. Eocene-Oligocene Climatic and Biotic Evolution. Princeton, NJ: Princeton University Press. pp. 368–381.
2. McLeod SA, Whitmore Jr FC, Barnes LG (1993) Evolutionary relationships and classification. In: Burns JJ, Montague JJ, Cowles CJ, editors. The Bowhead Whale. Lawrence, KA: Society for Marine Mammalogy Special Publication. pp. 45–70.
3. Milinkovitch MC (1995) Molecular phylogeny of cetaceans prompts revision of morphological transformations. Trends in Ecology and Evolution 10: 328-334.
4. Fordyce RE, Barnes LG (1994) The evolutionary history of whales and dolphins. Annual Review of Earth and Planetary Sciences 22: 419-455.
5. Nikaido M, Matsuno F, Hamilton H, Brownell RL, Jr., Cao Y, et al. (2001) Retroposon analysis of major cetacean lineages: the monophyly of toothed whales and the paraphyly of river dolphins. Proceedings of the National Academy of Sciences of the United States of America 98: 7384-7389.
6. Waddell VG, Milinkovitch MC, Berube M, Stanhope MJ (2000) Molecular phylogenetic examination of the delphinoidea trichotomy: congruent evidence from three nuclear *loci* indicates that porpoises (Phocoenidae) share a more recent common ancestry with white whales (Monodontidae) than they do with true dolphins (Delphinidae). Molecular Phylogenetics and Evolution 15: 314-318.
7. Muizon Cd (1988) Les relations phylogenetiques des Delphinida (Cetacea, Mammalia). Annales de Paléontologie 74: 159-227.

Table S1. Microsatellite polymorphism summary by locality.

Areas		Microsatellite loci				
		GATA098	MK9	GATA053	MK6	EV92Mn
North Pacific (n=14)	<i>Na</i>	10	9	6	15	8
	<i>Ho</i>	0.86	0.86	0.71	0.99	0.50
	<i>He</i>	0.92	0.87	0.67	0.93	0.90*
North Atlantic (n=45)	<i>Na</i>	11	13	9	21	13
	<i>Ho</i>	0.83	0.78	0.60	0.92	0.52
	<i>He</i>	0.87	0.88	0.62	0.91	0.84***
Mediterranean (n=79)	<i>Na</i>	9	9	8	15	9
	<i>Ho</i>	0.80	0.73	0.32	0.91	0.62
	<i>He</i>	0.84	0.77*	0.53***	0.85	0.83***

* Significant (* $p < 0.05$, ** $p < 0.01$ *** $p < 10^{-5}$) deviation from Hardy-Weinberg expectation

(100172 Monte Carlo steps)

Microsatellite polymorphism is shown within each sampling area and given the microsatellite.

Na is the total number of alleles at each locus, *Ho* the observed heterozygosity and *He* the expected heterozygosity.

Table S2. Mantel test for matrix correlation between genetic similarity/log (genetic similarity) and geographic distance/log(geographic distance).

	Z	r	p(null : r ≥ 0)	p(null : r ≤ 0)	r ²
ANALYSIS (a) raw data (M vs dist.)	68854.187175	-0.070791	0.403000	0.597000	5.0114182051e-03
ANALYSIS (b) M vs log(dist.)	295.866900	-0.104326	0.315000	0.685000	1.0883897294e-02
ANALYSIS (c) log(M) vs dist.	-27244.695966	0.433784	0.927000	0.073000	1.8816875300e-01
ANALYSIS (d) log(M) vs log(dist.)	-102.312455	0.327163	0.881000	0.119000	1.0703539681e-01

Z-statistics is calculated as $\sum_{i,j} A_{ij} B_{ij}$ where A is the genetic distance matrix and B the geographic distance matrix ; r is a standardised Z ; one-tailed p-values for null hypothesis of negative or positive correlation are provided ; r² is the correlation coefficient of the Reduced Major Axis regression.

26	-G...CG.....C.....G...T...C.T.....G.....C.....T.....TCTC.TT..G.....	1	2	3
27	-G.T..C.....T.T.A.....T..T.G.....T..CTC.....T.....-G..	1	1	2
28	-G.T..C.....T.T.A.....T..T.G.....T..CTC.....T.....	1		1
29	-G..T.C.....G.....T..A.....C.....T.....C..-G..	1	1	2
30	-G..T.C.....G.....T..A.....T.....C..-G..		1	1
31	-G..T.C.....G.....T..AT.....T.....C..-G..	1	1	1
32	-G..T.C.....T.....C.T.T..G..C..C.....T..GTT.....-G..	1	1	1
33	-C..G..-.....C.....A.....T.TT.G..C.C.C.....T.T.GTT.....-G..	1		1
34	-G..C.G.....C.....G.A.....TT.TT.G..C.C.C.....G..TCTC.....T.....-G..	1	1	1
35	-G.T..C.....-C...T.....G.A.....C.T.....G.....CTC.....T.....-G..	1		1
36	-G.T..C.....-C...T.....G.A.....C.T.....G.....TCT.....T.....-G..	1		1
37	-G..T.C.....G.....TG.A.....C.T..G.....TCT.....T.....	1		1
38	-G..T.C.....G.....G.....T..G.....CT.....TCT.....T.....-G..	1		1
39	-G..T.C.....G.....G.....C..T.TG.....C.....T.....C..-G..	1		1
40	-G..T.C.....G.....GT.....C.....T..A.....C.....AT.....C..-G..	1		1
41	-G..C.G.....G.A.....TT.T..G..C.C.C.....TCTC.....T..C..-G..	1		1
42	-G..C.G.....G.A.....TT.T..G..C.C.C.....TCTC.....T.....-G..	1		1
43	A..A..G..T.C.....G.A..C..T..T..G..C..A.....AT.TC.....C..-G..	1		1	
44	-A..G..T.CG.....G.....T.....C..T..T..A.....C.C.....T.TC.....T.....	1		1
45	-G..C.G.....G.A.T.....TT.T..G..C.C.C.....TCTC.....T.....-G..	1		1
46	-G..T.CG.....C..T.....T..C.T..T..G..C.....T..ATCTC.TT..G.....	1		1
47	-T.....TG.....T..C.T..T..G..C.....T..TCTC.TT..G.....	1		1
48	-G..CG.....C..TG.....T.C.C.T..T..G..CA.....T..TCTC.TT..G.....	1		1
49	-G..C.....C..TG..TT..C.T..T..G..C.....T..TCTCGTT.....	1		1
50	-G..CG.....C..TG.....T..C.T..T..G..C.....T..TCTC.TT..G.....	1		1
51	-G..C.....CA.....T..C.T..T..G..C.....T..TCTC.TT..G.....	1		1
52	-GT..T.C.....C.....A.....TC.T..G..C.....T..T.....-G..	1		1
53	-GT..T.C.....C.....A.....TC.T..G.....T..T.....-G..	1		1
54	-G..T.C.....AG.....G.A.T..C.T..T..A..C..C.....T.T.GTT.....-G..	2	2	2
55	-G..TT.C.....AC.....G.A.T..C.T..TT..A..C.A.C.....T.T..TT.....-G..	1	1	1
56	-G..C.....GA.....T..G.T..C..GC.....T..C.TTC.....-G..	1	1	1
57	-G.T..C.GC.....A.....T.T.....G.....C.....TCT.....	1	1	1
58	-G.T..GC.....A.T.....T.....G.....C.A.....C.TCT.....	1	1	1
59	-G..T.C.....G.A..C..T.....G.....C.....TCTC.....T.....-G..	1	1	1

Table S4. Gene flow Support Intervals (95%) among Striped dolphin populations given by Bayesian estimation from microsatellite and mtDNA markers.

	Pacific			Atlantic			Group 1			Group 2		
	P>A	P>I	P>2	A>P	A>I	A>2	1>P	1>A	1>2	2>P	2>A	2>I
microsatellites	2.48-5.80	5.75-9.65	10.00-29.77	2.08-10.43	2.19-12.14	5.2-14.85	2.87-9.27	2.69-10.82	4.39-15.25	2.15-15.03	1.44-4.13	0.85-1.63
mtDNA	0.01-26.64	0.01-29.24	8.73-71.77	0.01-52.27	0.65-128.84	17.92-193.45	0.01-32.89	0.01-58.11	0.31-46.70	0.01-152.72	12.24-189.34	0.01-172.89
overall	2.43-12.80	1.59-2.94	1.78-33.23	1.78-3.40	1.91-3.77	10.31-55.33	0.97-2.34	0.96-2.47	1.41-3.08	0.67-1.78	0.95-12.14	0.95-1.67

Gene flow ($M_{ij} = m_{ij}/\mu$) between the sample i and the sample j are given as estimated by microsatellite, mitochondrial control region and whole (“overall”) marker dataset analysis. $i>j$ means gene flow from the population i towards the population j (P=Pacific, A=Atlantic, 1=Group 1, 2=Group 2).

C.3) Conclusions

Ce travail a permis de distinguer deux lignées haplotypiques mitochondriales principales au niveau mondial dont une n'est présente qu'en Méditerranée (« haplogroupe 1 »). Le temps de coalescence de ces deux lignées évalué par deux méthodes est considérablement élevé, entre -11 et -5 millions d'années, et pourrait faire remonter la séparation des deux groupes à la Crise de Salinité du Messinien (env. -5 M.a.). En d'autres termes, l'haplogroupe 1 a divergé plus précocement que d'autres genres-espèces du clade *Tursiops-Stenella-Delphinus*. Le groupe d'individus porteurs de l'haplogroupe 1 (= groupe 1) est caractérisé par une diversité nucléotidique et haplotypique faible au regard des autres populations.

L'origine du processus de divergence pourrait se situer dans l'isolement méditerranéen lors de la Crise de Salinité du Messinien. La coexistence sympatrique de ces deux groupes au sein de la Méditerranée aurait pu persister jusqu'à maintenant sous l'influence de facteurs comportementaux (animaux d'écotypes différents par exemple comme observé chez le Grand Dauphin - (Hoelzel *et al.*, 1998) -), culturels (évolution matrilineale comme observé chez l'Orque - (Hoelzel & Dover, 1991)) ou encore génétiques : outbreeding avoidance - (Thornhill, 1993) - ou incompatibilité fœto-maternelle - (Laitinen *et al.*, 1993b) - lorsque la fécondation a bien eu lieu).

La seconde lignée (« haplogroupe 2 »), présente dans le monde entier, sauf en Méditerranée orientale, pourrait représenter un groupe ayant réinvesti la Méditerranée occidentale à la faveur de la réouverture du détroit de Gibraltar, en provenance de l'Atlantique. L'analyse des microsatellites montre en outre une différenciation génétique plus importante entre les deux groupes méditerranéens qu'entre bassins océaniques.

Chapitre IV : Étude du DRB chez le Dauphin bleu et blanc et perspectives générales

I. Introduction et buts de l'étude

Pour une plus ample description de l'évolution du système immunitaire adaptatif, cf. **Chapitre I.IV**.

I. A. Diversité du CMH et *balancing selection*

A.1) Evolution non-neutre et polymorphisme élevé

Les polymorphismes nucléotidique et haplotypique élevés observés au sein des *loci* exprimés du CMH sont parmi les plus élevés de tout le génome chez l'Homme. Il est communément admis que cette diversité élevée est conditionnée par l'intervention d'une **sélection « balancée »** (*balancing selection* - maintien d'un polymorphisme équilibré au sein d'une population, facteur principal de maintien de la diversité génétique). A l'aune de cette hypothèse, la persistance des allèles est plus longue que prédite par l'hypothèse de **neutralité sélective**, c'est-à-dire lorsque la probabilité de transmission d'une génération à l'autre de tous les allèles est la même et ne dépend donc que du hasard. Les mécanismes permettant le maintien du polymorphisme sont divers. Citons par exemple la **sélection fréquence-dépendante**, mécanisme par lequel la sélection peut préserver plus de deux états phénotypiques par la diminution du *fitness* du génotype lorsque la fréquence de ce dernier devient trop élevée. Un autre mécanisme est l'**avantage des hétérozygotes**, avantage sélectif d'un allèle récessif à l'état hétérozygote alors qu'il présente un désavantage à l'état homozygote, ce qui a pour conséquence une augmentation de la fréquence de cet allèle au sein de la population. Ce mécanisme appelé aussi surdominance ou *overdominance*, est celui à l'œuvre dans la drépanocytose.

Les processus de sélection agissant sur le CMH pourraient être à l'origine du rôle carrefour du système immunitaire entre *fitness* et comportement. Cette évolution peut être dirigée par des pathogènes, des phéromones induisant une sélection sexuelle (appariement assorti - *assortative mating* - ou discordant - *disassortative mating*) ou encore l'incompatibilité foëto-maternelle (**Hedrick, 1998**). L'étude de la dynamique évolutive du CMH s'est donc imposée dans des disciplines aussi variées

que l'écologie comportementale, la génétique des populations, la biologie de la conservation et la toute récente médecine évolutive. Ainsi, la plupart des recherches menées sur le CMH ont-elles porté sur des espèces d'intérêt médical (Homme, Rat, Souris, Primates non humains) en particulier pour l'étude des transplantations.

A.2) Le polymorphisme trans-spécifique

Le DRB est une région complexe comportant des *loci* dupliqués en nombre inconnu chez les Cétacés. Une étude sur le Béluga (**Murray & White, 1998**) a cependant laissé supposer un nombre de deux *loci* mais ce nombre est basé sur une phylogénie des allèles et non sur une localisation cytogénétique des séquences, indispensable pour déterminer le nombre de *loci*.

I. B. Evolution médiée par les pathogènes

B.1) Cadre général

Le succès d'un agent pathogène consiste à submerger ou contourner les défenses de l'hôte. Ces pathogènes ont donc développé des stratégies leur permettant d'infecter un hôte en évitant la détection ou la destruction par le système immunitaire, c'est ce qu'on appelle l'évasion immunitaire (**Alcami & Koszinowski, 2000**). Cette évasion passe par la localisation intracellulaire (*Salmonella, Plasmodium, Leishmania, Mycobacterium, Chlamydia ...*), la sécrétion de composés immunosuppresseurs (*Pseudomonas, Burkholderia ...*) ou la production de protéines leurres pour les anticorps (*Streptococcus, Staphylococcus ...*) (**Cantin et al., 2005; Finlay & Falkow, 1997**). Une des stratégies les plus efficaces consiste en la variation antigénique, c'est à dire changer constamment les cibles du système immunitaire adaptatif (*Trypanosoma, VIH ...*) (**Burton et al., 2005**).

Le système immunitaire adaptatif (cf. **Chapitre I.IV**) a ainsi évolué en réponse à ces pathogènes afin de maintenir l'homéostasie de l'individu. On peut donc penser qu'il existe des co-évolutions entre hôtes et pathogènes dont l'élément central est à rechercher auprès des molécules du Complexe Majeur d'Histocompatibilité.

B.2) Cas des *Paramyxoviridae*

Certains travaux (**Ovsyannikova et al., 2004**) suggèrent un lien entre DRB et présentation du virus de la Rougeole de l'Homme (Morbillivirus comme le Dolphin Morbillivirus).

I. C. Le DRB chez les Cétacés/Cétartiodactyles

Les molécules du Complexe Majeur d'Histocompatibilité ont longtemps été considérées comme moins polymorphes chez les Mammifères marins (**Trowsdale et al., 1989**) en raison d'une pression pathogène supposée moins importante en milieu marin. Cette affirmation a été mise en doute par des études plus récentes montrant des niveaux moyens de diversité génétique tant au sein des molécules de Classe I que de Classe II (**Flores Ramirez et al., 2004; Hayashi et al., 2006; Yang et al., 2005**). Cependant, la caractérisation du DRB par PCR-SSCP (*Single Strand Conformation Polymorphism*) chez le Béluga *Delphinapterus leucas* (**Murray & White, 1998**) ainsi que chez un certain nombre d'espèces de Mysticètes **Baker et al. (2006)** ont montré une diversité inférieure à celle des Mammifères terrestres. La question reste donc ouverte.

I. D. Objectifs

Il nous a donc semblé d'une part que l'espèce étudiée et d'autre part la sensibilité des techniques de génotypage pourraient avoir une influence notable sur l'évaluation de la diversité. Il apparaît donc nécessaire d'explorer un éventail de modèles biologiques marins élargi mais aussi de faire appel à des techniques éventuellement plus sensibles.

En outre, le Dauphin bleu et blanc est intéressant car il a subi en 1991 et 2003 deux épisodes d'infection au Dolphin Morbillivirus. On estime que l'épisode de 1991 a pu conduire à une très forte mortalité au sein de la population méditerranéenne de cette espèce (**Domingo et al., 1990; Van-Bresseem et al., 2001**) mais on a toujours aucune idée sur l'éventuelle influence d'une prédisposition génétique ou du rôle des contaminants d'origine humaine (**Kannan et al., 1993a**).

Nous avons donc tenté d'évaluer la diversité de la séquence de l'exon 2 du DRB,

correspondant au PBR (*Peptide Binding Region*), partie « émergée » en-dehors de la surface cellulaire de la molécule, chez le Dauphin bleu et blanc. Notre projet initial était d'obtenir suffisamment d'échantillons afin d'une part d'évaluer la distribution relative des différents allèles au sein des populations méditerranéennes décrites au chapitre précédent (cf. **Chapitre III, IV. C**), et d'autre part de détecter d'éventuelles déviations au modèle de *balancing selection* en vigueur sur le CMH et qui pourraient laisser suspecter une évolution dirigée par un pathogène.

II. Matériel et méthodes

II. A. Matériel biologique et *in silico*

A.1) Dauphins bleu et blanc échoués

Un total de 17 Dauphins bleu et blanc ont été échantillonnés, ainsi qu'un Grand dauphin *Tursiops truncatus*, un Rorqual commun *Balænoptera physalus* et deux Baleines grises *Eschrichtius robustus*. Les échantillons ont été obtenus à partir d'animaux échoués sur les côtes méditerranéennes pour les Dauphins bleu et blanc et un prêt de tissus pour les autres espèces (cf. **Chapitre III, II**).

Parmi ces échantillons, 8 individus étaient porteurs d'un haplotype mitochondrial du groupe 1 et 8 d'un haplotype du groupe 2.

A.2) Séquences de DRB de Cétacés issus de GenBank

Les 34 séquences de DRB issues de l'étude de Baker *et al.* (2006) ont été téléchargées à partir de la base de données GenBank et ont été ajoutées à l'étude. Il s'agit des espèces suivantes : la Baleine franche australe *Eubalæna australis*, la Baleine franche du Groënland *Balæna mysticetus*, la Baleine grise *Eschrichtius robustus*, la Baleine à bosse *Megaptera novæangliæ*, le Rorqual bleu *Balænoptera musculus*, le Cachalot *Physeter macrocephalus*, le Mésoplodon de Hubbs *Mesoplodon carlhubbsi* et l'Hyperoodon arctique *Hyperoodon ampullatus*.

II. B. Electrophorèse sur Gel à Gradient de Dénaturation, DGGE

B.1) Principe

1(a) Une alternative au clonage

Afin de séparer les populations de bandes issues de l'amplification de *loci* « *n*-pliqués » (au plus $2n$ fragments de séquences différentes chez un individu hétérozygote pour l'ensemble des n *loci*), on utilise un gel d'acrylamide présentant un gradient d'agents dénaturants. En effet, en raison de la taille similaire des fragments, les gels classiques ne permettent pas de séparer les bandes. Théoriquement, une seule mutation peut conduire à des différences de caractéristiques de « dénaturabilité » de l'ADN. Ainsi, au cours de la migration au sein du gel de polyacrylamide, les fragments rencontrent des concentrations croissantes en agents dénaturants. Lorsqu'un fragment d'une séquence donnée rencontre ses conditions limites de dénaturation, les portions les plus faiblement cohésives de la séquence vont commencer à se dénaturer et sa migration va se ralentir brutalement.

Ces propriétés ont permis à la DGGE de s'imposer pour l'analyse de prélèvements microbiens environnementaux à la place du clonage (cf. **Figure 41**), long, coûteux et nécessitant des répliques afin de capturer la totalité de la diversité.

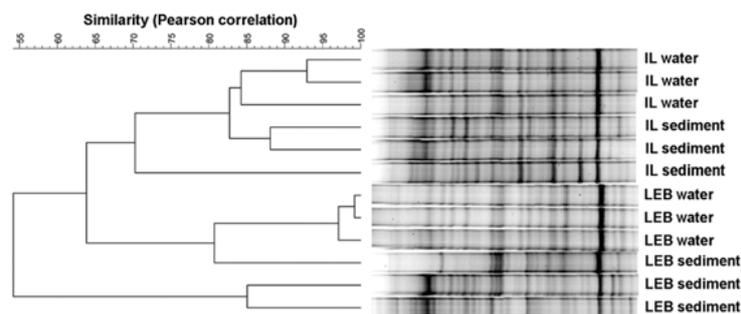


Figure 41. Etude par partitionnement - *clustering* - de profils de DGGE de populations d'*Escherichia coli* issues de prélèvements environnementaux.

1(b) Application au génotypage à moyenne échelle

La DGGE a permis d'effectuer des génotypages à des échelles moyennes ne nécessitant pas la création de puces. Ainsi, elle a été utilisée dans la caractérisation de communautés microbiennes - biofilms épilithiques (Lyautey *et al.*, 2005) ou

métagénomique du sol (Ascher *et al.*, 2009) - voire même dans la caractérisation de certaines (van der Hout *et al.*, 2006).

B.2) Mise en œuvre

2(a) Amorces

En dehors du PBR, la majeure partie des molécules du CMH de Classe II sont des domaines transmembranaires. Ces domaines évoluent donc dans un milieu peu variable et peu soumis à d'éventuelles pressions de sélection extérieure (membrane cellulaire dont les caractéristiques physiques évoluent peu). Seul l'exon 2, contenant la séquence codant le PBR, est hypervariable et il est flanqué de séquences hautement homologues comme on peut s'en apercevoir en alignant des séquences issues d'espèces éloignées.

Cela nous a permis d'utiliser des amorces hétérologues utilisées pour le typage du DRB chez les Primates (**Knapp et al., 1997**) :

5'MDRB

5'-GCCTCGAGTGTCCCCCAGCACGTTTC-3'

MDRB1+GC

5'-CGCCCGCCGCGCCCCGCGCCCGCCCGCCCGCCCGCCCGCCCGCAGCTTTCACCTCGCCGCTG-3'

2(b) PCR

Les conditions de la réaction de PCR sont les suivantes :

- composition du Mix (dans 50 µl) : 50 ng d'ADN génomique, 1.5 mM de MgCl₂, 2.5 mM de chaque dNTP, 25 pmol de chaque amorce et 2-3 unités de *Taq* polymérase ;
- programmation du thermocycleur : 95°C pendant 3 minutes, puis 30 cycles (94°C/90 s., 49°C/90 s., 74°C/90 s.), suivis d'une étape d'élongation à 72°C pendant 10 minutes.

Ces amorces conduisent à l'amplification d'un fragment de 320 paires de bases confirmé par la migration sur un gel à 1,5% d'agarose.

2(c) Migration

Le gel de DGGE consiste, de bas en haut, en :

- une « pile » (*stack*) de gel située en position déclinée (nous avons créé cette « pile » pour éviter les fuites lors du coulage du gel) présentant le maximum de concentration d'agents dénaturants (urée et formamide) ;
- un corps de gel contenant le gradient d'agents dénaturants. Dans notre cas, les deux solutions ont été titrées à 30 et 65% et sont distribuées dans le gel à l'aide d'un dispositif de distribution progressive ;
- une seconde « pile » de gel dépourvue de polyacrylamide dans laquelle sont moulés les puits de chargement.

Ce gel est immergé dans du TAE 1X, dans une cuve verticale (BIORAD D-GENE; Biorad, Richmond, CA ; cf. **Figure 42**). 30 µl de produit de PCR sont chargés dans chaque puits, mélangés à 15 µl de tampon de charge (*loading buffer*). La migration est effectuée à la tension de 300 Volts et 60°C de température. Après plusieurs essais, nous avons estimé que la durée optimale afin d'obtenir une bonne séparation des

bandes était de 5 heures.



Figure 42. Cuve de migration en action.

Pour chaque migration, 2 individus contrôles étaient fournis afin d'avoir un témoin de migration des bandes pour un génotypage normalisé. La révélation des bandes était effectuée après coloration au SYBR-Green (Invitrogen-Molecular Probes, Eugene, OR).

2(d) Réamplification et séquençage

En raison de la ténacité du gel et la proximité des bandes, chacune de celles-ci a été excisée à l'aide de cônes de 200 μ l à ouverture large (Continental Lab Products, San Diego, CA). Afin de provoquer une libération des amplicons dans la solution, les morceaux obtenus ont été immergés dans 50 μ l et soumis à trois cycles congélation-décongélation. La réamplification a alors porté sur 5 μ l de l'éluat obtenu.

La réaction de PCR « semi-nichée » a été menée dans les conditions suivantes :

- amorces : 5'MDRB et une amorce à l'intérieur de la séquence 3'DRB1M13 (5'-TGTAACGACGGCCAGTCGCTGCACTGTGAAGCTCTC-3') ;

- mélange réactionnel : 2.5 mM de MgCl₂, 2.5 mM de chacun des dNTP, 12.5 pmol par tube de chacune des amorces et 2 à 3 unités of Taq polymerase dans un volume final de 50 μ l ;

- programmation du thermocycleur : 30 cycles (94°C pendant 20 s., 65°C pendant 30 s., 72°C pendant 36 s.) puis 72°C pendant 10 min.

Après confirmation sur gel d'agarose, les fragments obtenus (210 pb) ont été purifiés à l'aide du kit Qiaquick PCR Purification Kit (Qiagen GmbH, Hilden, Germany, réf. 28106) et le séquençage effectué sur un séquenceur à capillaires CEQ™ 2000XL Sequencer (Beckman Coulter, Fullerton, CA) en utilisant le protocole Dye Terminator

Cycle Sequencing (Beckman Coulter, réf. 608120).

2(e) Analyse des séquences

e(i) Phylogénie des allèles

Les séquences ont été reconstituées, par recouvrement des fragments issus des deux différentes amorces, à l'aide du logiciel GENALYS. L'ensemble des séquences a par la suite été aligné en utilisant CLUSTALX avec 44 séquences déjà publiées d'exon 2 du DRB de 10 espèces de Cétacés (références GenBank n° DQ354667 à DQ354688 et AF012930 à AF012941) ainsi que 3 espèces supplémentaires utilisées comme groupe externe. Le modèle optimal de mutation a été déterminé de la même manière que pour la région de contrôle et l'amélogénine (Posada & Crandall, 1998).

e(ii) Statistiques sommaires et sélection

Nous avons analysé le polymorphisme à l'aide du logiciel DNASP (Rozas *et al.*, 2003). D'éventuels points de recombinaison ont été recherchés à l'aide du même logiciel, par calcul du paramètre de recombinaison $R = 4Nr$ (N = taille effective, r = taux de recombinaison) en raison des perturbations que ces points peuvent engendrer lors de l'analyse phylogénique. De même d'éventuelles signatures de conversion génique (recombinaison entre localisation non homologues) ont été recherchées par maximum de vraisemblance (Betran *et al.*, 1997).

III. Méthodologie : analyse des signatures de sélection

III. A. Evaluation de l'hétérogénéité des modèles de substitution entre les 3 positions de codons

A.1) Principe

Un moyen de détecter l'effet de la sélection sur une séquence codante est de tester si le modèle de mutation de la troisième position nucléotidique des codons est différent de celui des autres positions. En effet, si l'on se réfère au code génétique, chaque codon (ensemble de trois nucléotides) est censé correspondre à un acide aminé mais des redondances existent dans ce code et la troisième position est celle dont la variation détermine le plus souvent un changement d'acide aminé.

L'hypothèse est que si le modèle de mutation de la troisième position diffère de celui des deux autres positions, l'évolution n'est pas neutre et que cette portion du génome est soumise à sélection.

Afin d'estimer de façon différentielle les paramètres du modèle de mutation à chaque position, nous avons partitionné l'alignement en fonction de la position des nucléotides dans les codons (les premiers dans la première partition, etc.).

A.2) Implémentation logicielle : BEAST

L'estimation du modèle de mutation relatif de la troisième position de codon a été menée à l'aide du logiciel BEAST (Drummond *et al.*, 2002) conjointement avec l'estimation des T_{MRCA} de lignages d'allèles sélectionnés. Les paramètres autorisés à varier ont été : profondeur de l'arbre, topologie, taux de mutation et paramètres de population.

Le paramètre transition/tranversion κ du modèle HKY85 (Hasegawa *et al.*, 1985) a été pris en compte séparément pour chacune des positions de codon.

III. B. Analyse codon par codon : Vraisemblance par Effets-Fixés, *Fixed-Effect Likelihood*

Une méthode par maximum de vraisemblance approchée a été utilisée pour détecter et quantifier la sélection, positive et négative, à l'échelle de chaque codon individuel. Cette approche appelée *Fixed-Effect Likelihood* (FEL ; (Kosakovsky Pond *et al.*, 2005) consiste à ajuster, indépendamment à chaque position de l'alignement nucléotidique, un modèle *full-codon*, utilisant une matrice MG94 modifiée pour tenir compte des taux de mutations synonymes (α) et non-synonymes (β) (Muse & Gaut, 1994) sous deux hypothèses : d'une part H_0 l'hypothèse d'évolution neutre ($\alpha=\beta$), d'autre part H_1 , l'hypothèse alternative (α et β libres d'évoluer indépendamment). Les estimateurs du modèle sont la covariance des longueurs de branche et des paramètres de la matrice du modèle de substitution et sont déduits de l'alignement

dans son ensemble.

La probabilité postérieure $p_{i,j}$ d'observer les taux α_i et β_i au site s est calculée à partir de l'observation des vraisemblances conditionnelles calculées pour chaque site :

$$l_{i,j} = L(D_s | T, \Theta, \alpha_i, \beta_s = b_j)$$

soit d'après la règle de Bayes : $p_{i,j} = \Pr\{\alpha_s = a_i, \beta_s = b_j | D_s\} = \frac{l_{i,j}\{\alpha_s = a_i, \beta_s = b_j\}}{\sum_{m,n} l_{m,n}}$

Le Facteur de Bayes (*Bayes factor*) K est donné par le quotient des probabilités postérieures des modèles comparés deux à deux (e.g. M_1 et M_2) sur la base du vecteur de données x :

$$K = \frac{p(x | M_1)}{p(x | M_2)}$$

Le support de l'hypothèse est fonction de la grandeur de K (cf. Tableau 9).

On remarque aisément que cette démarche est le corrolaire du test de quotient des vraisemblances (*Likelihood Ratio Test*, cf. **Chapitre I.IV. B.1(c)**).

K	Support du modèle M₁ vs M₂
< 1:1	Négatif (supporte M ₂)
1:1 to 3:1	Faible support
3:1 to 10:1	Support Substantiel
10:1 to 30:1	Fort Support
30:1 to 100:1	Très Fort Support
>100:1	Support Décisif

Tableau 9. Intensité du support du modèle M₁ comparé au modèle M₂ en fonction du *Bayes Factor K*.

Lorsque l'hypothèse de neutralité est rejetée (les auteurs conseillent une différence significative déterminée par le logarithme du facteur de Bayes supérieur ou égal à 5).

IV. Principaux résultats

IV. A. Une variabilité notable du DRB chez le Dauphin bleu et blanc

A.1) Variabilité du nombre d'allèles

La **Figure 43** ci-après montre la variabilité des profils de DGGE obtenus :

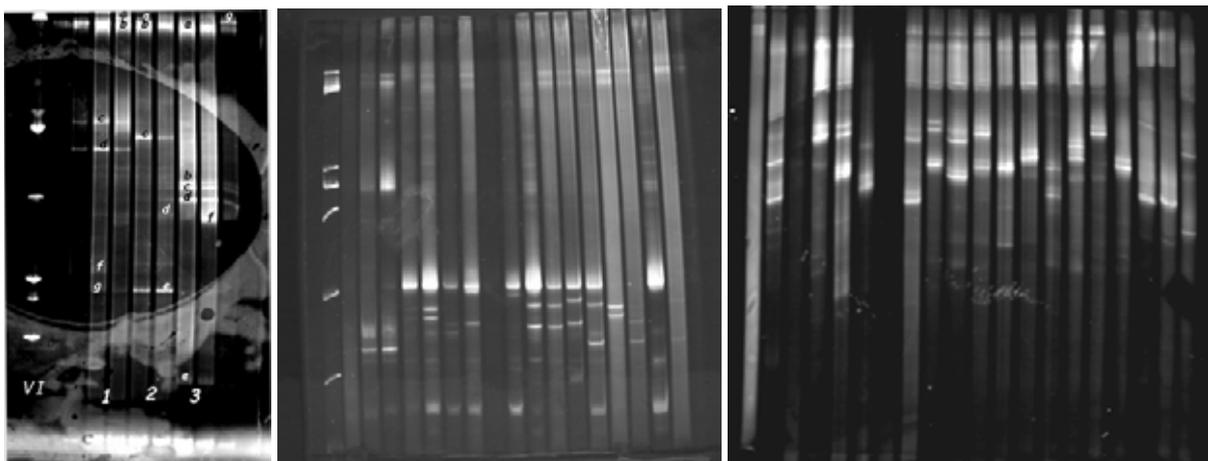


Figure 43. Profils de DGGE obtenus.

Les séquences obtenues ont été numérotées suivant les recommandations de l'IMGT (The international ImMunoGeneTics information system - <http://imgt.cines.fr/> -

(Klein *et al.*, 1990) de Stco-DRB#1 à Stco-DRB#15. Les séquences issues du Dauphin bleu et blanc ainsi que celles issues du Grand dauphin (5), le Rorqual commun (6) et la Baleine grise (4) sont représentées dans la **figure 1 de l'article n°5**. Le nombre d'allèles obtenu est important compte tenu de l'échantillon réduit.

A.2) Variabilité nucléotidique

Au sein des séquences Stco-DRB, on observe 33 sites polymorphes et une diversité nucléotidique de 0,0324 ($\pm 0,0001$). Aucun « point-chaud » de recombinaison n'a été mis en évidence ni aucune évidence de conversion génique, ce qui nous permet d'espérer une estimation de T_{MRCA} pas trop biaisée.

A.3) Phylogénie des allèles

L'arbre phylogénique du DRB observé est caractérisé par une structure fortement paraphylétique comme indiqué sur la **figure 3 de l'article n°5**. Un très important polymorphisme trans-spécifique est par ailleurs cohérent avec la théorie : des séquences de Mysticètes se retrouvent groupées avec des séquences d'Odontocètes à l'exception notable des séquences de *Monodontidæ*. Ce dernier point soulève le problème de l'identification de séquences du DRB chez cette espèce à partir de seules séquences nucléotidiques : s'agit-il vraiment de DRB et non de DQB ou d'autre *loci* (*cf.* description paragraphe IV.C) ?

Les T_{MRCA} estimés sont caractérisés par des intervalles de confiance larges probablement en raison du petit nombre de séquences étudié mais les valeurs sont consistantes avec ce que l'on sait de la phylogénie des Cétacés et les T_{MRCA} observés au chapitre précédent.

IV. B. Une signature de sélection au sein de l'exon 2

Des évidences de sélection balancée ont été observés sur 11 sites (codons n^{os} 1, 12, 19, 22, 29, 32, 39, 48, 58, 61 and 62) dont 4 (soulignés) correspondent à des acides-aminés à l'intérieur du PBR. De façon remarquable, des signes de sélection négative

on été observés sur un site (codon n° 15) à l'extérieur du PBR. Alors que l'hydrophobicité de l'acide aminé ne se voit pas affecté par cette substitution, elle l'est dans 75% des cas de substitution à l'intérieur du PBR et 100% de ces substitutions se retrouvent chez des animaux de la population 1 telle que définie au chapitre précédent sur la base de l'ADN mitochondrial.

Article n°5 en préparation : **Macé M** & Crouau-Roy B « Diversity and patterns of selection on the DRB of the Striped dolphin, *Stenella cæruleoalba* »

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

Title

Diversity and patterns of selection on the DRB of the Striped dolphin, *Stenella caeruleoalba*

Authors and Affiliations

Matthias Macé^{1,2§}, Brigitte Crouau-Roy¹

¹UMR 5174 UPS/CNRS EDB “Evolution et Diversité biologique”, Bât 4R3b2, Université Paul Sabatier, 118 route de Narbonne, 31062 TOULOUSE cedex 9, France

²Centre de Physiopathologie de Toulouse Purpan, INSERM U563, CHU Purpan, F-31300 Toulouse, France

§Corresponding author

Telephone: +33 561 556259

FAX : +33 561 557327

e-mail: matthias.mace@orange.fr

Introduction

The Major Histocompatibility Complex (MHC) is central to the vertebrate immune system. It is a multigene family encoding key receptor molecules whose primary function is to recognise and bind foreign peptides for presentation to the immune system (Klein & Figueroa, 1986). MHC class II molecules are mainly expressed on specialised antigen-presenting cells and bind endogenously derived antigens that are subsequently presented to helper T-lymphocytes to initiate an immune response. From an evolutionary perspective, the pre-eminent feature of the MHC is the extreme diversity that is observed at expressed *loci*: in Mammals, MHC generally exhibits high heterozygosity values and in humans, its coding genes are among the most variable known *loci* (Gaudieri *et al.*, 2000). This high polymorphism and the haplotypes which contain them are thought to be achieved by balancing selection. Under this hypothesis, alleles persist much longer than predicted by neutrality, the polymorphism and haplotypes being maintained by frequency-dependent selection process, heterozygote advantage, etc. Most of the selective pressures acting on MHC may be the result of the critical role in the immune system linking it to fitness and behaviour. This could be driven by pathogens, pheromone driven negative-assortative mating or maternal-foetal interaction (Hedrick, 1998). As such, the evolutionary dynamics of the MHC has become of relevance in behavioural ecology, population biology and conservation biology, and a paradigm for adaptive evolution. Most of the MHC studies focused on human and species involved in medical research (model for human diseases and organ transplantation) because of the complex organisation of the MHC: over the half of them deal with human and mouse. However, studying non-classical models helps not only in a phylogenetic perspective but also in physiological issues (Bernatchez & Landry, 2003) and increases understanding in relationships between molecular evolution and ecological processes (Edwards *et al.*, 1995b).

Pathogen-driven selection has been hypothesised to explain the remarkable polymorphism in the MHC. This complex has long been considered less variable in marine Mammals, especially in Cetacean, as a result of a putative lower pathogen pressure in the marine environment (Trowsdale *et al.*, 1989). But this assertion was challenged by recent studies on Cetacean MHC using molecular approaches showing averages or slightly high genetic diversities (Flores Ramirez *et al.*, 2004; Hayashi *et al.*, 2006; Yang *et al.*, 2005). However, DRB class II complex characterised by PCR-SSCP displays, in Beluga (Murray & White, 1998) and in some Mysticete (Baleen whales) species (Baker *et al.*, 2006), a diversity lower than for terrestrial mammal. It seems therefore that no general conclusion on MHC diversity

50 in Marine environment could be drawn because at least the technique used but also the
51 biological model might influence the level of polymorphism found in Cetacean.

52

53 As a contribution, the present study gives new insights on the variability in MHC class II
54 DRB in a marine species. The MHC class II DRB is a complex region of gene duplicates
55 whose number is unknown in any Cetacean species. Although one study in the Beluga
56 showed two *loci* but relying only on sequence phylogeny rather than cytogenetics (**Murray &**
57 **White, 1998**), no precise assessment was made using *e.g.* molecular probes. We therefore
58 aimed at characterising DRB in one widely distributed Cetacean species, the Striped dolphin
59 (*Stenella coeruleoalba*), a common Cetacean species that inhabits warm and temperate waters
60 all around the globe. This dolphin was subjected to a Dolphin Morbillivirus (DMV,
61 *Paramyxoviridae*) epizootics in 1991-92 that led to huge mortality (**Domingo et al., 1990;**
62 **Van-Bressem et al., 2001**). It is not known if immunogenetic predisposition and/or
63 contaminants (**Kannan et al., 1993b**) played a role in the spread of the infection given that
64 only Striped dolphin from the Mediterranean suffered such a great die-off (around thousands
65 individuals over a total number of several tens of thousands).

66

67 As some works suggests a link between DRB and Measles virus (*Paramyxoviridae*) antigen
68 presentation (**Ovsyannikova et al., 2004**), we decided to characterise by sequencing this
69 region's exon 2 in this species. We first sought to evaluate the diversity of the Peptide
70 Binding Region (exon 2) of DRB *loci* by sequencing a Mediterranean sample of Striped
71 dolphin. By studying allele phylogeny, we determined allele groups whose repartition we
72 tried to link to the existence of two sympatric populations within the Mediterranean. We
73 finally examined if any departure to the general MHC pattern of balancing selection could led
74 to suspect pathogen-driven selection such as Morbillivirus. In this study, we included new
75 sequences from 3 other Cetacean species, the Bottlenosed dolphin *Tursiops truncatus*, the Fin
76 whale *Balaenoptera physalus* and the Gray whale *Eschrichtius robustus*, and we added 34
77 published sequences from 8 other Cetacean species.

78

79 **Material and Methods**

80 A total of 17 Striped dolphin were sampled for DRB sequencing as well as one Bottlenosed
81 dolphin *Tursiops truncatus*, one Fin whale *Balaenoptera physalus*, and two Gray whales
82 *Eschrichtius robustus*. Samples were obtained from stranded animals from Mediterranean
83 coasts as well as from a loan of the University of Alaska Museum of the North (2 Gray whale
84 samples). Among Striped dolphin samples, 8 individuals showed group 1 haplotype, 7
85 showed group 2 haplotype as defined by (**Macé et al., submitted to PLoS ONE**). Samples
86 were stored at -20°C. Genomic DNA was then extracted using Phenol-chloroform protocol
87 (**Sambrook & Russell, 2001**).

88

89 *Polymerase Chain Reaction (PCR) and Denaturing Gradient Gel Electrophoresis (DGGE)*

90 Because of the highly conserved class II molecules outside the PBR, oligonucleotide primer
91 sequences originally designed for Primates were used: 5'MDRB 5'-
92 GCCTCGAGTGTCCCCCAGCACGTTTC-3' and slightly modified MDRB1+GC 5'-
93 CGCCCGCCGCGCCCCGCGCCCGGCCCGCCCGCCCCGCCCCGAGCTTTCACCT
94 CGCCGCTG-3' (**Knapp et al., 1997**). They amplify a 320-base pair fragment of the MHC
95 class II DRB exon 2 including the peptid binding region (PBR). Polymerase Chain Reaction
96 (PCR) was as following: ca. 50 ng of genomic DNA in 50 µl of the following mix: 1.5 mM
97 MgCl₂, 2.5 mM of each dNTP, 25 pmol of each primer and 2-3 unit of *Taq* polymerase.
98 Cycling conditions were: initiation 95°Cx3 min., followed by 30 times 94°C /90 s.
99 denaturation, 49°C/90 s. annealing, 74°C /90 s. extension, followed by 72°C/10 min. final

100 extension. PCR products were run on a 1,5% agarose gel to confirm amplification.
101 To separate DRB alleles differentiated by one or more substitutions, 30 µl of the PCR
102 products were mixed with 15 µl of loading buffer and run on a 30-65% parallel denaturing
103 gradient gel according to (Myers *et al.*, 1988; Sheffield *et al.*, 1989). The 12,6% acrylamide
104 gel was electrophoresed in a BIORAD D-GENE (Biorad, Richmond, CA) apparatus in
105 1xTAE buffer at 300 Volts and 60°C temperature. Time of migration for optimal band
106 separation was evaluated to 5 hours. For each run, the same 2 individuals were provided as
107 controls for normalised genotyping. Individual bands on the gel were visualised using SYBR-
108 Green stain (Invitrogen-Molecular Probes, Eugene, OR) and UV transillumination.

109

110 *Reamplification and sequencing*

111 Each band was excised by removing gel plugs using wide-bore pipet tips (Continental Lab
112 Products, San Diego, CA) and dipped into 50 µl distilled water for 3 freezing-thawing cycles
113 for DNA elution. To generate PCR product for sequencing, 5 µl of the eluate was reamplified
114 using the hemi-nested primer pair 5'MDRB and 3'DRB1M13 (5'-
115 TGTAACGACGGCCAGTCGCTGCACTGTGAAGCTCTC-3'). PCR was performed as
116 follows: 2.5 mM MgCl₂, 2.5 mM of each dNTP, 12.5 pmol by tube of each primer and 2-3
117 unit of Taq polymerase in a final volume of 50 µl. Cycling was: 30 cycles of 94°C /20 s.
118 denaturation, 65°C/30 s. annealing, 72°C /36 s. extension, followed by 72°C/10 min. final
119 extension. After confirmation of amplification by gel electrophoresis, PCR products were
120 purified using Qiaquick PCR Purification Kit (Qiagen GmbH, Hilden, Germany, #28106).
121 Sequencing was made on a CEQ™ 2000XL Sequencer (Beckman Coulter, Fullerton, CA)
122 with the following conditions: 30 ng of PCR product, 4 µl Dye Terminator Cycle Sequencing
123 kit (Beckman Coulter, #608120), 1 µM primer in a final volume of 20 µl.

124

125 *Sequence analysis*

126 DNA sequences in both directions were edited using GENALYS software (Masazumi
127 Takahashi, <http://software.cng.fr/docs/genalys.html>), assembled to obtain a consensus
128 sequence for each sample and aligned using CLUSTALX (Thompson *et al.*, 1997). 44
129 previously published sequences of 10 other Cetacean DRB exon 2 were incorporated into this
130 analysis (GenBank #DQ354667-DQ354688 and AF012930-AF012941) plus 3 species used as
131 outgroup. Nucleotide - p - and haplotype - h - diversities were estimated using DNASP (Rozas
132 *et al.*, 2003). Mutation model was selected for best fit to control region data using hierarchical
133 likelihood ratio test and Akaike Information Criterion (AIC) as implemented in the Modeltest
134 software (Posada & Crandall, 1998). As recombination and gene conversion may disturb
135 selection detection methods and Time to the Most Recent Common Ancestor (T_{MRCA})
136 estimates, searching for recombination breakpoints was performed using DNASP (Rozas *et al.*,
137 2003) by computing the recombination parameter $R = 4Nr$ (Hudson, 1987),
138 (r =recombination rate and N =effective population size). A maximum likelihood approach was
139 also used using HYPHY package (Kosakovsky Pond *et al.*, 2005). Gene conversion
140 assessment was made using the algorithm of (Betran *et al.*, 1997) implemented in DNASP.
141 Third codon position relative substitution model parameters were assessed using BEAST
142 (Bayesian Evolutionary Analysis Sampling Tree; (Drummond *et al.*, 2002)) along with the
143 T_{MRCA} of selected lineages (all *Stenella cæruleoalba* lineages, Mediterranean lineage, whole
144 Cetacea dataset). BEAST uses MCMC to average over tree space, so that each tree is
145 weighted proportional to its posterior probability. Parameters allowed to vary are: tree depth,
146 branch lengths (tree topology), mutation rate and population parameters.
147 Phylogenetic inference was performed using a Bayesian Monte Carlo Markov Chain (MCMC)
148 approach (MrBayes; (Ronquist & Huelsenbeck, 2003)).
149 A maximum likelihood method was used to detect and quantify positive and negative

150 selection. The fixed-effects likelihood (FEL) approach ((**Kosakovsky Pond et al., 2005**)) was
151 employed to test for selective pressure at a given site. Maximum likelihood estimates of
152 branch lengths and nucleotide substitution rate parameters were derived from the entire
153 alignment. A full-codon model, using a modified MG94 ((**Muse & Gaut, 1994**)) rate matrix
154 with site-specific instantaneous synonymous (alphas) and nonsynonymous (betas) rates was
155 then fitted independently to every codon position in the data, under two hypotheses: H_0 ,
156 neutral evolution (alphas equal betas); H_1 , non-neutral evolution (alphas and betas are free to
157 vary independently). When the hypothesis of neutrality was rejected at site s , it was called
158 positively selected if beta was estimated to be greater than alpha and negatively selected if
159 alpha was estimated to be greater than beta. The difference was considered significant when
160 the log Bayes factor ((**Kosakovsky Pond et al., 2005**)) was greater or equal than 5 following
161 the formula. The FEL method is implemented in the HYPHY package.

162

163 *MHC Stco-DRB alleles nomenclature*

164 Locus and allele names were given in accordance with the rules designated by the nonhuman
165 primate MHC nomenclature committee ((**Klein et al., 1990**)). New sequences have been
166 deposited in GenBank.

167

168 **Results**

169 *Allelic and nucleotide diversity in Striped dolphin DRB*

170 We tested 60 Striped dolphins for DRB amplification but only 20 amplified due to different
171 sample conservation states. Striped and Bottlenosed dolphin, Fin and Gray whales individuals
172 were successfully amplified with the primers initially designed for Primates and gave
173 satisfactory results regarding the size of amplified products (340 bp). The PCR products were
174 analysed by means of DGGE. Profiles were complex with 22 different bands and each
175 individual bearing 2 to 8 bands corresponding to alleles of DRB duplicates. These 22 bands
176 were subjected to sequence analysis for alleles/duplicates identification. Only 18 bands
177 reamplified successfully leading to the characterisation of 15 different alleles. These
178 sequences were labelled Stco-DRB*01 to *15 (GenBank accession # XX) following IMGT
179 instructions according to (**Klein et al., 1990**). The same kind of profiles was obtained for the
180 three other species. These striped dolphin sequences, and the 5, 6 and 4 sequences from the
181 Bottlenosed dolphin, Fin whale and Gray whales respectively are shown in figure 1 (GenBank
182 accession # XX).

183

184 Within Stco-DRB sequences, there are 33 polymorphic sites (31 are parsimonious
185 informative). Global nucleotide diversity of these sequences averages 0.0324 (± 0.0001) and
186 the mean number of substitutions is 8.998. Tajima's D is positive (1.967) and signs either
187 balancing selection or population decrease.

188 Recombination breakpoints were not found by neither the computation of $R = 4Nr$ (**Hudson,**
189 **1987**), (with r the recombination rate and N the effective population size), nor the maximum-
190 likelihood approach. As well, gene conversion seemed not to be present at least in the Striped
191 dolphin sequences. We could therefore assume that T_{MRCA} estimates will not be strongly
192 biased towards the past.

193 The mutation model parameters (κ) showed a significant difference between 3rd and the other
194 codon positions (see figure 2). This is consistent with selection signal.

195

196 *Allele phylogeny*

197 Given this different mutation rate observed on the third compared to other codon positions,
198 we assessed gene phylogeny by removing third codon position.

199 The DRB alleles tree is characterised by a paraphyletic structure.

200 Striped dolphin alleles clustered into 3 groups (Figure 3). Stco-DRB*01 to Stco-DRB*05 and
201 Stco-DRB*12 (hereafter referred as haplogroup 1) had a basal position, clustering with any
202 other lineage (B.P. < 0.6 for Stco-DRB*04 and 05). The main cluster then grouped the
203 following: Stco-DRB*06 to 11 and Stco-DRB*13-14. This group also included alleles from
204 the Bottlenosed dolphin (Tutr-DRB*01 to 03) and the Gray whale (Esro-DRB*01). However,
205 its Posterior Probability is weak (0.57) and it is itself subdivided in two groups: Stco-DRB*11
206 and Stco-DRB*14 (BP=0.66) versus all other (BP=0.51 but rising to 0.85 in the following
207 step of the subtree). The Stco-DRB*15 together with Tutr-DRB*04-05 and all of our Baph-
208 DRB alleles clustered with the majority of the published Mysticete sequences (BP=1.0).
209 Finally, 3 among 4 Gray whale sequences clustered together (BP=0.77). Noticeably,
210 Monodontid (Beluga and Narwhal) published sequences set apart from all other sequences but
211 the only Sperm whale sequence (BP=0.77).

212 Bayesian estimates of *Stenella* sequences T_{MRCA} gave 67 millions of years (Myrs; CI 95% 19-
213 115 Myrs) for the haplogroup 1 lineage, 27 Myrs (CI 95% 6-55 Myrs) for the main lineage (2
214 Myrs [0.15-7] for the minor lineage characterized by a BP=0.66) and 67 Myrs (CI 95% 19-
215 115 Myrs) for all Striped dolphin alleles, whereas whole Cetacean dataset TMRCA was about
216 80 Myrs (CI 95% 30-112 Myrs) for a prior allowing Cetacean-Artiodactyla-Carnivores
217 TMRCA to vary between 80 and 120 Myrs.

218

219 *Selection signal inference*

220 Considering the 15 Striped dolphin sequences data set, balancing selection was significantly
221 detected (log Bayes Factor ≥ 5) on 11 sites (codons # 1, 12, 19, 22, 29, 32, 39, 48, 58, 61 and
222 62 of the sequence aligned with previous Cetacea DRB sequences), 4 of which (underlined)
223 corresponding to PBR sites on HLA-DR homologous sequences; negative selection was
224 detected on one site (# 15) putatively out of binding pocket.

225 Whereas hydrophobicity remained unchanged at the codon #15, it changed at the sites located
226 in the presentation groove for 3/4 sites and these aminoacid replacement were shared by the
227 sequences of the haplogroup 1.

228 Considering alone a lineage including these sequences, no obvious signal of selection was
229 found. Interestingly, when considering all available Cetacea DRB sequences, the pattern was
230 similar to those of Striped dolphin considered as a whole (positive selection inferred on the
231 same sites as well as for negative selection). Figure 4 shows the three patterns observed.

232

233 *Allele repartition within Mediterranean populations*

234 Haplogroup 1 sequences seem overrepresented in individuals bearing group A mitochondrial
235 haplotype (85%) than group B (57%) but the difference is not significant ($p=0.23$) given the
236 paucity of the sample.

237

238 **Discussion**

239 *Striped dolphin DRB shows medium levels of diversity*

240 This study revealed 2 to 8 DRB bands per individual. It is therefore obvious that there are
241 several duplicated *loci*, at least four, since we used heterologous primers and are not sure to
242 have amplified all DRB duplicates. The phylogenic signal of these duplicates is weak since
243 the allele phylogeny only reveals 3 groups. This could be due to a recent divergence between
244 the sequences of two duplicate while the sequences of the other two would have diverged
245 earlier. This could explain the low BP of the main cluster (0.57) along with the higher values
246 of the two main branches inside this cluster (0.66 and 0.85). However, the very low value of
247 the TMRCA of the sequences whose BP is 0.66 would imply that the duplication event have
248 arisen after the Striped dolphin speciation.

249 However, the lack of these haplogroup in some individuals (1 over 7 in mitochondrial group
250 A individuals, 3 over 7 in group B individuals) could be explained by a variable number of
251 DRB copies and a difference in frequencies of this absence in Mediterranean groups.

252
253 We were able to characterise 15 DRB alleles over 20 individuals. This is much greater than
254 the 9 DRB alleles over 313 Belugas (*Delphinapterus leucas*; (Murray & White, 1998)),
255 comparable to DRB in *Megaptera novaeangiae* (7 alleles/5 individuals; (Baker et al., 2006))
256 but lesser than DQB diversity in this species (23/5) or in Baiji (*Lipotes vexillifer*; 43/18)
257 (Yang et al., 2005). This last authors agreed with (Trowsdale et al., 1989) arguing that
258 pathogen pressure could be greater in freshwaters than in marine environment. At the light of
259 our results, this argument cannot be considered alone as causal factor since the allele number
260 is also of same magnitude as in terrestrial mammals studied so far (e.g. 17 DRB in European
261 wolf (Seddon & Ellegren, 2002); 21 in Bighorn sheep (Gutierrez-Espeleta et al., 2001); 11
262 DRA alleles in Equidae (Brown et al., 2004)). The same conclusions can be drawn from
263 comparison with domestic mammals (e.g. Pig or Dog (Brunsberg et al., 1996; Seddon &
264 Ellegren, 2002)).with special attention to the greater number of alleles in domestic cat
265 compared to wild species (Yuhki & O'Brien, 1997), although they usually present more
266 variability than wild counterpart (e.g. Cattle vs Buffalo (Ellis & Ballingall, 1999; Sena et al.,
267 2003)). This could be the result of severe bottlenecks undergone by wild species (e.g.
268 Eurasian beaver, (Babik et al., 2005) or South African Antelope, (Van der Walt et al.,
269 2001)) vs Holocene demographic expansion of domestic Mammals. It is however less than the
270 numbers found in human (IMGT/HLA database, <http://www.ebi.ac.uk/imgt/hla>) and non-
271 human Primates ((Lukas et al., 2004; Suárez et al., 2006)). See table 1 for details.

272 Maybe the moderately high diversity found in the Striped dolphin DRB may result of the
273 PCR/DGGE sensitivity since many other studies relies on non-exhaustive cloning-sequencing.

274

275 *Trans-specific polymorphism and non-neutral evolution*

276 The Striped dolphin's DRB alleles phylogeny shows trans-specific clustering as usually
277 observed in MHC studies e.g. in Fishes (Figueroa et al., 2000) or Mammals (Go et al., 2002).
278 Indeed the main part of the sequences clustered with other Odontocete species belonging to
279 other genus, another (Stco-DRB*15) clustered with Mysticete (the other order of extant
280 Cetacea) species and the last remained basal in the phylogeny (haplogroup 1).

281 Inferring allele phylogeny by removing the third codon position allows minimizing the
282 potential misleading relationships arising from repeated positive selection on functional
283 codons with exons creating convergence (Takahata & Nei, 1990). In our Bayesian analysis
284 of phylogeny, it took less markov chains to reach convergence (average standard deviation of
285 split frequencies ≤ 0.1) for data without the third codon position (around 83.000) than without
286 removing this position (around 170.000). This could be the result from less phylogenetic noise
287 arising from selection.

288 Trans-species polymorphism is considered as a result of the non-neutral retention of alleles.
289 Independently similar sequence motifs can repeatedly arise as a result of positive selection
290 acting on functional codons within exons, thereby leading to convergence between non-
291 related alleles (Kriener et al., 2000). In the case of the Striped dolphin, some tracks can be
292 explored in order to find pathogens that could drive selection.

293

294

295 *Different patterns of selection and perspectives in « populations pathology »*

296 A higher level of inbreeding was inferred by the Internal Relatedness factor (Amos et al.,
297 2001) in individuals dead during the Morbilliviral outbreak (Valsecchi et al., 2004b). An
298 overall lower variability compared to other populations was also outlined by our previous

299 work (**Macé *et al.*, under revision**). However, even in the case of a low autosomal variability,
300 the MHC can go on keeping a great variability as it has been observed in the San Nicolas
301 Island fox (*Urocyon littoralis dickeyi*) (**Aguilar *et al.*, 2004**)?

302 Balancing selection on our data is mainly focused on PBR codons: this a likely explanation of
303 the moderately high diversity observed and Stco-DRB *loci*. But purifying selection at one
304 codon is also observed in all but one Striped dolphin DRB lineage. Purifying selection has
305 only been yet reported in immune genes in *Drosophila melanogaster* Parasite Recognition
306 Protein (**Jiggins & Hurst, 2003**) but to our knowledge not in any MHC genes. Haplogroups
307 with negative selection is more frequently observed in the Mediterranean group B (atlantic-
308 related as considered in (**Macé *et al.*, under revision**)) individuals. This could result of
309 selective sweep from a pathogenic agent. Although not significant (it would require 30
310 individuals from each group), the less extend of this signal in the Mediterranean group A
311 could be the signal of a “naïve” population with alleles that would be deleterious regarding
312 the pathogen responsible for selective sweep or at least of a very recent contact with the
313 pathogen that drove this selection. This phenomena could be likely in a population
314 characterised by a lower diversity with probably lesser DRB repertoire at the individual level.
315 The Dolphin Morbillivirus could represent a good candidate for a pathogen driving DRB
316 evolution since the binding of Measles Virus (MV) by the HLA-DRB1 molecule, with the
317 elicitation of a cell-mediated response, was shown in human blood.

318 Moreover, the fact that only the Mediterranean Striped dolphins developed an epizootics
319 whereas serologies outlined the Morbillivirus’ passing through in the Atlantic supports the
320 idea of a particular susceptibility if the Mediterranean Striped dolphin. The temporally-stable
321 regional differences in Morbillivirus prevalence in North seas Harbour seals (**Härkönen *et***
322 ***al.*, 2006**) also argue in favour of genetic susceptibility.

323 Finally, *in silico* 3D modelisation of the exon 2 in this lineage showed striking structural
324 differences in the PBR groove compared to the other lineages (data not shown) and this is in
325 agreement with the differences in replaced aminoacids hydrophobicity observed between
326 haplogroup 1 and others.

327
328 Emerging infectious diseases are often the results of host shift (**Brooks & Ferrao, 2005**) but
329 Dolphin Morbillivirus (DMV) seems to be an old virus (**van de Bildt *et al.*, 2005**). So either
330 Mediterranean Striped dolphin population has remained relatively isolated from the virus, or
331 previous viral episodes may not have left signatures in the Striped dolphin genome. In this last
332 case, the virulence of the Morbillivirus might be lesser than previously or contact too scarce
333 to reach a form of Nash equilibrium (**Sasaki & Iwasa, 1991**).

334 Much more than for mitochondrial DNA or microsatellite studies, MHC genotyping must rely
335 on good quality DNA. Given the forensic origin of our samples, we were only able to yield
336 quality genotyping for 20 individuals. More material will be needed in the future, even from
337 fresh carcasses or from live animals biopsy-sampling. In addition of at least 30 samples from
338 each Mediterranean population, the same number of samples should be obtained from
339 individuals that were diagnosed and died from Morbilliviral infection in order to reliably test
340 for disease association.

341

342 **Figures and tables.**

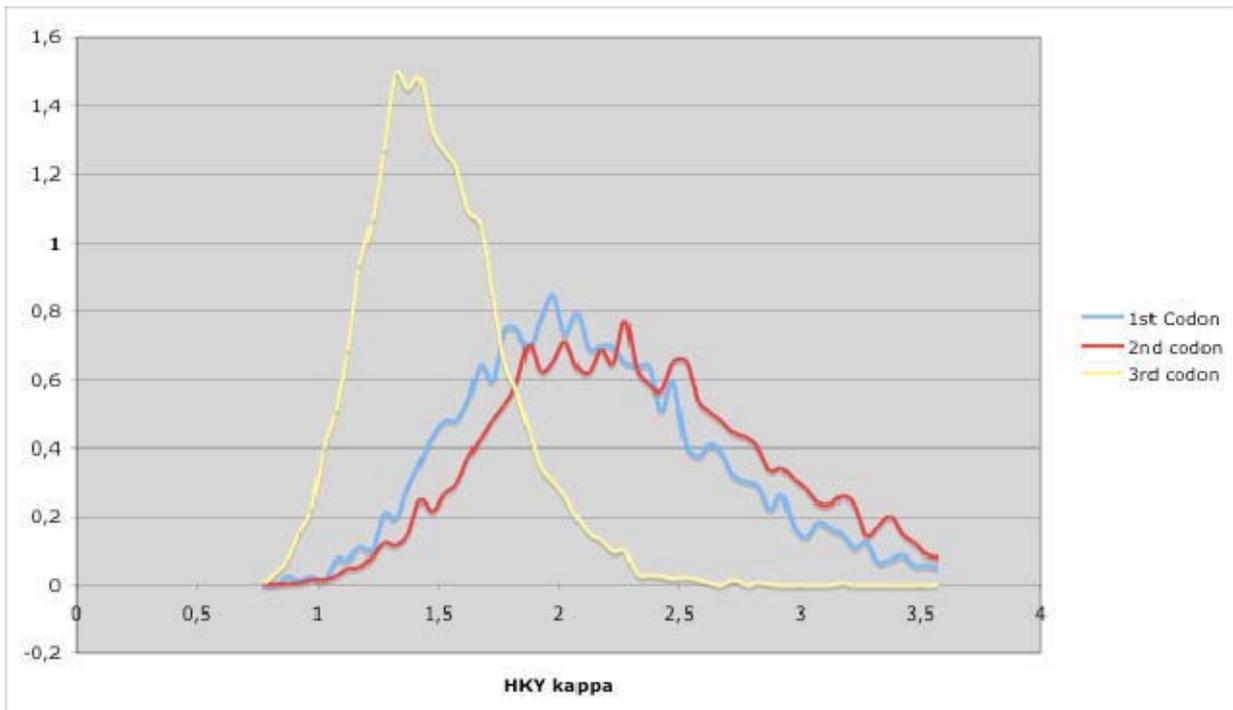


Fig. 2 Posterior distribution of HKY model's kappa from the three codons positions. Probability density is on the y-axis, kappa is on the x-axis.

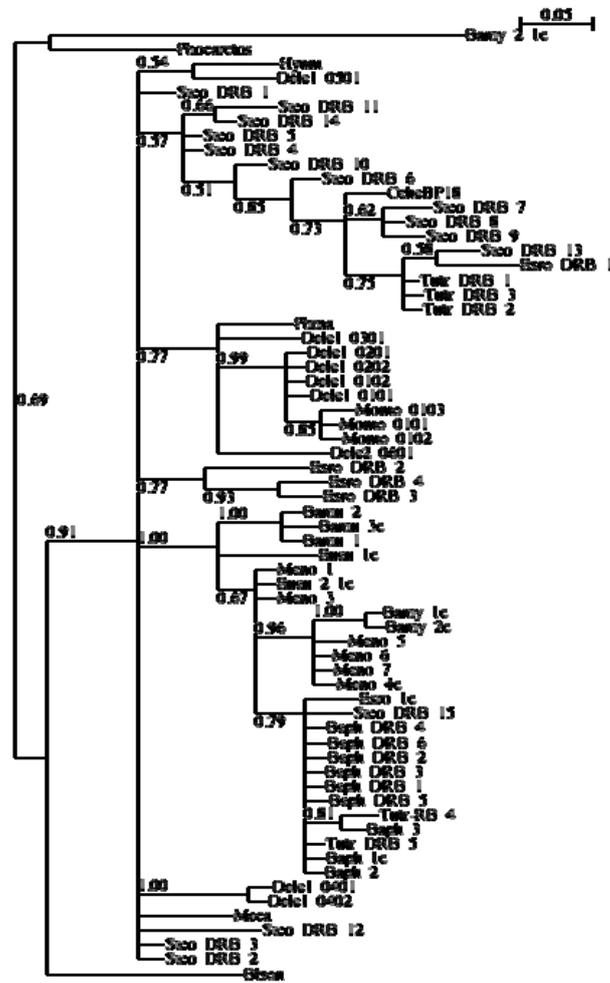


Fig. 3 Phylogeny of known alleles from Cetacea species' DRB (Esro: *Eschrichtius robustus*; Euau: *Eubalaena (glacialis) australis*; Bamy: *Balaena mysticetus*; **Baph:** *Balaenoptera physalus*; Meno: *Megaptera novaeangliae*; Dele: *Delphinapterus leucas*; Momo: *Monodon monoceros*; **Stco:** *Stenella coeruleoalba*; **Tutr:** *Tursiops truncatus* – bold indicating this study's new sequences)

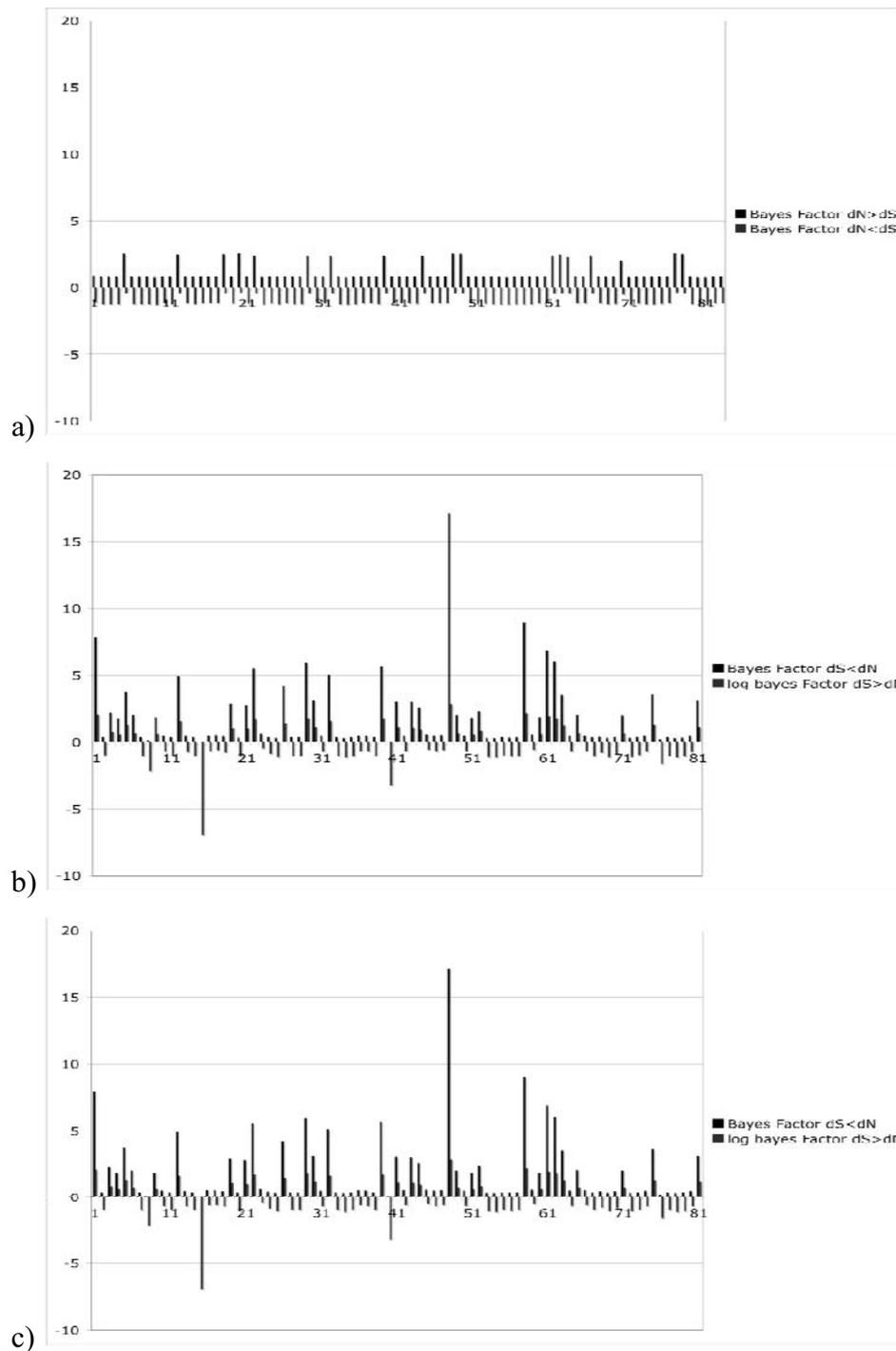


Fig. 4 Diagram showing sites under positive selection (Factor > 5) and negative selection (Factor < -5) following three sequences grouping : a) *Stenella's* divergent clade sequences, b) *Stenella coeruleoalba* sequences, c) whole cetacea dataset.

Species	Taxon	Haplotype diversity (# alleles/ individuals)	Nr (%) of variable nucleotide positions	References
<i>Microcebus murinus</i>	Primata/Prosimia	0.061 (14/228)	71 (41.5)	(Schad <i>et al.</i> , 2004) (Schad <i>et al.</i> , 2005)
<i>Microcebus berthae</i>	Primata/Prosimia	0.214 (9/42)	46 (26.9)	Sommer <i>et al.</i> , unpublished data
<i>Apodemus sylvaticus</i>	Rodentia	0.319 (38/119)	71 (32.7)	(Musolf <i>et al.</i> , 2004)
<i>Apodemus flavicollis</i>	Rodentia	0.184 (27/146)	49 (22.6)	(Froeschke & Sommer, 2005).
<i>Leopoldamys sabanus</i>	Rodentia	0.571 (28*/49)	85 (49.7)	Lenz <i>et al.</i> , unpublished data
<i>Gerbillurus paeba</i>	Rodentia	0.85 (34*/40)	68 (39.8)	(Harf & Sommer, 2005)
<i>Rhabdomys pumilio</i>	Rodentia	0.344 (20/58)	43 (25.1)	(Froeschke & Sommer, 2005)
<i>Rattus rattus</i>	Rodentia	0.224 (13/58)	72 (40.7)	Sommer, unpublished data
<i>Hypogeomys antimena</i>	Rodentia	0.021 (5/229)	37 (17.1)	(Sommer, 2003)
<i>Equus przewalskii</i>	perissodactyla	0.428 (6*/14)	52 (20.8)	(Hedrick <i>et al.</i> , 1999) ⁰
<i>Oryx leucoryx</i>	artiodactyla	0.052 (3/57)	35 (14.8)	(Hedrick <i>et al.</i> , 2000)
<i>Damaliscus pygargus</i>	artiodactyla	0.133 (6/45)	21 (8.4)	(Van der Walt <i>et al.</i> , 2001)
<i>Delphinapterus leucas</i>	Cetacea/Odontoceti	0.029 (9/313)		(Murray & White, 1998)
<i>Monodon monoceros</i>	Cetacea/Odontoceti	0.272 (3/11)		(Murray & White, 1998)
<i>Megaptera novaeangliae</i>	Cetacea/Mysticeti			
<i>Castor fiber</i>	Rodentia	(10/76)		(Babik <i>et al.</i> , 2005)
<i>Canis sp.</i>	Carnivora	ca. 0.075		(Seddon & Ellegren, 2002)
<i>Gorilla gorilla</i>	Primata/Catarrhini			
<i>Aotus nancymaae</i>	Primata/Platyrrhini	0.94(67/71)		(Suárez <i>et al.</i> , 2006)
<i>Aotus vociferans</i>	Primata/Platyrrhini	1.3(13/10)		(Suárez <i>et al.</i> , 2006)
<i>Aotus nigriceps</i>	Primata/Platyrrhini	2.0(30/15)		(Suárez <i>et al.</i> , 2006)
<i>Homo sapiens</i>	Primata/Catarrhini	545 alleles over 9 loci		http://ebi.ac.uk/imgt
<i>Bos taurus</i>	artiodactyla	63 on DRB3 locus		(Ellis & Ballingall, 1999)
<i>Bubalus bubalis</i>	artiodactyla	8 on DRB3 locus		(Sena <i>et al.</i> , 2003)

Table 1. Variability at DRB loci in selected species of Mammals.

References

- Aguilar, A., Roemer, G., Debenham, S., Binns, M., Garcelon, D., and Wayne, R. K.: High MHC diversity maintained by balancing selection in an otherwise genetically monomorphic mammal. *Proceedings of the National Academy of Sciences of the United States of America* 101: 3490-4, 2004
- Amos, W., Wilmer, J. W., Fullard, K., Burg, T. M., Croxall, J. P., Bloch, D., and Coulson, T.: The influence of parental relatedness on reproductive success. *Proceedings of the Royal Society of London series B Biological Sciences* 268: 2021-7, 2001
- Babik, W., Durka, W., and Radwan, J.: Sequence diversity of the MHC DRB gene in the Eurasian beaver (*Castor fiber*). *Molecular ecology* 14: 4249-57, 2005
- Baker, C. S., Vant, M. D., Dalebout, M. L., Lento, G. M., O'Brien, S. J., and Yuhki, N.: Diversity and duplication of DQB and DRB-like genes of the MHC in baleen whales (suborder: Mysticeti). *Immunogenetics* 58: 283-296, 2006
- Bernatchez, L. and Landry, C.: MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? *J Evol Biol* 16: 363-77., 2003
- Betran, E., Rozas, J., Navarro, A., and Barbadilla, A.: The estimation of the number and the length distribution of gene conversion tracts from population DNA sequence data. *Genetics* 146: 89-99, 1997
- Brooks, D. R. and Ferrao, A. L.: The historical biogeography of co-evolution: emerging infectious diseases are evolutionary accidents waiting to happen. *Journal Of Biogeography* 32: 1291-1299, 2005
- Brown, J. J., Thomson, W., Clegg, P., Eyre, S., Kennedy, L. J., Matthews, J., Carter, S., and Ollier, W. E.: Polymorphisms of the equine major histocompatibility complex class II DRA locus. *Tissue Antigens* 64: 173-9, 2004
- Brunsborg, U., Edfors-Lilja, I., Andersson, L., and Gustafsson, K.: Structure and organization of pig MHC class II DRB genes: evidence for genetic exchange between *loci*. *Immunogenetics* 44: 1-8, 1996
- Domingo, M., Ferrer, L., Pumarola, M., Marco, A., Plana, J., Kennedy, S., McAliskey, M., and Rima, B. K.: Morbillivirus in dolphins. *Nature* 348: 21, 1990
- Drummond, A. J., Nicholls, G. K., Rodrigo, A. G., and Solomon, W.: Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. *Genetics* 161: 1307-20, 2002
- Edwards, S. V., Wakeland, E. K., and Potts, W. K.: Contrasting histories of avian and mammalian Mhc genes revealed by class II B sequences from songbirds. *Proc Natl Acad Sci U S A* 92: 12200-4, 1995
- Ellis, S. A. and Ballingall, K. T.: Cattle MHC: evolution in action? *Immunological reviews* 167: 159-68, 1999
- Figuerola, F., Mayer, W. E., Sultmann, H., O'HUigin, C., Tichy, H., Satta, Y., Takezaki, N., Takahata, N., and Klein, J.: Mhc class II B gene evolution in East African cichlid fishes. *Immunogenetics* 51: 556-75, 2000
- Flores Ramirez, S., Miller, R. D., and Urban Ramirez, J.: Major histocompatibility complex I polymorphism in a cetacean: The gray whale (*Eschrichtius robustus*). *Marine Mammal Science* 20: 262-273, 2004
- Froeschke, G. and Sommer, S.: MHC class II DRB variability and parasite load in the striped mouse (*Rhabdomys pumilio*) in the Southern Kalahari. *Molecular Biology and Evolution* 22: 1254-9, 2005

- Gaudieri, S., Dawkins, R. L., Habara, K., Kulski, J. K., and Gojobori, T.: SNP profile within the human major histocompatibility complex reveals an extreme and interrupted level of nucleotide diversity. *Genome research* 10: 1579-86, 2000
- Go, Y., Satta, Y., Kawamoto, Y., Rakotoarisoa, G., Randrianjafy, A., Koyama, N., and Hirai, H.: Mhc-DRB genes evolution in lemurs. *Immunogenetics* 54: 403-17, 2002
- Gutierrez-Espeleta, G. A., Hedrick, P. W., Kalinowski, S. T., Garrigan, D., and Boyce, W. M.: Is the decline of desert bighorn sheep from infectious disease the result of low MHC variation? *Heredity* 86: 439-50, 2001
- Harf, R. and Sommer, S.: Association between major histocompatibility complex class II DRB alleles and parasite load in the hairy-footed gerbil, *Gerbillurus paeba*, in the southern Kalahari. *Molecular Ecology* 14: 85-91, 2005
- Härkönen, T., Dietz, R., Reijnders, P., Teilmann, J., Harding, K., Hall, A., Brasseur, S., Siebert, U., Goodman, S. J., Jepson, P. D., Dau Rasmussen, T., and Thompson, P.: The 1988 and 2002 phocine distemper virus epidemics in European harbour seals. *Diseases of Aquatic Organisms* 68: 115-30, 2006
- Hayashi, K., Yoshida, H., Nishida, S., Goto, M., Pastene, L. A., Kanda, N., Baba, Y., and Koike, H.: Genetic Variation of the MHC DQB Locus in the Finless Porpoise (*Neophocaena phocaenoides*). *Zoological science* 23: 147-53, 2006
- Hedrick, P.: Balancing selection and MHC. *Genetica* 104: 207-214, 1998
- Hedrick, P. W., Parker, K. M., Gutierrez Espeleta, G. A., Rattink, A., and Lievers, K.: Major histocompatibility complex variation in the Arabian oryx. *Evolution* 54: 2145-2151, 2000
- Hedrick, P. W., Parker, K. M., Miller, E. L., and Miller, P. S.: Major Histocompatibility Complex Variation in the Endangered Przewalski's Horse. *Genetics* 152: 1701-1710, 1999
- Hudson, R. R.: Estimating the recombination parameter of a finite population model without selection. *Genetic Research* 50: 245-50, 1987
- Jiggins, F. M. and Hurst, G. D.: The evolution of parasite recognition genes in the innate immune system: purifying selection on *Drosophila melanogaster* peptidoglycan recognition proteins. *J Mol Evol* 57: 598-605, 2003
- Kannan, K., Tanabe, S., Borrell, A., Aguilar, A., Focardi, S., and Tatsukawa, R.: Isomer-specific analysis and toxic evaluation of polychlorinated biphenyls in striped dolphins affected by an epizootic in the western Mediterranean sea. *Archives of environmental contamination and toxicology* 25: 227-33, 1993
- Klein, J., Bontrop, R. E., Dawkins, R. L., Erlich, H. A., Gyllensten, U. B., Heise, E. R., Jones, P. P., Parham, P., Wakeland, E. K., and Watkins, D. I.: Nomenclature for the major histocompatibility complexes of different species: a proposal. *Immunogenetics* 31: 217-9, 1990
- Klein, J. and Figueroa, F.: Evolution of the major histocompatibility complex. *Critical reviews in immunology* 6: 295-386, 1986
- Knapp, L. A., Cadavid, L. F., Eberle, M. E., Knechtle, S. J., Bontrop, R. E., and Watkins, D. I.: Identification of new mamu-DRB alleles using DGGE and direct sequencing. *Immunogenetics* 45: 171-9, 1997
- Kosakovsky Pond, S. L., Frost, S. D., and Muse, S. V.: HyPhy: hypothesis testing using phylogenies. *Bioinformatics Oxford, England* 21: 676-9, 2005
- Kriener, K., O'HUigin, C., Tichy, H., and Klein, J.: Convergent evolution of major histocompatibility complex molecules in humans and New World monkeys. *Immunogenetics* 51: 169-78, 2000

- Lukas, D., Bradley, B. J., Nsubuga, A. M., Doran-Sheehy, D., Robbins, M. M., and Vigilant, L.: Major histocompatibility complex and microsatellite variation in two populations of wild gorillas. *Molecular ecology* 13: 3389-402, 2004
- Macé, M., Bourret, V., and Crouau-Roy, B.: Polyphyly in the Mediterranean Striped dolphin (*Stenella coeruleoalba*) and the Messinian Salinity Crisis. submitted to PLoS ONE
- Macé, M., Bourret, V., and Crouau-Roy, B.: Polyphyly in the Mediterranean Striped dolphin (*Stenella coeruleoalba*) and the Messinian Salinity Crisis. *PLoS ONE*, under revision
- Murray, B. W. and White, B. N.: Sequence variation at the major histocompatibility complex DRB loci in beluga (*Delphinapterus leucas*) and narwhal (*Monodon monoceros*). *Immunogenetics* 48: 242-52, 1998
- Muse, S. V. and Gaut, B. S.: A likelihood approach for comparing synonymous and nonsynonymous nucleotide substitution rates, with application to the chloroplast genome. *Molecular Biology and Evolution* 11: 715-24, 1994
- Musolf, K., Meyer-Lucht, Y., and Sommer, S.: Evolution of MHC-DRB class II polymorphism in the genus *Apodemus* and a comparison of DRB sequences within the family Muridae (Mammalia: Rodentia). *Immunogenetics* 56: 420-6, 2004
- Myers, R. M., Sheffield, V. C., and Cox, D. R.: Detection of single base changes in DNA: ribonuclease cleavage and denaturing gradient gel electrophoresis. In K. Davies (ed.): *Genome Analysis: A practical approach*, IRL Press, Oxford, 1988
- Ovsyannikova, I. G., Johnson, K. L., Muddiman, D. C., Vierkant, R. A., and Poland, G. A.: Identification and characterization of novel, naturally processed measles virus class II HLA-DRB1 peptides. *Journal of virology* 78: 42-51, 2004
- Posada, D. and Crandall, K. A.: MODELTEST: testing the model of DNA substitution. *Bioinformatics Oxford, England* 14: 817-8, 1998
- Ronquist, F. and Huelsenbeck, J. P.: MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics Oxford, England* 19: 1572-4, 2003
- Rozas, J., Sanchez-DelBarrio, J. C., Messeguer, X., and Rozas, R.: DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics Oxford, England* 19: 2496-7, 2003
- Sambrook, J. and Russell, D. W.: *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 2001
- Sasaki, A. and Iwasa, Y.: Optimal growth schedule of pathogens within a host: switching between lytic and latent cycles. *Theoretical population biology* 39: 201-39, 1991
- Schad, J., Ganzhorn, J. U., and Sommer, S.: Parasite burden and constitution of major histocompatibility complex in the Malagasy mouse lemur, *Microcebus murinus*. *Evolution Int J Org Evolution* 59: 439-50, 2005
- Schad, J., Sommer, S., and Ganzhorn, J. U.: MHC variability of a small lemur in the littoral forest fragments of southeastern Madagascar. *Conservation Genetics* 5: 299-309, 2004
- Seddon, J. M. and Ellegren, H.: MHC class II genes in European wolves: a comparison with dogs. *Immunogenetics* 54: 490-500, 2002
- Sena, L., Schneider, M. P., Brenig, B., Honeycutt, R. L., Womack, J. E., and Skow, L. C.: Polymorphisms in MHC-DRA and -DRB alleles of water buffalo (*Bubalus bubalis*) reveal different features from cattle DR alleles. *Animal genetics* 34: 1-10, 2003
- Sheffield, V. C., Cox, D. R., Lerman, L. S., and Myers, R. M.: Attachment of a 40-base-pair G + C-rich sequence (GC-clamp) to genomic DNA fragments by the polymerase chain reaction results in improved detection of single-base changes. *Proc Natl Acad Sci U S A* 86: 232-6, 1989
- Sommer, S.: Effects of habitat fragmentation and changes of dispersal behaviour after a recent population decline on the genetic variability of noncoding and coding DNA of a monogamous Malagasy rodent. *Molecular Ecology* 12: 2845-51, 2003

- Stern, L. J. and Wiley, D. C.: Antigenic peptide binding by class I and class II histocompatibility proteins. *Structure London, England* 2: 245-51, 1994
- Suárez, C. F., Patarroyo, M. E., Trujillo, E., Estupiñán, M., Baquero, J. E., Parra, C., and Rodríguez, R.: Owl monkey MHC-DRB exon 2 reveals high similarity with several HLA-DRB lineages. *Immunogenetics* 58: 542-58, 2006
- Takahata, N. and Nei, M.: Allelic genealogy under overdominant and frequency-dependent selection and polymorphism of major histocompatibility complex *loci*. *Genetics* 124: 967-78, 1990
- Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G.: The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic acids research* 25: 4876-82, 1997
- Trowsdale, J., Groves, V., and Arnason, A.: Limited MHC polymorphism in whales. *Immunogenetics* 29: 19-24, 1989
- Valsecchi, E., Amos, W., Raga, J. A., Podesta, M., and Sherwin, W.: The effects of inbreeding on mortality during a morbillivirus outbreak in the Mediterranean striped dolphin (*Stenella coeruleoalba*). *Animal Conservation* 7: 139-146, 2004
- van de Bildt, M. W., Kuiken, T., and Osterhaus, A. D.: Cetacean morbilliviruses are phylogenetically divergent. *Arch Virol* 150: 577-83, 2005
- Van der Walt, J. M., Nel, L. H., and Hoelzel, A. R.: Characterization of major histocompatibility complex DRB diversity in the endemic South African antelope *Damaliscus pygargus*: a comparison in two subspecies with different demographic histories. *Molecular Ecology* 10: 1679-1688, 2001
- Van-Bresseem, M., Waerebeek, K. V., Jepson, P. D., Raga, J. A., Duignan, P. J., Nielsen, O., Di-Beneditto, A. P., Siciliano, S., Ramos, R., Kant, W., Peddemors, V., Kinoshita, R., Ross, P. S., Lopez-Fernandez, A., Evans, K., Crespo, E., and Barrett, T.: An insight into the epidemiology of dolphin morbillivirus worldwide. *Veterinary microbiology* 81: 287-304, 2001
- Yang, G., Yan, J., Zhou, K., and Wei, F.: Sequence variation and gene duplication at MHC DQB *loci* of baiji (*Lipotes vexillifer*), a Chinese river dolphin. *Journal of heredity, The* 96: 310-7, 2005
- Yuhki, N. and O'Brien, S. J.: Nature and origin of polymorphism in feline MHC class II DRA and DRB genes. *Journal of Immunology* 158: 2822-33, 1997

V. Perspectives

Les résultats présentés sont parcellaires et soulignent la nécessité de l'élargissement de l'échantillonnage au sein de l'espèce et dans d'autres espèces de la triade *Stenella-Delphinus-Tursiops*.

Il sera en outre nécessaire d'obtenir beaucoup plus de prélèvements dont le statut virologique (Dolfin Morbillivirus) est certain afin d'obtenir des niveaux de significativité dans les comparaisons des profils du DRB chez les deux groupes (30-50 cas + 30-50 contrôles ?).

Il serait enfin intéressant de comparer les TMRCA des différents Morbillivirus et de les rapporter aux TMRCA des différents lignages de DRB. Par ailleurs, l'étude de la structure spatiale de la molécule du DRB de chacun des lignages pourrait ouvrir la voie à des études fonctionnelles.

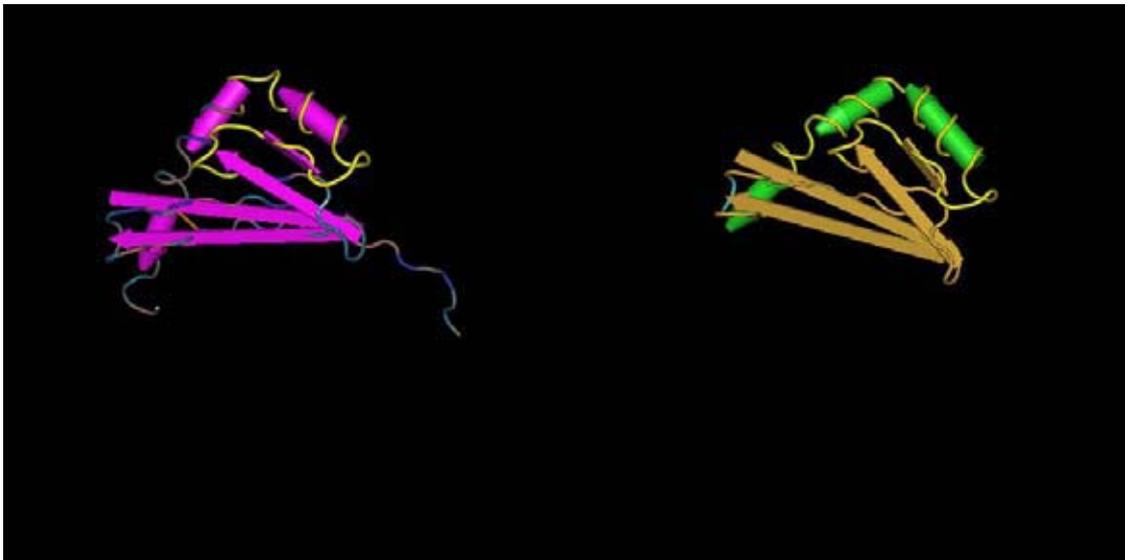


Figure 44. Modélisation de la structure tridimensionnelle de l'exon 2 du DRB (à gauche alleles Stco#2,3,12, à droite les autres allèles).

Chapitre V : Discussion Générale

SYNTHÈSE GENERALE

L'objectif général de ce travail était d'étudier le statut populationnel d'une espèce pouvant s'avérer un modèle biologique et épidémiologique accessible d'un Vertébré supérieur marin : le Dauphin bleu et blanc en Méditerranée. En effet, nous avons été amenés à nous poser des questions sur ses caractéristiques évolutives, en portant un intérêt particulier, quoique modeste, à son immunité spécifique. Pour cela, nous avons tenté d'approcher le problème comme suit :

1°) mise au point un marqueur de séquence sur le chromosome Y (gène de l'amélogénine) qui pourra être utilisé en biologie évolutive (phylogénie et génétique des populations) ;

2°) diversité génétique de l'espèce dans le cadre méditerranéen ;

3°) relation cette diversité génétique et la sensibilité au *Dolphin Morbillivirus*.

Afin de mener à bien cette étude, nous avons eu recours à une population d'échantillons de Dauphins échoués au cours des années 1990 à 2004 (n = 137) ainsi que des échantillons et séquences d'autres espèces. Outre les méthodes classiques de biologie moléculaire (PCR, séquence, génotypage, clonage, DGGE), ce travail a nécessité en particulier des approches statistiques variées, de type fréquentiste en génétique des populations (« F-statistics », AMOVA, régression distance génétiques/spatiales ...) ainsi que des approches bayésiennes appliquées à la phylogénie ou l'étude du mélange des populations (« *admixture* »).

Cependant, étant donné la difficulté de capture des individus, le nombre relativement restreint d'individus étudiés, bien que conséquent en Méditerranée, nous sommes conscients que notre échantillonnage tant sur le plan quantitatif que qualitatif (état des échantillons, localisation) peut soulever quelques questions en ce qui concerne la structuration des populations. Nous avons donc discuté nos résultats en tenant compte bien sûr de ces limites.

Dans un premier chapitre, nous avons étudié l'applicabilité d'une portion majoritairement intronique du gène de l'amélogénine comme marqueur d'origine paternelle sur notre modèle biologique. Pour cela, nous avons séquencé cette région chez une dizaine d'individus de Dauphins bleu et blanc ainsi que chez 4 autres espèces de Cétacés. L'analyse phylogénique de ces séquences, combinée à celles issues d'autres espèces de Cétartiodactyles, a été conçue pour prendre en compte les indels, fréquents sur le chromosome Y. D'une question spécifique à l'espèce, nous avons finalement élargi le champ à l'ensemble des Cétartiodactyles.

Les principaux résultats que nous avons pu dégager sont :

- la copie du gène présente sur le chromosome Y est porteuse d'une insertion d'environ 500 paires de bases par rapport à la copie présente sur le chromosome X chez les Cétartiodactyles à l'exception des Suioïdea ;

- cette insertion sur le chromosome Y montre une diversité génétique de substitutions notable chez le Dauphin bleu et blanc, au point qu'on puisse l'utiliser comme marqueur en génétique des populations (cf. [article n°1](#)).

Cette insertion a vraisemblablement eu lieu chez les mâles d'un ancêtre commun aux Cétartiodactyles, après la divergence précoce des Suioïdea et probablement des Tylopodes (non étudiés ici), c'est à dire vers la fin du Crétacé, à un moment où le goulet d'étranglement, causé par l'extinction en masse Crétacé/Tertiaire, a été tel qu'un certain nombre de mutations ont pu se fixer comme celles observées dans la descendance, c'est à dire chez la plupart des *Cetartiodactyla*. L'extraordinaire radiation des Mammifères et plus particulièrement des Cétartiodactyles qui a suivi a ainsi pu répercuter la mutation chez les espèces que nous connaissons aujourd'hui. Ce marqueur devrait donc permettre d'étudier des questions en rapport avec la radiation évolutive précoce des Cétartiodactyles.

D'un point de vue pratique, tant biologique (investigation et conservation) que vétérinaire (zootechnie et technologies alimentaires), le dimorphisme sexuel observé au niveau du marqueur étudié est utilisable pour le sexage d'espèces de Cétacés, Ruminants et Hippopotamidés.

Dans un deuxième chapitre, nous nous sommes intéressés à ce que l'on considérait jusqu'à présent comme la population de Dauphin bleu et blanc méditerranéenne. Nous avons tout d'abord reconsidéré la question de l'isolement relatif de cette population vis-à-vis de la population atlantique.

En effet, d'autres auteurs s'étaient déjà penchés sur la question à l'aide d'une analyse de restriction sur l'ADN mitochondrial et ont mis en évidence une différenciation génétique entre leurs échantillons atlantique et méditerranéen (**García-Martínez *et al.*, 1999**). Il nous est alors apparu intéressant de poser la même question en utilisant des microsatellites, marqueurs beaucoup plus polymorphes. Cela nous a permis d'envisager une revue de l'ensemble des microsatellites utilisables sur les Cétacés qui recense 246 *loci* (en majorité des dinucléotides) isolés de 18 espèces de cétacés et amplifiables dans 51 espèces représentant 11 familles- cf. [article n°2](#)).

L'échantillon considéré (78 individus de Méditerranée, 45 de l'Atlantique nord-est, 14 du

Pacifique nord) nous a permis de nuancer les conclusions de l'étude précédente (**García-Martínez et al., 1999**). En effet, cet échantillon méditerranéen montre des particularités (**déficit hétérozygotes** et **faible diversité allélique** ainsi que des déséquilibres de liaison importants entre les marqueurs), qui suggèrent une structuration et une **hétérogénéité** de la « population méditerranéenne » qui pourrait être la conséquence d'un échantillonnage d'individus de plusieurs populations (« effet Wahlund », cf. [article n°3](#)).

Nous avons donc décidé de combiner ces données de génotypages avec l'étude de la séquence de la région de contrôle de l'ADN mitochondrial.

Cela nous a permis de :

- rejeter l'hypothèse de panmixie de la population méditerranéenne (déviations à l'équilibre de Hardy-Weinberg, déséquilibre de liaison, F_{IS} différent d'une zone d'échantillonnage à l'autre mais non significativement différent de zéro si l'on considère toutes ces zones ensemble, pas de différence significative de F_{ST} d'une zone à l'autre, variance intra-zone supérieure à la variance inter-zone). Ces résultats montrant l'existence d'individus inbred ainsi qu'une structuration vont dans le sens d'un « effet Wahlund » ;

- rejeter l'hypothèse d'isolation par la distance ainsi que celle d'une structuration géographique tout en ne rejetant pas l'hypothèse de l'existence d'une barrière significative au flux génique entre le bassin occidental et le bassin oriental, mais uniquement sur la base de l'ADN mitochondrial. Ce résultat va dans le même sens que précédemment mais rajoute une probable notion de phylopatricie ;

Nous avons alors envisagé de repartir de l'échantillon méditerranéen dans son ensemble afin de dégager la structure à l'origine des observations précédentes par d'une part une approche phylogénétique et d'autre part une approche populationnelle Bayésienne basée sur des « coefficients d'ascendance ». Ces deux approches ont permis de délimiter deux groupes (« Groupe 1 » et « Groupe 2 ») nettement différenciés au sein de la Méditerranée. Afin de discuter de possibles origines, nous avons déterminé l'âge probable de divergence des deux lignées mitochondriales présentes (Miocène récent) à l'aide de deux méthodes, l'une fréquentiste, l'autre Bayésienne.

Notre tentative a été aussi d'évaluer les scénarios pouvant mener à l'actuel statut de l'espèce en relation avec les particularités géologiques de la Méditerranée. Cette approche globale dont la conclusion principale est une **polyphylie du Dauphin bleu et blanc en Méditerranée** nous a permis d'une part de discuter de l'origine géologique possible de celle-ci (Crise de Salinité du Messinien) mais aussi des facteurs endogènes qui ont permis la limitation de l'*admixture* entre

les deux groupes depuis 5 millions d'années.

Dans un troisième chapitre, nous avons tenté à notre modeste niveau une approche d'« épidémiologie évolutive » à l'instar des travaux de médecine évolutive qui commencent à prendre leur essor (Trevathan *et al.*, 1999). Nous n'avons fait qu'effleurer cette question, le projet étant probablement trop ambitieux pour le cadre d'une thèse en parallèle de l'étude du statut évolutif de l'espèce. Cependant, cette partie du travail nous a permis non seulement d'aborder des liens possibles avec la biologie cellulaire mais aussi de faire le lien avec notre formation initiale, médicale.

Nous avons dans un premier temps reconstitué la phylogénie des allèles d'une composante essentielle du Complexe Majeur d'Histocompatibilité dans la reconnaissance des allo-antigènes (comme les virus) : la molécule DR-B. Dans un deuxième temps, nous avons approché la pression de sélection s'exerçant de façon globale d'un part et sur chaque codon pris isolément d'autre part.

Les principaux enseignements de cette étude sont :

- d'une part, comme précédemment pour l'étude populationnelle, l'existence de plusieurs lignées évolutives au sein de l'espèce *Stenella caeruleoalba* et une notable diversité de cette portion du C.M.H., en désaccord avec l'ensemble des auteurs décrivant encore récemment une faible diversité dont l'explication était « environnementale » (milieu marin) ;

- d'autre part, les signatures de sélection observées ne mettent pas seulement en jeu une sélection « balancée » (observée sur la globalité de la séquence), mais aussi une sélection directionnelle (négative) sur un codon au sein d'une des deux lignées de DR-B, celle observée sur les individus du groupe et uniquement sur les individus du Groupe 1.

Les faiblesses de l'échantillonnage, évoquées en filigrane tout au long de la thèse, ne nous ont malheureusement pas permis de conclure à propos de la significativité de la différence de proportions d'individus du Groupe 1, ou de ceux porteurs d'un allèle de la lignée particulière du DR-B, au sein des individus suspects d'infection à DMV. Ces comparaisons auraient nécessité beaucoup plus d'échantillons ainsi qu'une plus grande précision du phénotype « malade » en raison justement de l'erreur de type II (faux positifs éventuels de l'examen nécropsique). On observe cependant une sur-représentation sub-significative d'individus du Groupe 1 au sein des individus dont l'examen nécropsique est consistant avec un mort par infection au DMV. On observe une tendance dans le même sens (sur-représentation de la lignée « méditerranéenne » du DR-B) au sein des individus supposés malades.

A la lumière de ces résultats, il nous semble intéressant de discuter les quelques points qui suivent et de dresser des perspectives futures possibles d'investigation.

DISCUSSION GENERALE ET PERSPECTIVES

Il nous semble tout d'abord opportun de discuter certaines observations à la lumière de nos résultats.

En premier lieu, il serait intéressant de se demander si les différences de taille observées entre l'Atlantique et la Méditerranée (**Di Méglia *et al.*, 1996**) ne seraient pas dues à un phénotype différent (plus petit ?) des individus du groupe 1. Il en est de même pour ce qui est des observations de différence de taille au sein de la Méditerranée (**Calzada & Aguilar, 1995**).

Certains auteurs (**García-Martínez *et al.*, 1995**), disent trouver une grande diversité haplotypique sur l'ensemble du génome mitochondrial (0,789) ; cependant nos valeurs sont toutes nettement supérieures, comprises entre 0,940 (groupe A) et 1.0 (Pacifique). Cette différence est selon toute vraisemblance due au fait que les enzymes de restriction utilisées par ces auteurs ne permettaient pas de mettre au jour les nombreux sites variables de la région de contrôle mitochondriale, sous-estimant de cette manière le nombre d'haplotypes. De même, les faibles valeurs de divergence de séquence (0,148% et 0,919%) sont à ramener à la totalité du génome mitochondrial qui est en moyenne beaucoup moins variable que ladite région de contrôle. En revanche, il paraît surprenant que ces auteurs n'aient pas mis en évidence de subdivision de population. En effet, vu l'ancienneté de la séparation des lignages mitochondriaux que nous avons observée, des sites variables synapomorphiques de chacun des deux groupes devraient être observés en-dehors de la région de contrôle.

PARAPATRIE OU SYMPATRIE ?

Il a été montré que la division morphologique traditionnelle entre variétés à long bec et à bec court de Dauphin commun *Delphinus delphis* n'est pas directement supportée par les données de séquences mitochondriales (**Natoli *et al.*, 2006**) et que l'on pourrait être en présence de convergences évolutives. D'autre part, une différenciation entre deux écotypes - « inshore » et « offshore » - de Dauphin tacheté de l'Atlantique *Stenella frontalis* a été récemment démontrée (**Adams & Rosel, 2006**). En outre, dans une investigation des relations entre taille des groupes

sociaux et divers paramètres en fonction de l'espèce, au sein des Delphinoïdea, des corrélations positives entre ouverture de l'habitat et taille des groupes au sein des espèces pélagiques (*Orcinus orca*, *Stenella caeruleoalba*, *Delphinus delphis*) ont été mises en évidence (Gygax, 2002). Les mêmes auteurs ont cependant observé des écarts à cette corrélation chez des espèces dont l'unicité fait débat (*Tursiops truncatus*, *Delphinus delphis*) et supposent que l'existence d'écotypes peut induire un biais dans la prédiction de la taille des groupes en fonction de l'ouverture de l'habitat. Ces mêmes écarts n'étant pas observés chez le Dauphin bleu et blanc, cela pourrait réfuter l'existence d'écotypes et renforcer notre hypothèse de coexistence sympatrique des deux groupes méditerranéens, et donc renforcer l'idée d'une moindre compatibilité fœto-maternelle entre des géniteurs issus de ces deux groupes et/ou de comportements reproducteurs suffisamment différents pour expliquer un écart à la panmixie.

Nos résultats sont en accord avec ceux d'auteurs observant une faible assignation géographique des individus (Valsecchi *et al.*, 2004a). Ainsi, les différences de coefficient d'*inbreeding* (IR) observées par ces auteurs entre les individus morts pendant et hors des périodes d'épizootie pourraient être dues à la différence que nous observons dans la répartition des deux groupes méditerranéens entre les individus supposés morts d'infection au DMV et les autres individus. En effet, il ne serait pas étonnant que le groupe 1 surreprésenté dans le bassin oriental fasse augmenter notablement l'IR dans cette région, vu la plus faible diversité de ces individus. Par ailleurs, ces mêmes auteurs ont observé en marge de leur étude une faible assignation géographique ce qui, combiné à nos observations, pourrait impliquer que le groupe 2, en étant plus philopatrick, représente la fraction assignable de leur échantillon. En effet, l'absence de structuration du groupe 1 pourrait avoir pour origine un hypothétique *habitus* pélagique de leurs ancêtres du Messinien : la Méditerranée n'ayant alors persisté qu'au niveau des fosses les plus profondes, il n'y avait pas de place pour un habitat de type talus continental. Cette hypothèse, bien que spéculative, pourrait être en accord avec le fait que le Grand dauphin n'a pas subi la même évolution et que la population méditerranéenne est le fruit d'une colonisation récente (pas de polyphyllie, gradient est-ouest ; (Natoli *et al.*, 2005)). Le Grand dauphin ayant un *habitus* plus côtier, il est possible que la perte de productivité des écosystèmes côtiers lors de la Crise de Salinité du Messinien ait provoqué la disparition d'ancêtres des Grands dauphins alors que les espèces pélagiques ont pu persister. De la même manière, on observe un gradient est-ouest (ouest-Atlantique/Méditerranée) modéré ainsi que la faible structure chez le Rorqual commun *Balænoptera physalus* (Bérubé *et al.*, 1998; Palsbøll *et al.*, 2004), ce qui dénote une colonisation récente de la Méditerranée et est en accord avec

l'origine récente (Pléistocène ; **Zane et al., 2000**) de la population Méditerranéenne de Krill (Crustacea: *Meganyctiphanes norvegica*), principale proie des Rorquals. Ainsi, on pourrait émettre l'hypothèse que les Mysticètes ayant un régime moins opportuniste n'aient pas pu persister comme l'ont fait les ancêtres des Dauphins bleu et blanc lors de la Crise Messinienne.

POPULATION, SOUS-ESPECE OU QUASI-ESPECE ?

Des doutes sur la stabilité du clade *Tursiops-Stenella(Prodelphinus)-Delphinus* ont été émis de longue date ainsi que sur les synapomorphies des genres le composant (**True, 1889**). De même, un siècle plus tard, certains auteurs (**Perrin et al., 1981**) ont remarqué qu'il existait un ensemble de caractères crâniens non partagés par tous les membres du genre *Stenella*. Ainsi, la monophylie du genre *Stenella* ne peut être assurée : il résulte vraisemblablement d'un assemblage artificiel. Les ressemblances entre les membres actuels du clade *Tursiops-Stenella(Prodelphinus)-Delphinus* nous semblent dues à des homoplasies (e.g. **Natoli et al., 2006**) ou à des simplésiomorphies qui ne résultent donc pas forcément du partage d'un ancêtre commun récent.

Seules deux études cladistiques ont été menées à bien sur un large éventail d'espèces de Delphinidae (**Barnes, 1990; Muizon, 1988**) et une seule par des techniques moléculaires (**Leduc et al., 1999**). Cependant, les deux études morphologiques n'ont utilisé qu'un petit nombre de caractères, les méthodologies n'ont pas été décrites (appréciation personnelle des caractères) et la monophylie des divers genres a été supposée et non testée. L'étude moléculaire en revanche a bien démontré la polyphylie des genres *Stenella* et *Tursiops* ainsi que la non-monophylie réciproque des genres *Delphinus*, *Stenella* et *Tursiops*. Le maintien d'une relative fluidité entre ces espèces pourrait résulter du même nombre de chromosomes et de morphologies proches combinées à une *inbreeding avoidance* marquée chez ces espèces (**Amos et al., 2001**).

Ainsi, bien sûr, la question se pose du statut taxonomique du Groupe 1. En effet, le faible flux de gènes observé entre les populations méditerranéennes pourrait élever le statut de ces groupes au-dessus du rang de population, au niveau subsppécifique. Dans un tel cas, la persistance d'un patron de coloration de type *cæruleoalba* pourrait résulter de la fixation d'un (plusieurs) gène(s) dominant(s) ou dont la fréquence aurait pu être portée à 1 dans les deux groupes même en présence d'un flux de gènes faible, en particulier si celui(ceux)-ci est(sont) soumis à une forte pression de sélection fréquence-dépendante qui pourrait résulter du comportement sexuel (**Felsenstein, 1976**). Il convient en outre de rappeler la grande variabilité des « patrons » de coloration chez toutes les espèces du clade *Tursiops-Stenella-Delphinus*

(Archer II & Perrin, 1999b; Perrin, 1975; Perrin, 1997) qui pourrait résulter d'homoplasies. Par ailleurs, les nombreux hybrides inter-spécifiques et fertiles mentionnés tant en captivité que sur le terrain (Baird *et al.*, 1998; Dhermain, 1995-2006; Herman, 1988; Reyes, 1996; Ridgway, 1999; Zornetzer & Duffield, 2003) pourraient n'être que la partie émergée d'un problème plus important : d'éventuels flux de gènes inter-spécifiques favorisés par les nombreuses associations de groupes polyspécifiques (Mallet, 2005; Mallet, 2007). Ceci pourrait expliquer le faible support des nœuds dans la reconstruction phylogénique des rapports entre petits *Delphinidae* (Leduc *et al.*, 1999) en raison de l'existence d'un certain degré de polyphylie au sein de mêmes genres et espèces. Cette opposition entre deux forces, diversifiante et homogénéisante, ne pourrait alors s'expliquer que si les gènes impliqués dans certains caractères morphologiques comme les patrons de coloration ou la taille du bec par exemple n'ont qu'une faible pénétrance ou bien à considérer un polymorphisme global de ces gènes à l'échelle du groupe des *Delphininae* (permettant ainsi un recouvrement interspécifiques de certains traits phénotypiques).

LES ANIMAUX ÉCHOUÉS SONT-ILS PLUS « CONSANGUINS » (*INBRED*) ?

Les travaux précédents (Valsecchi *et al.*, 2004a), montrant une tendance des animaux échoués à être plus « consanguins » (*inbred*), devraient être revus afin de prendre en compte l'existence des deux groupes méditerranéens. En effet, la valeur élevée de l'indice utilisé (IR = « *internal* » *relatedness*) pour les animaux échoués pourrait être due en particulier à la différence d'écologie dans les deux groupes : si l'un est plus côtier, il se pourrait que les individus aient plus de probabilité de s'échouer que ceux de l'autre, augmentant ainsi le degré de parenté entre les individus échoués. Une autre conclusion de leur étude serait peut-être à nuancer : l'IR élevé pourrait signifier que les individus échoués avaient une susceptibilité plus grande à diverses pathologies. Or cela pourrait s'avérer probable si l'on observait un biais dans la mesure d'IR en fonction de l'âge puisque ce genre de sélection ne s'observerait que sur les très jeunes individus, ce qui n'a pas été testé. Cependant, la diversité allélique inférieure du DR-B que l'on tend à observer chez les individus du groupe 1, ainsi que la faible diversité nucléaire et surtout mitochondriale, peut porter à croire que les individus de ce groupe sont non seulement plus *inbred* mais aussi moins largement immunocompétents donc éventuellement plus sensibles au DMV.

Par ailleurs, les individus collectés par (Valsecchi *et al.*, 2004a) étaient groupés en fonction de la période (« pendant » l'épisode et « après » l'épisode) sans diagnostic de certitude d'infection à

DMV. Cette absence de certitude rend le test de l'hypothèse d'une influence de l'IR sur la mortalité due au DMV difficile, du moins avec les faibles échantillonnages dont nous disposons.

Nous sommes conscient des limites de ce travail en particulier en ce qui concerne l'échantillonnage. Par exemple, de possibles biais ont pu provenir d'une différence dans la probabilité d'arrivée à la côte de population *inshore* et *offshore*, possibilité d'échouages de groupes liés familialement, etc.

En conséquence, et si l'on veut examiner ces hypothèses de susceptibilité au DMV différente chez les deux groupes, il conviendra dans le futur non seulement d'augmenter l'échantillonnage en termes quantitatifs mais aussi en termes qualitatifs, c'est à dire en récupérant d'une part des individus échoués dont le statut par rapport au DMV est connu, et d'autre part en collectant un maximum d'informations sur la localisation et le groupe dans lequel l'animal évolue (pour les animaux vivants).

LEXIQUE

Admixture, coefficient d'admixture	Terme de langue anglaise désignant spécifiquement le résultat du croisement de plusieurs populations précédemment isolées.
Ancêtre commun	En biologie de l'évolution, le concept de dernier ancêtre commun à deux lignées d'êtres vivants correspond à l'organisme le plus récent que ces deux taxons ont en commun. Il s'agit d'une entité théorique qui correspond au dernier nœud de l'arbre phylogénétique à partir duquel divergent les branches de chacune des lignées en question mais cela ne correspond pas forcément à un individu particulier ayant réellement existé, on le dit alors hypothétique.
Apomorphe	(gr. « forme lointaine »). En taxinomie, se dit d'un trait qui ne variera pas au cours de l'évolution.
Asymétrie fluctuante	Se dit d'une propriété d'un caractère qui est asymétrique « par accident » dans ses observations mais qui est symétrique si l'on considère toute une population.
Bottleneck	(angl. « col de bouteille »). Le goulet d'étranglement de population est, dans l'étude de l'évolution biologique, un évènement correspondant au fait qu'une proportion substantielle de la population d'une espèce est éliminée, ou empêchée de se reproduire. Il existe un type sensiblement différent de goulet d'étranglement démographique appelé effet fondateur. Celui-ci se produit lorsqu'un petit groupe d'une population est, pour la reproduction, isolé du groupe principal. Les goulets d'étranglement de population réduisent la diversité génétique, et donc la capacité des populations à s'adapter à de nouvelles contraintes de l'environnement.
Chromosome	(gr. « corps coloré »). Un chromosome est un élément microscopique constitué de molécules d'ADN.
Cluster de gènes	De <i>Grappe</i> en anglais. Ensemble de gènes regroupés physiquement et codant des produits similaires ou voisins.
CMH	Région du génome des Gnathostomes, qui regroupe l'essentiel de l'information génétique codant des protéines cellulaires de surfaces spécifiques de chaque individu.
DRB	Nom acronymique d'une région du CMH codant la chaîne beta du récepteur

	membranaire des cellules présentatrices d'antigène.
Cladistique	(gr. « (étude des) branches »). Système de classification phylogénétique basé sur les relations de parenté, établies sur la base du partage de caractères, (ou synapomorphie, ou aussi homologie secondaire), de façon à obtenir des groupes monophylétiques (ou clade).
Consanguinité	(lat. « lien du sang »). Se dit d'une propriété commune à des individus issus de reproduction entre individus apparentés. Par extension, ce terme désigne la proportion d'allèles identiques entre deux individus (personnes, animaux, plantes).
Déséquilibre de liaison	Se dit de l'association préférentielle d'allèles de gènes différents. Le résultat est que la fréquence des gamètes porteurs des allèles de deux locus différents A et B est différente du produit des fréquences des allèles.
Electrophorèse	(gr. « porté par l'ambre » / N.B. : l'ambre a été le support de la démonstration de l'électricité dite statique). Technique utilisée pour séparer et caractériser des molécules telles que des protéines ou des acides nucléiques (ADN, ARN, etc.) basée sur le fait que dans un milieu donné soumis à un champ électrique, la migration et la séparation des molécules se fait en fonction de leur charge électrique et pour des charges identiques, en fonction de leur taille et de leur forme.
Eluer	(lat. « rincer »). Terme chimique signifiant « balayer avec un fluide (éluant) des espèces chimiques variées que l'on souhaite faire migrer de manière à les séparer ».
Euteleostomi	(gr. « vraie bouche terminale »). Se dit Vertébrés possédant une mandibule.
Eucaryotes	(gr. littéralement « noyau vrai »). Se dit d'être vivants. caractérisé par le fait que la majorité du matériel génétique cellulaire est contenu dans un noyau constitué par une membrane nucléaire (un « vrai noyau »), par opposition aux procaryotes.
Fitness	(angl. « aptitude »). La valeur sélective (fitness, succès reproducteur ou valeur adaptative) est un concept central en théorie de l'évolution. Elle décrit la capacité d'un individu d'un certain génotype à se reproduire. C'est une mesure de la sélection naturelle.

Fréquentiste	Se dit d'approches statistiques privilégiant une interprétation relative à la fréquence dans un calcul de probabilités.
Gène	(gr. « origine »). Unité de base de l'information biologique qui se transmet de génération en génération et dont le support est l'A.D.N.
Gnathostome	Les Gnathostomes (littéralement « bouche pourvue d'une mâchoire ») sont un groupe rassemblant les vertébrés pourvus d'une mâchoire. Il est bien établi que les Gnathostomes constituent un groupe monophylétique.
Haplotype	(angl. haploid – gr. « simple » +genotype – gr. « empreinte de l'origine »). Un haplotype est un groupe d'allèles de différents gènes situés sur un même chromosome et habituellement transmis ensemble.
Heuristique (Méthode)	(gr. « trouver »). Méthode de résolution de problèmes non fondée sur un modèle formel et qui n'aboutit pas nécessairement à une solution optimale.
Hétérodimérisation	(gr. « deux parties différentes »; lat. suffixe d'action -atio). Capacité d'une protéine à se combiner de manière non covalente à une autre protéine non homologue.
HLA	(angl. <i>Human Leukocyte Antigen</i>). Le Complexe Majeur d'Histocompatibilité de l'Homme.
Homologue	(gr. littéralement « discours - similaire »). Se dit de gènes descendants d'un même gène ancestral.
Horloge moléculaire	En génétique, l'hypothèse de l'horloge moléculaire stipule que les mutations génétiques s'accumulent dans un génome à une vitesse globalement proportionnelle au temps géologique. Elle permet ainsi théoriquement, en reliant le taux de mutation des gènes au rythme de diversification de leur espèce, d'établir une échelle d'ordonnement, voire chronologique, d'évolution et de lien des espèces entre elles.
IBD	L'isolement par la distance (<i>Isolation By Distance</i>) est la tendance d'individus à se reproduire avec des individus issus de populations voisines plutôt que de populations distantes. Elle résulte en une distribution « clinale » des traits au sein d'une région géographique.
Immunoglobuline	(lat. littéralement « exemption » ,« petite

	boule ») Une glycoprotéine membranaire ou soluble impliquée dans les phénomènes de reconnaissance, de liaison et d'adhésion des cellules comme dans le système immunitaire.
Implémenter	(angl. « mettre en œuvre »). Terme de programmation informatique signifiant « traduire un algorithme dans un langage de programmation ».
Indels	(angl.). Néologisme pour « insertion-deletion ».
Lysosome	(gr. « corps de dissolution »). Organite cellulaire effectuant la digestion intracellulaire, et présent dans le cytosol de toutes les cellules eucaryotes animales, à l'exception des hématies.
Orthologue	(gr. littéralement « discours - droit »). Se dit de gènes homologues chez des espèces différentes et ayant donc un lien évolutif.
Paralogue	(gr. littéralement « discours - à côté »). Se dit de gènes homologues issus d'un événement de duplication.
Peptide	(gr. « digéré »). Polymère biochimique composé d'un faible nombre (moins de 10) d'acides aminés reliés par des liaisons peptidiques.
Phylogenie	La phylogénie (littéralement « naissance des tribus/groupes ») est l'étude des relations de parentés entre différents êtres vivants en vue de comprendre l'évolution des organismes vivants. On peut étudier la phylogénie d'un groupe d'espèces mais également, à un niveau intraspécifique, la généalogie entre populations ou entre individus.
Placentaire	(lat. « de galette »). Mammifère dont le fœtus est pourvu de placenta.
Plésiomorphe	(gr. « forme voisine »). En taxinomie, se dit d'un trait qui ne variera pas au cours de l'évolution.
Polymorphisme	(gr. « plusieurs formes »). Se dit de la propriété des espèces animales qui se présentent sous plusieurs formes différentes, à un même stade de développement ou plus généralement, à la fin de leur développement. Plus spécifiquement dans le cas de l'ADN, il s'agit de variations de la séquence nucléotidique.
polyphylétisme	(gr. « plusieurs groupes », suffixe de substantif - isme,). Se dit du fait d'être composé de

	plusieurs espèces ou taxa (taxon/ groupe polyphylétique).
Probabilités conditionnelles	En théorie des probabilités, la probabilité conditionnelle d'un événement A, sachant qu'un autre événement B de probabilité non nulle s'est réalisé (ou probabilité de A, sachant B) est le nombre noté défini par le quotient entre la probabilité de l'événement « A et B » et la probabilité de B.
Programme/version parallélisé(e)	Un programme (ou une version d'un programme) « parallélisé » est un programme dont l'exécution peut s'effectuer sur plusieurs processeurs en même temps. Son exécution est ainsi écourtée par rapport à une version « séquentielle ».
Recombinaison	La recombinaison génétique est « le phénomène conduisant à l'apparition, dans une cellule ou dans un individu, de gènes ou de caractères héréditaires dans une association différente de celle observée chez les cellules ou individus parentaux (définition historique retenue dans le J.O. de la République Française 22/10/2000).
Région de contrôle	Une région de contrôle du locus est une séquence d'ADN qui va stimuler la transcription d'ADN en fonction du contexte moléculaire de la cellule (lié au nombre de copies d'ARNm de ce gène déjà transcrite, permet le maintien d'un nombre suffisant de copies d'ARN). Plus spécifiquement dans le cas de l'ADN mitochondrial, il s'agit d'une région que en outre de sa fonction régulatrice est très polymorphe (intérêt en phylogénie).
TMRCA	Temps séparant le présent du dernier ancêtre commun.
Transition	En génétique, se dit d'une mutation impliquant le remplacement, l'addition ou la délétion d'un nucléotide.
Transversion	En génétique, se dit d'une mutation impliquant le remplacement d'une purine (A ou G) par une pyrimidine (C ou T) ou vice versa.
Vraisemblance, fonction de vraisemblance	Notion statistique due à Ronald Fisher (1922). La vraisemblance (<i>Likelihood</i>) mesure l'adéquation entre un échantillon aléatoire et une loi de probabilité que nous estimons être la plus proche de la distribution observée (en principe au regard d'un graphique). Elle prend

	la forme d'une fonction de probabilités conditionnelles qui décrit les valeurs d'une loi statistique en fonction des paramètres supposés connus. Elle s'exprime à partir de la fonction de densité.
--	---

REFERENCES

- E. Aasen & J. F. Medrano** (1990) Amplification of the ZFY and ZFX genes for sex identification in humans, cattle, sheep and goats. *Biotechnology (N Y)* **8**(12): 1279-81. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=1369448 <http://dx.doi.org/10.1038/nbt1290-1279>
- L. Abi Rached, M. F. McDermott & P. Pontarotti** (1999) The MHC big bang. *Immunological Reviews* **167**: 33-44. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10319249 <http://dx.doi.org/10.1038/nbt1290-1279>
- L. Abi-Rached, A. Gilles, T. Shiina, P. Pontarotti & H. Inoko** (2002) Evidence of en bloc duplication in vertebrate genomes. *Nature Genetics* **31**(1): 100-5. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=11967531 <http://dx.doi.org/10.1038/ng855>
- L. D. Adams & P. E. Rosel** (2006) Population differentiation of the atlantic spotted dolphin (*Stenella frontalis*) in the western North Atlantic, including the Gulf of Mexico. *Marine biology* **148**(3): 671-681. <http://dx.doi.org/10.1007/s00227-005-0094-2>
- A. Aguilar** (1991) Calving and early mortality in the western Mediterranean striped dolphin, *Stenella coeruleoalba*. *Canadian journal of zoology* **69**(5): 1408-1412. <http://dx.doi.org/10.1139/z91-199>
- A. Aguilar & J. A. Raga** (1993) The striped dolphin epizootic in the Mediterranean Sea. *Ambio* **22**: 528-528.
- A. Aguilar & A. Borrell** (1994) Abnormally high polychlorinated biphenyl levels in striped dolphins (*Stenella coeruleoalba*) affected by the 1990-1992 Mediterranean epizootic. *Science of the Total Environment* **154**(2-3): 237-47. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=7973610
- A. Aguilar** (2000a) Population biology, conservation threats and status of Mediterranean striped dolphins (*Stenella coeruleoalba*). *Journal of Cetacean Research and Management* **2**(1): 17-26.
- A. Aguilar** (2000b) Population biology, conservation threats and status of Mediterranean striped dolphins (*Stenella coeruleoalba*). *Journal of Cetacean Research and Management* **2**(1): 17-26.
- A. Aguilar, G. Roemer, S. Debenham, M. Binns, D. Garcelon & R. K. Wayne** (2004) High MHC diversity maintained by balancing selection in an otherwise genetically monomorphic mammal. *Proceedings of the National Academy of Sciences of the United States of America* **101**(10): 3490-4.
- Y. Aida, M. Niimi, M. Asahina, K. Okada, Y. Nakai & K. Ogimoto** (1995) Identification of a new bovine MHC class II DRB allele by nucleotide sequencing and an analysis of phylogenetic relationships. *Biochemical and Biophysical Research Communications* **209**(3): 981-8. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=7733993 <http://dx.doi.org/10.1006/bbrc.1995.1594>
- H. Akaike** (1974) A new look at the statistical model identification. *IEEE Transactions on Automatic Control* **AC-19**(6): 716-723. <http://dx.doi.org/10.1109/TAC.1974.1100705>
- A. Alcamí & U. H. Koszinowski** (2000) Viral mechanisms of immune evasion. *Immunology Today* **21**(9): 447-455. [http://dx.doi.org/10.1016/S0167-5699\(00\)01699-6](http://dx.doi.org/10.1016/S0167-5699(00)01699-6) [http://10.1016/S0167-5699\(00\)01699-6](http://10.1016/S0167-5699(00)01699-6)
- B. M. Aldridge, S. M. McGuirk, R. J. Clark, L. A. Knapp, D. I. Watkins & D. P. Lunn** (1998) Denaturing gradient gel electrophoresis: a rapid method for differentiating BoLA-DRB3 alleles. *Animal Genetics* **29**(5): 389-94.

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9800329 <http://dx.doi.org/10.1046/j.1365-2052.1998.295361.x>

V. C. Almada, R. F. Oliveira, E. J. Gonçalves, A. J. Almeida, R. S. Santos & P. Wirtz (2001) Patterns of diversity of the north-eastern Atlantic blenniid fish fauna (Pisces : Blenniidae). *Global Ecology and Biogeography* **10**(4): 411-422.

J. R. Alvarado Bremer, B. Stequert, N. W. Robertson & B. Ely (1998) Genetic evidence for inter-oceanic subdivision of bigeye tuna (*Thunnus obesus*) populations. *Marine Biology* **132**(4): 547-557.

W. Amos, J. W. Wilmer, K. Fullard, T. M. Burg, J. P. Croxall, D. Bloch & T. Coulson (2001) The influence of parental relatedness on reproductive success. *Proceedings of the Royal Society of London series B Biological Sciences* **268**(1480): 2021-7. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=11571049 <http://dx.doi.org/10.1098/rspb.2001.1751>

H. A. Anderson, E. M. Hiltbold & P. A. Roche (2000) Concentration of MHC class II molecules in lipid rafts facilitates antigen presentation. *Nature Immunology* **1**(2): 156-62. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=11248809 <http://dx.doi.org/10.1038/77842>

G. Andersson (1998) Evolution of the human HLA-DR region. *Frontiers in Biosciences* **3**: d739-45.

G. Andersson, A. C. Svensson, N. Setterblad & L. Rask (1998) Retroelements in the human MHC class II region. *Trends in Genetics* **14**(3): 109-14. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9540408 [http://dx.doi.org/10.1016/S0168-9525\(97\)01359-0](http://dx.doi.org/10.1016/S0168-9525(97)01359-0)

P. Andre, J. C. Cambier, T. K. Wade, T. Raetz & W. F. Wade (1994) Distinct structural compartmentalization of the signal transducing functions of major histocompatibility complex class II (Ia) molecules. *Journal of Experimental Medicine* **179**(2): 763-8. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8294884 <http://dx.doi.org/10.1084/jem.179.2.763>

R. C. Andrews (1911) Description of an apparently new Porpoise of the genus *Tursiops*, with remarks upon a Skull of *Tursiops gilli* Dall. *Bulletin of the American Museum of Natural History* **30**: 233-237. <http://hdl.handle.net/2246/1787>

Anonymous (1999) Agreement on the creation of a Mediterranean Sanctuary for marine mammals. *Final agreement signed in Roma on 25 November 1999 by the Ministers of Italy, France and Monaco*

T. Aoba (1996) Recent observations on enamel crystal formation during mammalian amelogenesis. *Anatomical Record* **245**(2): 208-18. [http://dx.doi.org/10.1002/\(SICI\)1097-0185\(199606\)245:2%3C208::AID-AR8%3E3.0.CO;2-S](http://dx.doi.org/10.1002/(SICI)1097-0185(199606)245:2%3C208::AID-AR8%3E3.0.CO;2-S)

F. Archer II & W. F. Perrin (1999a) *Stenella cœruleoalba*. *Mammalian Species*(603): 1-9.

F. I. Archer II & W. F. Perrin (1999b) *Stenella coeruleoalba*. *Mammalian Species* **603**: 1-9. <http://dx.doi.org/10.2307/3504476>

U. Árnason, M. Hoglund & B. Widegren (1984) Conservation of highly repetitive DNA in cetaceans. *Chromosoma* **89**(3): 238-242. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=6325102 <http://dx.doi.org/10.1007/BF00295005>

J. Ascher, M. T. Ceccherini, O. L. Pantani, A. Agnelli, F. Borgogni, G. Guerri, P. Nannipieri & G. Pietramellara (2009) Sequential extraction and genetic fingerprinting of a forest soil metagenome. *Applied Soil Ecology* **42**(2): 176-181. <http://dx.doi.org/10.1016/j.apsoil.2009.03.005>

H. Augier, L. Benkoel, A. Chamlian, W. K. Park & C. Ronneau (1993) Mercury, zinc and

selenium bioaccumulation in tissues and organs of Mediterranean striped dolphins *Stenella coeruleoalba* Meyen. Toxicological results of their interaction. *Cellular and molecular biology* **39**(6): 621-634.

J. Avise (2000) *Phylogeography: the History and Formation of Species*. Harvard University Press, Cambridge, MA.

W. Babik, W. Durka & J. Radwan (2005) Sequence diversity of the MHC DRB gene in the Eurasian beaver (*Castor fiber*). *Molecular Ecology* **14**(14): 4249-57.

R. W. Baird, P. M. Willis, T. J. Guenther, P. J. Wilson & B. N. White (1998) An intergeneric hybrid in the family Phocoenidae. *Canadian Journal of Zoology* **76**(1): 198-204. <http://dx.doi.org/10.1111/j.1365-294X.2005.02751.x>

C. S. Baker, M. D. Vant, M. L. Dalebout, G. M. Lento, S. J. O'Brien & N. Yuhki (2006) Diversity and duplication of DQB and DRB-like genes of the MHC in baleen whales (suborder: Mysticeti). *Immunogenetics* **58**(4): 283-96. <http://dx.doi.org/10.1007/s00251-006-0080-y>

F. M. Balfour (1880) On the spinal nerves of *Amphioxus*. *Quart. Journal of Microbiological Science Also* **s2-20**(77): 90-91. <http://jcs.biologists.org/cgi/reprint/s2-20/77/90>

P. Balvanera, G. C. Daily, P. R. Ehrlich, T. H. Ricketts, S. A. Bailey, C. Kark & H. Pereira (2001) Conserving biodiversity and ecosystem services. *Science* **291**(5511): 2047. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=11256386 [10.1126/science.291.5511.2047](http://dx.doi.org/10.1126/science.291.5511.2047)

K. T. Ballingall, A. Luyai, G. J. Rowlands, J. Sales, A. J. Musoke, S. P. Morzaria & D. J. McKeever (2004) Bovine leukocyte antigen major histocompatibility complex class II DRB3*2703 and DRB3*1501 alleles are associated with variation in levels of protection against *Theileria parva* challenge following immunization with the sporozoite p67 antigen. *Infection and Immunity* **72**(5): 2738-41. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15102783

L. G. Barrett-Lennard & G. M. Ellis (2001) Population Structure and Genetic Variability in Northeastern Pacific Killer whale: Towards an Assessment of Population Viability. in A report to Fisheries and Oceans Canada, Ottawa, 35 pp. http://www.dfo-mpo.gc.ca/csas/Csas/DocREC/2001/RES2001_065e.pdf

L. Bargelloni, J. A. Alarcon, M. C. Alvarez, E. Penzo, A. Magoulas, C. Reis & T. Patarnello (2003) Discord in the family Sparidae (Teleostei): divergent phylogeographical patterns across the Atlantic-Mediterranean divide. *Journal of evolutionary biology* **16**(6): 1149-58. <http://dx.doi.org/10.1046/j.1420-9101.2003.00620.x>

L. G. Barnes (1976) Outline of Eastern North Pacific Fossil Cetacean Assemblages. *Systematic Zoology* **25**(4): 321-343. <http://www.jstor.org/stable/2412508> <http://dx.doi.org/10.2307/2412508>

L. G. Barnes, D. P. Domming & C. E. Ray (1985) Status of studies on fossil marine mammals. *Marine Mammal Science* **1**(2): 15-53. <http://dx.doi.org/10.1111/j.1748-7692.1985.tb00530.x>

L. G. Barnes (1990) The fossil record and evolutionary relationships of the genus *Tursiops*. in *The bottlenose dolphin*. (S. L. a. R. R. Reeves Ed.). Academic Press Inc., San Diego. p 3-26.

N. H. Barton & B. Charlesworth (1998) Why sex and recombination? *Science* **281**(5385): 1986-1990. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9748151 <http://dx.doi.org/10.1126/science.281.5385.1986>

M. Bayes & M. Price (1763) LII. An Essay towards Solving a Problem in the Doctrine of Chances. By the Late Rev. Mr. Bayes, FRS Communicated by Mr. Price, in a Letter to John Canton, AMFRS. *Philosophical Transactions of the Royal Society of London* **53**: 370-418. <http://dx.doi.org/10.1098/rstl.1763.0053>

G. Bearzi (2002) Interactions between cetaceans and fisheries: Mediterranean Sea. *in* Cetaceans of the Mediterranean and Black Seas: state of knowledge and conservation strategies. A report to the ACCOBAMS Secretariat, Monaco, February A report to ACCOBAMS Secretariat, Monaco, February 2002. 20 p.

G. Bearzi, R. R. Reeves, G. Notarbartolo Di Sciara, E. Politi, A. Canadas, A. Frantzis & B. Mussi (2003) Ecology, status and conservation of short-beaked common dolphins *Delphinus delphis* in the Mediterranean Sea. *Mammal Review* **33**(3-4): 224-252. <http://dx.doi.org/10.1046/j.1365-2907.2003.00032.x>

G. Bearzi, E. Politi, S. Agazzi & A. Azzellino (2006) Prey depletion caused by overfishing and the decline of marine megafauna in eastern ionian sea coastal waters (central Mediterranean). *Biological conservation* **127**(4): 373-382. <http://dx.doi.org/10.1016/j.biocon.2005.08.017>

P. Beaubrun (1995) Preliminary atlas of distribution of the Mediterranean cetaceans (*P. B. Editor Ed.*). Oceanographic Museum of Monaco/CIESM, Monaco, Monaco. 87 p. pp.

P. Beerli & J. Felsenstein (2001) Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *Proceedings of the National Academy of Sciences of the United States of America* **98**(8): 4563-4568. <http://dx.doi.org/10.1073/pnas.081068098> <http://dx.doi.org/10.1073/pnas.081068098>

L. Bernatchez & C. Landry (2003) MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? *J Evol Biol* **16**(3): 363-77. <http://dx.doi.org/10.1046/j.1420-9101.2003.00531.x>

P. Berrebi (1995) Speciation of the genus *Barbus* in the north Mediterranean basin: Recent advances from biochemical genetics. *Biological conservation* **72**(2): 237-249. [http://dx.doi.org/10.1016/0006-3207\(94\)00086-6](http://dx.doi.org/10.1016/0006-3207(94)00086-6) [http://dx.doi.org/10.1016/0006-3207\(94\)00086-6](http://dx.doi.org/10.1016/0006-3207(94)00086-6)

M. Bérubé, A. Aguilar, D. Dendanto, F. Larsen, G. Notarbartolo di Sciara, R. Sears, J. Sigurjonsson, R. J. Urban & P. J. Palsbøll (1998) Population genetic structure of North Atlantic, Mediterranean Sea and Sea of Cortez fin whales, *Balaenoptera physalus* (Linnaeus 1758): analysis of mitochondrial and nuclear loci. *Molecular ecology* **7**(5): 585-599. <http://dx.doi.org/10.1046/j.1365-294x.1998.00359.x> <http://dx.doi.org/10.1046/j.1365-294x.1998.00359.x>

E. Betran, J. Rozas, A. Navarro & A. Barbadilla (1997) The estimation of the number and the length distribution of gene conversion tracts from population DNA sequence data. *Genetics* **146**(1): 89-99. <http://scholar.google.com/scholar?q=The+estimation+of+the+number+and+the+length+distribution+of+gene+conversion+tracts+from+population+DNA+sequence+data&hl=ca&btnG=Cerca>

G. Bianucci (2001) A new genus of Kentriodontid (Cetacea : Odontoceti) from the Miocene of south Italy. *Journal of Vertebrate Paleontology* **21**(3): 573-577. [http://dx.doi.org/10.1671/0272-4634\(2001\)021%5B0573:ANGOKC%5D2.0.CO;2](http://dx.doi.org/10.1671/0272-4634(2001)021%5B0573:ANGOKC%5D2.0.CO;2)

G. Bianucci & W. Landini (2002) Change in diversity, ecological significance and biogeographical relationships of the Mediterranean Miocene toothed whale fauna. *Geobios* **35**(Mémoire Spécial n°24): 19-28. [http://dx.doi.org/10.1016/S0016-6995\(02\)00045-1](http://dx.doi.org/10.1016/S0016-6995(02)00045-1)

O. R. Bininda-Emonds, M. Cardillo, K. E. Jones, R. D. MacPhee, R. M. Beck, R. Grenyer, S. A. Price, R. A. Vos, J. L. Gittleman & A. Purvis (2007) The delayed rise of present-day mammals. *Nature* **446**(7135): 507-12. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17392779 <http://dx.doi.org/10.1038/nature05634>

A. Birkun, Jr., T. Kuiken, S. Krivokhizhin, D. M. Haines, A. D. Osterhaus, M. W. van de Bildt, C. R. Joiris & U. Siebert (1999) Epizootic of morbilliviral disease in common dolphins

(*Delphinus delphis ponticus*) from the Black sea. *The Veterinary Record* **144**(4): 85-92. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10097312 <http://dx.doi.org/10.1136/vr.144.4.85>

P. L. Blanc (2000) Of sills and straits: a quantitative assessment of the Messinian Salinity Crisis. *Deep Sea Research, Part I Oceanographic Research Papers* **47**(8): 1429-60. [http://dx.doi.org/10.1016/S0967-0637\(99\)00113-2](http://dx.doi.org/10.1016/S0967-0637(99)00113-2)

C. Blanco, J. Aznar & J. A. Raga (1995) Cephalopods in the diet of *Stenella coeruleoalba* (Meyen,1833) from the Western Mediterranean during an epizootic in 1990. *Journal of Zoology, London* **237**: 151-158. <http://dx.doi.org/10.1111/j.1469-7998.1995.tb02753.x>

D. Bloch, G. Desportes, A. Petersen & J. Sigurjonsson (1996) Strandings of striped dolphins (*Stenella coeruleoalba*) in Iceland and the Faroe Islands and sightings in the Northeast Atlantic, North of 50°N latitude. *Marine Mammal Science* **12**(1): 125-132. <http://dx.doi.org/10.1111/j.1748-7692.1996.tb00310.x>

A. J. Bohonak (2002) IBD (Isolation by Distance): A Program for Analyses of Isolation by Distance. *Journal of Heredity* **93**(2): 153-154. <http://dx.doi.org/10.1093/jhered/93.2.153>

J. R. Boisserie, F. Lihoreau & M. Brunet (2005) The position of hippopotamidae within cetartiodactyla. *Proceedings of the National Academy of Sciences of the United States of America* **102**(5): 1537-1541. <http://dx.doi.org/10.1073/pnas.0409518102>

G. Bombace (2001) Influence of climatic changes on stocks, fish species and marine ecosystems in the Mediterranean sea. *Archivio di Oceanogra e Limnologia* **22**: 67-72.

C. L. Bonaparte (1836) Tavola analitica dei chelonii o testuggini. *Giornale Arcadico di Scienze Lettere ed Arti* **69**: 54-64. <http://books.google.fr/books?id=HD0oAAAAYAAJ&ots=GcAb2qLeyx&dq=Tavola analitica dei chelonii o testuggini&lr&hl=ca&pg=PA33-v=onepage&q=Tavola%20analitica%20dei%20chelonii%20o%20testuggini&f=false>

W. A. Bonass, J. Kirkham, S. J. Brookes, R. C. Shore & C. Robinson (1994) Isolation and characterisation of an alternatively-spliced rat amelogenin cDNA: LRAP--a highly conserved, functional alternatively-spliced amelogenin? *Biochimica et Biophysica Acta* **1219**(3): 690-2. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=7948026 [http://dx.doi.org/10.1016/0167-4781\(94\)90228-3](http://dx.doi.org/10.1016/0167-4781(94)90228-3)

A. Borrell, A. Aguilar, S. Corsolini & S. Focardi (1996) Evaluation of toxicity and sex-related variation of PCB levels in Mediterranean striped dolphins affected by an epizootic. *Chemosphere* **32**(12): 2359-69. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8653381 [http://dx.doi.org/10.1016/0045-6535\(96\)00143-9](http://dx.doi.org/10.1016/0045-6535(96)00143-9)

P. Borsa & J. P. Quignard (2001) Systematics of the Atlantic-Mediterranean soles *Pegusa impar*, *P. lascaris*, *Solea aegyptiaca*, *S. senegalensis*, and *S. solea* (Pleuronectiformes: Soleidae). *Canadian Journal of Zoology* **79**(12): 2297-2302. <http://dx.doi.org/10.1139/cjz-79-12-2297>

A. Bortolotto, L. Casini & L. A. Stanzani (1992) Dolphin mortality along the southern Italian coast (June-September 1991). *Aquatic Mammals* **18**(2): 56-60. http://www.aquaticmammalsjournal.org/share/AquaticMammalsIssueArchives/1992/AquaticMammals_18_2/Bortolotta.pdf

M. Bouillon & W. M. Mourad (2003) Complexe majeur d'histocompatibilité de classe II: diversité fonctionnelle[Major histocompatibility complex (MHC) class II: are lipid rafts the missing link?]. *Médecine/Sciences* **19**(10): 988-93. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=14613012 <http://dx.doi.org/10.1051/medsci/20031910988>

V. Bourret, M. Macé & B. Crouau-Roy (2007) Genetic variation and population structure of western Mediterranean and northern Atlantic *Stenella coeruleoalba* populations inferred from microsatellite data. *Journal of the Marine Biological Association of the United Kingdom* **87**: 265-269.

<http://dx.doi.org/10.1017/S0025315407054859>

N. Boury-Esnault, M. Pansini & M. J. Uriz (1992) A new Discorhabdella (Porifera, Demospongiae), a new Tethyan relict of pre-Messinian biota? *Journal of Natural History* **26**(1): 1-7.

<http://dx.doi.org/10.1516/U35T-9815-672W-7H30>
<http://dx.doi.org/10.1080/00222939200770011>

L. E. Bradbury, G. S. Kansas, S. Levy, R. L. Evans & T. F. Tedder (1992) The CD19/CD21 signal transducing complex of human B lymphocytes includes the target of antiproliferative antibody-1 and Leu-13 molecules. *Journal of Immunology* **149**(9): 2841-50.

<http://www.ncbi.nlm.nih.gov/gate2.inist.fr/pubmed/1383329>

F. M. Briand & A. Maldonado (1996) Transformations and evolution of the Mediterranean coastline in Bulletin de l'Institut Océanographique. Commission Internationale pour l'Exploration Scientifique de la Mer Méditerranée, Monaco. 243 pp.

<http://books.google.com/books?id=hQaAAAAAMAAJ&hl=ca&pgis=1>

A. D. Brisson (1762) Regnum Animale in Classes IX Distributum, Sive Synopsis Methodica (E. a. auctior. Ed.). Theodorum Haak, Leiden, Netherlands.

http://openlibrary.org/books/OL24145697M/Regnum_animale_in_classes_IX_distributum_sive_Synopsis_methodica

F. Broili (1913) Unser Wissen über die ältesten Tetrapoden. *Fortschritte der naturwissenschaftlichen:* 51-93.

http://books.google.fr/books?id=RrwWAQAIAAJ&q=Unser+Wissen+uber+die+altesten+Tetrapoden&dq=Unser+Wissen+uber+die+altesten+Tetrapoden&hl=ca&sa=X&ei=rh6T_i2GdPs8QOQvsi5CA&redir_esc=y

L. Bromham, A. Rambaut, R. Fortey, A. Cooper & D. Penny (1998) Testing the Cambrian explosion hypothesis by using a molecular dating technique. *Proceedings of the National Academy of Sciences of the United States of America* **95**(21): 12386-12389.

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9770496 <http://dx.doi.org/10.1073/pnas.95.21.12386>

D. R. Brooks & A. L. Ferrao (2005) The historical biogeography of co-evolution: emerging infectious diseases are evolutionary accidents waiting to happen. *Journal Of Biogeography* **32**(8): 1291-1299.

<http://dx.doi.org/10.1111/j.1365-2699.2005.01315.x>

C. J. Brown, L. Carrel & H. F. Willard (1997) Expression of genes from the human active and inactive X chromosomes. *American journal of human genetics* **60**(6): 1333-1343.

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9199554 <http://dx.doi.org/10.1086/515488>

J. J. Brown, W. Thomson, P. Clegg, S. Eyre, L. J. Kennedy, J. Matthews, S. Carter & W. E. Ollier (2004) Polymorphisms of the equine major histocompatibility complex class II DRA locus. *Tissue Antigens* **64**(2): 173-9.

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15245372 <http://dx.doi.org/10.1111/j.1399-0039.2004.00269.x>

U. Brunsberg, I. Edfors-Lilja, L. Andersson & K. Gustafsson (1996) Structure and organization of pig MHC class II DRB genes: evidence for genetic exchange between loci. *Immunogenetics* **44**(1): 1-8.

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8613138 <http://dx.doi.org/10.1007/BF02602651>

D. R. Burton, R. L. Stanfield & I. A. Wilson (2005) Antibody vs. HIV in a clash of evolutionary titans. *Proceedings of the National Academy of Sciences of the United States of America* **102**(42): 14943-14948.

<http://dx.doi.org/10.1073/pnas.0505126102>
<http://dx.doi.org/10.1073/pnas.0505126102>

O. Bütschli (1910) Einleitung, vergleichende Anatomie der Protozoen, Integument und Skelet der Metazoen. Springer, 401 pp.

<http://books.google.com/books?id=ELRIOwAACAAJ&hl=ca>

A. J. Cain & G. A. Harrison (1960) Phyletic weighting. *Proceedings of the Zoological Society of London* **135**: 1-31. <http://dx.doi.org/10.1111/j.1469-7998.1960.tb05828.x>

N. Calzada, C. H. Lockyer & A. Aguilar (1994) Age and sex composition of the striped dolphin die-off in the western Mediterranean. *MARINE MAMMAL SCIENCE* **10**(3): 299-310. <http://dx.doi.org/10.1111/j.1748-7692.1994.tb00484.x>

N. Calzada & A. Aguilar (1995) Geographical variation of body size in Western Mediterranean striped dolphins (*Stenella coeruleoalba*). *Zeitschrift für Säugetierkunde* **60**(5): 257-264.

http://scholar.google.com/scholar?hl=ca&q=%22Geographical+variation+of+body+size+in+Western+Mediterranean+striped+dolphins+%28Stenella+coeruleoalba%29%22&btnG=Cerca&as_ylo=&as_vis=0

N. Calzada & A. Aguilar (1996) Flipper development in the Mediterranean striped dolphin (*Stenella coeruleoalba*). *Anatomical Record* **245**(4): 708-14. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8837729 [http://dx.doi.org/10.1002/\(SICI\)1097-0185\(199608\)245:4%3C708::AID-AR11%3E3.0.CO;2-T](http://dx.doi.org/10.1002/(SICI)1097-0185(199608)245:4%3C708::AID-AR11%3E3.0.CO;2-T)

R. D. Campbell & J. Trowsdale (1993) Map of the human MHC. *Immunology Today* **14**(7): 349-52.

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8363724 [http://dx.doi.org/10.1016/0167-5699\(93\)90234-C](http://dx.doi.org/10.1016/0167-5699(93)90234-C)

R. Cantin, S. Methot & M. J. Tremblay (2005) Plunder and Stowaways: Incorporation of Cellular Proteins by Enveloped Viruses. *Journal of virology* **79**(11): 6577-6587. <http://dx.doi.org/10.1128/jvi.79.11.6577-6587.2005> <http://dx.doi.org/10.1128/jvi.79.11.6577-6587.2005>

J. Carlsson, J. R. McDowell, P. Diaz-Jaimes, J. E. Carlsson, S. B. Boles, J. R. Gold & J. E. Graves (2004) Microsatellite and mitochondrial DNA analyses of Atlantic bluefin tuna (*Thunnus thynnus thynnus*) population structure in the Mediterranean Sea. *Molecular Ecology* **13**(11): 3345-56. <http://dx.doi.org/10.1111/j.1365-294X.2004.02336.x>

G. Carnevale, W. Landini & G. Sarti (2006) Mare versus Lago-mare: marine fishes and the Mediterranean environment at the end of the Messinian Salinity Crisis. *Journal of the Geological Society* **163**(1): 75-80. <http://jgs.lyellcollection.org/cgi/content/abstract/163/1/75> <http://dx.doi.org/10.1144/0016-764904-158>

D. R. Carrier, S. M. Deban & J. Otterstrom (2002) The face that sank the Essex: potential function of the spermaceti organ in aggression. *Journal of Experimental Biology* **205**(12): 1755-1763. <http://jeb.biologists.org/cgi/content/abstract/205/12/1755>

I. Cassens, S. Vicario, V. G. Waddell, H. Balchowsky, D. Van-Belle, W. Ding, C. Fan, R. S. Mohan, P. C. Simoes-Lopes, R. Bastida, A. Meyer, M. J. Stanhope & M. C. Milinkovitch (2000) Independent adaptation to riverine habitats allowed survival of ancient cetacean lineages. *Proceedings of the National Academy of Sciences of the United States of America* **97**(21): 11343-7. <http://dx.doi.org/10.1073/pnas.97.21.11343>

D. Cebrian (1995) The striped dolphin *Stenella coeruleoalba* epizootic in Greece, 1991-1992. *Biological Conservation* **74**(2): 143-145. [http://dx.doi.org/10.1016/0006-3207\(95\)00024-X](http://dx.doi.org/10.1016/0006-3207(95)00024-X)

F. Cigala-Fulgosi & G. Pilleri (1985) The lower Serravalian cetacean fauna of Visiano (Northern Apennines, Parma, Italy). *Investigations on Cetacea* **17**: 55-93.

G. Clauzon, J.-P. Suc, S.-M. Popescu, M. Marunteanu, J.-L. Rubino, F. Marinescu & M. C. Melinte (2005) Influence of Mediterranean sea-level changes on the Dacic Basin (Eastern Paratethys) during the late Neogene: the Mediterranean Lago Mare facies deciphered. *Basin Research* **17**(3): 437-462. <http://www.blackwell-synergy.com/doi/abs/10.1111/j.1365-2117.2005.00269.x> <http://dx.doi.org/10.1111/j.1365-2117.2005.00269.x>

J. Clutton-Brock (1999) A natural history of domesticated mammals (2nd edition). Cambridge University Press, Cambridge. 238 pp. <http://books.google.fr/books?id=cgL-EbbB8a0C>; <http://catdir.loc.gov/catdir/samples/cam032/98037220.pdf>

R. B. Corley, N. J. LoCascio, M. Ovnicek & G. Houghton (1985) Two separate functions of class II (Ia) molecules: T-cell stimulation and B-cell excitation. *Proceedings of the National Academy of Sciences of the United States of America* **82**(2): 516-20. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=3155862 <http://dx.doi.org/10.1073/pnas.82.2.516>

R. Costanza, R. d'Arge, R. de Groot, S. Farber, M. Grasso, B. Hannon, K. Limburg, S. Naeem, R. V. O'Neill, J. Paruelo, R. G. Raskin, P. Sutton & M. van den Belt (1997) The value of the world's ecosystem services and natural capital. *Nature* **387**(6630): 253-260. <http://dx.doi.org/10.1038/387253a0> <http://dx.doi.org/10.1038/387253a0>

A. M. Crawford, K. G. Dodds & J. C. McEwan (2000) DNA markers, genetic maps and the identification of QTL: general principles in Breeding for disease resistance in farm animals. (S. C. B. R.F.E. Axford, F.W. Nicholas and J.B. Owen Ed.). CAB International, p 3-26. <http://bookshop.cabi.org/Uploads/Books/PDF/9780851993256/9780851993256.pdf> http://books.google.fr/books/about/Breeding_for_disease_resistance_in_farm.html?id=tmiQgAACAAJ&redir_esc=y

G. Cuvier (1812) Sur un nouveau rapprochement à établir entre les classes qui composent le Règne Animal. *Annales du Museum d'Histoire Naturelle* **19**: 73-84. http://books.google.com/books?id=qEgI8rTje60C&pg=PA73&hl=ca&source=gbs_toc_r&cad=0_0

G. Cuvier & A. Valenciennes (1828) Histoire naturelle des poissons. 490 pp. <http://archive.org/details/histoirenaturell111836cuvi>

C. H. Chandler, P. C. Phillips & F. J. Janzen (2009) The evolution of sex-determining mechanisms: lessons from temperature-sensitive mutations in sex determination genes in *Caenorhabditis elegans*. *Journal of evolutionary biology* **22**(1): 192-200. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=19120819 <http://dx.doi.org/10.1111/j.1420-9101.2008.01639.x>

B. Charlesworth (1996) The evolution of chromosomal sex determination and dosage compensation. *Current Biology* **6**(2): 149-62. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8673462 [http://dx.doi.org/10.1016/S0960-9822\(02\)00448-7](http://dx.doi.org/10.1016/S0960-9822(02)00448-7)

P. C. Cheng, M. L. Dykstra, R. N. Mitchell & S. K. Pierce (1999) A role for lipid rafts in B cell antigen receptor signaling and antigen targeting. *Journal of Experimental Medicine* **190**(11): 1549-60. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10587346 <http://dx.doi.org/10.1084/jem.190.11.1549>

S. A. Church & D. R. Taylor (2002) The evolution of reproductive isolation in spatially structured populations. *Evolution* **56**(9): 1859-62. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12389731 [http://dx.doi.org/10.1554/0014-3820\(2002\)056%5B1859:TEORII%5D2.0.CO;2](http://dx.doi.org/10.1554/0014-3820(2002)056%5B1859:TEORII%5D2.0.CO;2)

E. Danchin, V. Vitiello, A. Vienne, O. Richard, P. Gouret, M. F. McDermott & P. Pontarotti (2004) The major histocompatibility complex origin. *Immunological Reviews* **198**: 216-32. <http://dx.doi.org/10.1111/j.0105-2896.2004.00132.x>

E. G. Danchin, L. Abi-Rached, A. Gilles & P. Pontarotti (2003) Conservation of the MHC-like region throughout evolution. *Immunogenetics* **55**(3): 141-8. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12734695 <http://dx.doi.org/10.1111/j.0105-2896.2004.00132.x>

C. Darwin (1859) On the origin of species (*J. Murray Ed.*). John Murray, London. 502 pp. <http://darwin-online.org.uk/contents.html> - origin

T. Davit-Beal, A. S. Tucker & J. Y. Sire (2009) Loss of teeth and enamel in tetrapods: fossil record, genetic data and morphological adaptations. *Journal of anatomy* **214**(4): 477-501. <http://dx.doi.org/10.1111/j.1469-7580.2009.01060.x> <http://dx.doi.org/10.1111/j.1469-7580.2009.01060.x>

C. de Muizon (2001) Walking with whales. *Nature* **413**(6853): 259-260. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=11565010 <http://dx.doi.org/10.1038/35095137>

T. de Oliveira, O. G. Pybus, A. Rambaut, M. Salemi, S. Cassol, M. Ciccozzi, G. Rezza, G. C. Gattinara, R. D'Arrigo, M. Amicosante, L. Perrin, V. Colizzi & C. F. Perno (2006) Molecular epidemiology: HIV-1 and HCV sequences from Libyan outbreak. *Nature* **444**(7121): 836-7. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17171825 <http://dx.doi.org/10.1038/444836a>

S. Delgado, M. Girondot & J. Y. Sire (2005) Molecular evolution of amelogenin in mammals. *Journal of molecular evolution* **60**(1): 12-30. <http://dx.doi.org/10.1007/s00239-003-0070-8>

F. Dhermain, J. M. Bompar, G. Chappuis, M. Folacci & F. Poitevin (1994) The Morbillivirus epidemic of the striped dolphin *Stenella coeruleoalba* in the Mediterranean sea. *Recueil de Medecine Veterinaire (France)* **170**(2-3): 85-92. <http://agris.fao.org/agris-search/search/display.do?f=1995%2FFR%2FFR95008.xml%3BFR9504617>

F. Dhermain (1995-2006) Facteurs de mortalité chez les Cétacés. *Cours E.P.H.E., Sète*,

G. Di Bernardo, S. Del Gaudio, U. Galderisi, A. Cascino & M. Cipollaro (2009) Ancient DNA and Family Relationships in a Pompeian House. *Annals of Human Genetics* **9999**(9999)

G. Di Guardo, G. Marruchella, U. Agrimi & S. Kennedy (2005) Morbillivirus infections in aquatic mammals: a brief overview. *Journal of Veterinary Medicine Series A Physiol Pathol Clin Med* **52**(2): 88-93. <http://dx.doi.org/10.1111/j.1439-0442.2005.00693.x> <http://dx.doi.org/10.1111/j.1439-0442.2005.00693.x>

N. Di Méglio, R. Romero-Alvarez & A. Collet (1996) Growth comparison in striped dolphins, *Stenella coeruleoalba*, from the Atlantic and Mediterranean coasts of France. *Aquatic Mammals* **22**(1): 11-21. http://aquaticmammalsjournal.org/share/AquaticMammalsIssueArchives/1996/AquaticMammals_22-01/22-01_DiMeglio.pdf

E. Dias Neto, F. R. Santos, S. D. Pena & A. J. Simpson (1993) Sex determination by low stringency PCR (LS-PCR). *Nucleic Acids Research* **21**(3): 763-4. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8441697 <http://dx.doi.org/10.1093/nar/21.3.763>

Divers (1999) Cours de mastère "Géosciences et environnements marins".

P. Doherty (1996) Cell Mediated Immunity in Virus Infections. *Nobel Lecture in Medicine*. http://www.nobelprize.org/nobel_prizes/medicine/laureates/1996/doherty-lecture.pdf

M. Domingo, L. Ferrer, M. Pumarola, A. Marco, J. Plana, S. Kennedy, M. McAliskey & B. K. Rima (1990) Morbillivirus in dolphins. *Nature* **348**(6296): 21. <http://dx.doi.org/10.1038/348021a0>

M. Domingo, J. Visa, M. Pumarola, A. J. Marco, L. Ferrer, R. Rabanal & S. Kennedy (1992) Pathologic and immunocytochemical studies of morbillivirus infection in striped dolphins (*Stenella coeruleoalba*). *Veterinary Pathology* **29**(1): 1-10. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=1557861

M. Domingo, M. Vilafranca, J. Visa, N. Prats, A. Trudgett & I. Visser (1995) Evidence for chronic morbillivirus infection in the Mediterranean striped dolphin (*Stenella coeruleoalba*).

Veterinary microbiology 44(2-4): 229-39. [http://dx.doi.org/10.1016/0378-1135\(95\)00016-4](http://dx.doi.org/10.1016/0378-1135(95)00016-4)

C. J. Douady, F. Delsuc, Y. Boucher, W. F. Doolittle & E. J. P. Douzery (2003) Comparison of Bayesian and maximum likelihood bootstrap measures of phylogenetic reliability. *Molecular Biology and Evolution* 20(2): 248-254. <http://dx.doi.org/10.1093/molbev/msg042>

E. Douzery & F. M. Catzeflis (1995) Molecular evolution of the mitochondrial 12S rRNA in Ungulata (mammalia). *Journal of molecular evolution* 41(5): 622-636. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=7490777 <http://dx.doi.org/10.1007/BF00175821>

A. Drummond, R. Forsberg & A. G. Rodrigo (2001) The inference of stepwise changes in substitution rates using serial sequence samples. *Molecular Biology And Evolution* 18(7): 1365-1371. <http://dx.doi.org/10.1093/oxfordjournals.molbev.a003920>

A. J. Drummond, G. K. Nicholls, A. G. Rodrigo & W. Solomon (2002) Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. *Genetics* 161(3): 1307-20. <http://www.ncbi.nlm.nih.gov/pubmed/12136032>

A. J. Drummond, A. Rambaut, B. Shapiro & O. G. Pybus (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular biology and evolution* 22(5): 1185-1192. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15703244 <http://dx.doi.org/10.1093/molbev/msi103>

L. Du Pasquier (2000) The phylogenetic origin of antigen-specific receptors. *Current Topics in Microbiology and Immunology* 248: 160-85. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10793478 <http://www.ncbi.nlm.nih.gov/pubmed/10793478>

S. Duggen, K. Hoernle, P. van-den-Bogaard, L. Rupke & J. P. Morgan (2003) Deep roots of the Messinian salinity crisis. *Nature* 422(6932): 602-6. <http://dx.doi.org/10.1038/nature01553>

P. J. Duignan, J. R. Geraci, J. A. Raga & N. Calzada (1992) Pathology of morbillivirus infection in striped dolphins (*Stenella coeruleoalba*) from Valencia and Murcia, Spain. *Revue Canadienne de Recherche Vétérinaire / Canadian Journal of Veterinary Research* 56(3): 242-8. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1263546/>

P. J. Duignan, C. House, J. R. Geraci, N. Duffy, B. K. Rima, M. T. Walsh, G. Early, D. J. St Aubin, S. Sadove, H. Koopman & et al. (1995) Morbillivirus infection in cetaceans of the western Atlantic. *Veterinary Microbiology* 44(2-4): 241-9. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8588318 [http://dx.doi.org/10.1016/0378-1135\(95\)00017-5](http://dx.doi.org/10.1016/0378-1135(95)00017-5)

J. Dulčić & B. Grbec (2000) Climate change and Adriatic ichthyofauna. *Fisheries Oceanography* 9(2): 187-191.

I. Dupanloup, S. Schneider & L. Excoffier (2002) A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology* 11(12): 2571-2581. <http://dx.doi.org/10.1046/j.1365-294X.2002.01650.x>

C. Dupuis (1984) Willi Henning's Impact on Taxonomic Thought. *Annual Review of Ecology and Systematics* 15(1): 1-25. <http://arjournals.annualreviews.org/doi/abs/10.1146/annurev.es.15.110184.000245> <http://dx.doi.org/doi:10.1146/annurev.es.15.110184.000245>

J. D. Durand, P. G. Bianco, J. Laroche & A. Gilles (2003) Insight Into the Origin of Endemic Mediterranean Ichthyofauna: Phylogeography of Chondrostoma Genus (Teleostei, Cyprinidae). *Journal of Heredity* 94(4): 315-328. <http://jhered.oxfordjournals.org/cgi/content/abstract/94/4/315> <http://dx.doi.org/10.1093/jhered/esg074>

S. Edwards, T. Barrett, B. Liess, P.-P. Pastoret & J. Pohlenz (Editors) (1995a) Morbillivirus

S. V. Edwards, E. K. Wakeland & W. K. Potts (1995b) Contrasting histories of avian and mammalian Mhc genes revealed by class II B sequences from songbirds. *Proceedings of the National Academy of Sciences of the United States of America* 92(26): 12200-4.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8618869 <http://dx.doi.org/10.1073/pnas.92.26.12200>

V. W. Ekman (1906) Ströme und Stromsysteme, die vom Winde unter dem Einfluss der Kontinente verursacht werden. *Annalen der Hydrographie und Maritimen Meteorologie*(527-540)

H. Ellegren, C. J. Davies & L. Andersson (1993) Strong association between polymorphisms in an intronic microsatellite and in the coding sequence of the BoLA-DRB3 gene: implications for microsatellite stability and PCR-based DRB3 typing. *Animal Genetics* 24(4): 269-75.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8239071 <http://dx.doi.org/10.1111/j.1365-2052.1993.tb00310.x>

S. A. Ellis & K. T. Ballingall (1999) Cattle MHC: evolution in action? *Immunological reviews* 167: 159-68. <http://dx.doi.org/10.1111/j.1600-065X.1999.tb01389.x>

C. C. Emig (1989) Distributional patterns along the mediterranean continental margin (Upper Bathyal) using *Gryphus vitreus* (Brachiopoda) densities. *Palaeogeography-palaeoclimatology-palaeoecology* 71(3-4): 253-256. [http://dx.doi.org/10.1016/0031-0182\(89\)90053-9](http://dx.doi.org/10.1016/0031-0182(89)90053-9)

C. C. Emig & P. Geistdoerfer (2004) Faune profonde en Mer Méditerranée : les échanges historiques, géographiques et bathymétriques. *Carnets de Géologie* 3: 1-10.
http://paleopolis.rediris.es/cg/CG2004_A01_CCE-PG/index-fr.html

S. Ennis & T. F. Gallagher (1994) A PCR-based sex-determination assay in cattle based on the bovine amelogenin locus. *Animal Genetics* 25(6): 425-7.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=7695123 <http://dx.doi.org/10.1111/j.1365-2052.1994.tb00533.x>

S. Escorza-Treviño, A. Lang & A. E. Dizon (2002) Genetic differentiation and intraspecific structure of Eastern Tropical Pacific Spotted dolphins, *Stenella attenuata*, revealed by Mitochondrial And Microsatellite DNA Analyses. in A report to South-West Fisheries Science Center, La Jolla, CA, 20 pp. http://instructional1.calstatela.edu/sescorz/Lab/LJ_02_38.pdf

S. Escorza-Treviño, F. Archer, M. Rosales, A. Lang & A. Dizon (2005) Genetic differentiation and intraspecific structure of Eastern Tropical Pacific spotted dolphins, *Stenella attenuata*, revealed by DNA analyses. *Conservation Genetics* 6(4): 587-600.
<http://dx.doi.org/10.1007/s10592-005-9013-9> <http://dx.doi.org/10.1007/s10592-005-9013-9>

J. A. Estes & P. D. Steinberg (1988) Predation, herbivory, and kelp evolution. *Paleobiology* 14(1): 19-36. <http://www.jstor.org/stable/2400895>

J. A. Estes & D. O. Duggins (1995) Sea otters and kelp forests in Alaska : generality and variation in a community ecological paradigm. *Ecological-monographs* 65(1): 75-100.
<http://dx.doi.org/10.2307/2937159>

J. A. Estes, M. T. Tinker, T. M. Williams & D. F. Doak (1998) Killer whale predation on sea otters linking oceanic and nearshore ecosystems. *Science* 282(5388): 473-476.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9774274 <http://dx.doi.org/10.1126/science.282.5388.473>

J. E. Estes, N. S. Smith & J. F. Palmisano (1978) Sea Otter Predation and Community Organization in the Western Aleutian Islands, Alaska. *Ecology* 59(4): 822-833.
<http://www.esajournals.org/doi/abs/10.2307/1938786> <http://dx.doi.org/10.2307/1938786>

European Commission (2007) Climate Change Impacts on the Water Cycle, Resources and Quality. in Research-Policy Interface Scientific and Policy (P. Q. M. Moren-Abat, L. Feyen, A. Heiskanen, P. Noges, A. Solheim, E. Lipiatou Ed.). A report to European Commission, Brussels,

149. http://ec.europa.eu/research/environment/pdf/kina22422ens_web_water_and_cc.pdf
J. Felsenstein (1976) The theoretical population genetics of variable selection and migration. *Annual Reviews in Genetics* **10**: 253-80.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=797310 <http://dx.doi.org/10.1146/annurev.ge.10.120176.001345>
- J. Felsenstein** (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evolution* **17**(6): 368-76. <http://dx.doi.org/10.1007/BF01734359>
- J. Felsenstein** (1989) PHYLIP-Phylogeny Inference Package (version 3.2). *Cladistics* **5**(2): 164-166. <http://fx.doi.org/10.1111/j.1096-0031.1989.tb00562.x>
- J. Felsenstein** (2004a) Theoretical Evolutionary Genetics, syntèse des cours de Génétique Evolutive de Joseph Felsenstein à la Washington University (version 2004). Seattle. 383 pp. <http://evolution.genetics.washington.edu/pgbook/pgbook.html>
- J. Felsenstein** (2004b) Inferring Phylogenies. Sinauer associates, Sunderland, MS. 664 pp. <http://books.google.com/books?id=GV3cHQAACAAJ>
- F. Figueroa, W. E. Mayer, H. Sultmann, C. O'HUigin, H. Tichy, Y. Satta, N. Takezaki, N. Takahata & J. Klein** (2000) Mhc class II B gene evolution in East African cichlid fishes. *Immunogenetics* **51**(7): 556-75. <http://dx.doi.org/10.1007/s002510000181>
- B. B. Finlay & S. Falkow** (1997) Common themes in microbial pathogenicity revisited. *Microbiology and Molecular Biology Reviews* **61**(2): 136-169. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9184008
- R. A. Fisher** (1912) On an absolute criterion for fitting frequency curves. *Messenger of Mathematics* **41**: 155-160. <http://digital.library.adelaide.edu.au/coll/special//fisher/1.pdf>
- R. A. Fisher** (1922) On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society of London Ser. A* **222**: 309-368.
- R. A. Fisher** (1930) The genetical theory of natural selection. Clarendon Press, Oxford. <http://archive.org/details/geneticaltheoryo031631mbp>
- S. Flores Ramirez, R. D. Miller & J. Urban Ramirez** (2004) Major histocompatibility complex I polymorphism in a cetacean: The gray whale (*Eschrichtius robustus*). *Marine Mammal Science* **20**(2): 262-273. <http://dx.doi.org/10.1111/j.1748-7692.2004.tb01155.x>
- W. H. Flower** (1867) Description of the skeleton of *Inia geoffrensis*. *Transactions of the Zoological Society of London* **6**: 87-116. <http://dx.doi.org/10.1111/j.1096-3642.1867.tb00572.x>
- E. H. Follmann, G. W. Garner, J. F. Evermann & A. J. McKeirnan** (1996) Serological evidence of morbillivirus infection in polar bears (*Ursus maritimus*) from Alaska and Russia. *Veterinary Record* **138**(25): 615-8. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8807784 <http://dx.doi.org/10.1136/vr.138.25.615>
- J. Forcada & P. Hammond** (1998) Geographical variation in abundance of striped and common dolphins of the western Mediterranean. *Journal of sea research* **39**(3-4): 313-325. [http://dx.doi.org/10.1016/S1385-1101\(97\)00063-4](http://dx.doi.org/10.1016/S1385-1101(97)00063-4)
- R. E. Fordyce** (1992) Cetacean evolution and Eocene/Oligocene environments in Eocene-Oligocene Climatic and Biotic Evolution. (D. R. Prothero & W. A. Berggren Ed.). Princeton University Press, Princeton, NJ. p 368-381. <http://www.mendeley.com/research/cetacean-evolution-and-eoceneoligocene-environments/>
- R. E. Fordyce & L. G. Barnes** (1994) The evolutionary history of whales and dolphins. *Annual Review of Earth and Planetary Sciences* **22**: 419-455. <http://dx.doi.org/10.1146/annurev.ea.22.050194.002223>
- M. C. Fossi & L. Marsili** (2003) Topic 4.4: Effects of endocrine disruptors in aquatic mammals. *Pure and Applied Chemistry* **75**(11/12): 2235-2248.

<http://dx.doi.org/10.1351/pac200375112235>

P. Francour, C. F. Boudouresque, J. G. Harmelin, M. L. Harmelin-Vivien & J. P. Quignard (1994) Are the Mediterranean waters becoming warmer? Information from biological indicators. *Marine pollution bulletin* **28**(9): 523-526. http://portail.unice.fr/jahia/webdav/site/myjahiasite/users/doly/public/1994_Marine_Pollution_Bulletin.pdf [http://dx.doi.org/10.1016/0025-326X\(94\)90071-X](http://dx.doi.org/10.1016/0025-326X(94)90071-X)

A. Frantzis & D. L. Herzing (2002) Mixed-species associations of striped dolphins (*Stenella coeruleoalba*), short-beaked common dolphins (*Delphinus delphis*), and Risso's dolphins (*Grampus griseus*) in the Gulf of Corinth (Greece, Mediterranean Sea). *Aquatic Mammals* **28**(2) http://www.pelagosinstitute.gr/gr/pelagos/pdfs/Mixed_species.PDF

F. C. Fraser & B. A. Noble (1970) Variation of pigmentation pattern in Meyen's dolphin, *Stenella coeruleoalba* (Meyen). *Investigations on Cetacea* **2**: 147-164.

G. Froeschke & S. Sommer (2005) MHC class II DRB variability and parasite load in the striped mouse (*Rhabdomys pumilio*) in the Southern Kalahari. *Molecular Biology and Evolution* **22**(5): 1254-9. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15703235 <http://dx.doi.org/10.1093/molbev/msi112>

Y. X. Fu (1994) A phylogenetic estimator of effective population size or mutation rate. *Genetics* **136**(2): 685-692. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8150291

D. J. Funk & K. E. Omland (2003) Species-level paraphyly and polyphyly: Frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology and Systematics* **34**: 397-423. <http://dx.doi.org/10.1146/annurev.ecolsys.34.011802.132421>

A. Galov, G. Lauc, N. Nikolić, Z. Šatović, T. Gomerčić, M. Đuras Gomerčić, I. Kocijan, B. Šeol & H. Gomerčić (2008) Records and genetic diversity of striped dolphins (*Stenella coeruleoalba*) from the Croatian coast of the Adriatic Sea. *JMBA2 - Biodiversity Records* Published online(5873): 1-6. <http://dx.doi.org/10.1017/S1755267209000888>

M. D. Ganfornina & D. Sanchez (1999) Generation of evolutionary novelty by functional shift. *Bioessays* **21**(5): 432-9. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10376014 [http://dx.doi.org/10.1002/\(SICI\)1521-1878\(199905\)21:5%3C432::AID-BIES10%3E3.0.CO;2-T](http://dx.doi.org/10.1002/(SICI)1521-1878(199905)21:5%3C432::AID-BIES10%3E3.0.CO;2-T)

A. Gannier (1999) Diel variations of the striped dolphin distribution off the French Riviera (northwestern Mediterranean Sea). *Aquatic Mammals* **3**(25): 123-134. http://www.aquaticmammalsjournal.org/share/AquaticMammalsIssueArchives/1999/AquaticMammals_25-03/25-03_Gannier.pdf

A. Gannier (2005) Summer distribution and relative abundance of delphinids in the Mediterranean Sea. *Revue d'écologie-Terre & Vie* **60**(3): 223-238. http://www.cetaces.org/wp-content/uploads/2009/07/Gannier05_Medelf.pdf

J. García-Martínez, E. Barrio, J. A. Raga & A. Latorre (1995) Mitochondrial DNA variability of striped dolphins (*Stenella coeruleoalba*) in the Spanish Mediterranean waters. *Marine Mammal Science* **11**(2): 183-199. <http://www.blackwell-synergy.com/gate1.inist.fr/doi/abs/10.1111/j.1748-7692.1995.tb00517.x> <http://dx.doi.org/10.1111/j.1748-7692.1995.tb00517.x>

J. García-Martínez, A. Moya, J. A. Raga & A. Latorre (1999) Genetic differentiation in the striped dolphin *Stenella coeruleoalba* from European waters according to mitochondrial DNA (mtDNA) restriction analysis. *Molecular Ecology* **8**(6): 1069-73. <http://dx.doi.org/10.1046/j.1365-294x.1999.00672.x>

S. Gaspari, A. Azzellino, S. Airoidi & A. R. Hoelzel (2007) Social kin associations and genetic structuring of striped dolphin populations (*Stenella coeruleoalba*) in the Mediterranean Sea. *Molecular Ecology* **16**(14): 2922-2933. <http://dx.doi.org/10.1111/j.1365-294X.2007.03295.x>

J. Gatesy, C. Hayashi, M. A. Cronin & P. Arctander (1996) Evidence from milk casein genes that cetaceans are close relatives of hippopotamid artiodactyls. *Molecular biology and evolution* **13**(7): 954-963.

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8752004

J. Gatesy (1997) More DNA support for a Cetacea/Hippopotamidae clade: the blood-clotting protein gene gamma-fibrinogen. *Molecular biology and evolution* **14**(5): 537-543.

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9159931

J. Gatesy, M. Milinkovitch, V. Waddell & M. Stanhope (1999) Stability of cladistic relationships between Cetacea and higher-level artiodactyl taxa. *Systematic biology* **48**(1): 6-20.

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12078645 <http://dx.doi.org/10.1080/106351599260409>

J. Gatesy & M. O' Leary (2001) Deciphering whale origins with molecules and fossils. *Trends in Ecology and Evolution* **16**(10): 562-570. [http://dx.doi.org/10.1016/S0169-5347\(01\)02236-4](http://dx.doi.org/10.1016/S0169-5347(01)02236-4)

S. Gaudieri, R. L. Dawkins, K. Habara, J. K. Kulski & T. Gojobori (2000) SNP profile within the human major histocompatibility complex reveals an extreme and interrupted level of nucleotide diversity. *Genome research* **10**(10): 1579-86. <http://dx.doi.org/10.1101/gr.127200>

J. H. Geisler & M. D. Uhen (2003) Morphological support for a close relationship between Hippos and Whales. *Journal of Vertebrate Paleontology* **23**(4): 991-996. <http://dx.doi.org/10.1671/32>

C. Gibson, E. Golub, R. Herold, M. Risser, W. Ding, H. Shimokawa, M. Young, J. Termine & J. Rosenbloom (1991) Structure and expression of the bovine amelogenin gene. *Biochemistry* **30**(4): 1075-9. <http://dx.doi.org/10.1021/bi00218a028>

J. M. Gili (2001) Biodiversitat Marina Mediterrània. *L'Atzavara* **9**: 25-35. <http://www.icm.csic.es/bio/ftp/atzavara01.pdf>

T. N. Gill (1872) Arrangement of the families of fishes or classes Pisces, Marsipobranchii, and Leptocardii. *Smithsonian Miscellaneous Collections* **247**: 1-49. <http://www.archive.org/details/arrangementfam00gillrich>

A. Gilles, R. Chappaz, L. Cavalli, M. Lortscher & E. Faure (1998) Genetic differentiation and introgression between putative subspecies of *Leuciscus soufia* (Teleostei : Cyprinidae) of the region of the Mediterranean Alps. *Canadian Journal of Fisheries and Aquatic Sciences* **55**(10): 2341-2354. <http://dx.doi.org/10.1139/f98-103>

P. D. Gingerich, M. Haq, I. S. Zalmout, I. H. Khan & M. S. Malkani (2001) Origin of whales from early artiodactyls: hands and feet of Eocene Protocetidae from Pakistan. *Science* **293**(5538): 2239-2242.

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=11567134 <http://dx.doi.org/10.1126/science.1063902>

M. J. Giphat, A. M. Groenen & J. J. van der Poel (1990) Homologies between the major histocompatibility complex of man and cattle: consequences for disease resistance and susceptibility. *The Veterinary Quarterly* **12**: 202-211.

<http://dx.doi.org/10.1080/01652176.1990.9694267>

M. Girondot & J. Y. Sire (1998) Evolution of the amelogenin gene in toothed and toothless vertebrates. *European Journal of Oral Sciences* **106** Suppl 1: 501-8. http://max2.ese.u-psud.fr/epc/conservation/Publi/abstracta/AE_EJOS98.pdf

<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation>

[&list_uids=9541263](#)

Y. Go, Y. Satta, Y. Kawamoto, G. Rakotoarisoa, A. Randrianjafy, N. Koyama & H. Hirai (2002) Mhc-DRB genes evolution in lemurs. *Immunogenetics* **54**(6): 403-17. <http://dx.doi.org/10.1007/s00251-002-0480-6>

J. A. Godoy, J. J. Negro, F. Hiraldo & J. A. Donazar (2004) Phylogeography, genetic structure and diversity in the endangered bearded vulture (*Gypaetus barbatus*, L) as revealed by mitochondrial DNA. *Molecular Ecology* **13**(2): 371-90. <http://dx.doi.org/10.1046/j.1365-294X.2003.02075.x>

D. Golani, L. Orsi Relini, E. Massutí, J.-P. Quignard, B. Galil, C. Frogli, P. Noël, A. Zenetos, S. Gofas, G. Russo, J. Templado, M. Verlaque, S. Ruitton, F. Mineur & C. F. Boudouresque (2002) CIESM Atlas of Exotic Species in the Mediterranean. Vol 1. Fishes (F. Briand Ed.). <http://www.ciesm.org/online/atlas/intro.htm>

J. Goudet (2002) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3.2). <http://www.unil.ch/izea/software/fstat.html>.

S. J. Gould & E. S. Vrba (1982) Exaptation-A Missing Term in the Science of Form. *Paleobiology* **8**(1): 4-15. <http://www.jstor.org/stable/2400563>

D. Graur & D. G. Higgins (1994) Molecular evidence for the inclusion of cetaceans within the order Artiodactyla. *Molecular biology and evolution* **11**(3): 357-364. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8015431

J. A. Graves (1996) Mammals that break the rules: genetics of marsupials and monotremes. *Annual review of genetics* **30**: 233-260. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8982455 <http://dx.doi.org/10.1146/annurev.genet.30.1.233>

J. E. Gray (1821) On the natural arrangement of vertebrate animals. *London Medical Repository* **15**(88): 296-310. http://www.rhinosourcecenter.com/index.php?s=6e1702a37daa82b9189ca8930e506980&act=refs&CODE=ref_detail&id=1165236852

J. E. Gray (1846) On the cetaceous animals. in *The zoology of the voyage of HMS Erebus and Terror under the command of Captain Sir James Clark Ross during the years 1839 to 1843.* (J. R. a. J. E. Gray Ed.). London. p 13-53. http://books.google.fr/books/about/On_the_cetaceous_animals.html?id=TmcQHAAACAAJ&redir_esc=y

J. E. Gray (1866) Catalogue of seals and whales in the British Museum. Order of the Trustees, British Museum (Natural History). in *A report to London*, 402. http://books.google.fr/books/about/Catalogue_of_seals_and_whales_in_the_Bri.html?id=fkEAAAAAQAAJ&redir_esc=y

K. Grobben (1908) Beiträge zur Kenntnis des Baues und der systematischen Stellung der Arguliden. *Königlich Preussische Akademie der Wissenschaften zu Berlin* **117**: 191-233. <http://www.worldcat.org/title/beitrage-zur-kenntnis-des-baues-und-der-systematischen-stellung-der-arguliden/oclc/63603724>

S. W. Guo & E. A. Thompson (1992) Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics* **48**(2): 361-72. <http://dx.doi.org/10.2307/2532296>

G. A. Gutierrez-Espeleta, P. W. Hedrick, S. T. Kalinowski, D. Garrigan & W. M. Boyce (2001) Is the decline of desert bighorn sheep from infectious disease the result of low MHC variation? *Heredity* **86**(Pt 4): 439-50. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=11520344 <http://dx.doi.org/10.1046/j.1365-2540.2001.00853.x>

L. Gyax (2002) Evolution of group size in the superfamily Delphinoidea (Delphinidae,

Phocoenidae and Monodontidae): a quantitative comparative analysis. *Behavioral Ecology* **13**(5): 583–590. <http://dx.doi.org/10.1093/beheco/13.5.583>

E. S. Gysels, B. Hellemans, T. Patarnello & F. A. M. Volckaert (2004) Current and historic gene flow of the sand goby *Pomatoschistus minutus* on the European continental shelf and in the Mediterranean Sea. *Biological Journal of the Linnean Society* **83**(4): 561-576. <http://dx.doi.org/10.1111/j.1095-8312.2004.00411.x> <http://dx.doi.org/10.1111/j.1095-8312.2004.00411.x>

E. Haeckel (1866) *Generelle Morphologie der Organismen: allgemeine Grundzüge der organischen Formen-Wissenschaft, mechanisch begründet durch die von Charles Darwin reformirte Descendenz-Theorie.* G. Reimer, Berlin. 461 pp. <http://books.google.com/books?id=-pk5AAAAcAAJ&hl=ca>

E. Haeckel (1874) Die Gastraea-Theorie, die phylogenetische Classification des Thierreichs und die Homologie der Keimblätter. *Jenaische Zeitschrift Für Naturwissenschaft* **8**: 1-55. <http://books.google.com/books?id=87UUAAAAYAAJ&hl=ca>

J. B. S. Haldane (1949) Disease and evolution. *La Ricerca Scientifica* **19**(suppl.): 68–76. http://dx.doi.org/10.1007/0-387-28295-5_9

W. D. Hamilton (1963) The Evolution of Altruistic Behavior. *The American Naturalist* **97**(896): 354-356. <http://www.jstor.org/stable/2458473> <http://dx.doi.org/10.1086/497114>

M. P. Hare, F. Cipriano & S. R. Palumbi (2002) Genetic Evidence On The Demography Of Speciation In Allopatric Dolphin Species. *Evolution* **56**(4): 804-816. [http://dx.doi.org/10.1554/0014-3820\(2002\)056%5B0804:GEOTDO%5D2.0.CO;2](http://dx.doi.org/10.1554/0014-3820(2002)056%5B0804:GEOTDO%5D2.0.CO;2)

R. Harf & S. Sommer (2005) Association between major histocompatibility complex class II DRB alleles and parasite load in the hairy-footed gerbil, *Gerbillurus paeba*, in the southern Kalahari. *Molecular Ecology* **14**(1): 85-91. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15643953 <http://dx.doi.org/10.1111/j.1365-294X.2004.02402.x>

T. Härkönen, R. Dietz, P. Reijnders, J. Teilmann, K. Harding, A. Hall, S. Brasseur, U. Siebert, S. J. Goodman, P. D. Jepson, T. Dau Rasmussen & P. Thompson (2006) The 1988 and 2002 phocine distemper virus epidemics in European harbour seals. *Diseases of Aquatic Organisms* **68**(2): 115-30. <http://www.int-res.com/articles/dao2005/68/d068p115.pdf> http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=16532603

J. A. Harton, A. E. Van Hagen & G. A. Bishop (1995) The cytoplasmic and transmembrane domains of MHC class II beta chains deliver distinct signals required for MHC class II-mediated B cell activation. *Immunity* **3**(3): 349-58. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=7552999 [http://dx.doi.org/10.1016/1074-7613\(95\)90119-1](http://dx.doi.org/10.1016/1074-7613(95)90119-1)

M. Hasegawa, H. Kishino & T. Yano (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution* **22**(2): 160-74. <http://dx.doi.org/10.1007/BF02101694>

W. K. Hastings (1970) Monte Carlo Sampling Methods Using Markov Chains and Their Applications. *Biometrika* **57**(1): 97-109. <http://dx.doi.org/10.2307/2334940> <http://dx.doi.org/10.1093/biomet/57.1.97>

L. Hauser, G. J. Adcock, P. J. Smith, J. H. Ramirez & G. R. Carvalho (2002) Loss of microsatellite diversity and low effective population size in an overexploited population of New Zealand snapper (*Pagrus auratus*). *Proceedings of the National Academy of Sciences of the United States of America* **99**(18): 11742-11747. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12185245 <http://dx.doi.org/10.1073/pnas.172242899>

K. Hayashi, H. Yoshida, S. Nishida, M. Goto, L. A. Pastene, N. Kanda, Y. Baba & H.

Koike (2006) Genetic Variation of the MHC DQB Locus in the Finless Porpoise (*Neophocaena phocaenoides*). *Zoological science* **23**(2): 147-53. <http://dx.doi.org/10.2108/zsj.23.147>

P. Hedrick (1998) Balancing selection and MHC. *Genetica* **104**(3): 207-214. <http://www.springerlink.com/openurl.asp?genre=article&id=doi:10.1023/A:1026494212540>
<http://dx.doi.org/10.1023/A:1026494212540>

P. W. Hedrick, K. M. Parker, E. L. Miller & P. S. Miller (1999) Major Histocompatibility Complex Variation in the Endangered Przewalski's Horse. *Genetics* **152**(4): 1701-1710. <http://www.genetics.org/cgi/content/abstract/152/4/1701>

P. W. Hedrick, K. M. Parker, G. A. Gutierrez Espeleta, A. Rattink & K. Lievers (2000) Major histocompatibility complex variation in the Arabian oryx. *Evolution* **54**(6): 2145-2151. <http://dx.doi.org/10.1111%2Fj.0014-3820.2000.tb01256.x>

W. Hennig (1965) Phylogenetic Systematics. *Annual Review of Entomology* **10**(1): 97-116. <http://arjournals.annualreviews.org/doi/abs/10.1146/annurev.en.10.010165.000525>
<http://dx.doi.org/10.1146/annurev.en.10.010165.000525>

S. L. Herch & D. A. Duffield (1990) Distinction between northwest Atlantic offshore and coastal bottlenose dolphins based on hemoglobin profile and morphology in The Bottlenose Dolphin. (S. L. a. R. Reeves Ed.). Academic Press, New York. p 129-142.

L. M. Herman (1988) Cetacean behavior : mechanisms and functions (Wiley edition)(L. M. Herman & W. N. Tavolga Ed.). Wiley, New York. 463 pp. http://books.google.fr/books/about/Cetacean_behavior.html?id=nx0XAQAIAAJ&redir_esc=y

M. Hernandez, I. Robinson, A. Aguilar, L. M. Gonzalez, L. F. Lopez-Jurado, M. I. Reyero, E. Cacho, J. Franco, V. Lopez-Rodas & E. Costas (1998) Did algal toxins cause monk seal mortality? *Nature* **393**(6680): 28-9. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9590687 <http://dx.doi.org/10.1038/29906>

G. M. Hewitt (2001) Speciation, hybrid zones and phylogeography - or seeing genes in space and time. *Molecular ecology* **10**(3): 537-49. <http://dx.doi.org/10.1046/j.1365-294x.2001.01202.x>

J. E. Heyning (1997) Sperm whale phylogeny revisited : Analysis of the morphological evidence. *Marine Mammal Science* **13**(4): 596-613. <http://dx.doi.org/10.1111/j.1748-7692.1997.tb00086.x>

S. Hiendleder, B. Kaupe, R. Wassmuth & A. Janke (2002) Molecular analysis of wild and domestic sheep questions current nomenclature and provides evidence for domestication from two different subspecies. *Proceedings of the Royal Society B: Biological Sciences* **269**(1494): 893-904. <http://dx.doi.org/10.1098/rspb.2002.1975>

A. R. Hoelzel & G. A. Dover (1991) Genetic differentiation between sympatric Killer whale populations. *Heredity* **66**(p.2): 191-195. <http://dx.doi.org/10.1038/hdy.1991.24>

A. R. Hoelzel, J. M. Hancock & G. A. Dover (1991) Evolution of the cetacean mitochondrial D-loop region. *Molecular biology and evolution* **8**(4): 475-93. <http://mbe.oxfordjournals.org/content/8/4/475.full.pdf>

A. R. Hoelzel, C. W. Potter & P. B. Best (1998) Genetic differentiation between parapatric 'nearshore' and 'offshore' populations of the bottlenose dolphin. *Proceedings of the Royal Society B: Biological Sciences* **265**(1402): 1177-1183. <http://dx.doi.org/10.1098/rspb.1998.0416>

C. M. Holmlund & M. Hammer (1999) Ecosystem services generated by fish populations. *Ecological Economics* **29**(2): 253-268. [http://dx.doi.org/10.1016/S0921-8009\(99\)00015-4](http://dx.doi.org/10.1016/S0921-8009(99)00015-4)
[http://dx.doi.org/10.1016/S0921-8009\(99\)00015-4](http://dx.doi.org/10.1016/S0921-8009(99)00015-4)

K. J. Hsü, W. B. Ryan & M. B. Cita (1973) Late miocene desiccation of the mediterranean. *Nature* **242**(5395): 240-4. <http://dx.doi.org/10.1038/242240a0>

- C. C. Hu, J. D. Bartlett, C. H. Zhang, Q. Qian, O. H. Ryu & J. P. Simmer (1996) Cloning, cDNA sequence, and alternative splicing of porcine amelogenin mRNAs. *Journal of Dental Research* **75**(10): 1735-41.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8955667 <http://dx.doi.org/10.1177/00220345960750100501>
- R. R. Hudson (1987) Estimating the recombination parameter of a finite population model without selection. *Genetic Research* **50**(3): 245-50.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=3443297 <http://dx.doi.org/10.1017/S0016672300023776>
- G. E. Hutchinson (1957) Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology* **22**(2): 415-427. <http://dx.doi.org/10.1101/SQB.1957.022.01.039>
<http://dx.doi.org/10.1101/SQB.1957.022.01.039>
- J. Huxley (1957) The Three Types of Evolutionary Process. *Nature* **180**(4584): 454-455.
<http://dx.doi.org/10.1038/180454a0> <http://dx.doi.org/10.1038/180454a0>
- N. J. B. Isaac, S. T. Turvey, B. Collen, C. Waterman & J. E. Baillie (2007) Mammals on the EDGE: Conservation Priorities Based on Threat and Phylogeny. *PLoS ONE* **2**(3): e296.
<http://dx.doi.org/10.1371/journal.pone.0000296>
- IUCN Conserving marine biodiversity in Mediterranean Marine Programme. 3 pp.
http://cmsdata.iucn.org/downloads/marine_conservation_lr.pdf
- C. S. G. IUCN (1996) *Stenella coeruleoalba*. 2004 IUCN Red list of threatened species. Downloaded from <http://www.redlist.org> on 24 June 2005
- M. Iwase, Y. Satta, Y. Hirai, H. Hirai, H. Imai & N. Takahata (2003) The amelogenin loci span an ancient pseudoautosomal boundary in diverse mammalian species. *Proceedings of the National Academy of Sciences of the United States of America* **100**(9): 5258-5263.
<http://dx.doi.org/10.1073/pnas.0931310100> <http://dx.doi.org/10.1073/pnas.0931310100>
- F. Jacob (1977) Evolution and tinkering. *Science* **196**(4295): 1161-6.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=860134 <http://dx.doi.org/10.1126/science.860134>
- D. Jaume & G. A. Boxshall (1996) The persistence of an ancient marine fauna in Mediterranean waters: new evidence from misophrioid copepods living in anchihaline caves. *Journal of Natural History* **30**(11): 1583-1595. <http://dx.doi.org/10.1080/00222939600770921>
<http://dx.doi.org/10.1080/00222939600770921>
- T. A. Jefferson & B. E. Curry (2003) *Stenella clymene*. *Mammalian Species* **726**: 1-5.
<http://dx.doi.org/10.1644/726>
- K. Jegalian & D. C. Page (1998) A proposed path by which genes common to mammalian X and Y chromosomes evolve to become X inactivated. *Nature* **394**(6695): 776-780.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9723615 <http://dx.doi.org/10.1038/29522>
- F. M. Jiggins & G. D. Hurst (2003) The evolution of parasite recognition genes in the innate immune system: purifying selection on *Drosophila melanogaster* peptidoglycan recognition proteins. *Journal of molecular evolution* **57**(5): 598-605.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=14738318 <http://dx.doi.org/10.1007/s00239-003-2506-6>
- W. L. Johannsen (1903) Über Erblichkeit in Populationen und in reinen Linien. Ein Beitrag zur Beleuchtung schwebender Selektionsfragen (G. Fischer Ed.). Jena, Deutschland.
- W. L. Johannsen (1908) Über Knospenmutation bei *Phaseolus*. *Molecular and General Genetics* **MGG** **1**(1): 1-10. <http://dx.doi.org/10.1007/BF01990553>
<http://dx.doi.org/10.1007/BF01990553>
- T. H. Jukes & C. R. Cantor (1969) Evolution of protein molecules in Mammalian Protein Metabolism. (H. N. Munro Ed.). Academic Press, New York. p 21-132. .

K. Kannan, S. Tanabe, A. Borrell, A. Aguilar, S. Focardi & R. Tatsukawa (1993a) Isomer-specific analysis and toxic evaluation of polychlorinated biphenyls in striped dolphins affected by an epizootic in the western Mediterranean sea. *Archives of Environmental Contamination & Toxicology* **25**(2): 227-33. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8368865 <http://dx.doi.org/10.1007/BF00212134>

K. Kannan, S. Tanabe, A. Borrell, A. Aguilar, S. Focardi & R. Tatsukawa (1993b) Isomer-specific analysis and toxic evaluation of polychlorinated biphenyls in striped dolphins affected by an epizootic in the western Mediterranean sea. *Archives of environmental contamination and toxicology* **25**(2): 227-33.

S. Karuppiyah, A. Subramanian & J. P. Obbard (2005) Organochlorine residues in odontocete species from the southeast coast of India. *Chemosphere* **60**(7): 891-897. <http://dx.doi.org/10.1016/j.chemosphere.2005.01.013>

M. Kasahara, J. Nakaya, Y. Satta & N. Takahata (1997) Chromosomal duplication and the emergence of the adaptive immune system. *Trends in Genetics* **13**(3): 90-2. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9066265 [http://dx.doi.org/10.1016/S0168-9525\(97\)01065-2](http://dx.doi.org/10.1016/S0168-9525(97)01065-2)

T. Kasuya (1973) Systematic consideration of recent toothed whales based on the morphology of tympano-periotic bone. *Scientific Reports of the Whales Research Institute* **25**: 1-103.

T. Kasuya (1976) Reconsideration of life history parameters of the spotted and striped dolphins based on cemental layers. *Scientific Reports of the Whales Research Institute, Tokyo* **28**: 73-105.

T. Kasuya (1985) Effect of exploitation on reproductive parameters of the spotted and striped dolphins off the Pacific Coast of Japan. *Scientific Reports of the Whales Research Institute, Tokyo* **36**: 107-138.

R. Kaustuv, D. Jablonski & J. W. Valentine (2001) Climate change, species range limits and body size in marine bivalves. *Ecology Letters* **4**(4): 366-370. <http://www.blackwell-synergy.com/links/doi/10.1046/j.1461-0248.2001.00236.x> <http://dx.doi.org/10.1046/j.1461-0248.2001.00236.x>

M. Kimura (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution* **16**(2): 111-120.

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=7463489 <http://dx.doi.org/10.1007/BF01731581>

J. F. C. Kingman (1982a) The coalescent. *Stochastic Processes and their Applications* **13**(3): 235-248. [http://dx.doi.org/10.1016/0304-4149\(82\)90011-4](http://dx.doi.org/10.1016/0304-4149(82)90011-4)

J. F. C. Kingman (1982b) On the Genealogy of Large Populations. *Journal of Applied Probability* **19**: 27-43. <http://dx.doi.org/10.2307/3213548>

H. Kishino, J. L. Thorne & W. J. Bruno (2001) Performance of a divergence time estimation method under a probabilistic model of rate evolution. *Molecular Biology and Evolution* **18**(3): 352-361. <http://dx.doi.org/10.1093/oxfordjournals.molbev.a003811>

Y. Kitazoe, H. Kishino, P. J. Waddell, N. Nakajima, T. Okabayashi, T. Watabe & Y. Okuhara (2007) Robust time estimation reconciles views of the antiquity of placental mammals. *PLoS ONE* **2**: e384. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17440620 <http://dx.doi.org/10.1371/journal.pone.0000384>

J. Klein & F. Figueroa (1986) Evolution of the major histocompatibility complex. *Critical reviews in immunology* **6**(4): 295-386. <http://www.ncbi.nlm.nih.gov/pubmed/3536303>

J. Klein, R. E. Bontrop, R. L. Dawkins, H. A. Erlich, U. B. Gyllensten, E. R. Heise, P. P. Jones, P. Parham, E. K. Wakeland & D. I. Watkins (1990) Nomenclature for the major

histocompatibility complexes of different species: a proposal. *Immunogenetics* **31**(4): 217-9. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=2329006 <http://dx.doi.org/10.1007/BF00204890>

J. Klein & C. O'HUigin (1993) Composite origin of major histocompatibility complex genes. *Current Opinion in Genetics & Development* **3**(6): 923-30. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8118219

J. Klein & A. Sato (2000) Advances in immunology - The HLA system - First of two parts. *New England Journal Of Medicine* **343**(10): 702-709. <http://dx.doi.org/10.1056/NEJM200009073431006>

J. Klein, A. Sato & W. E. Mayer (2000) Jaws and AIS in Major Histocompatibility Complex: Evolution, Structure, and Function. (K. M Ed.). Springer, Tokyo. p 3-26.

L. A. Knapp, L. F. Cadavid, M. E. Eberle, S. J. Knechtle, R. E. Bontrop & D. I. Watkins (1997) Identification of new mamu-DRB alleles using DGGE and direct sequencing. *Immunogenetics* **45**(3): 171-9. <http://dx.doi.org/10.1007/s002510050186>

M. Kohn, F. Knauer, A. Stoffella, W. Schroder & S. Paabo (1995) Conservation genetics of the European brown bear--a study using excremental PCR of nuclear and mitochondrial sequences. *Molecular Ecology* **4**(1): 95-103. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=7711958 <http://dx.doi.org/10.1111/j.1365-294X.1995.tb00196.x>

D. Kordis, N. Lovsin & F. Gubensek (2006) Phylogenomic Analysis of the L1 Retrotransposons in Deuterostomia. *Syst Biol* **55**(6): 886-901. <http://dx.doi.org/10.1080/10635150601052637> <http://dx.doi.org/10.1080/10635150601052637>

S. L. Kosakovsky Pond, S. D. Frost & S. V. Muse (2005) HyPhy: hypothesis testing using phylogenies. *Bioinformatics Oxford, England* **21**(5): 676-9. <http://dx.doi.org/10.1093/bioinformatics/bti079>

K. Kriener, C. O'HUigin, H. Tichy & J. Klein (2000) Convergent evolution of major histocompatibility complex molecules in humans and New World monkeys. *Immunogenetics* **51**(3): 169-78. <http://dx.doi.org/10.1007/s002510050028>

W. Krijgsman, F. J. Hiigeni, I. Raffi, F. J. Sierro & D. S. Wilson (1999) Chronology, causes and progression of the Messinian salinity crisis. *Nature* **400**(6745): 652-655. <http://dx.doi.org/10.1038/23231>

B. T. Lahn & D. C. Page (1999) Four evolutionary strata on the human X chromosome. *Science* **286**(5441): 964-7. <http://dx.doi.org/10.1126/science.286.5441.964>

B. T. Lahn, N. M. Pearson & K. Jegalian (2001) The human Y chromosome, in the light of evolution. *Nature Reviews Genetics* **2**(3): 207-216. <http://dx.doi.org/10.1038/35056058> <http://dx.doi.org/10.1038/35056058>

T. Laitinen, S. Koskimies & P. Westman (1993a) Foeto-maternal compatibility in HLA-DR, -DQ, and -DP loci in Finnish couples suffering from recurrent spontaneous abortions. *European Journal of Immunogenetics, official journal of the British Society for Histocompatibility and Immunogenetics* **20**(4): 249-58.

T. Laitinen, S. Koskimies & P. Westman (1993b) Foeto-maternal compatibility in HLA-DR, -DQ, and -DP loci in Finnish couples suffering from recurrent spontaneous abortions. *European journal of immunogenetics* **20**(4): 249-58. <http://dx.doi.org/10.1111/j.1744-313X.1993.tb00140.x>

D. M. Lambert, P. A. Ritchie, C. D. Millar, B. Holland, A. J. Drummond & C. Baroni (2002) Rates of evolution in ancient DNA from Adelie penguins. *Science* **295**(5563): 2270-2273. <http://dx.doi.org/10.1126/science.1068105>

L. A. Lampson & R. Levy (1980) Two populations of Ia-like molecules on a human B cell line. *Journal of Immunology* **125**(1): 293-9.

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=6966655

P. M. Lavoie, J. Thibodeau, F. Erard & R. P. Sekaly (1999) Understanding the mechanism of action of bacterial superantigens from a decade of research. *Immunological Reviews* **168**: 257-69.

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10399079 <http://dx.doi.org/10.1111/j.1600-065X.1999.tb01297.x>

S. K. Lawrence, R. Srivastava, B. Rigas, M. J. Chorney, G. A. Gillespie, C. L. Smith, C. R. Cantor, F. S. Collins & S. M. Weissman (1986) Molecular approaches to the characterization of megabase regions of DNA: applications to the human major histocompatibility complex. *Cold Spring Harbor Symposium on Quantitative Biology* **51 Pt 1**: 123-30.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=3472708 <http://dx.doi.org/10.1101/SQB.1986.051.01.015>

S. Leatherwood & R. R. Reeves (1983) The Sierra Club handbook of whales and dolphins. Sierra Club Books San Francisco,
http://books.google.fr/books/about/The_Sierra_Club_handbook_of_whales_and_d.html?id=Kx4XAQAIAAJ&redir_esc=y

R. G. Leduc, W. F. Perrin & A. E. Dizon (1999) Phylogenetic relationships among the delphinid cetaceans based on full cytochrome B sequences. *Marine Mammal Science* **15**(3): 619-648. <http://dx.doi.org/10.1111/j.1748-7692.1999.tb00833.x>

P. Legendre & L. Legendre (1998) Numerical ecology in *Developments in Environmental Modelling*. Elsevier, Amsterdam.
[http://books.google.fr/books?id=KBoHuoNRO5MC&lpg=PP2&ots=9uZ3S7mAf0&dq=Numerical ecology in Developments in Environmental Modelling&hl=ca&pg=PP2 - v=onepage&q=Numerical%20ecology%20in%20Developments%20in%20Environmental%20Modelling&f=false](http://books.google.fr/books?id=KBoHuoNRO5MC&lpg=PP2&ots=9uZ3S7mAf0&dq=Numerical%20ecology%20in%20Developments%20in%20Environmental%20Modelling&hl=ca&pg=PP2-v=onepage&q=Numerical%20ecology%20in%20Developments%20in%20Environmental%20Modelling&f=false)

C. L veill , R. Al-Daccak & W. Mourad (1999a) CD20 is physically and functionally coupled to MHC class II and CD40 on human B cell lines. *European Journal of Immunology* **29**(1): 65-74. [http://dx.doi.org/10.1002/\(SICI\)1521-4141\(199901\)29:01%3C65::AID-IMMU65%3E3.0.CO;2-E](http://dx.doi.org/10.1002/(SICI)1521-4141(199901)29:01%3C65::AID-IMMU65%3E3.0.CO;2-E)

C. L veill , F. Chandad, R. Al-Daccak & W. Mourad (1999b) CD40 associates with the MHC class II molecules on human B cells. *European Journal of Immunology* **29**(11): 3516-26. [http://dx.doi.org/10.1002/\(SICI\)1521-4141\(199911\)29:11%3C3516::AID-IMMU3516%3E3.0.CO;2-D](http://dx.doi.org/10.1002/(SICI)1521-4141(199911)29:11%3C3516::AID-IMMU3516%3E3.0.CO;2-D)

B. S. Lieberman (2000) Paleobiogeography: Using Fossils to Study Global Change, Plate Tectonics, and Evolution in *Topics in Geobiology (N. H. L. a. D. S. Jones Ed.)*. Kluwer Academic/Plenum Publishers, New York. 208 pp.
http://books.google.com/books?id=fMCsfquuCrYC&pg=PR3&ots=Lbe59WhIuh&dq=paleobiogeography&sig=sPSxAKdgcMn_wizb6L2BTipnY-4

C. v. Linn us (1758) *Systema natur e per regna tria natur e, secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis*. Tomus I. Editio decima, reformata. Holmi e (Laurenti Salvii), Stockholm. 824 pp.
<http://www.biodiversitylibrary.org/item/15373>

T. P. Lipscomb, F. Y. Schulman, D. Moffett & S. Kennedy (1994) Morbilliviral disease in Atlantic bottlenose dolphins (*Tursiops truncatus*) from the 1987-1988 epizootic. *Journal of Wildlife Diseases* **30**(4): 567-71.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=7760492

A. Lowe, S. Harris & P. Ashton (2004) *Ecological Genetics: Design, Analysis, and Application*. Blackwell Publishing, Malden, MA. 164 pp.

<http://eu.wiley.com/WileyCDA/WileyTitle/productCd-1405100338,subjectCd-LS2K.html>

D. Lukas, B. J. Bradley, A. M. Nsubuga, D. Doran-Sheehy, M. M. Robbins & L. Vigilant (2004) Major histocompatibility complex and microsatellite variation in two populations of wild gorillas. *Molecular Ecology* **13**(11): 3389-402. <http://dx.doi.org/10.1111/j.1365-294X.2004.02353.x>

Z. Luo (2000) In search of the whales' sisters. *Nature* **404**(6775): 235-237. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10749193 <http://dx.doi.org/10.1038/35005194>

D. Lusseau (2003) The emergence of cetaceans: phylogenetic analysis of male social behaviour supports the Cetartiodactyla clade. *Journal of evolutionary biology* **16**(3): 531-535. <http://dx.doi.org/10.1046/j.1420-9101.2003.00541.x> <http://dx.doi.org/10.1046/j.1420-9101.2003.00541.x>

E. Lyautey, B. Lacoste, L. Ten-Hage, J.-L. Rols & F. Garabetian (2005) Analysis of bacterial diversity in river biofilms using 16S rDNA PCR-DGGE: methodological settings and fingerprints interpretation. *Water Research* **39**(2-3): 380-388. <http://dx.doi.org/10.1016/j.watres.2004.09.025>

L. Lybarger, X. Wang, M. Harris & T. H. Hansen (2005) Viral immune evasion molecules attack the ER peptide-loading complex and exploit ER-associated degradation pathways. *Current Opinions in Immunology* **17**(1): 71-8. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15653314 <http://dx.doi.org/10.1016/j.coi.2004.11.009>

J. Leonart & F. Maynou (2003) Fish stock assessments in the Mediterranean: state of the art. *Scientia Marina* **67**: 37-49. <http://www.vliz.be/imisdocs/publications/67464.pdf>

G. M. Mace, H. Masundire, J. Baillie, T. Ricketts, T. Brooks, M. Hoffmann & e. al (2005) Biodiversity in Millennium Ecosystem Assessment, 2005. Current State and Trends: Findings of the Condition and Trends Working Group. Ecosystems and human well-being (Vol. 1). Island Press, Washington, DC.

M. Macé, P.-O. Antoine, F. Duranthon, C. de Muizon & Y. Lignereux (1999) Description and systematics of a fossil Odontocete skull from Peruvian Miocene. *European Cetacean Society Annual Conference*.

M. Macé (2003) Did the Gray whale, *Eschrichtius robustus*, calve in the Mediterranean? *Lattara* **16**: 153-164. <http://www.daphne.cnrs.fr/daphne/viewNotice.html%3Bjsessionid=3C5507119E0391ED07AB094A8D106ABE?daphneIdentifier=FRQCCI-227498&sort=dateSort&dir=asc&page=0&maxPages=3&searchIndex=16&lastPage=false&nbResults=123&mode=full>

M. Macé & B. Crouau-Roy (2008) A highly polymorphic insertion in the Y-chromosome amelogenin gene can be used for evolutionary biology, population genetics and sexing in Cetacea and Artiodactyla. *BMC Genetics* **9**: 64. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=18925953 <http://dx.doi.org/10.1186/1471-2156-9-64>

M. Macé, V. Bourret & B. Crouau-Roy (submitted to PLoS ONE) Polyphyly in the Mediterranean Striped dolphin (*Stenella coeruleoalba*) and the Messinian Salinity Crisis.

M. Macé, V. Bourret & B. Crouau-Roy (under revision) Polyphyly in the Mediterranean Striped dolphin (*Stenella coeruleoalba*) and the Messinian Salinity Crisis. *PLoS ONE*

M. Maldonado & M. J. Uriz (1995) Biotic Affinities in a Transitional Zone Between the Atlantic and the Mediterranean: A Biogeographical Approach Based on Sponges. *Journal of biogeography* **22**(1): 89-110. <http://dx.doi.org/10.2307/2846075> <http://dx.doi.org/10.2307/2846075>

T. Malthus (1798) An Essay on the Principle of Population. J. Johnson, London. 134 pp. <http://www.esp.org/books/malthus/population/malthus.pdf>

- J. Mallet** (2005) Hybridization as an invasion of the genome. *Trends in Ecology and Evolution* **20**(5): 229-237. <http://dx.doi.org/10.1016/j.tree.2005.02.010>
- J. Mallet** (2007) Hybrid speciation. *Nature* **446**(7133): 279-283. <http://dx.doi.org/10.1038/nature05706> <http://dx.doi.org/10.1038/nature05706>
- D. Marchessaux** (1980) A review of the current knowledge of the cetaceans in the eastern Mediterranean Sea. *Vie marine* **2**: 59-66.
- J. T. Martinsohn, A. B. Sousa, L. A. Guethlein & J. C. Howard** (1999) The gene conversion hypothesis of MHC evolution: a review. *Immunogenetics* **50**(3-4): 168-200. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10602879 <http://dx.doi.org/10.1007/s002510050593>
- J. P. Masly, C. D. Jones, M. A. Noor, J. Locke & H. A. Orr** (2006) Gene transposition as a cause of hybrid sterility in *Drosophila*. *Science* **313**(5792): 1448-1450. <http://dx.doi.org/10.1126/science.1128721> <http://dx.doi.org/10.1126/science.1128721>
- K. Matsubara, H. Tarui, M. Toriba, K. Yamada, C. Nishida-Umehara, K. Agata & Y. Matsuda** (2006) Evidence for different origin of sex chromosomes in snakes, birds, and mammals and step-wise differentiation of snake sex chromosomes. *Proceedings of the National Academy of Sciences of the United States of America* **103**(48): 18190-18195. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17110446 <http://dx.doi.org/10.1073/pnas.0605274103>
- E. Mayr** (1942) Systematics and the origin of species, from the viewpoint of a zoologist. Columbia University Press, New York. 334 pp. <http://books.google.com/books?id=s2EGAAAAMAAJ&hl=ca>
- E. Mayr** (1965) Classification and Phylogeny. *American Zoologists* **5**(1): 165-174. <http://icb.oxfordjournals.org> <http://dx.doi.org/10.1093/icb/5.1.165>
- H. McDevitt** (1995) Evolution of MHC class II allelic diversity. *Immunological Reviews* **143**: 113-22. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=7558072 <http://dx.doi.org/10.1111/j.1600-065X.1995.tb00672.x>
- S. A. McLeod, F. C. Whitmore Jr & L. G. Barnes** (1993) Evolutionary relationships and classification in *The Bowhead Whale*. (J. J. Burns, J. J. Montague & C. J. Cowles Ed.). Society for Marine Mammalogy Special Publication, Lawrence, KA. p 45-70.
- J. G. Mead & R. L. J. Brownell** (1993) Order Cetacea in *Mammal Species of the World*. (D. E. Wilson & D. M. Reeder Ed.). Smithsonian Institution Press, Washington, DC. p 349-364. <http://books.google.com/books?id=jgAMbNSt8ikC&lpg=PP1&hl=ca&pg=PA723>
- K. Mehindate, J. Thibodeau, M. Dohlsten, T. Kalland, R. P. Sekaly & W. Mourad** (1995) Cross-linking of major histocompatibility complex class II molecules by staphylococcal enterotoxin A superantigen is a requirement for inflammatory cytokine gene expression. *Journal of Experimental Medicine* **182**(5): 1573-7. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=7595227 <http://dx.doi.org/10.1084/jem.182.5.1573>
- C. Meotti & M. Podestà** (1997) Stomach contents in striped dolphins *Stenella coeruleoalba* (Meyen, 1833) from the Western Ligurian Sea. *Atti della Società Italiana di Scienze Naturali e del Museo Civico di Storia Naturale di Milano* **137**(1): 5-15.
- N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller & E. Teller** (1953) Equation of state calculations by fast computing machines. *Journal of Chemical Physics* **21**: 1087-1092. <http://dx.doi.org/10.1063/1.1699114>
- F. J. F. Meyen** (1833) Beitrage zur Zoologie gesammelt auf einer Reise um die Erde. Saugethiere. *Nova Acta Academiae Caesareae Leopoldino-Carolinae* **16**: 551-610. <http://www.worldcat.org/title/beitrage-zur-zoologie-gesammelt-auf-einer-reise-um-die-erde/oclc/018354885>

J. Michaux, A. Reyes & F. Catzeflis (2001) Evolutionary History of the Most Speciose Mammals: Molecular Phylogeny of Muroid Rodents. *Molecular Biology and Evolution* **18**(11): 2017-2031. <http://mbe.oxfordjournals.org/cgi/content/abstract/18/11/2017>
<http://dx.doi.org/10.1093/oxfordjournals.molbev.a003743>

M. C. Milinkovitch (1995) Molecular phylogeny of cetaceans prompts revision of morphological transformations. *Trends in Ecology and Evolution* **10**(8): 328-334. [http://dx.doi.org/10.1016/S0169-5347\(00\)89120-X](http://dx.doi.org/10.1016/S0169-5347(00)89120-X)

M. C. Millinkovitch & J. G. Thewissen (1997) Even-toed fingerprints on whale ancestry. *Nature* **388**(6643): 622-624. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9262391 [http://dx.doi.org/10.1016/S0169-5347\(00\)89120-X](http://dx.doi.org/10.1016/S0169-5347(00)89120-X)

S. K. Mills & J. H. Beatty (1979) The propensity interpretation of fitness. *Philosophy of Science* **46**(2): 263-286. <http://www.jstor.org/stable/18704> <http://dx.doi.org/10.1086/288865>

E. D. Mitchell (1970) Pigmentation pattern evolution in delphinid cetaceans: an essay in adaptive coloration. *Canadian Journal of Zoology* **48**: 717-740. <http://dx.doi.org/10.1139/z70-129>

N. Miyazaki, T. Kasuya & M. Nishiwaki (1974) Distribution and migration of two species of *Stenella* in the Pacific coast of Japan. *Scientific Reports of the Whales Research Institute, Tokyo* **26**: 227-243.

N. Miyazaki & M. Nishiwaki (1978) School structure of the Striped dolphin off the Pacific coast of Japan. *Scientific Reports of the Whales Research Institute, Tokyo* **30**: 65-116.

N. Miyazaki (1984) Further analyses of reproduction in the striped dolphin, *Stenella coeruleoalba*, off the Pacific coast of Japan. *Reports of the International Whaling Commission Special Issue* **6**: 343-353.

C. Montgelard, F. M. Catzeflis & E. Douzery (1997) Phylogenetic relationships of artiodactyls and cetaceans as deduced from the comparison of cytochrome b and 12S rRNA mitochondrial sequences. *Molecular biology and evolution* **14**(5): 550-9. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9159933

C. Montixi, C. Langlet, A. M. Bernard, J. Thimonier, C. Dubois, M. A. Wurbel, J. P. Chauvin, M. Pierres & H. T. He (1998) Engagement of T cell receptor triggers its recruitment to low-density detergent-insoluble membrane domains. *Embo Journal* **17**(18): 5334-48. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9736612 <http://dx.doi.org/10.1093/emboj/17.18.5334>

C. d. Muizon (1988) Les relations phylogénétiques des Delphinida (Cetacea, Mammalia). *Annales de paléontologie* **74**(4): 159-227. <http://www.mendeley.com/research/les-relations-phylogénétiques-des-delphinida-cetacea-mammalia/>

D. W. Müller & K. J. Hsü (1987) Event stratigraphy and paleoceanography in the Fortuna basin (southeast Spain): A scenario for the Messinian Salinity Crisis. *Paleoceanography* **2**(6): 679-96. <http://dx.doi.org/10.1029/PA002i006p00679>

K. Müller (2006) Incorporating information from length-mutational events into phylogenetic analysis. *Molecular Phylogenetics and Evolution* **38**(3): 667-676. <http://dx.doi.org/10.1016/j.ympev.2005.07.011>

A. Munguia-Vega, Y. Esquer-Garrigos, L. Rojas-Bracho, R. Vazquez-Juarez, A. Castro-Prieto & S. Flores-Ramirez (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*). *Molecular ecology* **16**(19): 4051-4065. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17727623

W. J. Murphy, E. Eizirik, S. J. O'Brien, O. Madsen, M. Scally, C. J. Douady, E. Teeling, O. A. Ryder, M. J. Stanhope, W. W. de Jong & M. S. Springer (2001) Resolution of the Early

Placental Mammal Radiation Using Bayesian Phylogenetics. *Science* **294**(5550): 2348-2351.
<http://dx.doi.org/10.1126/science.1067179>

B. W. Murray & B. N. White (1998) Sequence variation at the major histocompatibility complex DRB loci in beluga (*Delphinapterus leucas*) and narwhal (*Monodon monoceros*). *Immunogenetics* **48**(4): 242-52. <http://dx.doi.org/10.1007/s002510050428>

S. V. Muse & B. S. Gaut (1994) A likelihood approach for comparing synonymous and nonsynonymous nucleotide substitution rates, with application to the chloroplast genome. *Molecular Biology and Evolution* **11**(5): 715-24.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=7968485

K. Musolf, Y. Meyer-Lucht & S. Sommer (2004) Evolution of MHC-DRB class II polymorphism in the genus *Apodemus* and a comparison of DRB sequences within the family Muridae (Mammalia: Rodentia). *Immunogenetics* **56**(6): 420-6.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15351919 <http://dx.doi.org/10.1007/s00251-004-0715-9>

N. Myers, R. A. Mittermeier, C. G. Mittermeier, G. A. da-Fonseca & J. Kent (2000) Biodiversity hotspots for conservation priorities. *Nature* **403**(6772): 853-8.
<http://dx.doi.org/10.1038/35002501>

R. M. Myers, V. C. Sheffield & D. R. Cox (1988) Detection of single base changes in DNA: ribonuclease cleavage and denaturing gradient gel electrophoresis in *Genome Analysis: A practical approach*. (K. Davies Ed.). IRL Press, Oxford.
<http://www.cabdirect.org/abstracts/19891606063.html>

A. Natoli, A. Birkun, A. Aguilar, A. Lopez & A. R. Hoelzel (2005) Habitat structure and the dispersal of male and female bottlenose dolphins (*Tursiops truncatus*). *Proceedings of the Royal Society of London. Series B. Biological Sciences* **272**(1569): 1217-26.
<http://dx.doi.org/10.1098/rspb.2005.3076>

A. Natoli, A. Canadas, V. M. Peddemors, A. Aguilar, C. Vaquero, P. Fernandez-Piqueras & A. R. Hoelzel (2006) Phylogeography and alpha taxonomy of the common dolphin (*Delphinus* sp.). *Journal of Evolutionary Biology* **19**(3): 943-54.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=16674590 <http://dx.doi.org/10.1111/j.1420-9101.2005.01033.x>

S. Nee, E. C. Holmes, A. Rambaut & P. H. Harvey (1995) Inferring population history from molecular phylogenies. *Proceedings of the Royal Society B: Biological Sciences* **349**(1327): 25-31.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8748016 <http://dx.doi.org/10.1098/rstb.1995.0087>

M. Nei, X. Gu & T. Sitnikova (1997) Evolution by the birth-and-death process in multigene families of the vertebrate immune system. *Proceedings of the National Academy of Sciences of the United States of America* **94**(15): 7799-806.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9223266 <http://dx.doi.org/10.1073/pnas.94.15.7799>

J. E. Neigel & J. C. Avise (1986) Phylogenetic relationships of mitochondrial DNA under various demographic models of speciation in *Evolutionary Processes and Theory*. (S. K. a. E. Nevo Ed.). Academic Press, Inc., New York. p 515-534.
<http://www.pnas.org/content/94/15/7799.full>

M. K. Newell, J. VanderWall, K. S. Beard & J. H. Freed (1993) Ligation of major histocompatibility complex class II molecules mediates apoptotic cell death in resting B lymphocytes. *Proceedings of the National Academy of Sciences of the United States of America* **90**(22): 10459-63.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8248132 <http://dx.doi.org/10.1073/pnas.90.22.10459>

G. Niedermann, R. Grimm, E. Geier, M. Maurer, C. Realini, C. Gartmann, J. Soll, S. Omura, M. C. Rechsteiner, W. Baumeister & K. Eichmann (1997) Potential immunocompetence of proteolytic fragments produced by proteasomes before evolution of the vertebrate immune system. *Journal of Experimental Medicine* **186**(2): 209-20. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9221750 <http://dx.doi.org/10.1084/jem.186.2.209>

E. E. Nielsen, M. M. Hansen, D. E. Ruzzante, D. Meldrup & P. Gronkjaer (2003) Evidence of a hybrid-zone in Atlantic cod (*Gadus morhua*) in the Baltic and the Danish Belt Sea revealed by individual admixture analysis. *Molecular Ecology* **12**(6): 1497-1508. <http://www.blackwell-synergy.com/doi/abs/10.1046/j.1365-294X.2003.01819.x> <http://dx.doi.org/10.1046/j.1365-294X.2003.01819.x>

M. Nikaido, A. P. Rooney & N. Okada (1999) Phylogenetic relationships among cetartiodactyls based on insertions of short and long interspersed elements: hippopotamuses are the closest extant relatives of whales. *Proceedings of the National Academy of Sciences of the United States of America* **96**(18): 10261-6. <http://dx.doi.org/10.1073/pnas.96.18.10261>

M. Nikaido, F. Matsuno, H. Hamilton, R. L. Brownell, Jr., Y. Cao, W. Ding, Z. Zuoyan, A. M. Shedlock, R. E. Fordyce, M. Hasegawa & N. Okada (2001) Retroposon analysis of major cetacean lineages: the monophyly of toothed whales and the paraphyly of river dolphins. *Proceedings of the National Academy of Sciences of the United States of America* **98**(13): 7384-9. <http://dx.doi.org/10.1073/pnas.96.18.10261>

K. S. Norris & T. P. Dohl (1988) The Structure and Functions of Cetacean Schools in Cetacean Behavior: mechanisms and functions. (L. M. Herman Ed.). Krieger, Malabar, FL. p 211-261.

G. Notarbartolo di Sciara (2002a) Section 3. Cetacean species occurring in the Mediterranean and Black Seas. in Cetaceans of the Mediterranean and Black Seas: state of knowledge and conservation strategies. A report to the ACCOBAMS Secretariat. (G. N. d. S. (Ed.) Ed.). A report to Monaco, February 2002. 17 p. <http://oceansdocs.org/bitstream/1834/840/1/Notarbartolo2.pdf>

G. Notarbartolo di Sciara (2002b) Cetacean species occurring in the Mediterranean and Black Seas. In: G. Notarbartolo di Sciara (Ed.), Cetaceans of the Mediterranean and Black Seas: state of knowledge and conservation strategies. in A report to A report to the ACCOBAMS Secretariat, Monaco, February 2002. 17.

R. M. Nowak (1999) Walker's mammals of the world (6th edition). Johns Hopkins University Press, Baltimore, MD. 1936 pp. <http://books.google.com/books?id=T37sFCI43E8C&hl=ca>

T. J. O'Shea, R. R. Reeves & A. K. Long (1998) Marine mammals and persistent ocean contaminants. *Proceedings of the Marine Mammal Commission Workshop*. http://www.whoi.edu/science/B/people/mhahn/Keystone_workshop.pdf

M. Oishi & Y. Hasegawa (1994) A list of fossil cetaceans in Japan. *The Island Arc* **3**(4): 493-505. <http://dx.doi.org/10.1111/j.1440-1738.1994.tb00128.x> <http://dx.doi.org/10.1111/j.1440-1738.1994.tb00128.x>

Y. Okada (1939) A study of Japanese Delphinidae (1). *Science Reports of the Tokyo Bunrika Daigaku, Section B* **3**(44): 1-16.

W. R. B. Oliver (1922) A Review of the Cetacea of the New Zealand Seas. *Proceedings of the Zoological Society of London* **92**(3) <http://dx.doi.org/10.1111/j.1096-3642.1922.tb02157.x>

J. Ollerton (2005) Speciation: flowering time and the Wallace effect. *Heredity* **95**(3): 181-182. <http://dx.doi.org/10.1038/sj.hdy.6800718>

R. Opgen-Rhein, L. Fahrmeir & K. Strimmer (2005) Inference of demographic history from genealogical trees using reversible jump Markov chain Monte Carlo. *BMC evolutionary biology* **5**(1): 6. <http://dx.doi.org/10.1186/1471-2148-5-6>

C. L. Organ & D. E. Janes (2008) Evolution of sex chromosomes in Sauropsida. *Integrative and Comparative Biology* 48(4): 512-519. <http://dx.doi.org/10.1093/icb/icn041>

A. Osterhaus, J. Groen, H. Niesters, M. van de Bildt, B. Martina, L. Vedder, J. Vos, H. van Egmond, B. Abou-Sidi & M. E. Barham (1997) Morbillivirus in monk seal mass mortality. *Nature* 388(6645): 838-9. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9278043 <http://dx.doi.org/10.1038/42163>

I. G. Ovsyannikova, K. L. Johnson, D. C. Muddiman, R. A. Vierkant & G. A. Poland (2004) Identification and characterization of novel, naturally processed measles virus class II HLA-DRB1 peptides. *Journal of virology* 78(1): 42-51. <http://dx.doi.org/10.1128/JVI.78.1.42-51.2004>

P. J. Palsbøll, M. Bérubé, A. Aguilar, G. Notarbartolo-Di-Sciara & R. Nielsen (2004) Discerning between recurrent gene flow and recent divergence under a finite-site mutation model applied to North Atlantic and Mediterranean Sea fin whale (*Balaenoptera physalus*) populations. *Evolution, International Journal of Organic Evolution* 58(3): 670-675. <http://dx.doi.org/10.1554/02-529>

T. J. Parker & W. A. Haswell (1897) A Text-Book of Zoology. Volume 2. Macmillan Press, London, London. http://books.google.fr/books/about/A_text_book_of_zoology.html?id=e18P2Oz7ZfsC&redir_esc=y

D. Pauly, A. W. Trites, E. Capuli & V. Christensen (1998) Diet composition and trophic levels of marine mammals. *ICES Journal of Marine Science* 55(3): 467-481. <http://dx.doi.org/10.1006/jmsc.1997.0280>

E. Penzo, G. Gandolfi, L. Bargelloni, L. Colombo & T. Patarnello (1998) Messinian salinity crisis and the origin of freshwater lifestyle in western Mediterranean gobies. *Molecular Biology and Evolution* 15(11): 1472-80. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12572610

W. F. Perrin, R. R. Warner, C. H. Fiscus & D. B. Holts (1973) Stomach contents of porpoise, *Stenella* spp., and yellowfin tuna, *Thunnus albacares*, in mixed-species aggregations. *Fishery Bulletin* 71(4): 1077-91. <http://swfsc.noaa.gov/publications/CR/1973/7327.PDF>

W. F. Perrin (1975) Variation of Spotted and Spinner Porpoise (genus *Stenella*) in the Eastern Pacific and Hawaii. *Bulletin of the Scripps Institution of Oceanography, University of California, San Diego La Jolla, California* 21: 1-206. <http://escholarship.org/uc/item/2dq444zq>

W. F. Perrin, E. D. Mitchell, J. G. Mead, D. K. Caldwell & P. J. H. van Bree (1981) *Stenella clymene*, a Rediscovered Tropical Dolphin of the Atlantic. *Journal of Mammalogy* 62(3): 583-598. <http://dx.doi.org/10.2307/1380405>

W. F. Perrin, C. E. Wilson & F. I. Archer (1994a) Striped dolphin. *Stenella coeruleoalba* (Meyen, 1833). in Handbook of marine mammals: the first book of dolphins. (S. H. R. a. R. Harrison Ed.). Academic Press, San Diego. p 129-159.

W. F. Perrin, C. E. Wilson & F. Archer II (1994b) Striped dolphin *Stenella coeruleoalba* (Meyen, 1833). Volume 5: *The First Book of Dolphins* 5: 129-159. http://books.google.fr/books/about/Handbook_of_marine_mammals.html?id=_oIXAQAAlA&redir_esc=y

W. F. Perrin (1997) Development and homologies of head stripes in the delphinoid cetaceans. *Marine Mammal Science* 13(1): 1-43. <http://dx.doi.org/10.1111/j.1748-7692.1997.tb00610.x>

W. F. Perrin (1998) *Stenella longirostris*. *Mammalian Species*(599): 1-7. <http://dx.doi.org/10.2307/3504456>

W. F. Perrin & S. L. Mesnick (2003) Sexual ecology of the Spinner dolphin, *Stenella*

longirostris: geographic variation in mating system. *Marine Mammal Science* **19**(3): 462-483.
<http://www.blackwell-synergy.com/doi/abs/10.1111/j.1748-7692.2003.tb01315.x>
<http://dx.doi.org/10.1111/j.1748-7692.2003.tb01315.x>

C. Pertoldi, M. Podesta, V. Loeschcke, S. Schandorff, L. Marsili, C. Mancusi, P. Nicolosi & E. Randi (2000) Effect of the 1990 die-off in the northern Italian seas on the developmental stability of the striped dolphin *Stenella coeruleoalba* (Meyen, 1833). *Biological Journal of the Linnean Society* **71**(1): 61-70. <http://dx.doi.org/10.1111/j.1095-8312.2000.tb01242.x>

E. Petit, F. Balloux & L. Excoffier (2002) Mammalian population genetics: why not Y? *Trends in Ecology and Evolution* **17**(1): 28-33. [http://dx.doi.org/10.1016/S0169-5347\(01\)02356-4](http://dx.doi.org/10.1016/S0169-5347(01)02356-4)

N. Phadnis & H. A. Orr (2009) A Single Gene Causes Both Male Sterility and Segregation Distortion in *Drosophila* Hybrids. *Science* **323**(5912): 376-379.
<http://dx.doi.org/10.1126/science.1163934> <http://dx.doi.org/10.1126/science.1163934>

H. Philippe & E. Douzery (1994) The pitfalls of molecular phylogeny based on four species, as illustrated by the Cetacea/Artiodactyla relationships. *Journal of Mammalian Evolution* **2**(2): 133-152. <http://dx.doi.org/10.1007/BF01464365>

M. Pidwirny (2006) "Concept of Ecological Niche". *Fundamentals of Physical Geography*. Number of Pages pp. <http://www.physicalgeography.net/fundamentals/contents.html>. Accessed the 2009-05-05.

G. Pilleri (1985) The Miocene Cetacea of the Belluno Sandstones. *Memorie di Scienze Geologiche* **36**: 1-250. http://www.aquila-books.com/belluno_main.htm

L. F. Pisano (1202) *Liber abaci*. Firenze.

R. L. Pitman, W. L. Perryman, D. LeRoi & E. Eilers (2007) A dwarf form of Killer Whale in Antarctica. *Journal of Mammalogy* **88**(1): 43-48. <http://dx.doi.org/10.1644/2007-MAMM-A-118R1.1>

S. Planes, R. Galzin, A. Garcia Rubies, R. Goni, J. G. Harmelin, L. Le Direach, P. Lenfant & A. Quetglas (2000) Effects of marine protected areas on recruitment processes with special reference to Mediterranean littoral ecosystems. *Environmental conservation* **27**(2): 126-143. <http://dx.doi.org/10.1017/S0376892900000175>

D. Posada & K. A. Crandall (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics Oxford, England* **14**(9): 817-8. <http://dx.doi.org/10.1093/bioinformatics/14.9.817>

D. Posada & K. A. Crandall (2001) Intraspecific gene genealogies: trees grafting into networks. *Trends in Ecology and Evolution* **16**(1): 37-45. [http://dx.doi.org/10.1016/S0169-5347\(00\)02026-7](http://dx.doi.org/10.1016/S0169-5347(00)02026-7)

J. K. Pritchard, M. Stephens & P. Donnelly (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**(2): 945-59. <http://www.genetics.org/content/164/4/1567.full>

A. Purvis (1995) A Composite Estimate of Primate Phylogeny. *Philosophical Transactions of the Royal Society B: Biological Sciences* **348**(1326): 405-421. <http://dx.doi.org/10.1098/rstb.1995.0078>

O. G. Pybus, A. Rambaut & P. H. Harvey (2000) An integrated framework for the inference of viral population history from reconstructed genealogies. *Genetics* **155**(3): 1429-1437. <http://www.genetics.org/content/155/3/1429>

D. C. Queller & K. F. Goodnight (1989) Estimating relatedness using genetic markers. *Evolution* **43**(2): 258-275. <http://dx.doi.org/10.2307/2409206>

J. P. Quignard & J. A. Tomasini (2000) Mediterranean fish biodiversity. *Biologia Marina Mediterranea* **7**: 1-66.

E. Radley, R. P. Alderton, A. Kelly, J. Trowsdale & S. Beck (1994) Genomic organization of HLA-DMA and HLA-DMB. Comparison of the gene organization of all six class II families in the human major histocompatibility complex. *Journal of Biological Chemistry* **269**(29): 18834-8. <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation>

[&list_uids=8034636](#)

F. Ramade, e. Sloff-Wilbert & e. De-Kruijf-Hans (1993) Environmental and ecotoxicological problems in the Mediterranean countries. Recent advances in ecotoxicology. *Science of the total environment* (SUP1): 97-107.

A. Rambaut & L. Bromham (1998) Estimating divergence dates from molecular sequences. *Molecular biology and evolution* **15**(4): 442-448. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9549094

L. Rendall & H. Whitehead (2001) Culture in whales and dolphins. *Behavioral and Brain Sciences* **24**(2): 309-382. <http://dx.doi.org/10.1017/S0140525X0100396X>

J. C. Reyes (1996) A possible case of hybridism in wild dolphins. *Marine Mammal Science* **12**(2): 301-307. <http://dx.doi.org/10.1111/j.1748-7692.1996.tb00581.x>

W. R. Rice (1996) Evolution of the Y Sex Chromosome in Animals. *Bioscience* **46**(5): 331-343 <http://www.jstor.org/stable/1312947> <http://dx.doi.org/10.2307/1312947>

M. Richards & V. Macaulay (2001) The mitochondrial gene tree comes of age. *American journal of human genetics* **68**(6): 1315-1320. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=11349234 <http://dx.doi.org/10.1086/320615>

S. H. Ridgway (1999) Handbook of marine mammals - The first book of dolphins and the porpoises in Handbook of marine mammals (S. H. Ridgway & R. J. Harrison Ed.). Academic Press, London. 486 pp. http://books.google.es/books/about/Handbook_of_marine_mammals.html?hl=es&id=5NytPwAACAAJ

M. Ridley (2004) Evolution (3rd edition). Blackwell Science, 751 pp.

L. H. Rieseberg, O. Raymond, D. M. Rosenthal, Z. Lai, K. Livingstone, T. Nakazato, J. L. Durphy, A. E. Schwarzbach, L. A. Donovan & C. Lexer (2003) Major ecological transitions in wild sunflowers facilitated by hybridization. *Science* **301**(5637): 1211-6. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12907807 <http://dx.doi.org/10.1126/science.1086949>

A. G. Rodrigo (1999) HIV evolutionary genetics. *Proceedings of the National Academy of Sciences of the United States of America* **96**(19): 10559-61. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10485860 <http://dx.doi.org/10.1073/pnas.96.19.10559>

A. G. Rodrigo & J. Felsenstein (1999) Coalescent approaches to HIV population genetics in Molecular evolution of HIV (K. Crandall Ed.). Johns Hopkins University Press, Baltimore, MD. 233-272 pp.

A. S. Romer (1955) Herpetichthyes, Amphibioidei, Choanichthyes or Sarcopterygii? *Nature* **176**(4472): 126-126. <http://dx.doi.org/10.1038/176126a0>

F. Ronquist & J. P. Huelsenbeck (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics Oxford, England* **19**(12): 1572-4. <http://dx.doi.org/10.1093/bioinformatics/btg180>

K. D. Rose & J. D. Archibald (2005) The rise of placental mammals: origins and relationships of the major extant clades. Johns Hopkins University Press, Baltimore, MD. 259 pp. http://books.google.com/books?id=DhchVG_rbQ8C&hl=ca

H. C. Rosenbaum, R. L. Brownell, M. W. Brown, C. Schaeff, V. Portway, B. N. White, S. Malik, L. A. Pastene, N. J. Patenaude, C. S. Baker, M. Goto, P. B. Best, P. J. Clapham, P. Hamilton, M. Moore, R. Payne, V. Rowntree, C. T. Tynan, J. L. Bannister & R. DeSalle (2000) World-wide genetic differentiation of *Eubalaena*: questioning the number of right whale species. *Molecular Ecology* **9**(11): 1793-802. <http://dx.doi.org/10.1046/j.1365-294x.2000.01066.x>

N. A. Rosenberg, S. Mahajan, S. Ramachandran, C. Zhao, J. K. Pritchard & M. W.

- Feldman** (2005) Clines, clusters, and the effect of study design on the inference of human population structure. *PLoS Genet* **1**(6): e70. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=16355252 <http://dx.doi.org/10.1371/journal.pgen.0010070>
- G. J. B. Ross** (1977a) The taxonomy of bottlenose dolphins *Tursiops* species in South African waters, with notes on their biology. *Annals of the Cape Provincial Museums. Natural history* **11**: 135-194. <http://md1.csa.com/partners/viewrecord.php?requester=gs&collection=ENV&recid=924858&q=&uid=1266178>
- G. J. B. Ross** (1977b) The Taxonomy of Bottlenosed Dolphins *Tursiops* Species in South African Waters, with Notes on Their Biology. *Annals of the Cape Provincial Museum* **11**: 135-194. http://books.google.fr/books/about/The_taxonomy_of_bottlenosed_dolphins_Tur.html?id=9sbAOQAACAAJ&redir_esc=y
- G. J. B. Ross** (1984) The smaller cetaceans of the south east coast of southern Africa. *Annals of the Cape Provincial Museums, Natural History* **15**: 173-410.
- G. J. B. Ross & V. G. Cockroft** (1990) Comments on Australian bottlenose dolphins and the taxonomic status of *Tursiops aduncus* (Ehrenberg, 1832) in *The Bottlenose Dolphin*. (S. Leatherwood & R. R. Reeves Ed.). Academic Press Inc., San Diego, CA. p 143-154.
- H. A. Ross, S. Murugan & C. Scott Baker** (2005) Are the species identities of cetacean sequences in GenBank correct? *Scientific Reports of the International Whaling Commission SC/57/SD4*
- J. Rozas, J. C. Sanchez-DelBarrio, X. Messeguer & R. Rozas** (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics Oxford, England* **19**(18): 2496-7. <http://dx.doi.org/10.1093/bioinformatics/btg359>
- N. Saitou & M. Nei** (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* **4**(4): 406-25. <http://www.ncbi.nlm.nih.gov/pubmed/3447015>
- E. C. Salido, P. H. Yen, K. Koprivnikar, L. C. Yu & L. J. Shapiro** (1992) The human enamel protein gene amelogenin is expressed from both the X and the Y chromosomes. *American Journal of Human Genetics* **50**(2): 303-16. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1682460/pdf/ajhg00073-0063.pdf>
- W. Salzburger, T. Mack, E. Verheyen & A. Meyer** (2005) Out of Tanganyika: genesis, explosive speciation, key-innovations and phylogeography of the haplochromine cichlid fishes. *BMC Evol Biol* **5**(1): 17. <http://dx.doi.org/10.1186/1471-2148-5-17>
- J. Sambrook & D. W. Russell** (2001) *Molecular Cloning: A Laboratory Manual* (3rd edition). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. <http://www.cshlpress.com/link/molclon3.htm>
- M. J. Sanderson** (2002) Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Molecular Biology and Evolution* **19**(1): 101-9. <http://dx.doi.org/10.1093/oxfordjournals.molbev.a003974>
- M. J. Sanderson** (2003) r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics Oxford, England* **19**(2): 301-2. <http://dx.doi.org/10.1093/bioinformatics/19.2.301>
- I. Sanmartin** (2003) Dispersal vs. vicariance in the Mediterranean: historical biogeography of the Palearctic Pachydeminae (Coleoptera, Scarabaeoidea). *Journal of biogeography* **30**(12): 1883-1897. <http://dx.doi.org/10.1046/j.0305-0270.2003.00982.x>
- M. B. Santos, G. J. Pierce, H. M. Ross, R. J. Reid & B. Wilson** (1994) Diets of small cetaceans from the Scottish coast. *ICES Journal of Marine Science* **1994/N:11**
- A. Sasaki & Y. Iwasa** (1991) Optimal growth schedule of pathogens within a host: switching between lytic and latent cycles. *Theoretical population biology* **39**(2): 201-39.

[http://dx.doi.org/10.1016/0040-5809\(91\)90036-F](http://dx.doi.org/10.1016/0040-5809(91)90036-F)

T. Sasaki, M. Nikaido, H. Hamilton, M. Goto, H. Kato, N. Kanda, L. Pastene, Y. Cao, R. Fordyce, M. Hasegawa & N. Okada (2005) Mitochondrial phylogenetics and evolution of mysticete whales. *Systematic biology* **54**(1): 77-90. <http://dx.doi.org/10.1080/10635150590905939>

P. G. Satchell, X. Anderton, O. H. Ryu, X. Luan, A. J. Ortega, R. Opamen, B. J. Berman, D. E. Witherspoon, J. L. Gutmann, A. Yamane, M. Zeichner-David, J. P. Simmer, C. F. Shuler & T. G. Diekwisch (2002) Conservation and variation in enamel protein distribution during vertebrate tooth development. *Journal of Experimental Zoology* **294**(2): 91-106. <http://dx.doi.org/10.1002/jez.10148>

J. Schad, S. Sommer & J. U. Ganzhorn (2004) MHC variability of a small lemur in the littoral forest fragments of southeastern Madagascar. *Conservation Genetics* **5**(3): 299-309. <http://dx.doi.org/10.1023/B:COGE.0000031137.50239.d3>

J. Schad, J. U. Ganzhorn & S. Sommer (2005) Parasite burden and constitution of major histocompatibility complex in the Malagasy mouse lemur, *Microcebus murinus*. *Evolution Int J Org Evolution* **59**(2): 439-50. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15807428 <http://dx.doi.org/10.1554/04-312>

M. R. Schick & S. Levy (1993) The TAPA-1 molecule is associated on the surface of B cells with HLA-DR molecules. *Journal of Immunology* **151**(8): 4090-7. <http://www.ncbi.nlm.nih.gov/pubmed/8409388>

S. Schneider, D. Roessli & L. Excoffier (2000) ARLEQUIN ver 2.0: a software for population genetics data analysis. Genetics and Biometry Laboratory, University of Geneva, Geneva. Geneva.

M. Schrope (2007) Food chains: killer in the kelp. *Nature* **445**(7129): 703-5. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17301765 <http://dx.doi.org/10.1038/445703a>

J. M. Seddon & H. Ellegren (2002) MHC class II genes in European wolves: a comparison with dogs. *Immunogenetics* **54**(7): 490-500. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12389097 <http://dx.doi.org/10.1007/s00251-002-0489-x>

U. Seifert, C. Maranon, A. Shmueli, J. F. Desoutter, L. Wesoloski, K. Janek, P. Henklein, S. Diescher, M. Andrieu, H. de la Salle, T. Weinschenk, H. Schild, D. Laderach, A. Galy, G. Haas, P. M. Kloetzel, Y. Reiss & A. Hosmalin (2003) An essential role for tripeptidyl peptidase in the generation of an MHC class I epitope. *Nature Immunology* **4**(4): 375-9. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12598896 <http://dx.doi.org/10.1038/ni905>

L. Sena, M. P. Schneider, B. Brenig, R. L. Honeycutt, J. E. Womack & L. C. Skow (2003) Polymorphisms in MHC-DRA and -DRB alleles of water buffalo (*Bubalus bubalis*) reveal different features from cattle DR alleles. *Animal genetics* **34**(1): 1-10. <http://dx.doi.org/10.1046/j.1365-2052.2003.00920.x>

B. Shapiro, A. J. Drummond, A. Rambaut, M. C. Wilson, P. E. Matheus, A. V. Sher, O. G. Pybus, M. T. Gilbert, I. Barnes, J. Binladen, E. Willerslev, A. J. Hansen, G. F. Baryshnikov, J. A. Burns, S. Davydov, J. C. Driver, D. G. Froese, C. R. Harington, G. Keddie, P. Kosintsev, M. L. Kunz, L. D. Martin, R. O. Stephenson, J. Storer, R. Tedford, S. Zimov & A. Cooper (2004) Rise and fall of the Beringian steppe bison. *Science* **306**(5701): 1561-5. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15567864 <http://dx.doi.org/10.1126/science.1101074>

V. C. Sheffield, D. R. Cox, L. S. Lerman & R. M. Myers (1989) Attachment of a 40-base-pair G + C-rich sequence (GC-clamp) to genomic DNA fragments by the polymerase chain reaction results in improved detection of single-base changes. *Proceedings of the National*

Academy of Sciences of the United States of America **86**(1): 232-6.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=2643100 <http://dx.doi.org/10.1073/pnas.86.1.232>

M. Shimamura, H. Yasue, K. Ohshima, H. Abe, H. Kato, T. Kishiro, M. Goto, I. Munechika & N. Okada (1997) Molecular evidence from retroposons that whales form a clade within even-toed ungulates. *Nature* **388**(6643): 666-670.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9262399 <http://dx.doi.org/10.1038/41759>

S. Sigurdardottir, C. Borsch, K. Gustafsson & L. Andersson (1992) Exon encoding the antigen-binding site of MHC class II beta-chains is divided into two subregions with different evolutionary histories. *Journal of Immunology* **148**(3): 968-73.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=1730884

L. Silvani, M. Gazo & A. Aguilar (1999) Spanish driftnet fishing and incidental catches in the western Mediterranean. *Biological conservation* **90**(1): 79-85.
[http://dx.doi.org/10.1016/S0006-3207\(98\)00079-2](http://dx.doi.org/10.1016/S0006-3207(98)00079-2)

G. G. Simpson (1945) The Principles of Classification and a Classification of Mammals. *Bulletin of the American Museum of Natural History* **85**: 1-350. <http://hdl.handle.net/2246/1104>

J. Y. Sire, S. Delgado, D. Fromentin & M. Girondot (2005) Amelogenin: lessons from evolution. *Archives of oral biology* **50**(2): 205-12.
<http://dx.doi.org/10.1016/j.archoralbio.2004.09.004>

M. Slatkin & L. Voelm (1991) FST in a hierarchical island model. *Genetics* **127**(3): 627-9.
<http://www.genetics.org/content/127/3/627.long>

M. Slatkin & L. Excoffier (1996) Testing for linkage disequilibrium in genotypic data using the Expectation-Maximization algorithm. *Heredity* **76**(Pt 4): 377-83.
<http://www.nature.com/hdy/journal/v76/n4/abs/6888150a.html>
<http://dx.doi.org/10.1038/hdy.1996.55>

E. J. Slijper (1936) Die Cetaceen, vergleichend-anatomisch und systematisch: Ein Beitrag zur vergleichenden Anatomie des Blutgefäß-, Nerven- und Muskelsystems, sowie des Rumpfskelettes der Säugetiere, mit Studien über die Theorie des Aussterbens und der Foetalisation in *Capita Zoologica (M. Nijhoff Ed.)*. The Hague, 350 pp.

J. J. Smith & S. R. Voss (2007) Bird and mammal sex-chromosome orthologs map to the same autosomal region in a salamander (*ambystoma*). *Genetics* **177**(1): 607-613.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17660573 <http://dx.doi.org/10.1534/genetics.107.072033>

M. L. Snead, E. C. Lau, M. Zeichner-David, A. G. Fincham, S. L. Woo & H. C. Slavkin (1985) DNA sequence for cloned cDNA for murine amelogenin reveal the amino acid sequence for enamel-specific protein. *Biochemical and Biophysical Research Communications* **129**(3): 812-8.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=4015654 [http://dx.doi.org/10.1016/0006-291X\(85\)91964-3](http://dx.doi.org/10.1016/0006-291X(85)91964-3)

J. Snoeks & G. R. Carvalho (2004) The cichlid diversity of Lake Malawi/Nyasa/Niassa: identification, distribution and taxonomy. Cichlid Press, El Paso, TX. 360 pp.
<http://books.google.com/books?id=xpTFAAAACAAJ&hl=ca>

A. M. Sole Cava, M. Klautau, N. Boury Esnault, R. Borojelic & J. P. Thorpe (1991) Genetic evidence for cryptic speciation in allopatric populations of two cosmopolitan species of the calcareous sponge genus *Clathrina*. *Marine Biology* **111**(3): 381-386.
<http://dx.doi.org/10.1007/BF01319410>

S. Sommer (2003) Effects of habitat fragmentation and changes of dispersal behaviour after a recent population decline on the genetic variability of noncoding and coding DNA of a monogamous Malagasy rodent. *Molecular Ecology* **12**(10): 2845-51.

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12969486 <http://dx.doi.org/10.1046/j.1365-294X.2003.01906.x>

B. Spring, C. Fonatsch, C. Muller, G. Pawelec, J. Kompf, P. Wernet & A. Ziegler (1985) Refinement of HLA gene mapping with induced B-cell line mutants. *Immunogenetics* **21**(3): 277-91.

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=3872842 <http://dx.doi.org/10.1007/BF00375380>

A. M. Springer, J. A. Estes, G. B. van Vliet, T. M. Williams, D. F. Doak, E. M. Danner, K. A. Forney & B. Pfister (2003) Sequential megafaunal collapse in the North Pacific Ocean: an ongoing legacy of industrial whaling? *Proceedings of the National Academy of Sciences of the United States of America* **100**(21): 12223-8.

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=14526101 <http://dx.doi.org/10.1073/pnas.1635156100>

C. Stamatis, A. Triantafyllidis, K. A. Moutou & Z. Mamuris (2004) Mitochondrial DNA variation in northeast atlantic and mediterranean populations of norway lobster, *Nephrops norvegicus*. *Molecular Ecology* **13**(6): 1377-1390. <http://dx.doi.org/10.1111/j.1365-294X.2004.02165.x>

L. J. Stern & D. C. Wiley (1994a) Antigenic peptide binding by class I and class II histocompatibility proteins. *Structure* **2**(4): 245-51.

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8087551 [http://dx.doi.org/10.1016/S0969-2126\(00\)00026-5](http://dx.doi.org/10.1016/S0969-2126(00)00026-5)

L. J. Stern & D. C. Wiley (1994b) Antigenic peptide binding by class I and class II histocompatibility proteins. *Structure London, England* **2**(4): 245-51.

P. Stockley (2003) Female multiple mating behaviour, early reproductive failure and litter size variation in mammals. *Proceedings of the Royal Society of London Series B Biological Sciences* **270**(1512): 271-8.

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12614576 <http://dx.doi.org/10.1098/rspb.2002.2228>

L. Stoltze, A. K. Nussbaum, A. Sijts, N. P. Emmerich, P. M. Kloetzel & H. Schild (2000) The function of the proteasome system in MHC class I antigen processing. *Immunology Today* **21**(7): 317-9.

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10950502 [http://dx.doi.org/10.1016/S0167-5699\(00\)01665-0](http://dx.doi.org/10.1016/S0167-5699(00)01665-0)

K. Strimmer & O. G. Pybus (2001) Exploring the demographic history of DNA sequences using the generalized skyline plot. *Molecular biology and evolution* **18**(12): 2298-2305.

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=11719579 <http://dx.doi.org/10.1093/oxfordjournals.molbev.a003776>

C. F. Suárez, M. E. Patarroyo, E. Trujillo, M. Estupiñán, J. E. Baquero, C. Parra & R. Rodriguez (2006) Owl monkey MHC-DRB exon 2 reveals high similarity with several HLA-DRB lineages. *Immunogenetics* **58**(7): 542-58. <http://dx.doi.org/10.1007/s00251-006-0127-0>

H. Sultmann, B. W. Murray & J. Klein (2000) Identification of seven genes in the major histocompatibility complex class I region of the zebrafish. *Scandinavian Journal of Immunology* **51**(6): 577-85.

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10849368

K. M. Sullivan, A. Mannucci, C. P. Kimpton & P. Gill (1993) A rapid and quantitative DNA sex test: fluorescence-based PCR analysis of X-Y homologous gene amelogenin. *Biotechniques* **15**(4): 636-8, 640-1.

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8251166

P. Taberlet, H. Mattock, C. Dubois-Paganon & J. Bouvet (1993) Sexing free-ranging brown bears *Ursus arctos* using hairs found in the field. *Molecular Ecology* **2**(6): 399-403. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8162229 <http://dx.doi.org/10.1111/j.1365-294X.1993.tb00033.x>

K. Takahashi, A. P. Rooney & M. Nei (2000) Origins and divergence times of mammalian class II MHC gene clusters. *The Journal of heredity* **91**(3): 198-204. <http://dx.doi.org/10.1093/jhered/91.3.198>

N. Takahata & M. Nei (1990) Allelic genealogy under overdominant and frequency-dependent selection and polymorphism of major histocompatibility complex loci. *Genetics* **124**(4): 967-78. <http://www.genetics.org/content/124/4/967>

K. Tamura (1992) Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G+C-content biases. *Molecular biology and evolution* **9**(4): 678-687.

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=1630306

K. Tamura & M. Nei (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular biology and evolution* **10**(3): 512-526.

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8336541

S. Tavaré (1986) Some probabilistic and statistical problems in the analysis of DNA sequences in Some mathematical questions in biology – DNA sequence analysis. Am Math Soc,, Providence, RI. p 57-86.

http://books.google.fr/books?id=8aI1phhOKhgC&pg=PA57&lpg=PA57&dq=Some+probabilistic+and+statistical+problems+in+the+analysis+of+DNA+sequences+in+Some+mathematical+questions+in+biology&source=bl&ots=r1NA5QAgQh&sig=mPu-zdK4A4WbROXtOB8vCPoZFUE&hl=ca&sa=X&ei=jyZ7T76VEc_c8gOLkOTBCA&redir_esc=y_v=onepage&q=Some%20probabilistic%20and%20statistical%20problems%20in%20the%20analysis%20of%20DNA%20sequences%20in%20Some%20mathematical%20questions%20in%20biology&f=false

J. G. Thewissen, E. M. Williams, L. J. Roe & S. T. Hussain (2001) Skeletons of terrestrial cetaceans and the relationship of whales to artiodactyls. *Nature* **413**(6853): 277-281. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=11565023 <http://dx.doi.org/10.1038/35095005>

J. G. M. Thewissen & E. M. Williams (2002) THE EARLY RADIATIONS OF CETACEA (MAMMALIA): Evolutionary Pattern and Developmental Correlations. *Annual Review of Ecology and Systematics* **33**(1): 73-90.

<http://arjournals.annualreviews.org/doi/abs/10.1146/annurev.ecolsys.33.020602.095426>
<http://dx.doi.org/10.1146/annurev.ecolsys.33.020602.095426>

J. D. Thompson, D. G. Higgins & T. J. 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 Research* **22**(22): 4673-4680.

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=7984417 <http://dx.doi.org/10.1093/nar/22.22.4673>

J. D. Thompson, T. J. Gibson, F. Plewniak, F. Jeanmougin & D. G. Higgins (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* **25**(24): 4876-82.

N. W. Thornhill (1993) The natural history of inbreeding and outbreeding: theoretical and empirical perspectives. University of Chicago press, Chicago, MI. 575 pp. http://books.google.com/books?id=7uPLc_Em6PYC&hl=ca

- T. Tregenza & N. Wedell** (2000) Genetic compatibility, mate choice and patterns of parentage: invited review. *Molecular Ecology* **9**(8): 1013-27.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10964221 <http://dx.doi.org/10.1046/j.1365-294x.2000.00964.x>
- W. Trevathan, E. O. Smith & J. J. McKenna** (1999) *Evolutionary Medicine* (1st edition). Oxford University Press US, New York. 496 pp.
<http://books.google.com/books?id=grrA421tRNkC&hl=ca>
- G. M. Troisi, K. Haraguchi, M. P. Simmonds & C. F. Mason** (1998) Methyl sulphone metabolites of polychlorinated biphenyls (PCBs) in cetaceans from the Irish and the Aegean seas. *Archives of environmental contamination and toxicology* **35**(1): 121-128.
<http://dx.doi.org/10.1007/s002449900358>
- J. Trowsdale, V. Groves & A. Arnason** (1989) Limited MHC polymorphism in whales. *Immunogenetics* **29**(1): 19-24. <http://dx.doi.org/10.1007/BF02341609>
- J. Trowsdale & R. D. Campbell** (1992) Complexity in the major histocompatibility complex. *European Journal of Immunogenetics* **19**(1-2): 45-55.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=1567814 <http://dx.doi.org/10.1111/j.1744-313X.1992.tb00047.x>
- J. Trowsdale** (1995) "Both man & bird & beast": comparative organization of MHC genes. *Immunogenetics* **41**(1): 1-17.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=7806269 <http://dx.doi.org/10.1007/BF00188427>
- F. W. True** (1889) Contributions to the natural history of the cetaceans. A review of the family Delphinidae. *Bulletin of the United States National Museum* **36**: 1-191.
<http://dx.doi.org/10.5479/si.03629236.36.1>
- J. R. True & S. B. Carroll** (2002) Gene co-option in physiological and morphological evolution. *Annual Reviews in Cellular Developmental Biology* **18**: 53-80.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12142278 <http://dx.doi.org/10.1146/annurev.cellbio.18.020402.140619>
- J. P. Truman, M. L. Ericson, C. J. Choqueux-Seebold, D. J. Charron & N. A. Mooney** (1994) Lymphocyte programmed cell death is mediated via HLA class II DR. *International Immunology* **6**(6): 887-96.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8086377 <http://dx.doi.org/10.1093/intimm/6.6.887>
- I. Tsur, B. Yakobson, D. Elad, D. Moffett & S. Kennedy** (1997) Morbillivirus infection in a bottlenose dolphin from the Mediterranean Sea. *European Journal of Veterinary Pathology* **3**: 83-85.
- S. Tudela & K. Short** (2005) Paradigm shifts, gaps, inertia, and political agendas in ecosystem-based fisheries management : Politics and socio-economics of ecosystem-based management of marine resources. *Marine Ecology Progress Series* **300**: 282-286.
<http://dx.doi.org/10.3354/meps300282>
- E. Valsecchi, W. Amos, J. A. Raga, M. Podesta & W. Sherwin** (2004a) The effects of inbreeding on mortality during a morbillivirus outbreak in the Mediterranean striped dolphin (*Stenella coeruleoalba*). *Animal Conservation* **7** Part **2**: 139-146.
<http://dx.doi.org/10.1017/S1367943004001325>
- E. Valsecchi, W. Amos, J. A. Raga, M. Podesta & W. Sherwin** (2004b) The effects of inbreeding on mortality during a morbillivirus outbreak in the Mediterranean striped dolphin (*Stenella coeruleoalba*). *Animal Conservation* **7**(2): 139-146.
- M. W. van de Bildt, T. Kuiken & A. D. Osterhaus** (2005) Cetacean morbilliviruses are phylogenetically divergent. *Arch Virol* **150**(3): 577-83.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15747165 <http://dx.doi.org/10.1007/s00705-004-0426-4>

A. H. van der Hout, A. M. van den Ouweland, R. B. van der Luijt, H. J. Gille, D. Bodmer, H. Bruggenwirth, I. M. Mulder, P. van der Vlies, P. Elfferich, M. T. Huisman, A. M. ten Berge, J. Kromosoeto, R. P. Jansen, P. H. van Zon, T. Vriesman, N. Arts, M. B. Lange, J. C. Oosterwijk, H. Meijers-Heijboer, M. G. Ausems, N. Hoogerbrugge, S. Verhoef, D. J. Halley, Y. J. Vos, F. Hogervorst, M. Ligtenberg & R. M. Hofstra (2006) A DGGE system for comprehensive mutation screening of BRCA1 and BRCA2: application in a Dutch cancer clinic setting. *Human Mutation* 27(7): 654-666. <http://dx.doi.org/10.1002/humu.20340>

J. M. Van der Walt, L. H. Nel & A. R. Hoelzel (2001) Characterization of major histocompatibility complex DRB diversity in the endemic South African antelope *Damaliscus pygargus*: a comparison in two subspecies with different demographic histories. *Molecular Ecology* 10(7): 1679-1688. <http://dx.doi.org/10.1046/j.0962-1083.2001.01321.x>

M. Van-Bresseem, K. V. Waerebeek, P. D. Jepson, J. A. Raga, P. J. Duignan, O. Nielsen, A. P. Di-Beneditto, S. Siciliano, R. Ramos, W. Kant, V. Peddemors, R. Kinoshita, P. S. Ross, A. Lopez-Fernandez, K. Evans, E. Crespo & T. Barrett (2001) An insight into the epidemiology of dolphin morbillivirus worldwide. *Veterinary microbiology* 81(4): 287-304. [http://dx.doi.org/10.1016/S0378-1135\(01\)00368-6](http://dx.doi.org/10.1016/S0378-1135(01)00368-6)

P. F. Verhulst (1838) Notice sur la loi que la population suit dans son accroissement. *Correspondance Mathématique et Physique* 10: 113-121. http://books.google.com/books?id=8GsEAAAAYAAJ&lr=&as_drrb_is=b&as_minm_is=0&as_miny_is=1830&as_maxm_is=0&as_maxy_is=1850&as_brr=0&hl=ca&pg=PA113

F. Veyrunes, P. D. Waters, P. Miethke, W. Rens, D. McMillan, A. E. Alsop, F. Grutzner, J. E. Deakin, C. M. Whittington, K. Schatzkamer, C. L. Kremitzki, T. Graves, M. A. Ferguson-Smith, W. Warren & J. A. Marshall Graves (2008) Bird-like sex chromosomes of platypus imply recent origin of mammal sex chromosomes. *Genome research* 18(6): 965-673. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=18463302 <http://dx.doi.org/10.1101/gr.7101908>

P. M. Vitousek, H. A. Mooney, J. Lubchenco & J. M. Melillo (1997) Human Domination of Earth's Ecosystems. *Science* 277(5325): 494. <http://dx.doi.org/10.1126/science.277.5325.494>

G. Voelker (1999) Dispersal, Vicariance, and Clocks: Historical Biogeography and Speciation in a Cosmopolitan Passerine Genus (*Anthus*: Motacillidae). *Evolution* 53(5): 1536-1552. <http://dx.doi.org/10.2307/2640899> <http://dx.doi.org/10.2307/2640899>

M. T. Vossen, E. M. Westerhout, C. Soderberg-Naucler & E. J. Wiertz (2002) Viral immune evasion: a masterpiece of evolution. *Immunogenetics* 54(8): 527-42. <http://dx.doi.org/10.1007/s00251-002-0493-1>

P. J. Waddell, N. Okada & M. Hasegawa (1999) Towards resolving the interordinal relationships of placental mammals. *Systematic biology* 48(1): 1-5. <http://www.jstor.org/stable/2585262> <http://dx.doi.org/10.1093/sysbio/48.1.1>

V. G. Waddell, M. C. Milinkovitch, M. Berube & M. J. Stanhope (2000) Molecular phylogenetic examination of the delphinoidea trichotomy: congruent evidence from three nuclear loci indicates that porpoises (*Phocoenidae*) share a more recent common ancestry with white whales (*Monodontidae*) than they do with true dolphins (*Delphinidae*). *Molecular Phylogenetics and Evolution* 15(2): 314-8. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10837160 <http://dx.doi.org/doi:10.1006/mpev.1999.0751>

A. Wallace (1889) Darwinism: An Exposition of the Theory of Natural Selection. MacMillan, London. 512 pp. <http://books.google.com/books?id=7HhIAAAAMAAJ&hl=ca>

R. Weijermars (1988) Neogene tectonics in the Western Mediterranean may have caused the Messinian Salinity Crisis and an associated glacial event. *Tectonophysics* 148(3-4): 211-19. [http://dx.doi.org/10.1016/0040-1951\(88\)90129-1](http://dx.doi.org/10.1016/0040-1951(88)90129-1)

B. S. Weir (1996) Genetic Data Analysis II: Methods for Discrete Population Genetic Data.

Sinauer Associates, Sunderland, MA. 376 pp. <http://www.ebooksdownloadfree.com/Science-Technology/Genetic-Data-Analysis-2-Methods-for-Discrete-Population-Gen-BI10360.html>

P. A. Weller, R. Critcher, P. N. Goodfellow, J. German & N. A. Ellis (1995) The human Y chromosome homologue of XG: transcription of a naturally truncated gene. *Human molecular genetics* **4**(5): 859-868. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=7633446 <http://dx.doi.org/10.1093/hmg/4.5.859>

K. M. Westover & A. L. Hughes (2001) Molecular evolution of viral fusion and matrix protein genes and phylogenetic relationships among the Paramyxoviridae. *Molecular Phylogenetics and Evolution* **21**(1): 128-134. <http://dx.doi.org/doi:10.1006/mpev.2001.0999> <http://dx.doi.org/10.1006/mpev.2001.0999>

H. Whitehead & J. Mann (2000) Female Reproductive Strategies and Social Bonds in Cetacean Societies: Field Studies of Dolphins and Whales. (R. C. C. J. Mann, P. L. Tyack & H. Whitehead Ed.). University of Chicago Press, Chicago. p 219-246.

R. H. Whittaker & L. Margulis (1978) Protist classification and the kingdoms of organisms. *Biosystems* **10**(1-2): 3-18. <http://www.ncbi.nlm.nih.gov/pubmed/418827> [http://dx.doi.org/10.1016/0303-2647\(78\)90023-0](http://dx.doi.org/10.1016/0303-2647(78)90023-0)

Wikipedia (2008) Cluster analysis, article from Wikipedia, the free online encyclopædia (English version). Online material maintained by at http://en.wikipedia.org/wiki/Data_clustering. Accessed the 2008-10-02.

T. Wilke (2003) Salenthydrobia gen. nov (Rissooidea : Hydrobiidae): a potential relict of the Messinian salinity crisis. *Zoological Journal Of The Linnean Society* **137**(2): 319-336. <http://dx.doi.org/doi:10.1046/j.1096-3642.2003.00049.x> <http://dx.doi.org/10.1046/j.1096-3642.2003.00049.x>

I. J. Wilson & D. J. Balding (1998) Genealogical inference from microsatellite data. *Genetics* **150**(1): 499-510. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9725864

R. Winchester (1994) The molecular basis of susceptibility to rheumatoid arthritis. *Advances in Immunology* **56**: 389-466. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=7521116 [http://dx.doi.org/10.1016/S0065-2776\(08\)60456-3](http://dx.doi.org/10.1016/S0065-2776(08)60456-3)

J. O. Wolff & D. W. Macdonald (2004) Promiscuous females protect their offspring. *Trends in Ecology & Evolution* **19**(3): 127-134. <http://dx.doi.org/10.1016/j.tree.2003.12.009>

B. Worm, E. B. Barbier, N. Beaumont, J. E. Duffy, C. Folke, B. S. Halpern, J. B. C. Jackson, H. K. Lotze, F. Micheli & S. R. Palumbi (2006a) Impacts of Biodiversity Loss on Ocean Ecosystem Services. *Science* **314**(5800): 787. <http://10.1126/science.1132294>

B. Worm, E. B. Barbier, N. Beaumont, J. E. Duffy, C. Folke, B. S. Halpern, J. B. C. Jackson, H. K. Lotze, F. Micheli, S. R. Palumbi, E. Sala, K. A. Selkoe, J. J. Stachowicz & R. Watson (2006b) Impacts of biodiversity loss on ocean ecosystem services. *Science* **314**(5800): 787-790. <http://dx.doi.org/10.1126/science.1132294>

S. Wright (1931) Evolution in Mendelian populations. *Genetics* **16**: 97-159. <http://www.esp.org/foundations/genetics/classical/holdings/w/sw-31.pdf>

B. Würsig, T. R. Kieckhefer & T. A. Jefferson (1990) Visual displays for communication in cetaceans in Sensory abilities of cetaceans. (J. Thomas & R. Kastelein Ed.). Plenum Press, New York. p 545-559.

M. Wurtz & D. Marrale (1991) On the stomach contents of striped dolphins (*Stenella coeruleoalba* Meyen 1933) from the Ligurian coast, central Mediterranean Sea. *European Research on Cetaceans, European Cetacean Society Conference*.

M. Wurtz, R. Poggi & M. R. Clarke (1992) Cephalopods from the stomachs of a risso's

dolphin (*Grampus griseus*) from the Mediterranean. *Journal of the Marine Biological Association of the United Kingdom* **72**(4): 861-867. <http://dx.doi.org/10.1017/S0025315400060094>

Y. Xiong, M. C. Brandley, S. Xu, K. Zhou & G. Yang (2009) Seven new dolphin mitochondrial genomes and a time-calibrated phylogeny of whales. *BMC evolutionary biology* **9**: 20.

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=19166626 <http://dx.doi.org/10.1186/1471-2148-9-20>

A. Xu, M. J. van Eijk, C. Park & H. A. Lewin (1993) Polymorphism in BoLA-DRB3 exon 2 correlates with resistance to persistent lymphocytosis caused by bovine leukemia virus. *Journal of Immunology* **151**(12): 6977-85.

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8258704

G. Yang, J. Yan, K. Zhou & F. Wei (2005) Sequence variation and gene duplication at MHC DQB loci of baiji (*Lipotes vexillifer*), a Chinese river dolphin. *Journal of heredity, The* **96**(4): 310-7. <http://dx.doi.org/10.1093/jhered/esi055>

P. L. Young, K. Halpin, P. W. Selleck, H. Field, J. L. Gravel, M. A. Kelly & J. S. Mackenzie (1996) Serologic evidence for the presence in Pteropus bats of a paramyxovirus related to equine morbillivirus. *Emerging Infectious Diseases* **2**(3): 239-40.

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8903239 <http://dx.doi.org/10.3201/eid0203.960315>

N. Yuhki & S. J. O'Brien (1997) Nature and origin of polymorphism in feline MHC class II DRA and DRB genes. *Journal of Immunology* **158**(6): 2822-33.

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9058818

L. Zane, L. Ostellari, L. Maccatrozzo, L. Bargelloni, J. Cuzin Roudy, F. Buchholz & T. Patarnello (2000) Genetic differentiation in a pelagic crustacean (*Meganyctiphanes norvegica*: Euphausiacea) from the North East Atlantic and the Mediterranean Sea. *Marine biology* **136**(2): 191-199. <http://dx.doi.org/10.1007/s002270050676>

J. A. Zeh & D. W. Zeh (1996) The Evolution of Polyandry I: Intra-genomic Conflict and Genetic Incompatibility. *Proceedings of the Royal Society of London Series B Biological Sciences* **263**(1377): 1711-1717. <http://dx.doi.org/10.1098/rspb.1996.0250>

A. Zenetos, M. E. Çınar, M. A. Pancucci-Papadopoulou, J. G. Harmelin, G. Furnari, F. Andaloro, N. Bellou, N. Streftaris & H. Zibrowius (2005) Annotated list of marine alien species in the Mediterranean with records of the worst invasive species. *Mediterranean Marine Science* **6**(2): 63. http://www.medit-mar-sc.net/contents/vol_6_2/a625.html

K. v. Zittel (1879) *Handbuch der Paläontologie*. Vol. 1. R. Oldenbourg,, München und Leipzig. 765 pp. <http://books.google.com/books?id=o7HZOgAACAAJ&hl=ca>

H. R. Zornetzer & D. A. Duffield (2003) Captive-born bottlenose dolphin x common dolphin (*Tursiops truncatus* x *Delphinus capensis*) intergeneric hybrids. *Canadian Journal of Zoology* **81**(10): 1755-1762. http://article.pubs.nrc-cnrc.gc.ca/ppv/RPViewDoc?_handler_=HandleInitialGet&journal=cjz&volume=81&calyLang=eng&articleFile=z03-150.pdf

<http://dx.doi.org/10.1139/z03-150>

Résumé

Le Dauphin bleu et blanc est l'espèce de Cétacé majoritaire en Méditerranée. L'espèce a subi une épizootie sévère à Morbillivirus en 1990-91 en Méditerranée uniquement.

L'objectif de ce travail était d'étudier l'histoire évolutive de cette espèce en Méditerranée et d'essayer de la lier à cet épisode infectieux. Pour cela, des marqueurs mitochondriaux (séquence de la région de contrôle), microsatellites, du chromosome Y (amélogénine) et du Complexe Majeur d'Histocompatibilité (DRB) ont été étudiés.

Dans un premier temps, l'analyse phylogénique de l'amélogénine comparée au sein des Cétartiodactyla a permis de montrer l'existence d'une synapomorphie (insertion) pouvant être utilisée à la fois en biologie évolutive et pour le sexage.

Dans un deuxième temps, la variabilité et de la région de contrôle et des microsatellites a été étudiée par statistique fréquentiste (F-statistiques) et bayésienne : coefficients individuels de mélange – admixture coefficients, flux de gènes entre groupes, estimation des temps de coalescence des divers clades avec et sans horloge moléculaire. Il en ressort en particulier que deux groupes fortement différenciés coexistent en Méditerranée. La divergence de ces groupes est ancienne (Miocène) et pourrait être due à une divergence allopatrique dont l'événement vicariant serait l'isolement de la Méditerranée lors de la Crise de Salinité du Messinien. Finalement, le DRB a été étudié ce qui constitue une première dans cette espèce sachant que cette molécule a fait l'objet de peu de travaux chez les Cétacés. Cette étude apporte des éléments descriptifs en termes de diversité et de profils de sélection. Le DRB s'avère finalement plus variable qu'initialement décrit dans diverses espèces de Mammifères marins.

Ce travail ouvre à la fois des perspectives quant aux mécanismes endogènes et exogènes permettant la coexistence sympatrique de populations/sous-espèces et en termes de médecine évolutive (susceptibilité accrue à des pathogènes). Le Dauphin bleu et blanc pourrait en conséquence s'avérer un bon modèle pour étudier les rapports entre évolution et pathologie.