



Intérêts et limites de l'approche moléculaire pour aborder la biogéographie et la spéciation : l'exemple de quelques Mammifères d'Afrique tropicale

Sophie Quéroutil

► To cite this version:

Sophie Quéroutil. Intérêts et limites de l'approche moléculaire pour aborder la biogéographie et la spéciation : l'exemple de quelques Mammifères d'Afrique tropicale. Ecologie, Environnement. Université Rennes 1, 2001. Français. NNT : . tel-00138122

HAL Id: tel-00138122

<https://theses.hal.science/tel-00138122>

Submitted on 23 Mar 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

N° Ordre
de la thèse : 2530

THESE
présentée
DEVANT L'UNIVERSITE DE RENNES 1
pour obtenir
le grade de : **DOCTEUR DE L'UNIVERSITE DE RENNES 1**
Mention : Biologie

PAR
Sophie QUEROUIL

Equipe d'accueil : UMR 6552 CNRS, Station Biologique de Paimpont
Ecole doctorale : Vie-Agro-Santé
Composante universitaire : UFR SVE

**INTERETS ET LIMITES DE L'APPROCHE MOLECULAIRE POUR ABORDER
LA BIOGEOGRAPHIE ET LA SPECIATION :
QUELQUES EXEMPLES CHEZ LES MAMMIFERES D'AFRIQUE TROPICALE**

SOUTENUE LE 10 Décembre 2001 devant la commission d'examen

COMPOSITION DU JURY :

Mme Christiane Denys - Pr, Muséum National d'Histoire Naturelle, Paris - Rapporteur
M. Daniel Lachaise - DR CNRS, Gif sur Yvette - Rapporteur
M. Jacques Blondel - DR CNRS, CEF, Montpellier - Examineur
M. Jean-Sébastien Pierre - Pr, Université de Rennes 1, Rennes - Examineur
M. Erik Verheyen - Chercheur, Institut Royal des Sciences Naturelles de Belgique, Bruxelles -
Examineur
Mme Annie Gautier-Hion - DR CNRS, Université de Rennes 1, Paimpont - Directeur

INTERETS ET LIMITES DE L'APPROCHE MOLECULAIRE POUR ABORDER LA BIOGEOGRAPHIE ET LA SPECIATION : QUELQUES EXEMPLES CHEZ LES MAMMIFERES D'AFRIQUE TROPICALE

RESUME

Parmi les approches qui visent à déterminer la façon dont les organismes se diversifient dans le temps et dans l'espace, la biogéographie tente de reconstruire l'histoire des peuplements à partir des distributions des taxons, tandis que la phylogéographie analyse l'évolution d'une lignée à la fois. Dans cette étude, nous nous sommes proposés de tester les hypothèses biogéographiques formulées pour la faune tropicale (théorie des refuges, des barrières fluviales, des gradients environnementaux...), en appliquant l'approche moléculaire à quelques taxons de Mammifères africains (Insectivores, Rongeurs, Primates). Nous avons cherché à déterminer l'apport et les limites de la technique moléculaire la plus couramment employée en biologie évolutive : le séquençage d'ADN mitochondrial.

Nous avons d'abord tenté d'obtenir une phylogénie moléculaire de quelques taxons potentiellement intéressants pour la biogéographie, dans le but de vérifier leur monophylie et de calibrer une horloge moléculaire. Puis, nous avons recherché et comparé les schémas phylogéographiques de quatre espèces de petits mammifères forestiers et d'une super-espèce de primates. Enfin, nous nous sommes intéressés aux processus évolutifs impliqués dans la spéciation. Nous avons évalué le mode géographique de spéciation et l'évolution de quelques traits d'histoire de vie chez les primates de la tribu des Cercopithecini.

Nos résultats phylogénétiques confirment que l'histoire des gènes n'est pas forcément celle des taxons et qu'il est important de prendre en compte plusieurs sources d'information indépendantes, telles que des gènes non liés sur la même molécule, la morphologie, l'écologie, et le comportement. L'étude des modalités de la spéciation, qui est tributaire de la fiabilité des analyses phylogénétiques, indique une prédominance de l'allopatrie et des événements de vicariance du Miocène et du Pliocène chez les Cercopithecini. Enfin, les analyses phylogéographiques ont révélé quatre scénarios phylogéographiques différents pour les quatre modèles retenus, ce qui peut refléter soit des distributions initiales différentes, soit une réponse différentielle aux mêmes événements selon les taxons. Ces scénarios présentent une certaine concordance avec les régions fauniques définies pour les forêts d'Afrique centrale, mais suggèrent que les événements de divergence intra-spécifique remonteraient au Plio-Pleistocène et seraient donc beaucoup plus anciens que les derniers épisodes de fragmentation de la forêt liés aux cycles glaciaires.

MENTION : Biologie

MOTS-CLES : Biogéographie, Phylogéographie, Spéciation, Insectivores, Rongeurs, Primates, Afrique tropicale

ADRESSE DU LABORATOIRE : UMR 6552 CNRS, Université Rennes 1, Station Biologique, 35380 PAIMPONT

POWERS AND LIMITATIONS OF THE MOLECULAR APPROACH FOR THE STUDY OF BIOGEOGRAPHY AND SPECIATION : EXAMPLES IN TROPICAL AFRICAN MAMMALS

SUMMARY

Among the sciences which study the patterns of diversification and distribution of organisms in space and time, biogeography investigates the history of biota based on organism distributions, and phylogeography focuses on one taxa at the time. Biogeographic studies carried out in tropical areas have led to a series of evolutionary hypotheses about faunal diversification and speciation, such as the refuge, the riverine barrier, and the environmental gradient theories. In the present work, we attempt to test these hypotheses by applying the molecular approach to selected mammalian taxa (Insectivora, Rodentia, and Primates) from tropical Africa. We try to estimate the interests and limits of mitochondrial DNA sequencing, the most widely used molecular technique in evolutionary biology, for addressing questions about faunal evolution.

First, we constructed a molecular phylogeny of taxa we intended to use for phylogeographic studies, in order to ascertain their monophyly and calibrate a molecular clock for divergence time estimates. Second, we analysed and compared the phylogeographic patterns of four forest-dwelling small mammal species and one primate super-species. Third, we evaluated the evolutionary processes involved in the speciation of cercopithecine primates, by testing their geographic mode of speciation and reconstructing evolutionary scenarios for some life-history traits.

The comparison of our molecular-based phylogenies with other sources of information confirms that gene history is not necessarily the same as organism history. Thus, mitochondrial DNA should be studied in combination with other independent data, such as nuclear genes, morphology, ecology and behaviour. The acquisition of reliable phylogenies is a pre-requisite for the study of speciation. In the case of the Cercopithecini, our results suggest that speciation has been predominantly allopatric and driven by Miocene and Pliocene vicariance events. The obtained phylogeographic patterns consistently suggest a role of Pliocene and Pleistocene vicariance events in the intra-specific diversification of small mammals. All phylogeographic patterns are different, which may be explained by differences in initial geographic and temporal conditions, or by different responses to the same events. However, there is a certain degree of consistency between these patterns and the faunal areas defined for the central African forest. The examination of additional taxa could potentially unveil a general picture despite individual differences.

MENTION : Biology

KEY-WORDS : Biogeography, Phylogeography, Speciation, Insectivora, Rodentia, Primates, tropical Africa

LABORATORY ADDRESS : UMR 6552 CNRS, University of Rennes 1, Biological Station, 35380 PAIMPONT

Orbital Consequences

The sun and the earth describe orbital changes
which drive climate cycles and modify ranges.
The shape of the land forms a number of places
that allow the survival of different races.
When enclaves advance with the ice in retreat
some form hybrid zones where two ranges meet.
Such regions are common and not very wide
so the mixing of genes affects neither side.
They divide up the range in a patchwork of pieces
with echoes and glimpses on the nature of species.
A brief rendezvous and the ice comes again.

When the glaciers melt so that ranges expand
some plants will spread quickly where there's suitable
land.
Those insects which eat them will follow this lead
some flying, some walking to establish their breed.
Those that try later meet a resident band,
they must somehow be better to make their own stand.
But the mixture will change as more types arrive
and warming conditions allow new species to thrive.
Some will move on to fresh places ahead,
those that remain must adapt, or are dead.
And then the tide turns and the ice comes again.

Each refuge could foster a deviant form,
new neighbours, chance changes and drift from the
norm.
When the warm breakout comes, those few in the van
disperse from the edge and breed where they can.
Pioneer pockets grow to large populations,
a very good place to strike new variations.
Some may not work well with their parental kind
so stopping the spread of those from behind.
Continental theatres provide plenty of chances
to establish new morphs in both retreats and advances.
New species may form when the ice comes again.

So what will you do when the ice comes again?
It could be quite quick, if the ice cores speak plain.
The great ocean currents that warm our green spring
may stop in a season should the salt balance swing.
Great civilizations in north temperate lands
must migrate south to the sun and the sands.
But past pollen and dust tell us these will be drier,
wet forests will shrink and population grow higher.
Our forebears hung on near a sea or a cave.
They fished and they painted, they dreamed, they were
brave .
So like Noah and Eric, we must adapt and survive.

G. M. Hewitt (2001)

REMERCIEMENTS

Je tiens à remercier tout particulièrement Marc Colyn, pour m'avoir proposé ce sujet de thèse et avoir co-dirigé mes recherches, et Erik Verheyen, pour m'avoir accueillie pendant deux ans au Laboratoire de Taxinomie et Systématique Biochimique de l'Institut Royal des Sciences Naturelles de Belgique et avoir co-encadré mes travaux. Je remercie Annie Gautier-Hion pour m'avoir accueillie à la Station Biologique de Paimpont au sein de l'UMR 6552 et éclairée de ses conseils avisés, depuis le DEA jusqu'à la rédaction de ce manuscrit. J'exprime toute ma reconnaissance à Pierre Deleporte pour ses réflexions théoriques et méthodologiques, ainsi qu'à Rainer Hutterer et Walter Verheyen, pour leur aide précieuse en matière de taxinomie des musaraignes et des rongeurs.

Je remercie très sincèrement les rapporteurs, Christiane Denys et Daniel Lachaise, et les autres membres du jury, Jacques Blondel, Jean-Sébastien Pierre et Erik Verheyen, qui ont accepté d'examiner ce travail.

Ces travaux n'auraient pas pu voir le jour sans tous ceux qui ont collecté des spécimens sur le terrain : Olivier Perpète, Patrick Barrière, Violaine Nicolas, Sylvain Dufour, et bien d'autres qui nous ont envoyé du matériel. Ils ont été financés par (1) le programme DSPT.5 – action Biodiversité (Biologie, Médecine et Santé) du Ministère de l'Enseignement Supérieur et de la Recherche ; (2) le programme Ecofac EU-DGVIII "Conservation et Utilisation Rationnelle des Ecosystèmes Forestiers en Afrique Centrale" (dirigé par AGRECO, GEIE, BDPA, SCETAGRI, SECA, CIRAD-FORET) ; (3) le "Programme de Développement de la Région Nord" EU-DGVIII-PDRN (NORCADEV) ; (4) le programme Biofac EU-DGVIII "Origine et maintien de la biodiversité en Afrique Centrale" (Université de Rennes 1 / CNRS, UMR 6552) ; (5) le projet WWF Dzanga-Sangha Reserve ; (6) le projet Ebola - Forêt de Taï, OMS Abidjan. Le travail de laboratoire a été financé en partie par le projet DWTC 31.64 intitulé "Evaluating the Effect of Pleistocene Climate Changes on Speciation Patterns in African Vertebrates".

Je remercie également Eric Petit pour ses commentaires scientifiques, Tine Dillen pour sa participation au séquençage, et Yves Picard pour son assistance technique. Merci aussi à Alexandre Hassanin et Simon Tillier qui m'ont reçue au Service de Systématique Moléculaire du Museum National d'Histoire Naturelle de Paris.

Enfin, j'adresse toute ma sympathie aux parents, collègues, amis, et copains qui m'ont soutenue pendant ces années : Marjorie, Neuciane, Dominique, Philippe, Patrick, Violaine, Dany, David, Stéphane, Jean-Marie et tous les autres... sans oublier Rémi qui n'aura pas réussi à m'épargner les moments difficiles d'une fin de thèse, mais qui me les aura rendus plus doux.

Pour finir, je dédie ce manuscrit à Maude, partie visiter le monde des Fées pour ne plus se heurter à celui des Hommes.

TABLE DES MATIERES

INTRODUCTION GENERALE	1
CHAPITRE 1 : CADRE THÉORIQUE DE L'ÉTUDE	5
I - BIOGEOGRAPHIE, PHYLOGEOGRAPHIE ET EVOLUTION DES ESPECES	5
1) Biogéographie et phylogéographie	5
2) Les moteurs de l'évolution	7
a) <i>Vicariance et dispersion</i>	7
b) <i>Expansion de l'aire de distribution et zones de contact secondaire</i>	9
c) <i>Sélection naturelle et compétition</i>	9
d) <i>Isolement reproductif</i>	10
3) Théories de l'évolution	11
a) <i>La théorie des gradients environnementaux</i>	11
b) <i>La théorie paléogéographique</i>	12
c) <i>La théorie des barrières fluviales</i>	12
d) <i>La théorie des refuges</i>	13
e) <i>La théorie des perturbations</i>	15
II - LE CAS DE L'AFRIQUE TROPICALE	16
1) Choix de la zone d'étude	16
2) Bilan des travaux biogéographiques portant sur les forêts d'Afrique centrale et occidentale	18
a) <i>Reconnaissance de régions fauniques</i>	18
b) <i>Barrières</i>	21
c) <i>Refuges</i>	21
d) <i>Variabilité des schémas de distribution selon les taxons</i>	22
3) Paléoclimats et conditions paléo-environnementales en Afrique tropicale	24
a) <i>Indices révélateurs des conditions paléoclimatiques et paléo-environnementales</i>	24
b) <i>Dynamique globale et répercussions en Afrique tropicale</i>	25
4) Chronologie des conditions paléo-environnementales en Afrique tropicale	27
a) <i>Secondaire et tertiaire</i>	27
b) <i>Pléistocène</i>	28
c) <i>Période récente (quarante derniers milliers d'années)</i>	29
5) Fragmentation forestière et évolution de la faune en Afrique tropicale	29
a) <i>Informations apportées par les fossiles</i>	29
b) <i>Preuves de l'effet de la fragmentation forestière actuelle sur la faune</i>	31
CHAPITRE 2 : MATÉRIEL ET CHOIX DES MÉTHODES	32
I - ZONE D'ETUDE	32
1) Situation géographique	32
2) Caractéristiques géophysiques de l'Afrique tropicale	32
a) <i>Géomorphologie</i>	32
b) <i>Hydrologie</i>	34
c) <i>Géologie et pédologie</i>	34
d) <i>Climatologie</i>	36

e) <i>Végétation</i>	40
f) <i>Occupation humaine</i>	43
3) Sites de collectes	43
II - MATERIEL BIOLOGIQUE	46
1) Choix des modèles	46
2) Matériel biologique utilisé	54
III - ACQUISITION DES DONNEES MOLECULAIRES	55
1) Choix de la technique moléculaire	55
a) <i>Les différentes techniques moléculaires</i>	55
b) <i>ADN mitochondrial ou ADN nucléaire ?</i>	56
c) <i>Choix des séquences mitochondriales</i>	58
2) Protocoles employés	61
a) <i>Extraction</i>	61
b) <i>Amplification</i>	62
c) <i>Purification</i>	63
d) <i>Séquençage</i>	64
IV - RECONSTRUCTIONS PHYLOGENETIQUES	66
1) Alignement	66
2) Les différentes méthodes employées en phylogénie	67
a) <i>Méthodes phénétiques (ou "de distances")</i>	67
b) <i>Méthodes basées sur les caractères</i>	68
c) <i>Comment choisir ?</i>	72
3) Amélioration de la qualité de la reconstruction	73
a) <i>Le problème de l'homoplasie et la saturation des substitutions</i>	74
b) <i>Le problème de l'évolution mosaïque</i>	76
c) <i>Le problème des longues branches</i>	77
d) <i>Le codage des indels</i>	77
e) <i>Le choix des groupes externes</i>	78
f) <i>Plus de taxons ou plus de nucléotides ?</i>	79
4) Combinaison de plusieurs jeux de données	79
a) <i>Analyses séparées et "super-arbres"</i>	79
b) <i>Analyse combinée</i>	80
5) Test qualitatif de l'arbre obtenu	81
a) <i>Tests de la fiabilité des branches</i>	81
b) <i>Mesure de la qualité globale de l'arbre</i>	82
c) <i>Tests d'arbres alternatifs</i>	83
6) Datation	84
a) <i>Principe d'horloge moléculaire</i>	84
b) <i>Linéarisation des arbres</i>	85
c) <i>Datation</i>	86
7) Reconstruction de scénarios évolutifs	86
V - ANALYSES PHYLOGEOGRAPHIQUES ET BIOGEOGRAPHIQUES	88
1) Phylogéographie	88
a) <i>Construction d'arbres des haplotypes</i>	88
b) <i>Reconstitution de l'histoire évolutive d'une lignée</i>	90
c) <i>Autres méthodes</i>	91
2) Phylogéographie comparée	91

3) Biogéographie cladistique	92
a) Définition des relations entre aires	92
b) Test de compatibilité entre un arbre théorique des aires et un arbre observé	95
c) Recherche des conditions géographiques ancestrales	95
d) Test des modalités géographiques de spéciation	96
CHAPITRE 3: ANALYSES PHYLOGÉNÉTIQUES	98
I - PHYLOGENIE DES MUSARAIGNES AFRICAINES (ARTICLE 1)	98
II - PHYLOGENIE DU GENRE <i>HYLOMYSCUS</i> , MURIDES FORESTIERS AFRICAINS	111
1) Matériel et méthodes	112
a) détermination spécifique	112
b) Phylogénie du genre <i>Hylomyscus</i>	114
2) Résultats	115
a) Détermination spécifique	115
b) Phylogénie du genre <i>Hylomyscus</i>	118
3) Discussion	124
CHAPITRE 4: PHYLOGÉOGRAPHIE	129
I - PHYLOGEOGRAPHIE COMPAREE DE DEUX MUSARAIGNES	129
1) Phylogéographie comparée de <i>Sylvisorex johnstoni</i> et <i>S. ollula</i> (ARTICLE 2)	129
2) Analyse complémentaire : analyse des groupements emboîtés	155
a) Matériel et méthodes	155
b) Résultats	155
c) Discussion	160
II - PHYLOGEOGRAPHIE COMPAREE DE DEUX RONGEURS	161
1) Matériel et méthodes	161
2) Résultats	163
3) Discussion	173
III - ETUDE DE LA SUPER-ESPECE <i>CERCOPITHECUS CEPHUS</i>	177
1) Matériel et méthodes	178
2) Résultats	179
3) Discussion	183
IV - COMPARAISON DES MODELES DANS UNE PERSPECTIVE BIOGEOGRAPHIQUE	186
1) Matériel et méthodes	186
2) Résultats	187
3) Discussion	197
CHAPITRE 5: EVOLUTION DE TRAITS D'HISTOIRE DE VIE ET SPECIATION	201
I - MODE GEOGRAPHIQUE DE SPECIATION CHEZ LES CERCOPITHEQUES (ARTICLE 3)	201
II - PHYLOGENIE ET EVOLUTION DE L'ARBORICOLIE CHEZ LES CERCOPITHEQUES (ARTICLE 4)	217

CHAPITRE 6 : SYNTHÈSE ET PERSPECTIVES	252
I - LIMITES METHODOLOGIQUES DES ANALYSES PHYLOGENETIQUES	252
1) Limites des données moléculaires, et autres sources potentielles d'information	252
2) Traitement des données moléculaires	253
a) <i>Comment obtenir une phylogénie plus fiable ?</i>	253
b) <i>Comment affiner les datations ?</i>	254
II - DE LA PHYLOGEOGRAPHIE A LA BIOGEOGRAPHIE	255
1) Diversité des schémas phylogéographiques	255
2) Tentative de synthèse biogéographique	257
3) Comparaison avec les hypothèses biogéographiques actuelles et avec d'autres modèles	257
III - MODALITES DE LA SPECIATION	259
1) Spéciation sympatrique ou allopatrique ?	259
2) Ancienneté des événements de spéciation et de divergence intra-spécifique	260
3) Retour sur les différentes théories de l'évolution	262
IV - APPLICATIONS ET PERSPECTIVES	264
1) Intérêt pour la conservation	264
a) <i>Diversité morphologique ou diversité génétique ?</i>	264
b) <i>Comment conserver la biodiversité d'Afrique centrale à long terme ?</i>	265
2) Implications biogéographiques	266
a) <i>Choix des modèles</i>	266
b) <i>Echantillonnage géographique</i>	267
c) <i>Approche multidisciplinaire</i>	268
REFERENCES BIBLIOGRAPHIQUES	270
GLOSSAIRE	292
ANNEXES	295

INDEX DES ANNEXES

Annexe 1. Echelle des temps géologiques	295
Annexe 2. Echantillons demandés à l'extérieur	296
Annexe 3. Protocoles d'extraction, amplification et séquençage	297
Annexe 4. Amorces utilisées	302
Annexe 5. Clef d'interprétation pour l'analyse des groupements emboîtés	304
Annexe 6. Réseaux parcimonieux d'haplotypes obtenus pour <i>S. johnstoni</i>	307
Annexe 7. Réseau parcimonieux d'haplotypes obtenu pour <i>S. ollula</i>	308
Annexe 8. Réseaux parcimonieux d'haplotypes obtenus pour <i>H. stella</i>	309
Annexe 9. Réseaux parcimonieux d'haplotypes obtenus pour <i>S. longicaudatus</i>	310

INDEX DES FIGURES

Fig. 1.1. Délimitation des zones d'étude A et B sur la carte de l'Afrique politique	17
Fig. 1.2. Régions fauniques et refuges forestiers basés sur la distribution des Mammifères, d'après Happold (1996), adapté de Happold et Delany (1971)	19
Fig. 1.3. Principales unités biogéographiques d'Afrique centrale et leur subdivision en régions et sous-régions fauniques, d'après M. Colyn (comm. pers.)	20
Fig. 1.4. Refuges forestiers envisagés pour le dernier maximum glaciaire (18.000 ans B.P.), d'après Maley (1996)	22
Fig. 1.5. Parallèle entre les fluctuations environnementales (forêts / savanes) et les cycles glaciaires, emprunté à deMenocal (1995).	25
Fig. 1.6. Reconstitution hypothétique de l'évolution du couvert forestier en Afrique au cours des derniers 18.000 ans, d'après Adams et Faure (1997)	30
Fig. 2.1. Le relief de l'Afrique	33
Fig. 2.2. Réseau hydrographique de l'Afrique, limites de la forêt tropicale et zone de mosaïque forêt-savane	35
Fig. 2.3. Les climats de l'Afrique	38
Fig. 2.4. Les phytorégions d'Afrique tropicale (Underwood <i>et al.</i> , 1998)	41
Fig. 2.5. Végétation de l'Afrique centrale (d'après Letouzey, 1968)	42
Fig. 2.6. Principales localités de collecte des spécimens analysés	44
Fig. 2.7. Organisation de la molécule d'ADN mitochondrial chez les Vertébrés (Simon <i>et al.</i> , 1994)	58
Fig. 2.8. Structure secondaire de la seconde moitié de l'ARN-r 16s chez l'Homme, d'après Gutell et Fox (1988)	60
Fig. 3.1. Comparaison entre arbres NJ basés sur l'ARNr 16s et le cytochrome <i>b</i> , pour l'ensemble des spécimens du genre <i>Hylomyscus</i>	116
Fig. 3.2. Courbes de saturation de l'ARN 16s et du cytochrome <i>b</i> , pour une sélection de spécimens du genre <i>Hylomyscus</i>	117
Fig. 3.3. Phylogénie du genre <i>Hylomyscus</i> basée sur l'ARN 16s : arbres obtenus par différentes méthodes, avec support bootstrap	119
Fig. 3.4. Phylogénie du genre <i>Hylomyscus</i> basée sur le cytochrome <i>b</i> : arbres obtenus par différentes méthodes, avec support bootstrap	121
Fig. 3.5. Phylogénie du genre <i>Hylomyscus</i> basée sur l'ARN 16s + le cytochrome <i>b</i> : arbres obtenus par différentes méthodes, avec support bootstrap	123
Fig. 4.1. Définition des groupements emboîtés sur l'arbre des haplotypes obtenu pour <i>S. johnstoni</i>	156
Fig. 4.2. Définition des groupements emboîtés sur l'arbre des haplotypes obtenu pour <i>S. ollula</i>	157
Fig. 4.3. Localités de collecte des spécimens de <i>H. stella</i> et <i>S. longicaudatus</i> , et unités, régions et sous-régions fauniques d'après M. Colyn	162

Fig. 4.4. Boxplots représentant les distances intra- et inter- localités pour <i>H. stella</i> et <i>S. longicaudatus</i>	164
Fig. 4.5. Arbre d'étendue minimale obtenu pour <i>H. stella</i>	166
Fig. 4.6. Arbre d'étendue minimale obtenu pour <i>S. longicaudatus</i>	166
Fig. 4.7. Arbre théorique des aires, et arbres MP non contraints avec valeurs de bootstrap pour <i>H. stella</i> et <i>S. longicaudatus</i>	167
Fig. 4.8. Définition des groupements emboîtés sur l'arbre des haplotypes obtenu pour <i>H. stella</i>	169
Fig. 4.9. Définition des groupements emboîtés sur l'arbre des haplotypes obtenu pour <i>S. longicaudatus</i>	170
Fig. 4.10. Relations phylogénétiques au sein de la super-espèce <i>C. cephus</i>	181
Fig. 4.11. Arbres des haplotypes de la super-espèce <i>C. cephus</i>	182
Fig. 4.12. Comparaison des groupements phylogénétiques observés chez les quatre modèles avec les phytorégions définies par Letouzey (1968)	188
Fig. 4.13. Traduction des événements de vicariance ayant affecté trois des modèles sous forme de séquence linéaire, selon la méthode de Hovenkamp	190
Fig. 4.14. Événements cladogénétiques partagés par au moins deux modèles, mis en évidence en appliquant la méthode de Hovenkamp, et régions fauniques redéfinies à l'issue des analyses biogéographiques	191
Fig. 4.15. Synthèse des résultats phylogéographiques obtenus pour les quatre modèles	188
Fig. 4.16. Arbres NJ avec longueur des branches exprimées en nombre de substitutions, pour chacun des quatre modèles	191
Fig. 4.17. Diagramme de Venn montrant la fraction de la diversité totale partagée par plusieurs régions ou unique à chaque région, et sa traduction sous forme de cladogramme des aires	193

INDEX DES DOCUMENTS

Doc. 2.1. <i>Sylvisorex johnstoni</i> (Dobson, 1888)	48
Doc. 2.2. <i>Sylvisorex ollula</i> Thomas, 1913	49
Doc. 2.3. <i>Hylomyscus stella</i> (Thomas, 1911)	50
Doc. 2.4. <i>Stochomys longicaudatus</i> (Tullberg, 1893)	51
Doc. 2.5. Super-espèce <i>Cercopithecus cephus</i> Linnaeus, 1758	52

INTRODUCTION GENERALE

Pour comprendre l'évolution du monde vivant, il est nécessaire de connaître la façon dont les organismes se diversifient dans le temps et dans l'espace. C'est l'objectif de la biogéographie, discipline qui étudie la distribution des taxons à l'échelle des peuplements, repère les caractères communs à ces distributions, et cherche à identifier les processus environnementaux, biologiques et historiques qui les ont façonnés (Myers et Giller, 1988). Une approche plus récente, baptisée phylogéographie, consiste à étudier l'évolution d'une lignée à la fois, en intégrant une dimension phylogénétique (Avice *et al.*, 1987). La comparaison des histoires individuelles ainsi obtenues tend à rejoindre les objectifs de la biogéographie.

La diversification des espèces a longtemps été étudiée en observant comment celles-ci se distribuent actuellement dans l'espace, et en utilisant les données chronologiques fournies par les fossiles, lorsqu'ils existent. Depuis quelques décennies, l'approche moléculaire permet un regard direct sur l'histoire des taxons, telle qu'elle est recelée par leur matériel génétique. Cette approche présente des avantages pratiques et théoriques (e.g. Avice, 2000). En particulier, elle génère des caractères supposés indépendants des conditions environnementales, et dont l'évolution peut avoir été suffisamment régulière pour qu'il soit possible de calibrer une "horloge moléculaire" et de dater les événements de divergence. La technique la plus prisée actuellement est celle du séquençage d'ADN mitochondrial, car elle permet d'obtenir rapidement un grand nombre de caractères variables, en s'affranchissant des problèmes de recombinaison et de polymorphisme individuel propres à l'ADN nucléaire.

Les études biogéographiques et phylogéographiques réalisées dans diverses régions du monde ont permis de proposer des hypothèses évolutives concernant la flore et la faune. La plupart des modèles évolutifs attribuent un rôle majeur à la fragmentation des aires de distribution (allopatrie). Plusieurs de ces modèles font intervenir les changements paléo-environnementaux liés aux mouvements tectoniques (théorie paléogéographique - Emsley, 1965) et aux oscillations climatiques (théorie des refuges - Haffer, 1969, qui attribue un rôle majeur aux zones de stabilité environnementale, et théorie des perturbations - Colinvaux, 1993, Bush, 1994, qui considère au contraire les modifications de l'habitat dans les zones d'écotone). Un autre modèle évolutif met l'accent sur la notion de barrière physique à la dispersion (théorie des barrières fluviales - Capparella, 1991, Ayres et Clutton-Brock, 1992). Des études récentes ont réhabilité le concept de diversification en présence de flux génique (sympatrie). Notamment, en ce qui concerne la faune, les variations locales de l'environnement floristique pourraient être suffisantes pour

initier une diversification des taxons par sélection naturelle (théorie des gradients environnementaux - Endler, 1982).

La faune des régions tropicales a suscité de nombreux travaux et hypothèses en raison de sa grande diversité. Elle connaît un regain d'intérêt actuellement, en raison de la déforestation croissante liée aux activités anthropiques. Cependant, l'essentiel des travaux récents qui ont permis d'alimenter le débat sur les modes de spéciation en milieu tropical se sont focalisés sur l'Amérique du Sud (Moritz *et al.*, 2000). La plupart des études réalisées en Afrique tropicale ne sont basées que sur la distribution des taxons et la reconnaissance de zones d'endémisme et de diversité élevée (e.g. Carcasson, 1964, Moreau, 1969, Grubb, 1982, Mayr et O'hara, 1986, Colyn, 1991). Elles se heurtent au fait que la taxinomie et la distribution des espèces sont souvent mal connues. De plus, il leur manque une dimension historique car elles ne prennent pas en compte les relations phylogénétiques entre taxons. Des travaux intégrant la notion de parenté entre assemblages faunistiques ont été menés par notre équipe sur la faune mammalienne d'Afrique centrale (Deleporte et Colyn, 1999 ; Colyn et Deleporte, sous presse). Ils ont permis d'affiner le découpage en régions fauniques qui avait été proposé auparavant. Ils ont révélé une structure complexe, avec un emboîtement de régions et de sous-régions fauniques séparées par des zones de mélange des faunes, qui semble pouvoir être expliquée par la théorie des refuges, combinée à celle des barrières fluviales.

Dans cette thèse, nous nous proposons d'évaluer l'intérêt de l'approche moléculaire pour tester les hypothèses évolutives proposées pour la faune forestière tropicale, en nous focalisant sur les forêts de plaine africaines. Nous avons privilégié une région de l'Afrique centrale pour laquelle nous disposons d'un échantillonnage conséquent et qui semble présenter à elle seule toute la complexité qui peut être observée à plus large échelle. Nous avons choisi nos modèles biologiques parmi trois Ordres de Mammifères : Insectivores, Rongeurs et Primates, parce qu'ils constituent un des sujets d'étude majeurs de notre équipe, et que leur taxinomie et leur distribution sont de mieux en mieux connues. Les Primates sont moins bien représentés dans nos collections, mais ont été beaucoup plus étudiés que les petits mammifères sur le plan biologique et écologique. Nous les avons retenus pour aborder plus en détail les mécanismes de la spéciation.

Nos objectifs s'organisent autour de deux pôles. Dans un premier temps, nous étudierons les variations géographiques de la diversité génétique pour quelques modèles, et chercherons quelles conclusions biogéographiques peuvent être tirées de la comparaison de ces schémas phylogéographiques. La détermination spécifique de certains des modèles présélectionnés étant délicate,

nous procéderons au préalable à des révisions taxinomiques et phylogénétiques. Chacun des schémas obtenus sera analysé individuellement, et comparé d'une part avec les barrières physiques et écologiques actuelles, et d'autre part avec les schémas plus généraux observés à l'échelle des peuplements. En particulier, nous chercherons s'il existe une corrélation avec les régions fauniques reconnues en Afrique centrale et avec les hypothèses qu'elles impliquent (refuges, barrières fluviales et corridor savanicole). Nous nous intéresserons aux mécanismes responsables des événements de divergence (vicariance, dispersion, isolement par la distance) et à la dimension temporelle de ces événements. Ensuite, nous tenterons de synthétiser l'information contenue dans les différents schémas phylogéographiques. Nous chercherons à savoir si la majorité des taxons a réagi de la même façon aux contraintes environnementales, comme le suppose le principe de base de la biogéographie.

Dans un second temps, nous soulèverons quelques questions directement liées à la façon dont procède l'évolution. Nous tenterons de savoir si certaines des prédictions associées aux théories évolutives qui ont été formulées pour la faune forestière tropicale sont vérifiées (synchronie entre différenciation intra-spécifique et périodes de fragmentation forestière, synchronie entre taxons, traces de contraction / expansion des populations). Nous évaluerons l'importance de la vicariance et de l'allopatrie, principes sur lesquels reposent la plupart des méthodes biogéographiques, à l'aide d'un test du mode géographique de spéciation. Nous aborderons également la question de la permanence des préférences écologiques au cours de l'évolution, qui est un postulat implicite à la théorie des refuges.

Cette thèse est organisée en six chapitres. Le premier chapitre présente le contexte théorique et conceptuel de l'étude. Il propose une synthèse des différentes théories évolutives et des processus qu'elles mettent en œuvre, puis fait le point sur les travaux biogéographiques portant sur la faune des forêts de plaine d'Afrique tropicale, et sur les connaissances actuelles en matière de changements paléoclimatiques et paléo-environnementaux.

Le second chapitre présente le matériel et les méthodes employés. Il décrit la zone d'étude dans le contexte actuel et explique le choix des modèles. Il fait le point sur les méthodes permettant d'obtenir des données moléculaires, de les traiter de façon à reconstruire des phylogénies, et d'en donner une interprétation phylogéographique et biogéographique. A chaque étape, les principales méthodes sont présentées avec leurs avantages et leurs inconvénients, de façon à justifier le choix méthodologique qui a été fait.

Les chapitres suivants présentent les résultats, regroupés selon trois pôles : phylogénie, phylogéographie et spéciation. Le chapitre 3 regroupe des reconstructions phylogénétiques, qui ont été

réalisées à partir de séquences d'ADN mitochondrial dans le but de choisir des taxons pertinents pour les analyses phylogéographiques et de calibrer une horloge moléculaire permettant d'envisager la datation des principaux événements.

Le chapitre 4 présente l'étude phylogéographique des modèles retenus et une comparaison des schémas phylogéographiques obtenus visant à déterminer s'il existe une histoire commune et, dans le cas contraire, à identifier les processus et les contraintes impliqués dans l'évolution des taxons.

Le chapitre 5 montre l'utilisation qui peut être faite des reconstructions phylogénétiques pour appréhender les processus évolutifs. Il tente d'évaluer la pertinence des différentes théories de l'évolution dans le cas de l'Afrique tropicale.

Le sixième et dernier chapitre fait la synthèse des résultats et propose des perspectives de recherche pour une meilleure compréhension de l'évolution des faunes en Afrique centrale.

CHAPITRE 1 : CADRE THEORIQUE DE L'ETUDE

I - BIOGEOGRAPHIE, PHYLOGEOGRAPHIE ET EVOLUTION DES ESPECES

1) Biogéographie et phylogéographie

Dès les XVIII^e et XIX^e siècles, des biologistes ont constaté que la biodiversité est irrégulièrement répartie à la surface de la planète (Buffon, 1761 ; Linnaeus, 1781) et qu'il existe des caractères communs aux distributions actuelles des espèces animales et végétales (Wallace, 1876). La reconnaissance de régions où la faune ou la flore sont relativement homogènes a donné lieu à la définition de régions biogéographiques. Ces régions auraient chacune une histoire (géologique, climatique...), qui serait reflétée par les taxons qui les habitent (Cracraft, 1988). La science qui étudie la distribution des taxons à l'échelle des peuplements dans le but d'identifier les processus qui les ont façonnés se nomme la **biogéographie** (Myers et Giller, 1988). Son principe de base est que des schémas de distribution communs à différents groupes d'organismes sont plus parcimonieusement expliqués par des événements climatiques ou géologiques extrinsèques que par des caractères intrinsèques à chaque taxon (Platnick et Nelson, 1978).

Les relations entre aires biogéographiques sont interprétées soit en fonction des conditions écologiques actuelles (biogéographie écologique) soit comme le résultat des changements paléo-environnementaux ayant eu lieu par le passé (biogéographie historique ; de Candolle, 1820). La biogéographie écologique cherche à expliquer la distribution actuelle des espèces en fonction de leurs capacités écologiques, de leur environnement physique (barrières géographiques actuelles, limites de zones climatiques ou phytologiques) et biotique (compétition, prédation, parasitisme). La biogéographie historique attribue un rôle majeur aux barrières géographiques qui peuvent résulter de variations paléogéographiques (variations eustatiques, dynamique fluviale, orogénèse, érosion, volcanisme) ou paléoécologiques (oscillations climatiques et changements floristiques ; Haffer, 1990). Elle intègre maintenant une dimension phylogénétique qui consiste à rechercher des schémas de parenté redondants entre taxons occupant les mêmes aires biogéographiques, de façon à mettre en évidence des événements de séparation (ou vicariance) ayant affecté la majeure partie des peuplements (Nelson et Platnick, 1981). En fait, l'approche écologique et l'approche historique sont nécessaires et interdépendantes, mais sont adaptées à des échelles de temps différentes (Morrone et Crisci, 1995). Les événements remontant à plusieurs millions d'années sont traités par la biogéographie historique et les événements couvrant les

dernières centaines ou milliers d'années (Holocène) sont analysés par la biogéographie écologique, tandis que l'étude des glaciations du Pléistocène fait appel aux deux approches.

Une des branches récentes de la biogéographie s'intéresse directement à l'histoire des taxons, plutôt qu'à celle des aires. Cette discipline, appelée **phylogéographie**, fournit de la matière première pour alimenter et tester les hypothèses biogéographiques. La phylogéographie étudie la distribution, dans le temps et dans l'espace, de groupes monophylétiques animaux ou végétaux (Avice *et al.*, 1987). Elle intègre une dimension phylogénétique qui permet de retracer l'histoire évolutive des taxons et de tester les hypothèses issues d'études biogéographiques. La comparaison de schémas phylogéographiques obtenus pour différents taxons (phylogéographie comparée) permet d'estimer s'ils dérivent vraisemblablement d'une histoire évolutive commune ou s'ils résultent de processus évolutifs indépendants (Sullivan *et al.*, 2000). Cependant, ce n'est pas parce que des schémas de distribution concordent pour différentes espèces qu'ils sont nécessairement imputables à une histoire commune (Cunningham et Collins, 1994). A l'inverse, ce n'est pas parce que les schémas de distribution diffèrent que les taxons n'ont pas été soumis aux mêmes contraintes. La concordance phylogénétique entre les généalogies de différents organismes dépend notamment de la durée de l'isolement (Avice *et al.*, 1987 ; Avice, 1989). Les schémas de distribution sont influencés par l'âge des taxons, leur distribution originelle (Zink, 1996), leurs caractéristiques écologiques, leurs capacités de dispersion (e.g. Wilson, 1974 ; Ditchfield, 2000) et la dynamique des communautés (Bennett, 1990).

Une étude phylogéographique peut se faire à différents niveaux phylétiques, allant des populations aux espèces ou groupes d'espèces. On se place le plus souvent au rang de l'espèce car c'est l'unité taxinomique minimale pour laquelle il n'y a théoriquement pas d'hybridation possible entre unités. Cependant, l'utilisation successive de taxons imbriqués permet d'aborder des échelles spatiales de plus en plus fines et des phénomènes de plus en plus récents.

Bien que biogéographie et phylogéographie étudient en premier lieu la distribution de la diversité biologique, ces disciplines posent des questions relatives à l'origine de la diversité (spéciation) et à son maintien (équilibre entre spéciation et extinction). Elles sont ainsi au cœur des débats sur l'évolution et alimentent les différentes théories de la spéciation.

2) Les moteurs de l'évolution

Il existe différents processus générateurs de diversité qui peuvent être regroupés en deux catégories : les facteurs historiques (vicariance, dispersion par delà une barrière, expansion de l'aire de distribution par dispersion naturelle entraînant la formation de zones de contact secondaire) et les facteurs écologiques (pression de sélection liée à l'habitat et à la compétition). Ces processus peuvent conduire à la formation de nouvelles espèces si des mécanismes d'isolement reproductif se développent. La spéciation peut alors se produire en allopatrie (distributions géographiques disjointes des populations parentes – ex : vicariance), parapatricie (distributions jointives séparées par une barrière – ex : dispersion) ou sympatricie (distributions recouvrantes – ex : pressions de sélection environnementales).

a) Vicariance et dispersion

Les deux principes les plus souvent évoqués en biogéographie sont la vicariance et la dispersion.

La **vicariance** résulte d'une fragmentation de l'aire de distribution du fait de l'émergence d'une barrière géographique liée à l'orogénèse, à des modifications du réseau hydrographique, ou à la fragmentation de l'habitat en relation avec des fluctuations climatiques. Il y a alors **rupture du flux génique** et différenciation des populations ainsi isolées. Lorsque le succès reproducteur varie fortement d'une femelle à l'autre, la différenciation entre populations isolées est accélérée par un phénomène de **sélection des lignées**. Ce phénomène favorise certains génotypes au détriment des autres et diminue ainsi la diversité au sein de chaque population. Il a été formalisé par la «**théorie de la coalescence** » (Kingman, 1982). Le terme coalescence désigne le moment où, en remontant dans le passé d'une généalogie de gènes, on retrouve l'ancêtre commun à tous les gènes de la génération actuelle (cf. Excoffier, 1997). L'intervalle de temps écoulé depuis un événement de coalescence est appelé temps de coalescence. Il est d'autant plus grand que la taille moyenne de la population est faible et le déséquilibre du taux de reproduction entre individus élevé.

Lors d'un événement de vicariance, si les aires isolées sont de tailles très différentes (péripatricie), il y a effet de fondation et dérive génétique, comme dans le cas de la dispersion. Pour certains auteurs, c'est le clivage inégal des populations, entraînant une accumulation de caractères dérivés dans les "isolats périphériques", qui serait le principal facteur de spéciation (Croizat *et al.*, 1974 ; Bush, 1975). Il se formerait un gradient, avec les espèces les plus ancestrales au centre et les plus dérivées en périphérie.

La **dispersion** ("dispersal" ou "jump dispersal", en anglais) désigne le franchissement d'une barrière géographique ou écologique préexistante à la disjonction. Cette définition de la dispersion doit

être distinguée de la notion de dispersion des jeunes loin de leurs parents ("dispersion" ou "range expansion", en anglais). La dispersion ne doit pas non plus être confondue avec la théorie du "dispersalisme" (Matthew, 1915) qui suppose que l'apparition d'espèces nouvelles a lieu dans un noyau central appelé centre d'origine ou centre de dispersion et que les espèces les plus ancestrales sont repoussées dans des zones périphériques.

La dispersion peut conduire à la spéciation du fait de la rupture du flux génique associée à un **effet de fondation** (lié au faible effectif des fondateurs) et à une **dérive génétique** (liée à la faible diversité génétique au sein de la nouvelle population). Cet effet a été décrit par Mayr (1942, 1963) qui considère que le moteur principal de l'effet de fondation serait la pression de sélection liée à l'isolement.

Le principe de dispersion peut être résumé ainsi : comme les fondateurs n'apportent qu'une fraction de la diversité génétique de la population souche, soit seulement certains allèles, le brassage génétique est faible et le taux d'homozygotie élevé ; en conséquence, la valeur sélective des allèles change et certains allèles sont éliminés (Mac Arthur et Wilson, 1967). La différenciation serait accélérée par une pression de sélection liée au nouvel environnement.

La dispersion a été considérée par le passé comme le principal moteur de l'évolution. Ce principe a ensuite été rejeté parce qu'il permettrait d'expliquer n'importe quelle distribution, et constitue donc une théorie non réfutable (Rosen, 1978 ; Nelson et Platnick, 1981). En revanche, la théorie de la vicariance est actuellement très prisée par les biogéographes qui se basent sur ce principe pour identifier des aires d'endémisme et proposer des scénarios évolutifs communs à plusieurs taxons (cf. Voelker, 1999). Cependant, certains proposent de reconsidérer le rôle de la dispersion, car elle semble impliquée dans bien des cas, notamment pour expliquer la colonisation de certaines îles (Voelker, 1999).

Pour savoir lequel de ces deux facteurs est responsable d'un événement de divergence, il faut dater cet événement ainsi que l'émergence de la barrière séparant les taxons divergents. Si la barrière est plus ancienne que la divergence, l'hypothèse de la dispersion est retenue. Si la barrière et la divergence sont d'âges comparables, l'hypothèse de la vicariance est retenue. Des études menées sur un large éventail de taxons indiquent que la vicariance pourrait prédominer sur la dispersion (Zink *et al.*, 2000).

b) Expansion de l'aire de distribution et zones de contact secondaire

Même lorsqu'il n'y a pas franchissement de barrière, l'extension de l'aire de distribution d'une espèce par dispersion naturelle des individus peut avoir des conséquences évolutives. Elle peut donner lieu à la rencontre de populations appartenant initialement à une même espèce dans des zones de contact

secondaire (Endler, 1982). Mayr et O'Hara (1986) envisagent cinq cas de figure lorsque deux populations allopatriques se rencontrent, à savoir, dans l'ordre croissant par rapport à la durée de la séparation et au degré d'isolement reproducteur : 1) fusion homogène ; 2) développement d'une frontière entre sous-espèces s'il y a divergence morphologique mais pas écologique ; 3) formation d'une zone hybride ; 4) parapatrie avec recouvrement minimum, avec ou sans hybridation selon que la spéciation est complète ou non ; 5) intrusion de l'une des espèces dans l'aire de répartition de l'autre avec un recouvrement important, créant une zone d'introgression. On parlera alors de colonisation multiple.

Dans les cas où il y a **hybridation**, il peut y avoir formation de nouvelles espèces, soit parce que les hybrides sont plus viables que les espèces parentes, soit parce que l'introgression de gènes d'une espèce dans le génotype d'une autre espèce augmente la variabilité génétique de la seconde et de ce fait ses possibilités de réaction à la sélection naturelle (Howard, 1998). Les hybrides peuvent au contraire être contre-sélectionnés, ce qui entraîne un **renforcement** de l'individualité de chaque espèce (Dobzhansky, 1940).

La dispersion naturelle des individus peut aussi mener à l'exploitation de niches écologiques vacantes pouvant être à l'origine d'une différenciation entre populations (Grubb, 1978) selon le principe décrit ci-après.

c) Sélection naturelle et compétition

L'environnement exercerait une pression de sélection sur les organismes favorisant l'émergence de phénotypes nouveaux, et sélectionnant les phénotypes qui confèrent un meilleur taux de survie aux individus qui les arborent. Ce mécanisme, connu sous le nom de **sélection naturelle**, a d'abord été proposé par Darwin (1859) qui n'était cependant pas en mesure d'expliquer l'héritabilité des caractères nouveaux. La "théorie synthétique de l'évolution" (Huxley, 1942), qui est à l'origine du Néo-darwinisme, a apporté le chaînon manquant. En effet, les phénotypes étant le produit de l'expression des génotypes dans un certain environnement, ils sont transmis génétiquement des parents à leurs descendants (e.g. Mayr, 1942, 1963 ; Dobzhansky, 1951). Ce mécanisme peut conduire à une différenciation phénotypique rapide et à la différenciation de nouvelles espèces, en l'absence de barrière au flux génique (e.g. Vitt *et al.*, 1997 ; Schneider *et al.*, 1999). Selon les auteurs, cette situation est qualifiée de sympatrique (e.g. Johnson et Gullberg, 1998), parapatrique (e.g. Endler, 1982) ou 'micro-allopatrique' (Mayr, 1963).

Selon la **théorie de la niche** (Grinnel, 1917 ; Elton, 1927 ; Hutchinson, 1957), c'est essentiellement la **compétition** pour des ressources limitantes du milieu qui favoriserait l'émergence de

caractères nouveaux, puis de formes nouvelles, par déplacement de caractères et spécialisation pour un nouveau type d'habitat. La colonisation d'une niche écologique vacante par des individus d'une espèce exploitant jusqu'alors une niche différente favoriserait l'émergence de traits phénotypiques nouveaux. Selon l'**hypothèse de la Reine Rouge**, ces traits phénotypiques tendraient à devenir de plus en plus marqués, certains traits non adaptatifs pouvant être favorisés par sélection sexuelle (Van Valen, 1973). Cette théorie a d'abord été formalisée en considérant que les communautés étaient des systèmes clos et que toutes les niches potentielles étaient occupées, c'est-à-dire que la libération d'une niche ne pouvait se faire que par extinction d'une espèce. Les communautés sont maintenant considérées comme des systèmes dynamiques, soumis à des variations stochastiques, et non saturés (cf. Blondel, 1995). Il n'en demeure pas moins qu'elles sont contraintes par la capacité d'accueil du milieu, et que seul un nombre limité d'espèces partageant les mêmes caractéristiques écologiques peut survivre au sein d'une même communauté. Par conséquent, la compétition intraspécifique devrait favoriser l'augmentation d'amplitude des niches, tandis que la compétition interspécifique devrait favoriser la diminution d'amplitude des niches et le déplacement de caractères (Barbault, 1995).

d) Isolement reproductif

Pour qu'il y ait formation de nouvelles espèces, il faut qu'un certain degré d'isolement reproductif soit atteint. L'isolement reproductif peut avoir une base génétique (accumulation de mutations génétiques ou remaniements chromosomiques – cf. modèle stasipatrique de spéciation de White, 1968), phénotypique (différences morphologiques liées au mode d'exploitation du milieu et conséquences sur la probabilité de rencontre des partenaires sexuels) ou comportementale (choix du partenaire sexuel). Il peut intervenir avant la fécondation, par sélection sexuelle et sociale ou isolement des partenaires sexuels dans différents habitats, ou après la fécondation en raison d'une baisse de fécondité ou de viabilité des hybrides. L'isolement reproductif a peu de chances de se développer du simple fait de l'allopatrie, mais peut être renforcé par la rencontre secondaire des populations sœurs (Dobzhansky, 1940) ou par des pressions de sélection environnementales (Schluter, 1998).

La rupture du flux génique due à un événement de vicariance (émergence d'une barrière dans l'aire de distribution d'une espèce) ou de dispersion (par delà une barrière pré-existante) favoriserait la différenciation génétique entre populations. La dispersion naturelle des individus entraînant la rencontre de populations dans des zones de contact secondaire accentuerait la différenciation entre populations préalablement isolées et pourrait éventuellement favoriser l'émergence de nouvelles espèces par

hybridation. Enfin, la colonisation de nouveaux habitats stimulerait la différenciation phénotypique, sous l'action de pressions environnementales sélectives et de la compétition interspécifique. Cependant, de nouvelles espèces ne peuvent se former que s'il se développe des mécanismes d'isolement reproductif entre populations préalablement différenciées.

3) Théories de l'évolution

Les mécanismes décrits ci-dessus sont à la base des différentes hypothèses évolutives qui cherchent à identifier les modalités de la spéciation animale. La plupart de ces hypothèses, que nous désignerons sous le nom de "théories", supposent que la spéciation se produit en allopatrie. Cependant, la théorie des gradients environnementaux prédit que la spéciation est possible en parapatricie (voire en sympatricie), sans qu'il y ait séparation effective des populations.

a) La théorie des gradients environnementaux

Cette théorie a été formulée par Endler (1977, 1982), pour qui les variations écologiques rencontrées le long de gradients environnementaux pourraient suffire à expliquer la différenciation des animaux en Amazonie. Elle a fait l'objet de nombreuses critiques (Mayr et O'Hara, 1986 ; Cracraft et Prum, 1988 ; Haffer, 1997), mais a été réhabilitée récemment pour expliquer des cas de variabilité morphologique intra-spécifique en présence de flux génique chez certaines espèces forestières tropicales (e.g. Smith *et al.*, 1997 ; Schluter, 1998). Elle serait particulièrement impliquée au niveau des écotones (Smith *et al.*, 1997). En Australie, des populations de lézards ayant été isolées géographiquement présentent une certaine diversité génétique mais aucune diversité morphologique (Schneider et Moritz, 1999). En revanche, des populations géographiquement proches mais se trouvant dans des habitats différents sont génétiquement très semblables et morphologiquement différenciées (Schneider *et al.*, 1999). Selon ces auteurs, des pressions de sélection environnementales différentes selon les milieux entraîneraient une diversification morphologique, qui pourrait être à l'origine de la formation d'espèces jumelles. La différenciation morphologique pourrait être plus directement et plus rapidement impliquée dans des événements de spéciation que la différenciation génétique.

La théorie des gradients prédit que des taxons frères devraient occuper des habitats distincts mais adjacents. Au sein d'une même espèce, la divergence phénotypique entre populations devrait être beaucoup plus importante entre habitats qu'au sein de chaque habitat (Moritz *et al.*, 2000).

b) La théorie paléogéographique

D'après Croizat (1976), les événements de vicariance liés aux mouvements tectoniques et aux fluctuations du niveau de la mer auraient constitué une condition nécessaire et suffisante à la différenciation des faunes en Amérique du Sud. Des travaux portant sur le bassin Amazonien (e.g. Da Silva et Patton, 1998 ; Clough et Summers, 2000 ; Loughheed *et al.*, 2000) et le rift africain (e.g. Denys *et al.*, 1986 ; Partridge *et al.*, 1995 ; Wieczorek *et al.*, 2000) semblent mettre en évidence le rôle de l'orogénèse et des mouvements tectoniques dans la formation de nouvelles espèces. En Amazonie, les mouvements de subsidence liés à la surrection des Andes auraient entraîné la formation de sous-bassins tectoniques dont les limites géographiques et la dynamique temporelle coïncident avec les schémas phylogéographiques trouvés au sein de plusieurs genres de rongeurs (Da Silva et Patton, 1998).

Cette théorie qui concerne les derniers 60 millions d'années (MA) est difficile à prouver à cause des événements de migrations qui se sont produits depuis. De plus, selon Haffer (1997), les changements géologiques n'ont pas été suffisamment complexes et fréquents pour expliquer l'évolution de la faune, et leur effet direct sur la faune est difficilement dissociable de celui des changements environnementaux qu'ils ont entraînés.

c) La théorie des barrières fluviales

L'impact des rivières sur la distribution des espèces animales a tout d'abord été remarqué en Amazonie par Wallace (1852). Au rôle de séparateur des faunes que cet auteur leur a reconnu s'ajoute un effet de barrière au flux génique qui peut avoir favorisé la divergence entre espèces. L'effet de barrière augmente avec le débit et la largeur de la rivière, donc de la source vers l'embouchure (e.g. Ayres et Clutton-Brock, 1992). Il varie selon les caractéristiques biologiques des espèces animales, notamment leur taille et leur capacité à franchir l'obstacle, à la nage ou en volant. Les capacités natatoires peuvent d'ailleurs varier au sein d'une même espèce (e.g. Hickman *et al.*, 1983). Trois cas de figure sont possibles : la rivière peut avoir fragmenté l'aire de distribution de l'espèce (vicariance), ou peut au contraire avoir existé préalablement à la dispersion de l'espèce et soit avoir empêché que celle-ci colonise l'autre rive (barrière effective), soit avoir été traversée (dispersion).

La théorie des barrières fluviales a été invoquée pour expliquer la différenciation des Oiseaux (Capparella, 1991), des Primates (Ayres et Clutton-Brock, 1992 ; Peres *et al.*, 1996) et d'autres Mammifères (Patton *et al.*, 2000), en Amazonie. Elle peut être testée en comparant la différenciation entre populations des rives opposées et populations d'une même rive. La prédiction à tester est que la divergence entre taxons frères est plus importante d'une rive à l'autre de la rivière que sur une même rive.

Cependant, plusieurs facteurs peuvent rendre ce test inopérant (Haffer, 1997 ; Patton *et al.*, 2000). Il y a un risque que des phénomènes d'isolement par la distance masquent l'effet de barrière. Il a pu se produire des transferts passifs de faune d'une rive à l'autre du fait des modifications assez fréquentes du cours des rivières, telles que la fermeture de méandres. Il arrive que de larges rivières ne constituent pas une barrière pour certains taxons tandis que des rivières plus petites délimitent leur aire de distribution. Durant les périodes glaciaires, le débit et la largeur des rivières a dû être beaucoup moins important. Pour finir, si les taxons présents de part et d'autre de la rivière ne sont pas des taxons frères, la rivière constitue vraisemblablement une zone de rencontre secondaire entre taxons ayant évolué ailleurs.

d) La théorie des refuges

Haffer (1969) a fait l'hypothèse que la diversité en milieu forestier tropical est le fruit de l'hétérogénéité spatiale et des alternances de phases de contraction et expansion de la forêt, en relation avec les variations climatiques. La principale cause des variations climatiques serait les cycles de Milankovitch, qui sont liés à l'attraction exercée par les planètes sur la terre, et à la modification subséquente de son orbite autour du soleil.

- Durant les périodes glaciaires (pléniglaciaires), la sécheresse aurait entraîné une diminution des forêts de plaine dans les régions tropicales, tellement marquée que seuls des lambeaux disjoints auraient persisté. Dans des conditions extrêmes de fragmentation, il y aurait eu situation de "refuges" forestiers. Les espèces forestières auraient été soumises à une forte vicariance à l'origine d'événements de spéciation allopatriques du fait de la rupture du flux génique (Booth, 1958a ; Moreau, 1969). On suppose que les zones historiques de refuge potentiel sont celles où la diversité et l'endémisme sont actuellement les plus élevés (Happold, 1996). Les zones de grande diversité où il n'y a pas d'endémisme marqué peuvent en revanche correspondre à des régions de confluence des faunes.

- Durant l'interglaciaire, au contraire, l'humidité aurait permis la transgression des forêts sur les savanes, avec expansion des aires de distribution. L'expansion forestière peut être à l'origine de la création de nouveaux habitats et peut donner l'occasion aux individus qui se dispersent d'occuper des niches vacantes (Grubb, 1978). Elle favorise également la rencontre entre faunes au niveau de zones de contact secondaire. Selon le degré d'isolement reproducteur des espèces, il peut y avoir hybridation, sympatrie, ou mélange entre des populations n'ayant que peu divergé l'une de l'autre.

En dépit de sa formulation première, cette théorie ne se limite pas aux effets des oscillations climatiques du Pléistocène, mais concerne aussi tous les événements climatiques et géologiques pouvant avoir entraîné une fragmentation de l'habitat durant le quaternaire, le tertiaire, et même au-delà jusqu'à

60 MA (Haffer, 1997). Elle peut être généralisée à d'autres types d'habitats que le milieu forestier, puisque les variations climatiques et les mouvements tectoniques créent des "refuges écologiques" pour les espèces qui y sont inféodées (Haffer, 1990). Elle peut aussi être combinée à la théorie des barrières fluviales. Ainsi, durant les phases glaciaires, il y aurait eu isolement des populations dans des refuges séparés par de grandes zones d'habitat défavorable et par le cours des rivières, surtout dans les régions aval (Colyn *et al.*, 1991 ; Haffer, 1993).

La théorie des refuges n'a de sens que si les espèces forestières sont inféodées au milieu forestier et que leur plasticité écologique est limitée. Cela suppose que les espèces forestières devraient être phylogénétiquement plus proches entre elles qu'elles ne le sont des espèces savanicoles (Lynch, 1988). La théorie des refuges n'est pas facilement testable, car elle ne fournit pas de modèle prédictif des relations entre aires. Par contre, elle prédit qu'il devrait y avoir une certaine concordance géographique et temporelle entre taxons. De plus, des espèces ayant divergé récemment devraient se trouver dans des régions adjacentes et, à l'échelle intra-spécifique, des événements d'expansion de populations avec contact secondaire devraient être identifiables (Moritz *et al.*, 2000).

La théorie des refuges a fait l'objet de nombreuses critiques, qui ont presque toutes été contrées :

- Pour certains, il n'est pas prouvé qu'il y ait eu fragmentation du bloc forestier Amazonien (Colinvaux, 1993 ; Bush, 1994 ; Müller *et al.*, 1995). Il est possible que la fragmentation ait été incomplète, les fragments de forêt ayant pu être reliés par des galeries forestières (Haffer, 1997). Des simulations réalisées pour l'Amazonie prédisent que, lors du dernier maximum glaciaire, la baisse de température aurait compensé l'effet de la baisse des précipitations, favorisant ainsi le maintien de la forêt. Malgré cela, l'hétérogénéité régionale des conditions environnementales et ses répercussions sur la densité de la canopée auraient constitué un facteur de vicariance (Cowling *et al.*, 2001).

- La fragmentation peut avoir généré l'apparition de nouvelles espèces, mais aussi la disparition d'autres espèces. La balance entre apparition et disparition d'espèces pourrait être déficitaire (Williams et Pearson, 1997).

- Les phénomènes de spéciation observés ne sont pas forcément synchrones et ne coïncident pas toujours avec les périodes glaciaires (Ridley, 1997). Cependant, en Afrique du moins, il semble que le renouvellement de certaines espèces de Mammifères ait été initié par les cycles de Milankovitch et par des changements climatiques globaux (Vrba, 1995 ; Partridge *et al.*, 1995)

- Lorsque les derniers événements de spéciation semblent remonter au tertiaire, ils ne peuvent être expliqués par les phases d'oscillations climatiques majeures, puisque celles-ci se sont produites au Pléistocène. Ce serait notamment le cas en Amazonie, non seulement pour la faune (e.g. Lynch, 1988 ; Joseph *et al.*, 1995 ; Moritz *et al.*, 2000), mais aussi pour la flore (Hooghiemstra, 1999). Cependant, la théorie des refuges ne se limite pas aux variations climatiques du Pléistocène (Haffer, 1997).

- D'après les plus fervents défenseurs de la biogéographie écologique (Mac Arthur, 1972 ; Endler, 1982), les conditions éco-géographiques sont suffisantes pour expliquer les schémas de distribution actuels. Cependant, elles ne semblent pas suffisantes pour expliquer l'origine des espèces (Haffer, 1997).

- Le modèle d'évolution impliqué ne fonctionne que si les populations étudiées ont effectivement été isolées assez longtemps pour qu'il y ait eu différenciation dans les refuges postulés. Seules les conséquences des derniers refuges simultanés peuvent être observées, alors que l'évolution est graduelle et que la succession des cycles climatiques a progressivement effacé ce que les cycles précédents avaient créé (Haffer, 1997).

e) La théorie des perturbations

Cette théorie, qui est inspirée de celle des "refuges disparaissants" (Vanzolini et Williams, 1981), considère que les variations climatiques auraient constitué le principal moteur de l'évolution. Cependant, contrairement au principe de la théorie des refuges, ce n'est pas l'aridité qui aurait été déterminante durant les périodes glaciaires, mais le froid. De même, ce ne sont pas les zones de stabilité environnementale qui auraient favorisé la spéciation, mais les zones de perturbation (Colinvaux, 1993 ; Bush, 1994). Dans les zones périphériques du bassin Amazonien, les changements climatiques auraient entraîné de constants remaniements de la végétation. Au niveau de ces écotones dynamiques, la micro-fragmentation de l'habitat aurait favorisé la spéciation en allopatrie. De plus, une forte compétition entre espèces, due à la disparition d'habitat favorable et à des invasions successives, aurait favorisé la spéciation directement, mais aussi indirectement par l'exclusion et l'isolement de certains taxons.

Cette théorie n'est pas directement testable mais peut éventuellement être invalidée puisqu'elle prédit qu'il devrait y avoir peu ou pas de concordance phylogéographique entre taxons. Selon Haffer (1997), elle ne s'appliquerait qu'au quaternaire, puisque au tertiaire ce n'est pas tant la température que l'humidité qui a varié.

Pour le milieu forestier tropical, plusieurs théories concurrentes mais non exclusives permettraient d'expliquer la différenciation intra-spécifique, et la spéciation après développement de mécanismes d'isolement génétique (Haffer, 1997) :

- la théorie des gradients environnementaux qui propose un modèle de spéciation parapatrique,
- la théorie de perturbation-vicariance,
- la théorie paléogéographique,
- la théorie des barrières fluviales,
- la théorie des refuges, qui est la théorie la plus souvent invoquée.

Ces théories tendent également à expliquer la distribution actuelle des espèces. Ainsi, la théorie des refuges explique les assemblages faunistiques actuels par la fragmentation de l'habitat au cours des derniers cycles glaciaires, tandis que la théorie des barrières fluviales explique les limites de distribution par le rôle de barrière joué par les rivières.

II - LE CAS DE L'AFRIQUE TROPICALE

1) Choix de la zone d'étude

Les questions de l'origine et du maintien de la biodiversité sont souvent abordées dans le contexte tropical, car c'est là que la diversité faunistique et floristique est la plus importante. Alors qu'il existe de nombreuses études phylogéographiques et biogéographiques focalisées sur l'Amérique du Sud, le continent africain, déjà bien étudié sur le plan biogéographique, commence tout juste à être exploré par des études phylogéographiques. Il nous a donc semblé intéressant d'aborder l'évolution en milieu forestier tropical sur le continent africain.

Dans un premier temps, nous nous sommes focalisés sur les forêts de plaines comprises dans une bande intertropicale délimitée par les 12 ° parallèles Nord et Sud et limitée à l'Est par la région du rift africain (**fig. 1.1**, zone A).

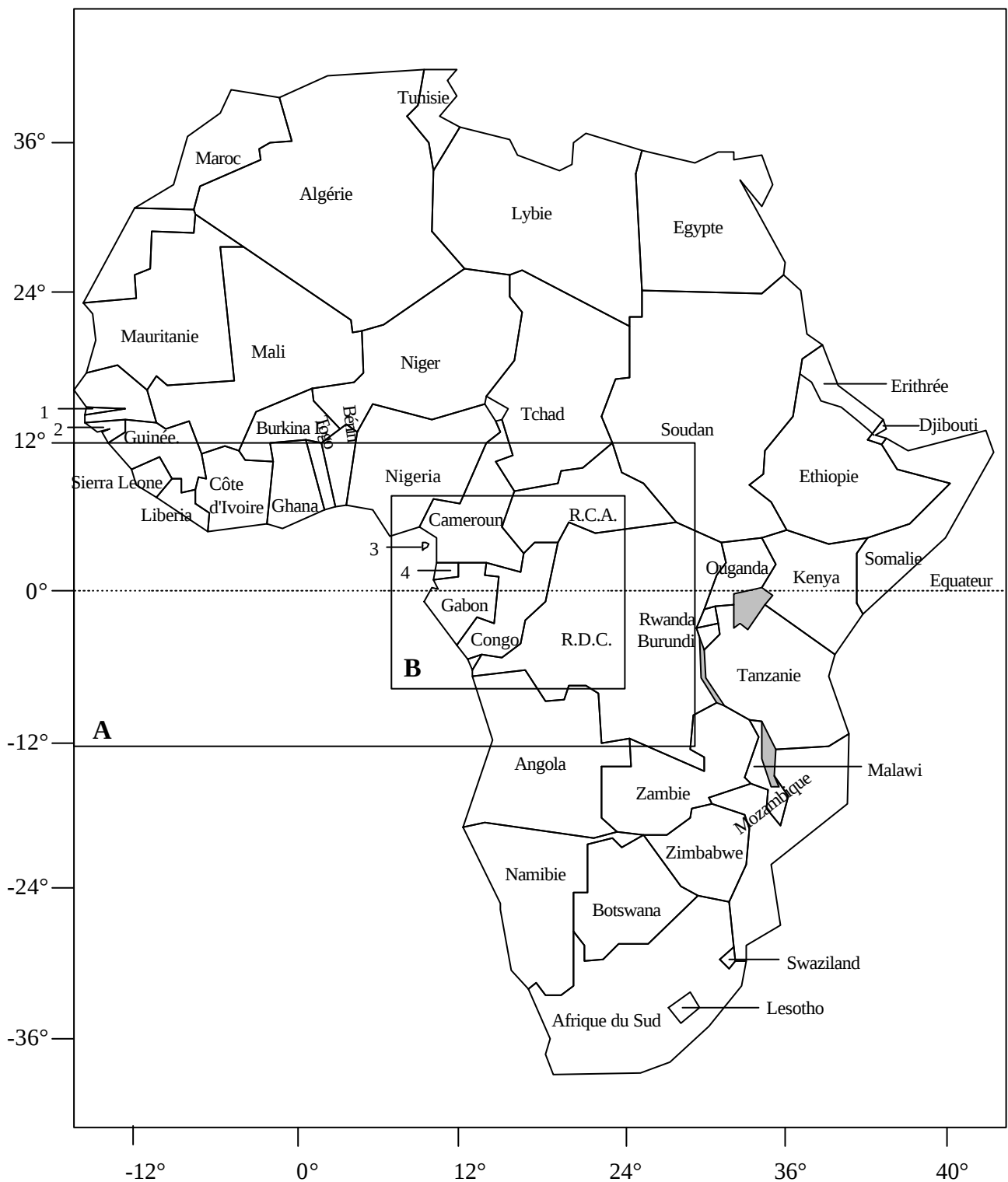


Figure 1.1. délimitation des zones d'étude A et B sur la carte de l'Afrique politique.

1 - Gambie ; 2 - Guinée Bissau ; 3 - Ile de Bioko ; 4 - Guinée Equatoriale

Dans un second temps, nous nous sommes concentrés sur une portion du bloc forestier d'Afrique centrale qui semble avoir eu une histoire particulièrement complexe. Cette région est délimitée par des latitudes comprises entre 6°N et 6°S et des longitudes comprises entre 8°E et 20°E. (**fig. 1.1**, zone B). Elle inclut le Gabon, la République du Congo, la moitié Sud du Cameroun, et une partie de la République Centrafricaine (RCA) et de la République Démocratique du Congo (RDC). Pour plus de simplicité, nous désignerons par la suite la République du Congo sous le nom de Congo, et la République Démocratique du Congo par l'abréviation RDC.

2) Bilan des travaux biogéographiques portant sur les forêts d'Afrique centrale et occidentale

a) Reconnaissance de régions fauniques

Depuis les années 50, de nombreux auteurs travaillant principalement sur les Oiseaux et les Mammifères, se sont intéressés à la distribution de la faune forestière en Afrique tropicale (e.g. Marchant, 1954 ; Booth, 1958a, 1958b). Carcasson (1964) et Moreau (1966, 1969), qui se sont intéressés respectivement aux Papillons et aux Oiseaux, ont défini trois régions d'endémisme en se basant sur la richesse spécifique et le taux d'endémisme local. Ces trois régions, le plus souvent appelées Ouest, Ouest Centrale et Est Centrale, correspondent respectivement aux forêts d'Afrique de l'Ouest, et aux parties Ouest et Est du bloc forestier principal. Elles ont été corroborées par des travaux ultérieurs portant sur les Amphibiens et les Reptiles (Laurent, 1973), les Mammifères (Kingdon, 1971 ; Grubb, 1978, 1982, 1990 ; Colyn, 1991) et les Oiseaux (Prigogine, 1988). Ces études ont reconnu une quatrième région majeure appelée Sud Centrale, située dans la portion de la cuvette congolaise située au Sud du fleuve Congo qui n'avait pas été prise en compte dans les premiers travaux.

D'autres auteurs ont proposé un découpage plus fin, en se basant non plus uniquement sur l'endémisme et la richesse spécifique, mais aussi sur les limites de distribution des taxons. Pour les Mammifères, une synthèse récente des travaux de Booth (1954, 1958a), Grubb (1978, 1982), Hamilton (1988, 1989) et Oates (1988) par Delany et Happold (1971) et reprise par Happold (1996) a permis de mettre en évidence sept sous-régions de relative uniformité faunistique, correspondant aux principaux blocs interfluviaux (**fig. 1.2**). Ce découpage est corroboré par des travaux portant sur les Oiseaux (Diamond et Hamilton, 1980) et les Bégonias (Sosef, 1994, 1996).

Des études menées par notre équipe en Afrique centrale indiquent que la réalité est encore plus complexe (Colyn, 1991 ; Colyn *et al.*, 1991 ; Deleporte et Colyn, 1999). Pour l'Afrique centrale, il est effectivement possible de distinguer une région Ouest centrale et une région Est centrale, mais celles-ci

peuvent être subdivisées en régions et même en sous régions fauniques (**fig. 1.3**). De plus, de nombreux travaux de révision des aires de distribution de différentes espèces (e.g. Colyn, 1993 ; Colyn et Van Rompaey, 1994a,b ; Verheyen *et al.*, 1996, 1997, 2000) ont permis de mettre en évidence des zones tampons entre les sous-régions, dans lesquelles il se produit un mélange des faunes en provenance des régions adjacentes. Ces zones sont qualifiées de "zones d'intergradation".

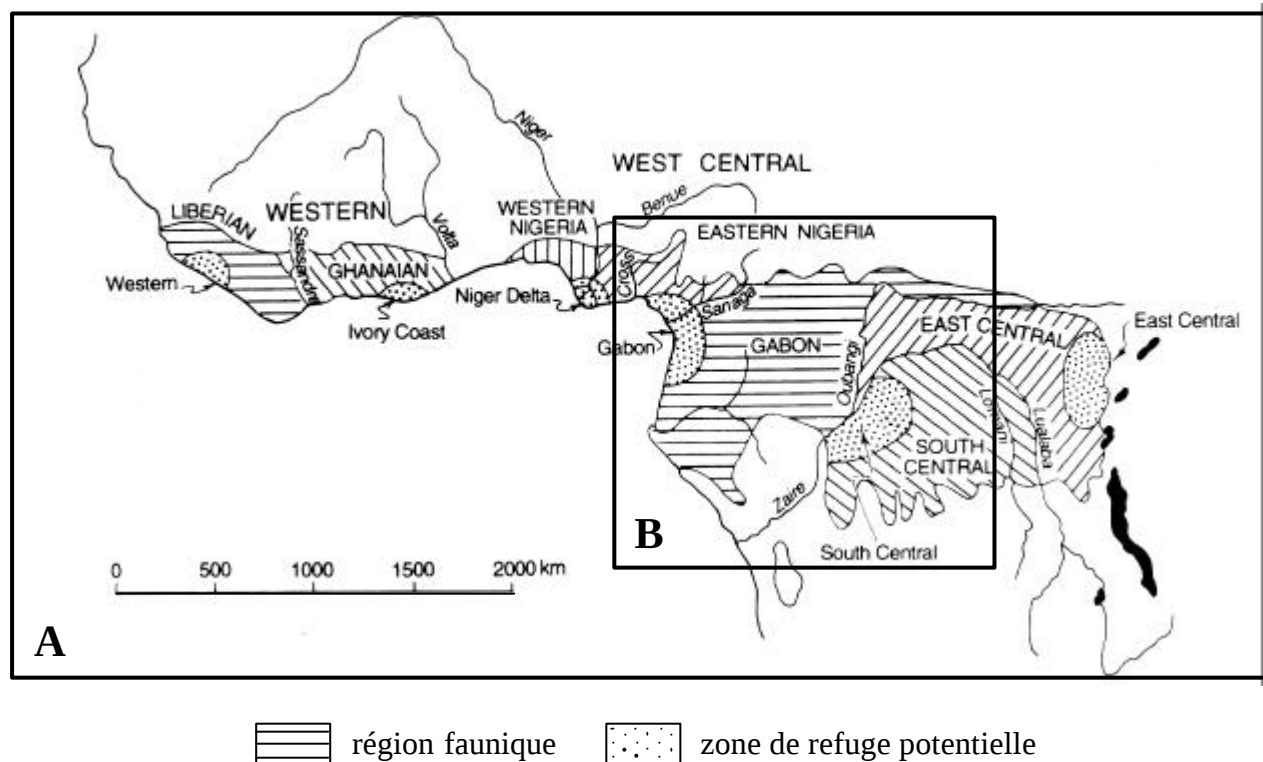


Figure 1.2. Régions fauniques et refuges forestiers basés sur la distribution des Mammifères, d'après Happold (1996), adapté de Happold et Delany (1971). Synthèse des travaux de Booth (1954, 1958a), Grubb (1978, 1982), Hamilton (1988, 1989) et Oates (1988).

Dans le cas de la région Ouest Centrale, qui est au cœur de cette étude, la majorité des auteurs distinguent une région Est Nigéria, incluant le Nord du Cameroun, et une région Gabon, comprenant le Gabon ainsi qu'une partie du Cameroun et du Congo (cf. synthèse de Delany et Happold, 1971, reprise par Happold, 1996 – **fig. 1.2**). D'autres auteurs reconnaissent plutôt une région continentale et une région côtière (e.g. Louette, 1992 pour les Oiseaux ; Deleporte et Colyn, 1999 pour les Mammifères). Cette distinction avait déjà été proposée par Marchant (1954) et apparaît aussi au niveau de la végétation (White, 1983). Elle est également supportée par des travaux phylogéographiques récents portant sur un petit nombre de taxons apparentés (e.g. Ohler, 1996 pour un genre d'Amphibiens ; Beresford et Cracraft, 1999 pour un genre d'Oiseaux).

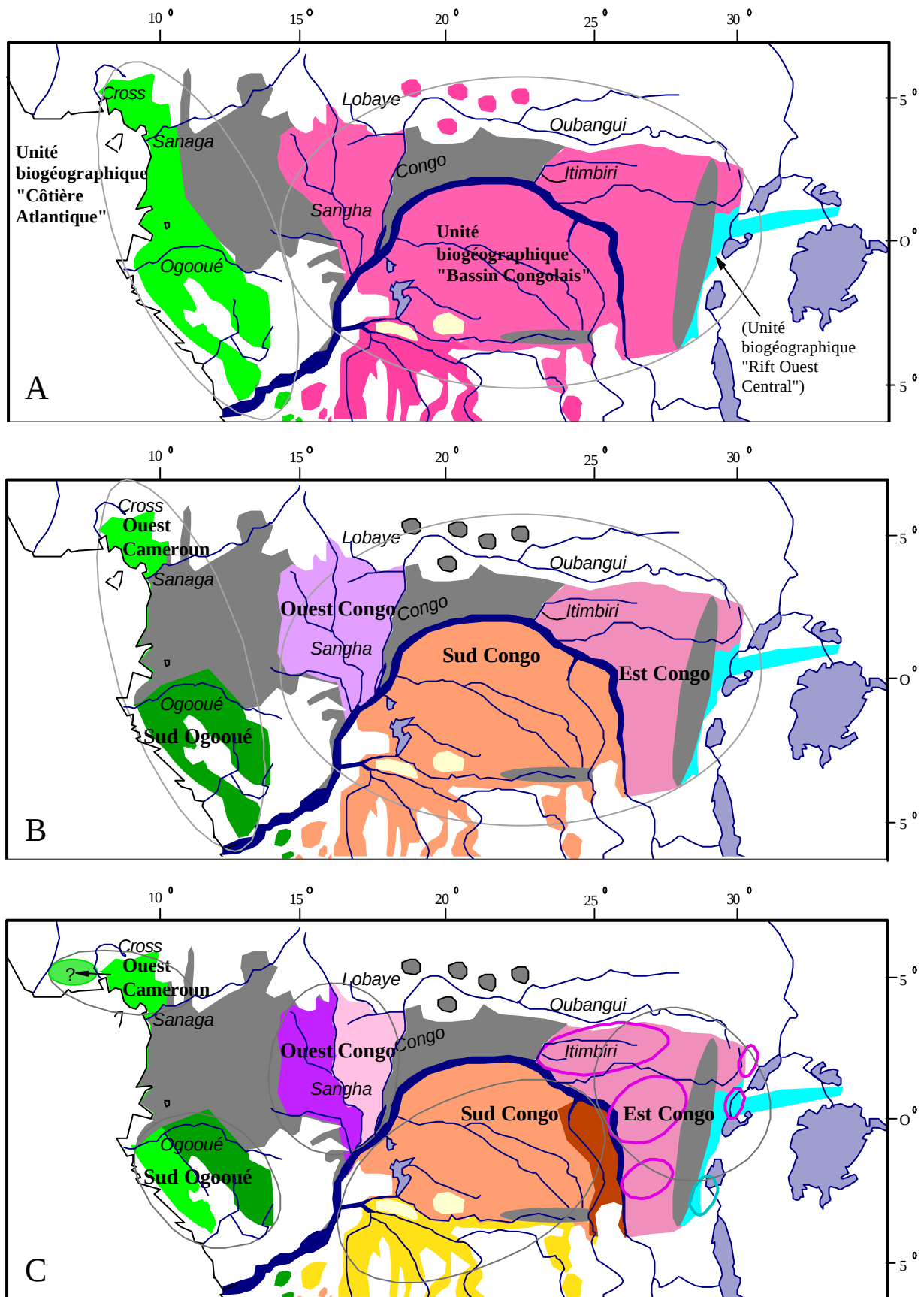


Figure 1.3. Principales unités biogéographiques d'Afrique centrale (A) et leur sub- division en régions (B) et sous-régions fauniques (C), d'après M. Colyn (comm. pers.). Les zones hachurées représentent les régions d'intergradation.

b) Barrières

Le rôle de barrière joué par les rivières est manifeste. Le fleuve Congo constitue une barrière majeure pour les Oiseaux (Prigogine, 1988 ; Louette, 1992) et affecte fortement la distribution des Primates (Colyn, 1987 ; Grubb, 1990) et d'autres Mammifères (Happold, 1996). L'Oubangui sépare les régions Ouest Centrale et Est Centrale, tandis que la Sassandra, le Niger et la Sanaga séparent les sous-régions définies pour les Mammifères (Happold, 1996).

Les forêts de l'Ouest sont séparées du bloc principal par une languette de savane appelée "Gap" du Dahomey, dont l'origine remonterait à au moins 3 MA (Robert et Chamley, 1987). Ce corridor de savane a dû être beaucoup plus conséquent qu'il ne l'est actuellement, durant les périodes glaciaires, mais des îlots forestiers ont pu s'y former puis disparaître, en fonction des fluctuations climatiques (Aubréville, 1949). Pour certains auteurs, ce seraient les deux grandes rivières qui l'encadrent, le Niger et la Volta, plutôt que le "Gap" lui-même, qui auraient joué un rôle majeur dans la séparation des faunes (e.g. Booth, 1958b et Robbins, 1978 pour les Mammifères ; Moreau, 1969 pour les Oiseaux).

Au sein de la région Ouest Centrale, la distinction entre une région côtière et une région continentale laisse supposer l'existence d'une barrière biogéographique passée séparant jadis deux régions entre lesquelles il n'y a aujourd'hui aucune frontière apparente (Deleporte et Colyn, 1999). Ce rôle de barrière pourrait avoir été assuré par un corridor de savane qui aurait traversé le bloc forestier du Nord au Sud pendant la majeure partie du Pléistocène (cf. Bigalke, 1978). Cette idée est renforcée par la présence de savanes incluses (de Foresta, 1990), de termitières d'espèces savaniques (Lanfranchi et Schwartz, 1990), d'un gravier très grossier qui se serait formé dans des conditions semi-arides (Maley, 1996) et de formations végétales de type pionnières telles que les formations clairsemées à marantacées (Letouzey, 1968), à la limite entre les régions côtières et continentales.

c) Refuges

La théorie des refuges est souvent invoquée pour expliquer l'existence des régions fauniques (e.g. Booth, 1958a ; Grubb, 1978 ; Diamond et Hamilton, 1980). Au sein de chaque région, des centres d'endémisme seraient les témoins de zones de refuge ayant existé en période de forte fragmentation forestière, et à partir desquelles chaque région aurait été recolonisée. Les premiers travaux prenaient en compte indistinctement les espèces de plaine et de montagne et localisaient les refuges dans des zones de montagne (e.g. Carcasson, 1964 ; Moreau, 1969). Cependant, il apparaît que l'histoire des faunes des forêts de plaine est indépendante de celle des faunes de forêt de montagne (Grubb, 1978 ; Colyn,

1991 ; Colyn *et al.*, 1991). C'est donc en plaine qu'il convient de chercher les refuges impliqués dans la différenciation des espèces de plaine.

La reconstitution des paléo-environnements et l'étude des zones d'endémisme végétal et animal permettent de localiser hypothétiquement quelques refuges de plaine (**fig. 1.4**, extraite de Maley, 1996) et d'envisager l'existence d'un refuge majeur le long du fleuve Congo (Colyn *et al.*, 1991).

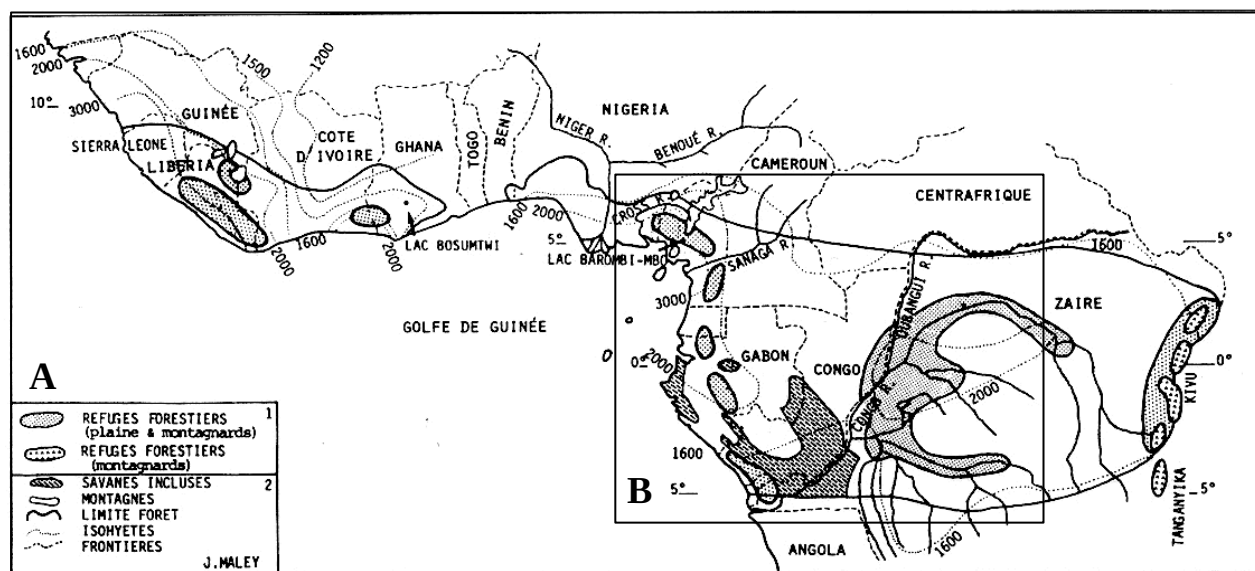


Figure 1.4. Refuges forestiers envisagés pour le dernier maximum glaciaire (18.000 ans B.P.), d'après Maley (1996).

La localisation des refuges est adaptée de Van Rompaey (1993, 1994) pour le Liberia et la Côte d'Ivoire, de Sosef (1994) pour le Gabon et le Congo, de Colyn (1987, 1991) pour le bassin du Congo, et Mayr et O'Hara (1986) pour la Haute-Guinée et l'Afrique de l'Est.

Les conditions actuelles (limite de la forêt, savanes incluses et isohyètes) sont adaptées de White (1983).

d) Variabilité des schémas de distribution selon les taxons

Bien qu'il existe des caractères généraux aux distributions des espèces en forêt tropicale africaine, les schémas de distribution diffèrent selon les taxons (Grubb, 1978 ; Happold, 1996). L'étendue des aires de distribution varie selon les espèces. Par exemple, chez les Primates, le singe de Brazza, *Cercopithecus neglectus*, a une large distribution en Afrique centrale, qui se superpose avec les aires de distribution de 2 espèces et 6 sous-espèces de la super-espèce *C. cephus* et de 2 espèces et 5 sous-espèces de la super-espèce *C. mona* (cf. Gautier-Hion *et al.*, 1999). Chez les mangoustes, l'espèce *Xenogale naso* ne présente aucune variation morphométrique inter-populationnelle sur l'ensemble de l'Afrique centrale (Colyn et Van Rompaey, 1994a), tandis qu'au sein du genre *Crossarchus*, cinq sous-espèces se partagent la même région (Colyn et Van Rompaey, 1994b). Ces différences pourraient

s'expliquer par l'occupation de refuges différents au cours de la dernière période glaciaire et par des capacités de colonisation plus ou moins développées selon les espèces. En effet, les capacités de colonisation et de franchissement des barrières géographiques sont déterminantes pour assurer la recolonisation du milieu en période de transgression forestière. Par exemple, la faculté de franchir les rivières, à la nage ou en vol, peut influencer fortement les distributions. Tandis que certains Primates (Grubb, 1990 ; Gautier-Hion *et al.*, 1999) et Rongeurs (Happold, 1996) présentent des allo-espèces séparées par des rivières, certaines mangoustes (Colyn et Van Rompaey, 1994a,b) et la plupart des Oiseaux (Louette, 1992) ne sont pas arrêtés par les rivières. C'est aussi le cas de *C. neglectus*, qui a des capacités de nage développées (Gautier-Hion et Gautier, 1978). Cependant, la faculté de franchir une barrière n'explique pas à elle seule le développement d'une population par delà l'obstacle. Des facteurs sociaux pourraient aussi intervenir. Suite à la dispersion de quelques individus, les espèces solitaires auraient de plus grandes chances de générer une population viable que les espèces grégaires (M. Colyn, comm. pers.).

La taille des aires de répartition peut aussi être interprétée en terme de sensibilité à la fragmentation de l'habitat (Grubb, 1978). En période de fragmentation forestière, les espèces les moins exigeantes et les plus plastiques sur le plan écologique ont pu survivre même dans des petits îlots forestiers, tout en maintenant éventuellement un certain flux génique. Ce pourrait être le cas pour certains rongeurs à large distribution, comme *Lophuromys sikapusi* et *Praomys tullbergi* (cf. Happold, 1996). D'autres espèces, plus inféodées au milieu forestier, ont dû disparaître localement et ne survivre que dans des régions isolées (cf. Haffer, 1969). Cet argument expliquerait les distributions isolées de certains Primates, comme *C. lhoesti*, *C. preussi* et *C. solatus* (Harrison, 1988). Les capacités de survie des espèces sont déterminées notamment par leurs possibilités de modulation de l'utilisation des niches écologiques en fonction de la pression de compétition (Happold, 1996). Elles dépendent aussi de leurs possibilités de modulation de la reproduction en réponse à des modifications à court terme de l'environnement. De ce point de vue, les espèces dont la durée de vie est de l'ordre d'une année, comme c'est le cas pour la plupart des petits mammifères, sont favorisées par rapport aux espèces longévives, comme les primates simiens.

Le bloc forestier d'Afrique central et de l'Ouest peut être divisé en plusieurs régions dites régions fauniques au sein desquelles la faune est relativement homogène. Les principaux facteurs impliqués dans l'isolement de ces régions semblent être :

- le fractionnement de la forêt au cours du Pléistocène, entraînant la formation de centres d'endémisme,

- les barrières fluviales : le fleuve Congo, la Sanaga, la Sangha, l'Oubangui, le Niger et la Volta,
- les barrières savanicoles : visibles actuellement, comme le gap du Dahomey, ou bien passés, comme celui qui aurait existé entre les sous-régions côtière et continentale de la région Ouest Centrale dans des conditions de fragmentation extrême.

Tous les taxons n'auraient pas réagi de la même façon à ces différents types de barrières, du fait de leurs préférences écologiques et de leurs capacités de survie et de colonisation.

3) Paléoclimats et conditions paléo-environnementales en Afrique tropicale

Sachant que la fragmentation de la forêt Amazonienne a été remise en cause et que les conséquences de la fragmentation forestière sur l'évolution de la faune ne sont pas prouvées, on peut s'interroger sur la réalité des postulats qui sous-tendent les diverses théories de l'évolution. Dans le cas de l'Afrique tropicale, de nombreuses études permettent de retracer l'évolution des conditions paléoclimatiques et paléo-environnementales. Elles mettent en évidence une relation entre les événements géomorphologiques et tectoniques majeurs, la dynamique climatique à l'échelle de la planète, les conditions climatiques en Afrique et la végétation en Afrique tropicale.

a) Indices révélateurs des conditions paléoclimatiques et paléo-environnementales

La température et l'insolation sont inférées des isotopes d'oxygène $\delta^{18}\text{O}$ contenus dans les dépôts de foraminifères des grands fonds marins (e.g. Shackleton *et al.*, 1984 ; Jahns *et al.*, 1998). Les glaciations sont reflétées par les preuves de variation du niveau de la mer et par les traces géologiques témoignant de l'importance de la calotte glaciaire. L'hygrométrie est déduite des variations du niveau de la mer, du niveau des lacs et du débit des rivières. L'aridité est attestée par de nombreux indices (Bennett, 1990 ; Jansen, 1990) : le taux de poussière éolienne, le rapport entre phytolithes transportées par le vent et diatomées des eaux douces dans les sédiments marins, le taux de dépôt de sédiments inorganiques (deux à quatre fois plus élevé en période d'aridité qu'en période humide), et le type de roches sédimentaires. Ainsi, dans des conditions arides, les smectites fortement cristallines, la calcite et le talc augmentent au détriment de la kaolinite et du quartz.

L'influence des paléo-climats sur la végétation peut être déterminée par l'étude des corrélations entre indices climatiques et type de végétation dans les dépôts marins ("deep-sea cores") et terrestres. Les sédiments marins sont abondants et couvrent de grandes échelles géographiques, alors que les

sédiments terrestres (lacustres) sont plus rares et couvrent des régions plus petites et de localisation plus précise. L'essentiel des informations dont nous disposons résulte de l'étude des pollens fossiles, mais quelques informations directes ont été apportées par l'étude des plantes fossiles (e.g. Bonnefille, 1987, Dechamps *et al.*, 1990). Certaines plantes sont caractéristiques d'un environnement sec (par exemple les graminées), des eaux en phase d'eutrophisation, des forêts de plaine ou des forêts de montagne (tel que *Podocarpus*). Seulement, comme on se base sur leur écologie actuelle, il faut prouver que c'était aussi le cas par le passé, car les relations entre climat et végétation peuvent avoir changé, notamment du fait des variations du taux de CO₂ atmosphérique (Maley et Brenac, 1998).

b) Dynamique globale et répercussions en Afrique tropicale

A l'échelle planétaire, il existe des cycles d'oscillations liés aux cycles de Milankovitch et qui remonteraient à 65 MA. On distingue trois types de cycles ayant des périodes différentes, avec prédominance de cycles de 23.000 ans avant 2,8 MA, 41.000 ans entre 2,8 et 1 MA et 100.000 ans depuis (Bennett, 1990). Ces cycles sont liés à l'attraction exercée par les planètes sur la terre, et à la modification subséquente de son orbite autour du soleil. Ils affectent le régime pluviométrique et notamment la mousson, ainsi que les radiations solaires reçues par la terre. Ils ont une influence directe sur le volume des calottes glaciaires (**Fig. 1.5**). Au cours des derniers 2,3 MA, on compte ainsi 20 cycles d'alternance climatique, avec une augmentation de l'amplitude glaciaire au cours du dernier million d'années (deMenocal, 1995).

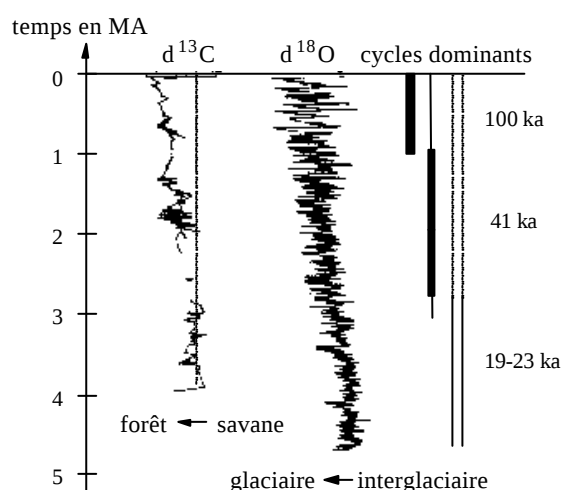


Figure 1.5. Parallèle entre les fluctuations environnementales (révélées par les isotopes de carbones, à gauche), **l'amplitude des cycles glaciaires** (représentée par les isotopes d'oxygène dans les Foraminifères des fonds marins, au milieu) **et leur périodicité** (à droite), **emprunté à deMenocal (1995).**

A ces cycles climatiques, il faut ajouter de grands changements à répercussion mondiale, liés à la tectonique des plaques. Ce sont, par exemple, les refroidissements liés à la fermeture du canal de Panama il y a 2,5 MA et aux dernières phase de surrection de l'Himalaya et de la cordillère des Andes, il y a 0,7 à 0,9 MA (cf. Ruddiman et Kutzbach, 1989 ; Hallam, 1994). D'autre part, le passage d'un système glaciaire unimodal (une seule calotte glaciaire : en Antarctique) à un système bimodal (deux calottes glaciaires : Antarctique et Arctique) vers 2,5 MA aurait aussi entraîné un refroidissement global (Ruddiman *et al.*, 1986).

Dans les régions tropicales, les données recueillies dans les sédiments lacustres et marins révèlent une diminution de la température et de la pluviosité durant les périodes glaciaires, et une augmentation de ces deux paramètres durant les inter-glaciaires (deMenocal, 1995 ; Maley, 1996).

Parallèlement, il y aurait eu un déclin des forêts de plaine et une expansion de la savane en période froide et sèche (e.g. Hooghiemstra et Agwu, 1988 ; Jahns *et al.*, 1998 – cf. **Fig. 1.5**), sans pour autant que l'un ou l'autre de ces milieux disparaisse complètement. Des macro-fossiles végétaux trouvés en bordure de la cuvette congolaise, au Nord du pays Bateke, au Congo, attestent de la persistance de zones de savane durant les périodes humides et de la persistance d'îlots forestiers durant les périodes arides, au cours des derniers 30.000 ans (Dechamps *et al.*, 1990).

La situation est inversée pour les forêts de montagne, qui sont actuellement fragmentées et auraient eu une extension supérieure pendant les périodes froides et sèches. On estime que la limite des forêts de montagne, qui se situe actuellement à 1500 m d'altitude, aurait subi un abaissement de 1000 m durant les périodes les plus froides (cf. Coetzee, 1964 ; Bonnefille *et al.*, 1990).

Au cours des temps géologiques, il y a eu de nombreux cycles d'oscillations climatiques doublés d'événements entraînant des changements climatiques globaux. Les périodes glaciaires auraient été associées à un climat froid et sec en Afrique tropicale, tandis que les inter-glaciaires auraient été associées à un climat chaud et humide. Parallèlement, il y aurait eu expansion de la savane durant les périodes glaciaires et expansion de la forêt durant les inter-glaciaires.

4) Chronologie des conditions paléo-environnementales en Afrique tropicale

a) Secondaire et tertiaire

Les données disponibles pour cette période ne permettent pas de décrire précisément les conditions climatiques pour l'Afrique tropicale. Nous présenterons donc une synthèse pour l'ensemble du continent africain.

Au début du Crétacé (cf. échelle des temps géologique en Annexe 1), l'Afrique faisait encore partie du Gondwana. Elle s'est séparée de l'Amérique du Nord au cours du Crétacé. De la fin du Crétacé, il y a environ 70 MA, jusqu'à la fin du Tertiaire, le continent africain se trouvait beaucoup plus au Sud que maintenant par rapport à l'Equateur (Guiraud et Maurin, 1991). Il est probable que toute la moitié Sud de l'Afrique, jusqu'au bassin du Congo, était aride et couverte par une formation sableuse. La moitié Nord aurait été occupée par une forêt humide bordant la Thétys, de taille fluctuante selon les conditions climatiques (Maley, 1996). Des sédiments très anciens et d'origine détritique, trouvés dans les fonds marins atlantiques (Robert, 1980 ; Robert et Chamley, 1987) révèlent selon les époques un climat sec (abondance de smectites) ou un climat chaud et humide (abondance de kaolinite). Jusqu'à l'Eocène, le climat africain était vraisemblablement assez sec. Il aurait par contre été plutôt humide du début de l'Oligocène au début du Miocène et de la moitié à la fin du Miocène.

Au début du Miocène, il y a environ 19 MA, époque à laquelle s'est produite la fermeture de la Téthys, des corridors forestiers devaient relier les principaux blocs forestiers africains et indo-australasiens (Thomas, 1985 ; Cox et Moore, 1993). Par la suite, ces blocs se sont trouvés isolés du fait de la surrection de montagnes en Arabie, en Turquie et au Moyen-Orient (il y a 5-10 MA, Dercourt *et al.*, 1986), de l'ouverture de la Mer Rouge (vers 3,5 ou 4 MA, Hsü *et al.*, 1977) et d'une augmentation de l'aridité (Cox et Moore, 1993).

Le rift africain a subi des phases de soulèvement (1500 m en moyenne, mais jusqu'à 4300 m dans les monts Ruwenzori) et d'effondrement. Son activité aurait commencé vers 45-30 MA, se serait poursuivie vers 18-14 MA, puis vers la fin du Miocène, de sorte que sa structure en graben se serait mise en place avant 3,5 MA (Le Turdu *et al.*, 1999). Il aurait continué son extension et connu des éruptions importantes par la suite, notamment vers 2,5 MA et à la charnière entre Pliocène et Pléistocène. Ces événements ont donné lieu à la formation de barrières, et auraient également eu un impact climatique (e.g. Partridge *et al.*, 1995).

Très peu d'informations sont disponibles pour le Pliocène, mais il semble qu'il se serait produit une transition vers une végétation herbacée à cycle en C4 entre 7 et 5 MA (Cerling *et al.*, 1993). Ce changement est suggéré par les variations de contenu en isotopes de carbone des paléosols et de

l'émail des dents de Mammifères fossiles, qui reflète la proportion relative de plantes à cycles en C3 (la plupart des arbres et herbacées des milieux tempérés actuels) et en C4 (caractéristique des herbacées des milieux arides). D'autres part, on dispose de quelques données palynologiques remontant jusqu'à 8 MA pour l'Est africain (Yemane *et al.*, 1985). Dans le créneau 4 à 4,5 MA, il semble qu'il y ait eu prédominance de la forêt humide. La forêt se serait maintenue jusque vers 3 MA, puis une période de transition aurait suivi, entre 2,5 et 2,8 MA. Cette phase de refroidissement aurait fortement affecté le biome africain (Vrba, 1995). La transition entre le Pliocène et le Pléistocène se caractérise par une augmentation de l'aridité et un abaissement de la température, ainsi que par une amplification des oscillations climatiques. Ce phénomène pourrait avoir résulté de la mise en place de la calotte glaciaire Arctique (Ruddiman *et al.*, 1986).

b) Pléistocène

Pour la période comprise entre 1 et 2 MA, il y aurait eu un pic d'aridification modérée il y a 1,7 MA (Shackleton *et al.*, 1984). Ce pic est corroboré par des données portant sur les isotopes d'oxygène, le taux de poussière éolienne, les pollens fossiles et par des modélisations basées sur les variations de température à la surface des océans (deMenocal, 1995).

Les données fournies par les dépôts des grands fonds apportent des informations sur les derniers 800.000 ans (e.g. Dupont et Agwu, 1992 ; Dupont *et al.*, 1996 ; Jahns, 1996 ; Jahns *et al.*, 1998). La sécheresse et l'amplitude des variations climatiques semblent avoir été plus prononcées au cours des derniers 800.000 ans qu'elles ne l'avaient été auparavant (Maley, 1996). Ce phénomène s'est encore accentué au cours des derniers 200-250.000 ans (Dupont et Agwu, 1992). Il y aurait eu un pic d'aridification très marqué vers 700-900.000 ans (Shackleton *et al.*, 1984) ou 600-700.000 ans (Jansen, 1990). Plus récemment, il y aurait eu une phase aride et froide entre 65 et 75.000 ans (van Andel et Tzedakis, 1996). En conséquence, au cours des derniers 800.000 ans, le couvert forestier a subi de grandes variations d'amplitude, passant d'une extension supérieure à la situation actuelle (10% du temps) à une situation de grande fragmentation qui peut être qualifiée de situation de "refuges" (10% du temps ; Maley, 1996).

Il semble que les fluctuations climatiques n'ont pas toujours été parfaitement synchrones à l'échelle de l'Afrique tropicale. Par exemple, la datation des pics de *Podocarpus* peut varier de 10.000 ans entre différents sites d'Afrique centrale et d'Afrique de l'Est (Dupont *et al.*, 1996).

c) Période récente et Holocène (quarante derniers milliers d'années)

Les informations dont nous disposons proviennent essentiellement des données recueillies dans les sédiments lacustres. Ils indiquent une alternance de phases climatiques de courte durée (e.g. Bonnefille *et al.*, 1990 ; Maley et Brenac, 1998 ; Vincens *et al.*, 1998). La synthèse réalisée par Maley (1990) en conclusion du chapitre sur les paléo-environnements continentaux en Afrique centrale atlantique fait apparaître trois grandes périodes :

- 0 à 12.000 ans B.P. (Kibangien) : période relativement chaude et humide qui peut être subdivisée en trois phases (0-3.500 ans ; 3.500-7.000 ans et 7.000-12.000 ans), avec extension maximale de la forêt vers 7-8.000 ans (**fig. 1.6 B à 1.6 D**).

- 12.0000 à 30.000 ans B.P. (Léopoldvillien) : phase relativement aride et froide (surtout entre 15.000 et 20.000 ans B.P.), abaissement de la température de 6 à 9°C, expansion de la savane, abaissement de la limite de la végétation montagnarde, avec extension forestière minimale il y a environ 18.000 ans (**fig. 1.6 A**).

- 30.0000 à 40.000 ans B.P. (Njilien) : phase relativement humide, transgression forestière.

En Afrique tropicale, le climat a dû être globalement assez sec jusqu'à l'Eocène, puis plutôt humide jusqu'à la fin du Pliocène, et à nouveau sec à partir du Pléistocène. La sécheresse et l'amplitude des variations climatiques semblent avoir été plus marquées au cours des derniers 800.000 ans (Maley, 1996) et encore plus durant les derniers 250.000 ans (Dupont et Agwu, 1992).

Au cours des derniers 18.000 ans, le couvert forestier a connu des variations extrêmes, allant d'un tissu extrêmement fragmenté il y a 18.000 ans à un tissu continu il y a 8.000 ans. L'extension actuelle de la forêt est proche de son maximum.

5) **Fragmentation forestière et évolution de la faune en Afrique tropicale**

a) Informations apportées par les fossiles

Si l'on exclut l'Est africain, peu de fossiles ont été trouvés en Afrique tropicale, en partie à cause du peu de fouilles réalisées, mais surtout du fait des conditions climatiques défavorables à la préservation des tissus osseux en milieu forestier (Van Neer, 1990). En 1990, seuls 3 sites camerounais, 1 site congolais et 6 sites situés en R.D.C., datés d'environ 40.000 à 2.000 ans B.P., avaient été inventoriés sur notre zone d'étude. Sur certains sites actuellement forestiers, la faune est caractéristique d'un environnement ouvert (exemple : Matupi, en R.D.C.).

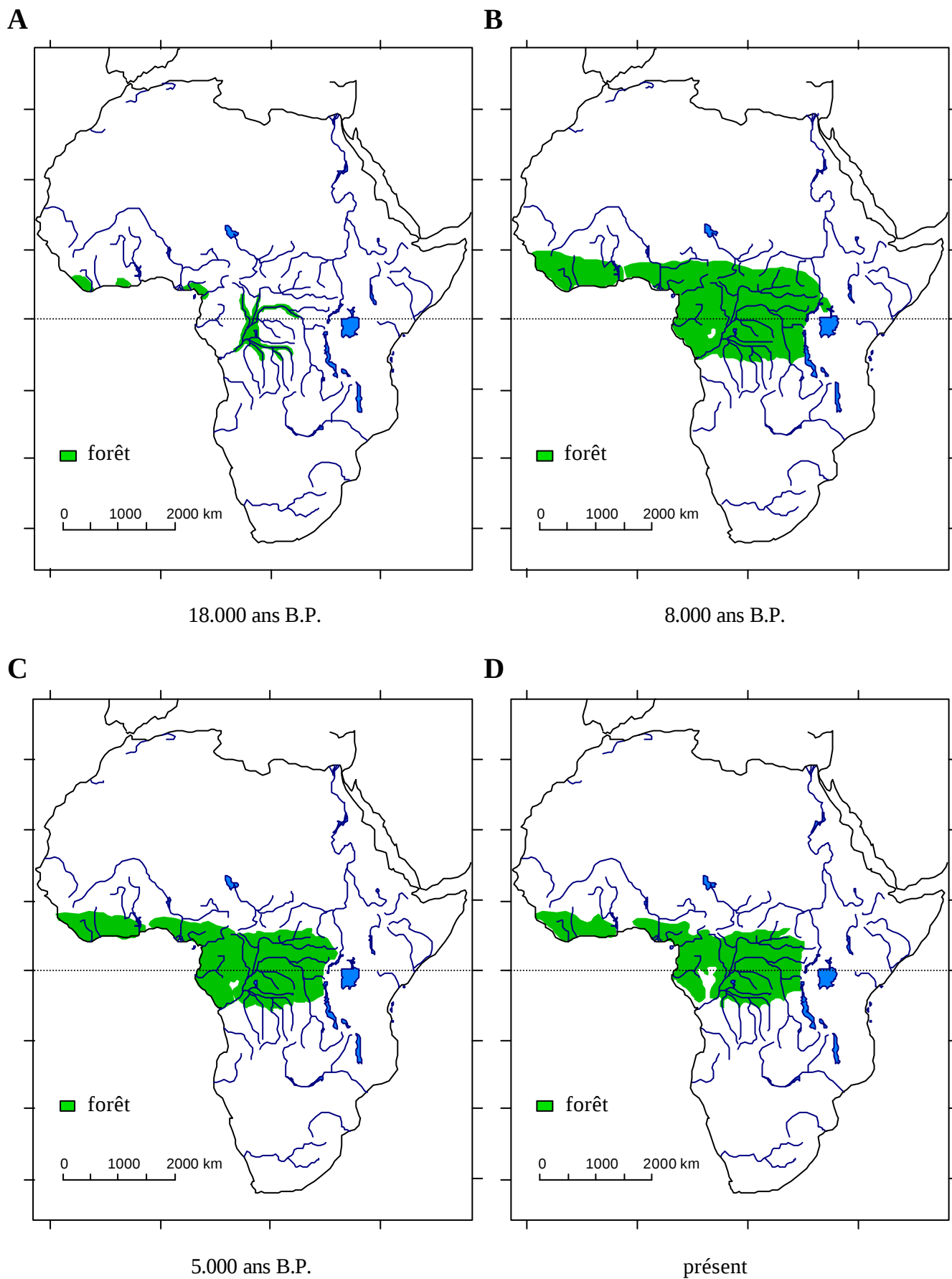


Figure 1.6. Reconstitution hypothétique de l'évolution du couvert forestier en Afrique au cours des derniers 18.000 ans, basée sur une revue de la littérature réalisée par Adams et Faure (1997).

Sur d'autres sites, il apparaît une transition entre faune savanicole et faune forestière à la fin du Pléistocène (exemple : Ntadi Yomba, au Congo). Ces changements dans les types de faune fossile sont interprétés comme étant la conséquence de changements paléo-environnementaux récents.

Sur une échelle temporelle beaucoup plus grande, des phases d'extinction et de renouvellement des faunes ont été mises en évidence en Afrique, à la charnière entre l'Eocène et l'Oligocène, au début du Miocène et entre le Pliocène et le Pléistocène (Maglio, 1978). Alors que Maglio (1978) explique les phases d'apparition de nouvelles espèces par des événements d'immigration de faunes en provenance d'Eurasie, Vrba (1985, 1995) les interprète comme une conséquence des événements climatiques globaux. Pour la période du Plio-Pléistocène, le refroidissement qui s'est produit il y a environ 2.5 à 2.7 MA aurait entraîné la régression du couvert forestier et l'apparition de formes savaniques chez les bovidae (Vrba, 1995) et les rongeurs (Wesselman, 1985).

b) Preuves de l'effet de la fragmentation forestière actuelle sur la faune

De plus en plus d'études tendent à montrer l'impact de la fragmentation forestière sur la diversité animale (e.g. Quinn et Harrison, 1988 ; Cornelius *et al.*, 2000 ; Goodman et Rakotondravony, 2000). La faune des îlots forestiers est souvent une réplique appauvrie de celle qui occupe le bloc forestier principal le plus proche, et le nombre d'espèces qui survivent dans des îlots forestiers décroît progressivement avec la taille des îlots. C'est notamment le cas pour les petits Mammifères insectivores du haut plateau central de Madagascar (Goodman et Rakotondravony, 2000). Même pour les Oiseaux qui ont pourtant la capacité de se déplacer par voie aérienne sur des distances parfois importantes, une étude réalisée dans le centre du Chili révèle que les îlots forestiers ne permettent pas le maintien de la diversité présente dans le bloc principal (Cornelius *et al.*, 2000). Le nombre d'espèces est fonction de la surface du lambeau forestier et de la présence ou non d'espèces occupant les écosystèmes avoisinants. Les espèces qui ne se maintiennent pas semblent être celles qui sont de plus grande taille, plus spécialisées et moins aptes à la dispersion.

Les fluctuations climatiques et les fluctuations environnementales qui en ont résulté ont conduit à des modifications dans la distribution de la faune et ont pu favoriser des événements de spéciation. La connaissance des phénomènes paléo-écologiques est donc nécessaire à l'interprétation des schémas biogéographiques et phylogéographiques.

I - ZONE D'ETUDE

1) Situation géographique

Nous rappelons que deux échelles géographiques seront considérées. Dans un premier temps, nous nous intéresserons à l'ensemble des forêts de plaine de la zone intertropicale, c'est-à-dire au bloc forestier guinéo-congolais dans son ensemble, ainsi qu'à la zone de mosaïque forêt-savane qui les borde (zone A ; **fig. 2.1**). Cette zone est presque entièrement comprise entre les 12 ° parallèles Nord et Sud et est limitée à l'Est par la région du rift africain, qui débute vers le 30 ° méridien. Dans un deuxième temps, nous nous concentrerons sur une portion du bloc forestier d'Afrique centrale située dans la région Cameroun-Gabon-Congo et comprise entre 6°N et 6°S et entre 8°E et 20°E (zone B ; **fig. 2.1**).

2) Caractéristiques géophysiques de l'Afrique tropicale

a) Géomorphologie

Pour l'ensemble de la zone A, le relief est seulement interrompu par des montagnes isolées à l'Ouest et au centre (**fig. 2.1**). Le relief très accidenté de la région du rift africain ne sera pas détaillé car il se trouve en bordure de la zone d'étude.

En Afrique centrale (zone B), le relief est peu découpé, et les massifs montagneux ne dépassent pas 1500 m, à l'exception du Mont Cameroun et de la chaîne du Mont Oku, situés au Cameroun, qui s'élèvent respectivement à 4070 m et 3008 m. En dehors de ce massif montagneux, deux ensembles régionaux se dessinent (Petit, 1990) :

- la cuvette congolaise, qui entoure la boucle formée par le fleuve Congo. Elle couvre une surface d'un million de km², dont un tiers est inondé au moins pendant une partie de l'année. Le centre de la cuvette est presque plat, et s'étale à une altitude comprise entre 270 et 400 m. La cuvette est dominée par un bourrelet périphérique de hautes terres composé de plaines et de plateaux atteignant 1000 m.

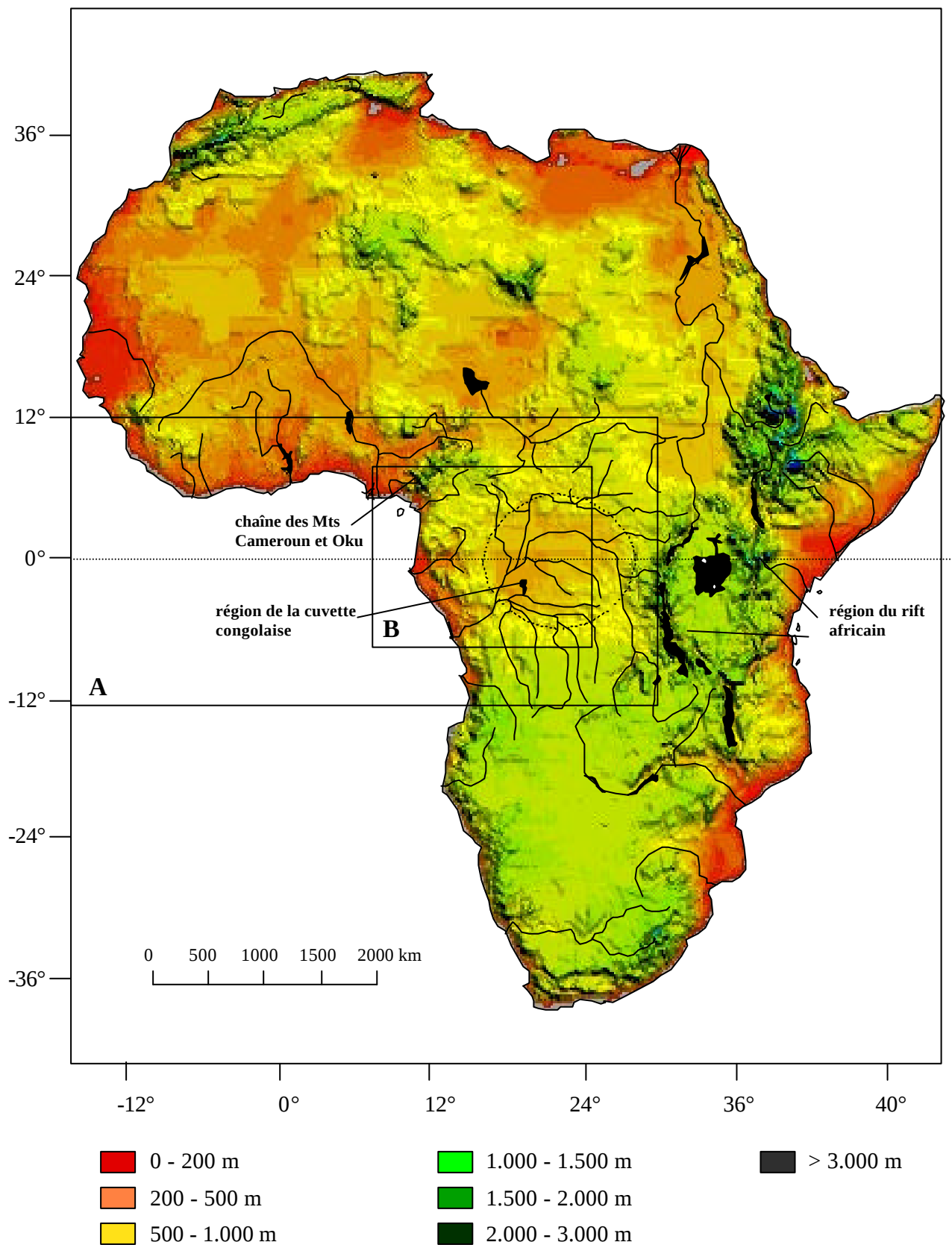


Figure 2.1. Le relief de l'Afrique (source : Encarta).

A et B : zones d'étude.

- la bordure atlantique, comprenant la chaîne du Mayombe et la retombée occidentale du socle cristallin au Nord. Elle se compose d'une façade littorale basse et sableuse et d'un arrière-pays montagneux. Au Sud, se dressent des chaînes montagneuses formant des ensembles bien définis, tels que la chaîne du Mayombe, le massif du Chaillu et la plate-forme fortement plissée du Niara Nyanga. Au Nord-Ouest, plusieurs pénéplaines se succèdent en une série de plateaux étagés : la pénéplaine du Chaillu, celle de l'Ivindo, puis celle du Ntem.

b) Hydrologie

Dans la zone tropicale A, le réseau hydrographique est très développé. Il comprend de nombreux fleuves et leurs affluents, au débit très élevé et variable selon les saisons (**fig. 2.2**).

Dans la zone B, on distingue trois grands bassins versants correspondant aux principaux fleuves (Bricquet, 1990) :

- la Sanaga, plus grand fleuve du Cameroun (1043 km), draine environ 131.000 km². Son débit annuel est de 65 milliards de m³, soit un débit moyen de 2.072 m³ par seconde à Edéa, près de l'estuaire.

- l'Ogooué, plus grand fleuve du Gabon (1200 km), draine environ 214.000 km². Son débit moyen est de 4.400 m³ par seconde à Lambaréné, juste avant la zone deltaïque. Son principal affluent, l'Ivindo, draine 63.000 km².

- le Congo, second fleuve du monde par son débit (41.000 m³ par seconde en moyenne, à Brazzaville) et cinquième par sa longueur (4.374 km), draine environ 3.684.000 km². Sa largeur dépasse localement 30 km, et est de 10 km à l'estuaire. Ses principaux affluents, tels que la Sangha et l'Oubangui, ont également un débit élevé.

c) Géologie et pédologie

Pour l'essentiel de la zone A, la roche mère correspond au bouclier pré-cambrien (Juo et Wilding, 1996). Les bassins de la Volta et du Niger-Benue, ainsi que la périphérie de la cuvette congolaise, reposent sur des roches remontant au moins au Crétacé. Le cœur de la cuvette congolaise est composé de sédiments lacustres et fluviatiles, déposés entre le Jurassique et le quaternaire, et atteignant localement 1.000 mètres d'épaisseur (Lepersonne, 1974).

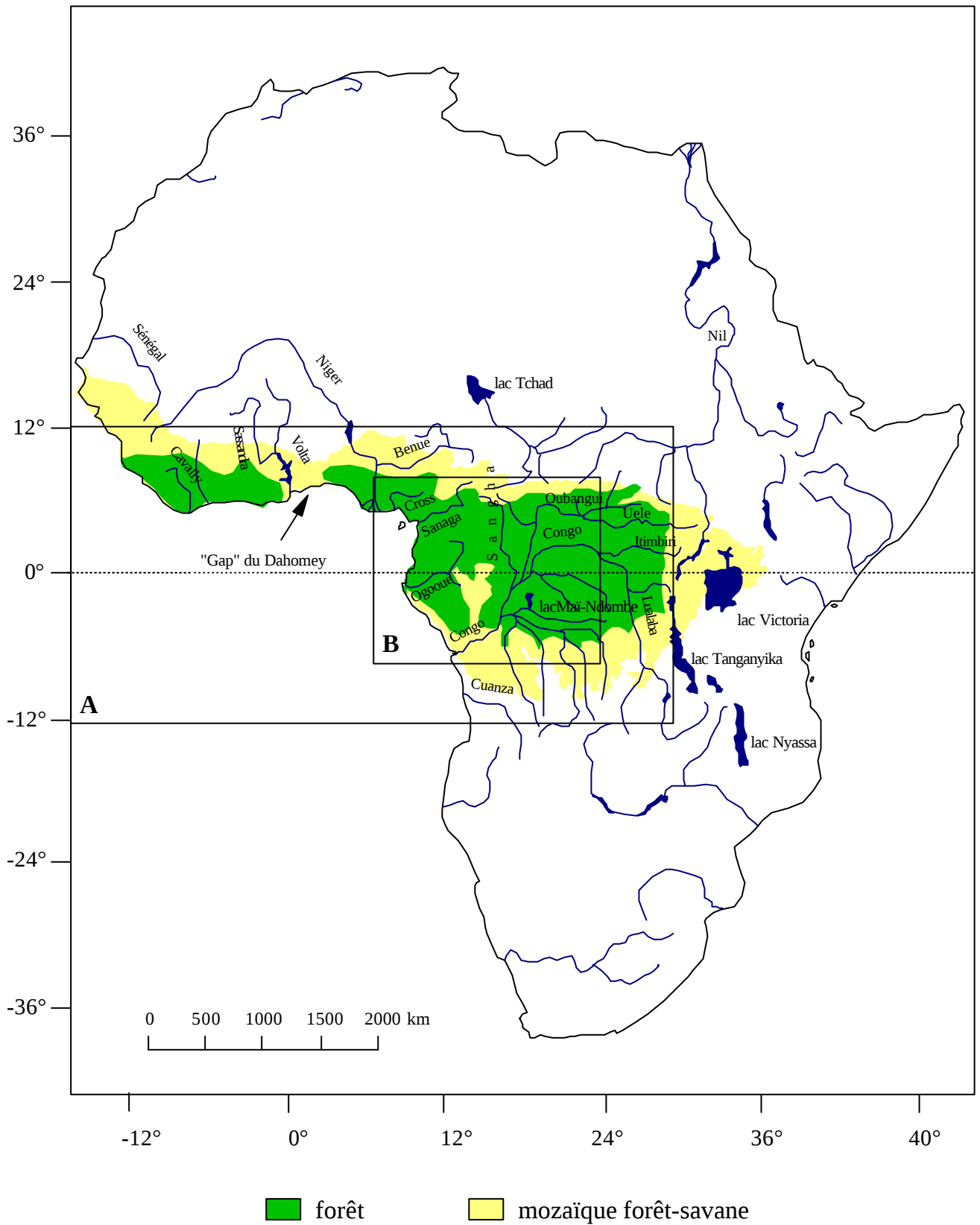


Figure 2.2. Réseau hydrographique de l'Afrique, limites de la forêt tropicale et zone de mosaïque forêt-savane.

Les reliefs de la région du Mont Cameroun sont constitués de roches volcaniques. Les sols sont partout de nature argileuse et pauvres en matière organique (Juo et Wilding, 1996). Ils diffèrent entre l'Afrique de l'Ouest, où l'on trouve des Alfisols, peu développés et modérément lessivés, et l'Afrique centrale, où les sols sont de type Oxisols (aussi appelés ferrasols ou sols ferralitiques), très développés et intensément lessivés.

Au sein de la zone B, la nature de la roche mère et des sols est assez homogène (Juo et Wilding, 1996). La roche mère est constituée par le bouclier pré-cambrien, recouvert par une couche de sédiments plus récents au niveau de la cuvette congolaise. Les sols sont essentiellement de type Oxisols et atteignent 20 mètres d'épaisseur en moyenne. Ils résultent d'une pédogenèse régie par un fort drainage et des températures élevées, conditions qui ont pour conséquences une altération complète des minéraux primaires (sauf le quartz), l'élimination d'une grande partie de la silice, et la formation d'oxydes de fer et d'aluminium (gibbsite, goethite et hématite) et d'argile de type kaolinite (Schwartz, 1990).

Il existe des variations locales de la composition des sols, en relation avec les conditions climatiques. Ainsi, les sols sont de type oxysols jaunes, plus acides, pâles et pauvres, dans les régions les plus arrosées (Juo et Wilding, 1996 ; Lawson, 1996), et ils sont rouges, du fait de leur richesse en kaolinite, dans les zones les plus sèches du Gabon (rapport EDICEF, 1983). Dans certaines régions géographiquement limitées, les oxysols sont associés à des sols hydromorphes minéraux ou organiques (Schwartz, 1990). Dans la plaine alluviale du Congo, les sols sont de type Inceptisol, moins développés et aux horizons moins différenciés (Juo et Wilding, 1996).

d) Climatologie

Régimes climatiques

Le climat tropical actuel est sous l'étroite dépendance des basses pressions intertropicales, des dépressions thermiques sahariennes et angolaises et des hautes pressions subtropicales (Mpounza et Samba-Kimbata, 1990). Leur position respective par rapport à l'Equateur météorologique, qui diffère un peu de l'Equateur géographique, détermine le régime des pluies. De décembre à février, le balancement des hautes pressions vers le Sud entraîne une période de sécheresse au Nord et de forte pluviosité au Sud. De juin à août, le rejet des hautes et des basses pressions vers l'hémisphère Nord est à l'origine d'une pluviosité accrue au Nord et d'une période de sécheresse au Sud. Les pluies sont déclenchées par la pénétration de la zone de convergence intertropicale dans celle des hautes pressions subtropicales. Comme celle-ci a lieu durant l'été de l'hémisphère d'implantation, c'est-à-dire durant une période de

forte insolation, il y a diminution de la pression dans la zone de convergence, ce qui entraîne une augmentation du gradient de pression entre cette zone et celle des hautes pressions, et est à l'origine de la mousson.

Le climat d'Afrique tropicale se caractérise par des températures élevées toute l'année (moyenne diurne $> 25^{\circ}\text{C}$), et par des pluies abondantes (Pagney, 1976). On distingue plusieurs régimes climatiques en fonction de la distribution annuelle des pluies (**fig. 2.3**) :

- climat équatorial : régime des pluies bimodal, avec deux minima aux solstices et deux maxima à l'équinoxe.
- climat subéquatorial : régime des pluies bimodal, mais les minima sont plus accusés, surtout le minimum principal qui peut durer jusqu'à trois mois consécutifs sans pluies.
- climat tropical humide : régime des pluies unimodal, avec un minimum centré sur l'hiver de l'hémisphère d'implantation, soit de décembre à février pour les régions à affinité Nord et en juin-juillet pour les régions à affinité Sud.
- climat sub-tropical : plus chaud et sec. Ce type de climat n'est représenté qu'en marge de la zone A.

Dans la zone B, quatre régimes climatiques sont représentés (Mpounza et Samba-Kimbata, 1990). Le climat est équatorial de type océanique le long des côtes du Cameroun, de Guinée Equatoriale, et du Nord du Gabon ; équatorial de type continental au Sud du Cameroun, au Sud-Ouest de la RCA, à l'Est du Gabon et dans la partie forestière du bassin central congolais ; subéquatorial en lisière du bloc forestier (Sud de la RCA, centre du Congo et Nord de la RDC) ; et tropical humide au Sud du Congo et au Nord de l'Angola. Nous ne détaillerons les différents facteurs climatiques que pour la zone B.

Insolation

L'insolation est faible (< 2000 heures) du fait d'une nébulosité presque permanente (Mpounza et Samba-Kimbata, 1990). Elle est particulièrement basse en milieu forestier côtier, et plus élevée en milieu forestier continental. La radiation solaire est maximale durant l'été et minimale durant l'hiver de l'hémisphère d'implantation.

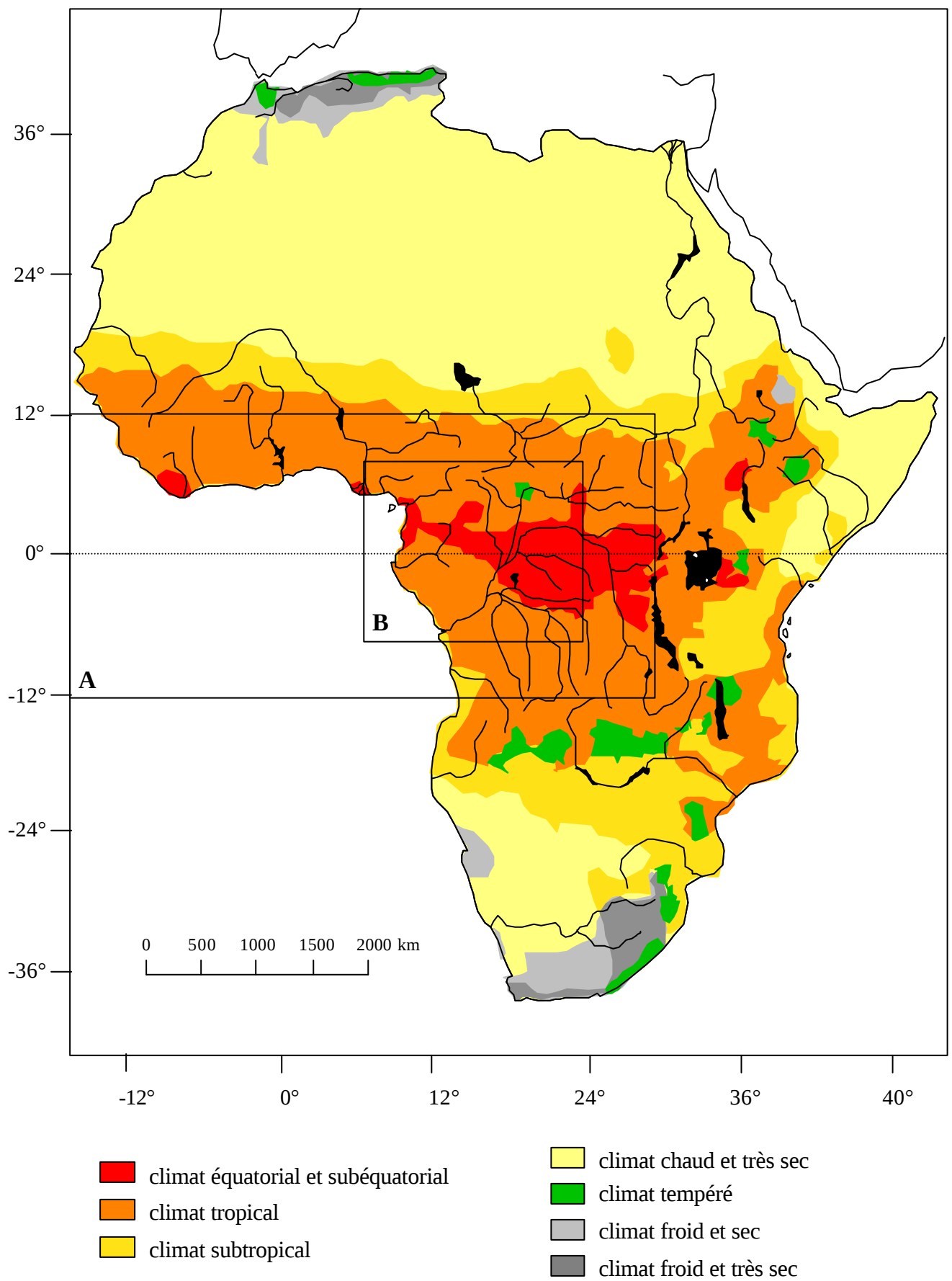


Figure 2.3. Les climats de l'Afrique (source : Encarta).

Températures

Les régimes thermiques sont peu contrastés (Mpounza et Samba-Kimbata, 1990). La température moyenne annuelle est de l'ordre de 26-27°C, aussi bien près de la côte qu'à l'intérieur des terres. L'amplitude thermique annuelle ne dépasse pas 4°C (à Brazzaville, Congo) et augmente avec la latitude. L'amplitude diurne est fonction de l'opacité moyenne : elle atteint 10°C lorsque le ciel est dégagé et 3°C lorsqu'il est couvert.

Humidité

L'humidité relative moyenne de l'air est partout très élevée (80-85%), en raison de la régularité des alizés et de la mousson atlantique (Mpounza et Samba-Kimbata, 1990). Elle présente cependant des variations régionales liées notamment à la présence de forêt, de lacs, de cours d'eau et d'étendues marécageuses. Elle est supérieure à 85% sur la côte camerounaise et la cuvette centrale congolaise. Elle descend en dessous de 80% au NW du Cameroun. L'amplitude saisonnière est faible (4 à 10%) alors que les écarts diurnes sont relativement élevés (20 à 40%).

Précipitations

Les secteurs les plus arrosés (plus de 2.000 mm de pluie par an) sont le centre de la cuvette congolaise et la frange côtière camerounaise et gabonaise. Dans la zone côtière, les maxima atteignent 11.000 mm au Mont Cameroun et les précipitations diminuent de la côte vers l'intérieur et du Nord vers le Sud. La plus grande partie de la zone d'étude reçoit des précipitations modérées comprises entre 1.400 et 2.000 mm de pluie par an, ce qui est peu comparé à la pluviosité des forêts tropicales d'Amérique du Sud et d'Asie (e.g. White, 1983). Les secteurs les moins arrosés (moins de 1.400 mm), situés en marge de la zone d'étude, correspondent à la frange côtière congolaise, aux dépressions du Nyari et de la Nyanga (au Congo) et de l'enclave de Cabinda (en RDC). La variabilité inter-annuelle oscille le plus souvent entre 10 et 15%, mais atteint 33% au NW du Congo. Le nombre de jours de pluie se situe globalement entre 100 et 200 jours par an, mais il peut atteindre 260 jours par an sur le versant SW du Mont Cameroun et sur les Monts de Cristal au Gabon. Dans de nombreux secteurs, tels que la frange côtière du Congo, de la RDC et de l'Angola, la région de Franceville au Gabon, le SW et le SE de la RCA et le Nord du Congo, le nombre de jours de pluie par an est inférieur à 100, ce qui est faible à ces latitudes.

e) Végétation

Régions phytogéographiques

L'Afrique tropicale est occupée en grande partie par une formation forestière de plaine, centrée autour de l'Equateur, et qui représente près d'un cinquième des forêts de la planète (Sayer *et al.*, 1992). Au Nord comme au Sud, la forêt est bordée par des zones de mosaïque de forêt et de savane, qui laissent progressivement place aux savanes soudaniennes, puis sahéliennes au Nord, et à des savanes arborées de type 'Miombo' au Sud (**fig. 2.4**). D'Ouest en Est, on distingue deux blocs forestiers principaux, séparés par une zone de mosaïque de forêt-savane appelée 'Gap du Dahomey'. A l'Est, la forêt de plaine est remplacée par des forêts de montagne dans la région des grands lacs, puis par une mosaïque de forêt-savane.

Pour la zone B, les travaux de Letouzey (1968) mettent en évidence 4 régions phytogéographiques forestières principales se succédant d'Ouest en Est : une formation littorale de type mangrove, une formation côtière Atlantique sempervirente et hygrophile à Caesalpiniacées, une formation continentale humide mixte, et une zone de forêt inondée abritant des formations hygrophiles (**fig. 2.5**). Elles sont bordées au Nord et au Sud par forêts semi-caducifoliées, et laissent place localement à des forêts clairsemées à Marantacées et Zingibéracées et à des forêts de montagne.

Les études locales qui ont été réalisées au Gabon (Caballé, 1978 ; Reitsma, 1988 ; White, 1992), au Cameroun (Letouzey, 1968), au Congo (de Foresta, 1990 ; Hecketsweiler *et al.*, 1991), et en République Démocratique du Congo (Evrard, 1968), révèlent la grande diversité des essences forestières et l'hétérogénéité de leur distribution. Il apparaît que le développement de la végétation est fortement influencé par les conditions climatiques, et notamment par les valeurs extrêmes et les variations saisonnières de la température et de la pluviosité. Actuellement, les conditions nécessaires au maintien des forêts tropicales semblent être une température élevée tout au long de l'année et une pluviosité élevée. La diversité floristique la plus forte se trouve dans les zones où la pluviosité est la plus élevée et répartie régulièrement sur l'année ; elle serait maximale sur les Monts de Cristal (Reitsma, 1988). Un autre facteur déterminant pour la végétation est le relief. Les régions d'altitude sont occupées par des plantes de montagne (e.g. White, 1983), tandis que dans les dépressions inondées pendant la majeure partie de l'année se développent des essences de sols hydromorphes, telles les raphiales de la cuvette du Congo (cf. Evrard, 1968).

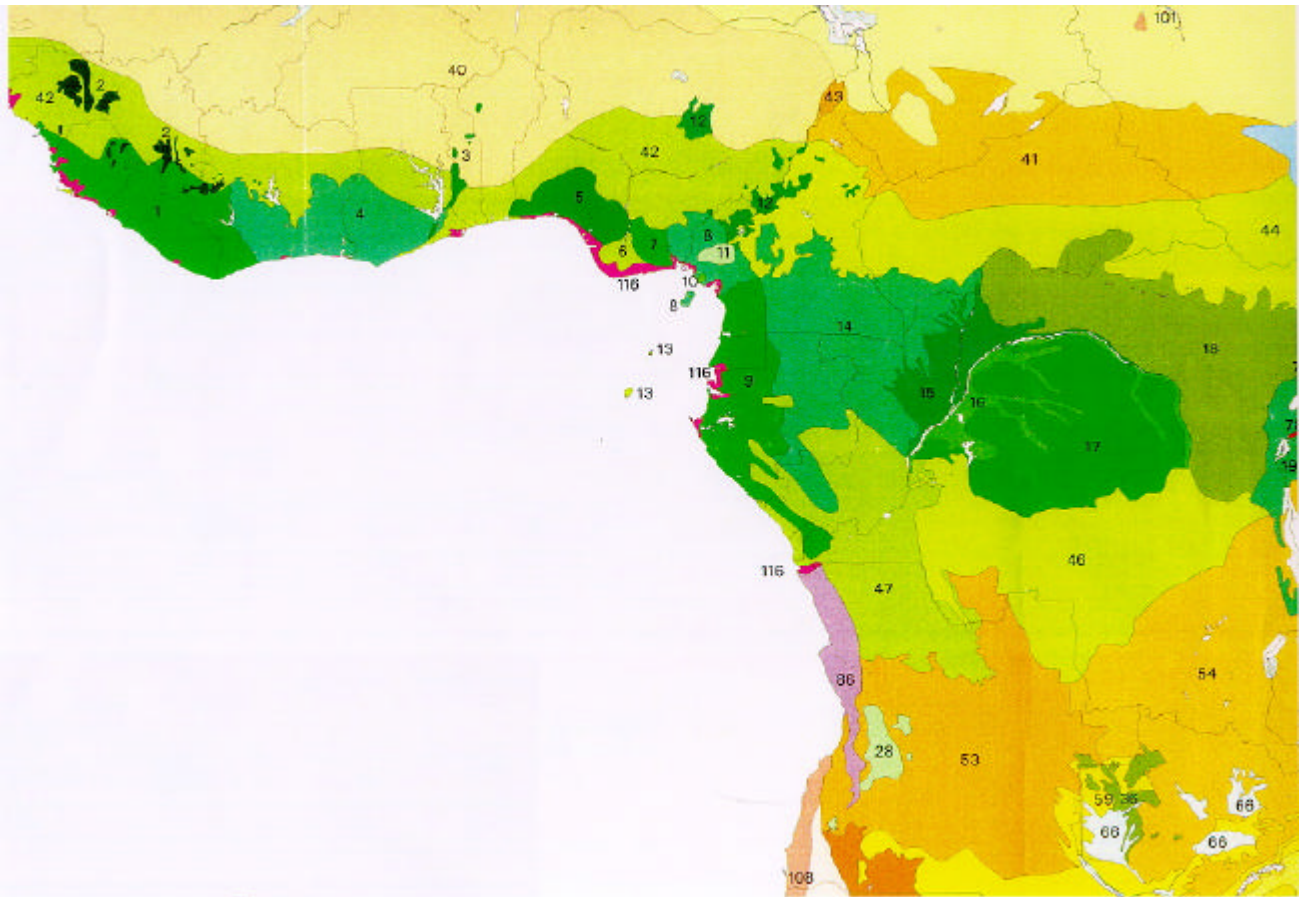


Figure 2.4. Les phytorégions d'Afrique tropicale (Underwood *et al.*, 1998).

Forêts tropicales et subtropicales humides

- 1 Forêt de plaine Ouest guinéenne
- 2 Forêt de montagne guinéenne
- 3 Forêt de montagne (highlands) du Togo
- 4 Forêt de plaine Est guinéenne
- 5 Forêt de plaine du Niger
- 6 Forêt marécageuse du delta du Niger
- 7 Forêt de transition Cross-Niger
- 8 Forêt côtière de la région Cross-Sanaga-Bioko
- 9 Forêt côtière équatoriale Atlantique
- 10 Forêt de montagne du Mont Cameroun et Bioko
- 11 Forêt de montagne Rumpi-Bakossi
- 12 Forêt de montagne (highlands) du Cameroun
- 13 Forêt humide de Sao Tome et Principe
- 14 Forêt de plaine de la région NW Congo
- 15 Forêt inondée de la région Ouest Congo
- 16 Forêt inondée de la région Est Congo
- 17 Forêt de plaine de la cuvette centrale congolaise
- 18 Forêt de plaine de la région NE Congo
- 19 Forêt de montagne du rift Albertine
- 28 Forêt de montagne d'Angola

Forêts tropicales et subtropicales sèches

- 36 Forêt sèche zambésienne à *Cryptosepalum*

Savanes herbeuses, arbustives et arborées tropicales et subtropicales

- 39 Savane à acacias sahélienne
- 40 Savane Ouest soudanienne
- 41 Savane Est soudanienne
- 42 Mosaïque de forêt-savane guinéenne
- 43 Mosaïque des plateaux Mandara
- 44 Mosaïque de forêt-savane de la région Nord Congo
- 46 Mosaïque de forêt-savane de la région Sud Congo
- 47 Mosaïque de forêt-savane de la région Ouest Congo
- 53 Savane arborée angolaise (Miombo)
- 54 Savane arborée zambésienne (Miombo)
- 59 Savane herbeuse Ouest zambésienne

Savanes herbeuses et arborées inondées

- 64 Savane inondée du lac Tchad
- 66 Savane herbeuse inondée zambésienne

Savanes herbeuses et arbustives de montagne

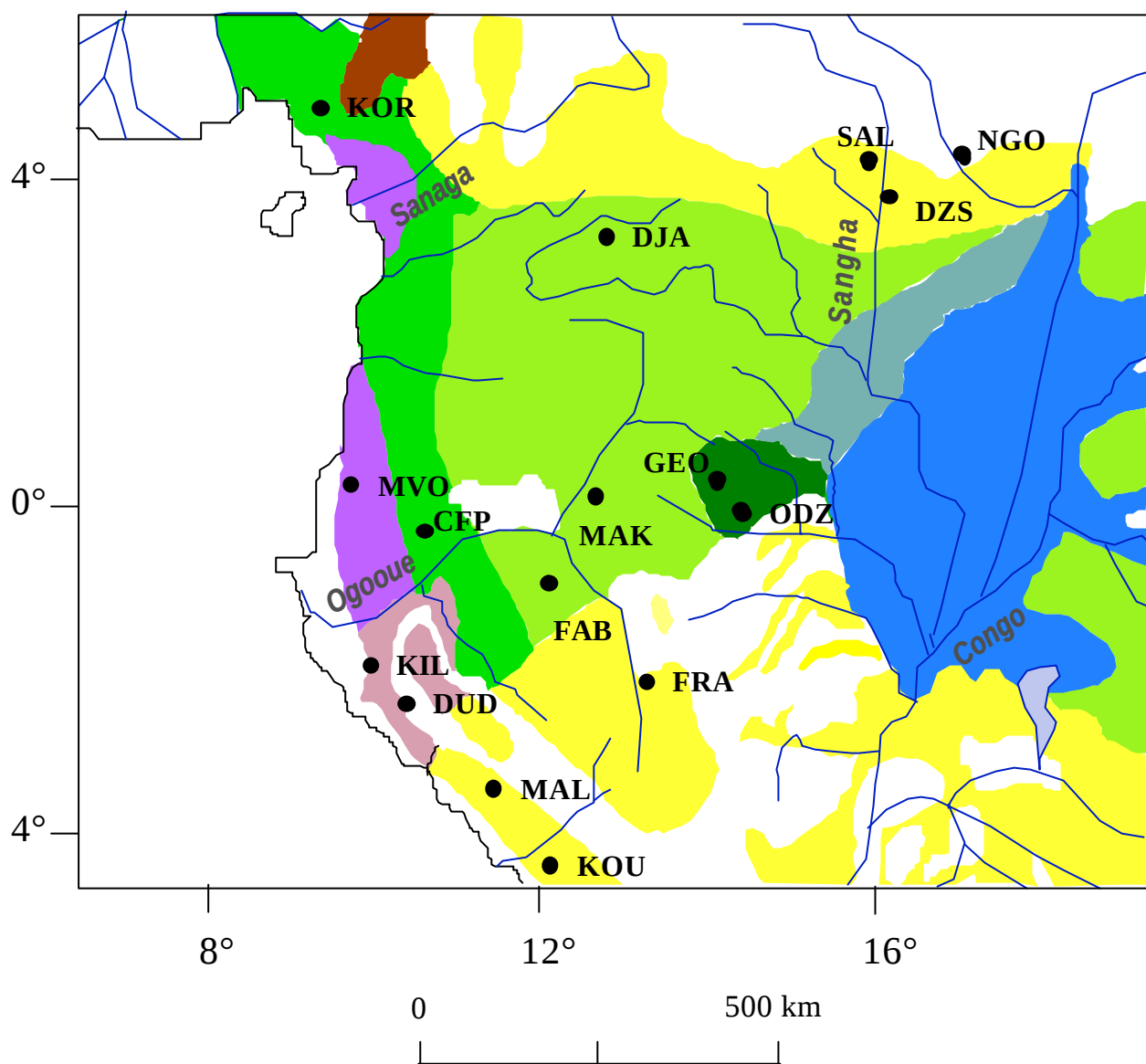
- 86 Savane et savane arborée des escarpements angolais

Savanes arbustives des régions désertiques et déserts

- 101 Savane arborée sèche des montagnes du Jebel Marra
- 108 Désert de Kaokoveld

Mangroves

- 115 Mangroves guinéennes
- 116 Mangroves d'Afrique centrale



	Forêts semi-caducifoliées		Zone inondée presque toute l'année, avec formations sempervirentes, raphiales...
	Forêts sempervirentes biafréennes et gabonaises à Caesalpiniacées		Forêts atlantiques littorales
	Végétation de montagne		Forêts clairsemées à Marantacées et Zingibéracées
	Forêts de type gabonais avec alternance ou mélange de formations sempervirentes et semi-caducifoliées		

Carte 2.5 : Végétation de l'Afrique Centrale (d'après Letouzey, 1968)

Les codes de localités font référence au tableau 2.1

Dynamique forêt-savanne

Il existe des zones de mosaïque forêt-savane, caractérisées par l'existence de savanes incluses en forêt ou de galeries forestières en milieu ouvert. La présence de savanes incluses dans des régions où le climat est favorable à la forêt semble être un témoin de l'extension de la savane lors de la dernière période aride (de Foresta, 1990). Certaines de ces savanes incluses se trouvent dans des zones de recolonisation forestière très récente (de l'ordre de quelques centaines d'années), comme l'atteste la présence d'espèces pionnières en lisière et de forêts clairsemées à marantacées en périphérie. Le maintien de ces savanes pourrait être expliqué par une aridification modérée au cours des derniers 3.000 ans et par l'action des herbivores, principalement celle des éléphants (e.g. Van de Koppel et Prins, 1998), dans des zones où le relief favorise des phénomènes d'inondation périodique. Elles sont parfois entretenues par l'homme, mais ne semblent pas être d'origine anthropique, du moins dans la région du Mayombe (de Foresta, 1990).

f) Occupation humaine

Il semble que l'Afrique centrale se soit peuplée plus tardivement que le reste de l'Afrique, sans doute lors de la phase climatique du Würm qui a commencé il y a 80.000 ans. Le refroidissement aurait entraîné une fragmentation de la forêt, facilitant sa pénétration par l'homme.

En périphérie de la cuvette congolaise, l'occupation humaine remonterait au moins à 60.000 ans (Bahuchet, 1986). Des outils datant de l'âge de pierre récent (16-50.000 ans) et moyen (<50.000 ans) ont été trouvés dans la cuvette, en bordure de la zone forestière (Congo, Gabon, RCA, RDC). Certains outils pourraient dater de l'âge de pierre ancien, époque à laquelle vivait *Homo erectus*. A cette période, il y a environ 100.000 ans, les sites concernés étaient vraisemblablement forestiers.

Seuls deux sites actuellement situés en forêt ont été découverts : la grotte de Hau au Nord-Ouest de la RDC (van Noten *et al.*, 1980) et un site du Sud du Cameroun (Omi et Kato, 1982).

Des traces d'occupation Néolithique (3.000-4.000 ans B.P.) sont présentes le long de la vallée de l'Ogooué (de Maret, 1990 ; Oslisly, 1995), mais les premières traces d'agriculture remontent seulement à environ 2.000 ans.

3) Sites de collectes

Des captures de petits mammifères et des prélèvements de biopsies ont été réalisés dans 22 localités d'Afrique centrale, ainsi qu'en Afrique de l'Ouest (1 localité) et en Afrique de l'Est (4 localités). Pour chaque localité, entre 1 et 21 sites voisins ont été échantillonnés (**fig. 2.6, tab. 2.1**).

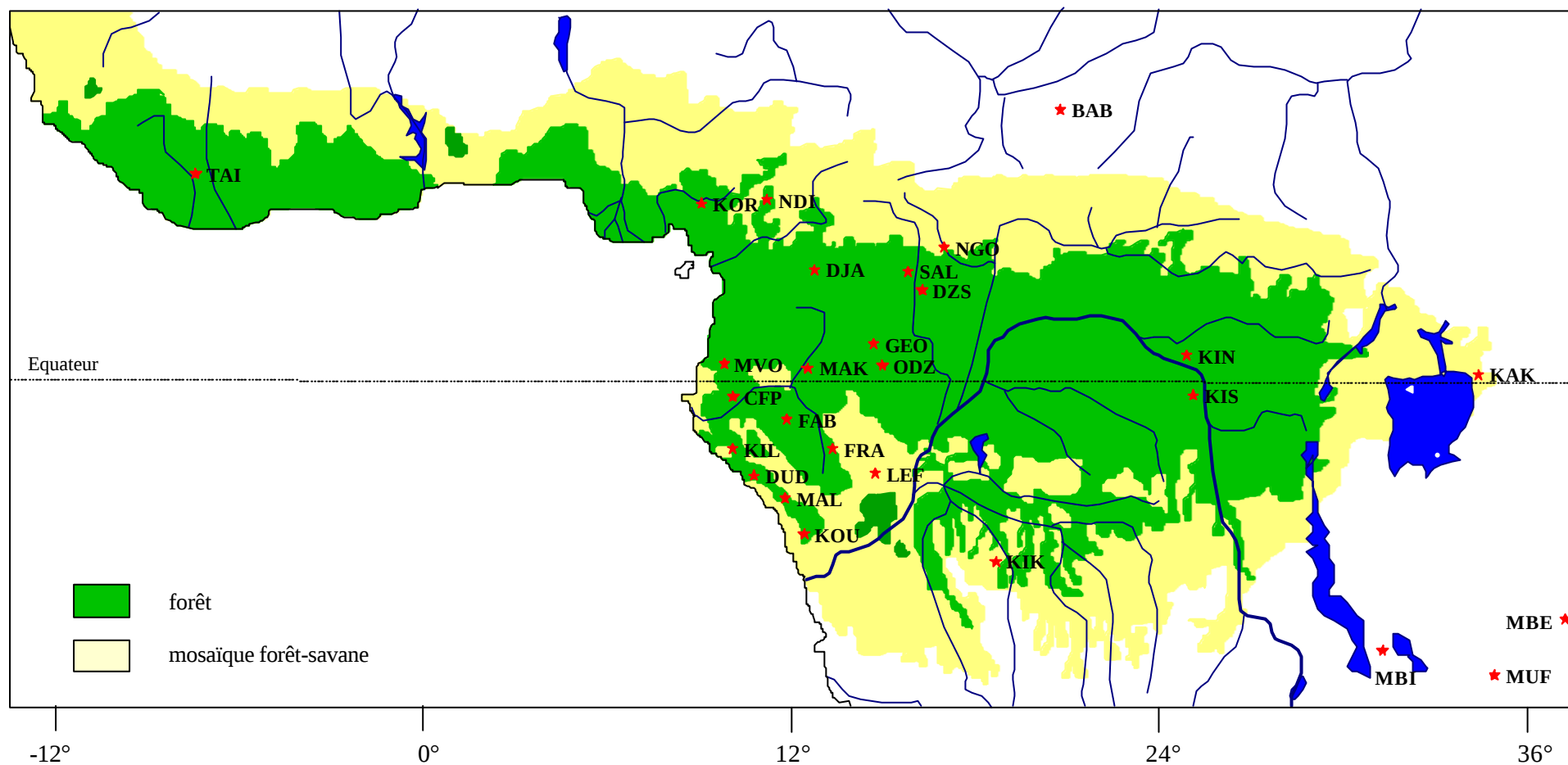


Figure 2.6. Principales localités de collecte des spécimens analysés.
Les codes de trois lettres sont ceux indiqués dans le tableau 2.1.

Tableau 2.1: Liste des principales localités de collectes avec nombre de sites distincts et coordonnées géographiques (F = forêt et S = savanne)

Code	Localité	Pays	Situation	N. sites	Latitude	Longitude
DJA	Dja	Cameroun	bloc forestier	11	3°20 N	13°00 E
KOR	Korup N.P.	Cameroun	bloc forestier	1	5°16 N	9°08 E
NDI	Nditam	Cameroun	mosaïque F-S	1	5°20 N	11°15 E
GEO	Escarpement	Congo	bloc forestier	1	1°04 N	14°29 E
KOU	Kouilou	Congo	mosaïque F-S	1	4°1- S	11°5- E
LEF	Léfini	Congo	mosaïque F-S	4	2°32 S	14°44 E
ODZ	Odzala	Congo	périphérie F	9	0°30 N	14°50 E
CFP	Four Place	Gabon	bloc forestier	2	0°15 S	10°20 E
DUD	Monts Doudou	Gabon	bloc forestier	1	2°18 S	10°30 E
FAB	Forêt des Abeilles	Gabon	bloc forestier	4	0°57 S	11°56 E
FRA	Franceville	Gabon	péri-urbain	1	1°40 S	13°31 E
KIL	Kili-Massana	Gabon	mosaïque F-S	2	1°35 S	10°10 E
MAK	Makokou	Gabon	bloc forestier	1	0°27 N	12°54 E
MAL	Malounga	Gabon	mosaïque F-S	1	3°09 S	10°45 E
MVO	Mvoum	Gabon	mosaïque F- péri-urbain	1	0°35 N	9°53 E
KAK	Kakamega	Kenya	F montagne	1	0°14 N	34°52 E
BAB	Bamingui-Brendja	RCA	isolat forestier	7	7°47 N	20°57 E
DZS	Dzanga-Sangha	RCA	bloc forestier	1	2°47 N	16°25 E
NGO	Ngotto	RCA	périphérie F	21	4°00 N	17°10 E
SAL	Salo	RCA	périphérie F	3	3°10 N	16°05 E
TAI	Tai	RCI	bloc forestier	2	5°52 N	7°27 W
KIK	Kikwit	RDC	mosaïque F-S	1	5°06 S	18°48 E
KIN	Kisangani Nord *	RDC	bloc forestier	2	0°50 N	25°15 E
KIS	Kisangani Sud *	RDC	bloc forestier	3	0°15 N	24°30 E
MBE	Mbete	Tanzanie	F montagne	1	6°52 S	37°41 E
MBI	Mbizi	Tanzanie	F montagne	1	7°5- S	31°4- E
MUF	Mufindi	Tanzanie	F montagne	1	8°35 S	35°20 E

* Kisangani Nord = rive droite et Kisangani Sud = rive gauche du fleuve Congo

II - MATERIEL BIOLOGIQUE

1) Choix des modèles

Les modèles ont été choisis au sein de trois Familles de mammifères de façon à prendre en compte l'hétérogénéité des caractéristiques biologiques et écologiques de la faune mammalienne. Ils incluent à la fois des petits mammifères (rongeurs, musaraignes) et des mammifères de plus grande taille (primates) ; des mammifères typiquement forestiers et des mammifères de l'écotone, voire de savane ; des animaux à tendance granivore (rongeurs), frugivore (primates) ou insectivore (musaraignes) ; des espèces longévives (primates) et des espèces dont la durée de vie n'excède généralement pas une année (micro-mammifères).

La plupart des spécimens ont été choisis parmi les petits et moyens Mammifères d'Afrique centrale collectés par des membres de notre équipe. En complément, 83 échantillons ont été obtenus sur demande auprès de chercheurs travaillant dans des laboratoires extérieurs, ou auprès de directeurs de Parcs Zoologiques (Annexe 2).

Les petits mammifères se prêtent bien aux analyses phylogéographiques car ils sont abondants, diversifiés et faciles à capturer. Leur étude pose cependant un problème éthique car, la détermination spécifique ne pouvant généralement pas se faire sur base de la morphologie externe, il est nécessaire d'euthanasier un certain nombre de spécimens. Les musaraignes, insectivores de la Famille des Soricidae, semblent constituer un modèle privilégié du fait de leur sensibilité aux modifications de l'environnement (Reumer, 1989). Elles présentent des adaptations à des conditions climatiques et biotiques très diverses (Wolsan et Wójcik, 1998). Ainsi, parmi les 9 genres et les 149 espèces présents en Afrique, les espèces du genre *Crocidura* occupent des habitats diversifiés, les *Sylvisorex* sont presque tous forestiers et les *Suncus* sont plutôt savanicoles (Wolsan et Hutterer, 1998).

Les rongeurs de la Famille des Muridae constituent également un modèle attractif car ils sont très diversifiés. Avec 1326 espèces reconnues, ils représentent près de 29% des espèces de Mammifères au monde (Musser et Carleton, 1993), contre 8% pour les musaraignes (Wolsan et Wójcik, 1998). En Afrique, on dénombre plus de 30 genres et 150 espèces (Kingdon, 1997), dont 9 genres et 32 à 35 espèces en milieu forestier (Dieterlen, 1990).

Un modèle complémentaire a été choisi au sein de l'ordre des primates qui est l'un des mieux connus en matière de taxinomie et de distribution. La Famille des Cercopithecidae est représentée en Afrique par au moins 13 genres et 39 espèces (Groves, 1993). Les espèces du genre *Cercopithecus* et des genres voisins, qui forment la tribu des Cercopithecini, sont principalement des espèces de forêt de

plaine, mais certaines habitent les forêts de montagne ou la savane (e.g. Kingdon, 1997).

Pour tester les scénarios proposés pour les Mammifères d'Afrique Centrale (Colyn, 1999 ; Delporte et Colyn, 1999), nous avons choisi de travailler au rang de l'espèce et de la super-espèce. Certains modèles ont été choisis parmi des genres monotypiques ou des espèces dont la détermination ne semblait pas poser de problème particulier (rongeurs de l'espèce *Stochomys longicaudatus*, primates de la super-espèce *Cercopithecus cephus*). Pour certains genres dont la taxonomie est délicate et au sein desquels il n'est pas toujours possible de déterminer les spécimens au rang de l'espèce dans l'état actuel des connaissances, une révision systématique a été menée sur des critères morpho-anatomiques et moléculaires, en collaboration avec des spécialistes (musaraignes des genres *Crocidura* et *Sylvisorex*, rongeurs du genre *Hylomyscus*). Les résultats de ces analyses préliminaires ont permis de choisir des espèces bien représentées dans la région ciblée et présentant une variabilité moléculaire suffisante. Ils ont aussi permis de mettre en évidence des mouvements géographiques à l'échelle continentale et de procéder à une datation des principaux événements.

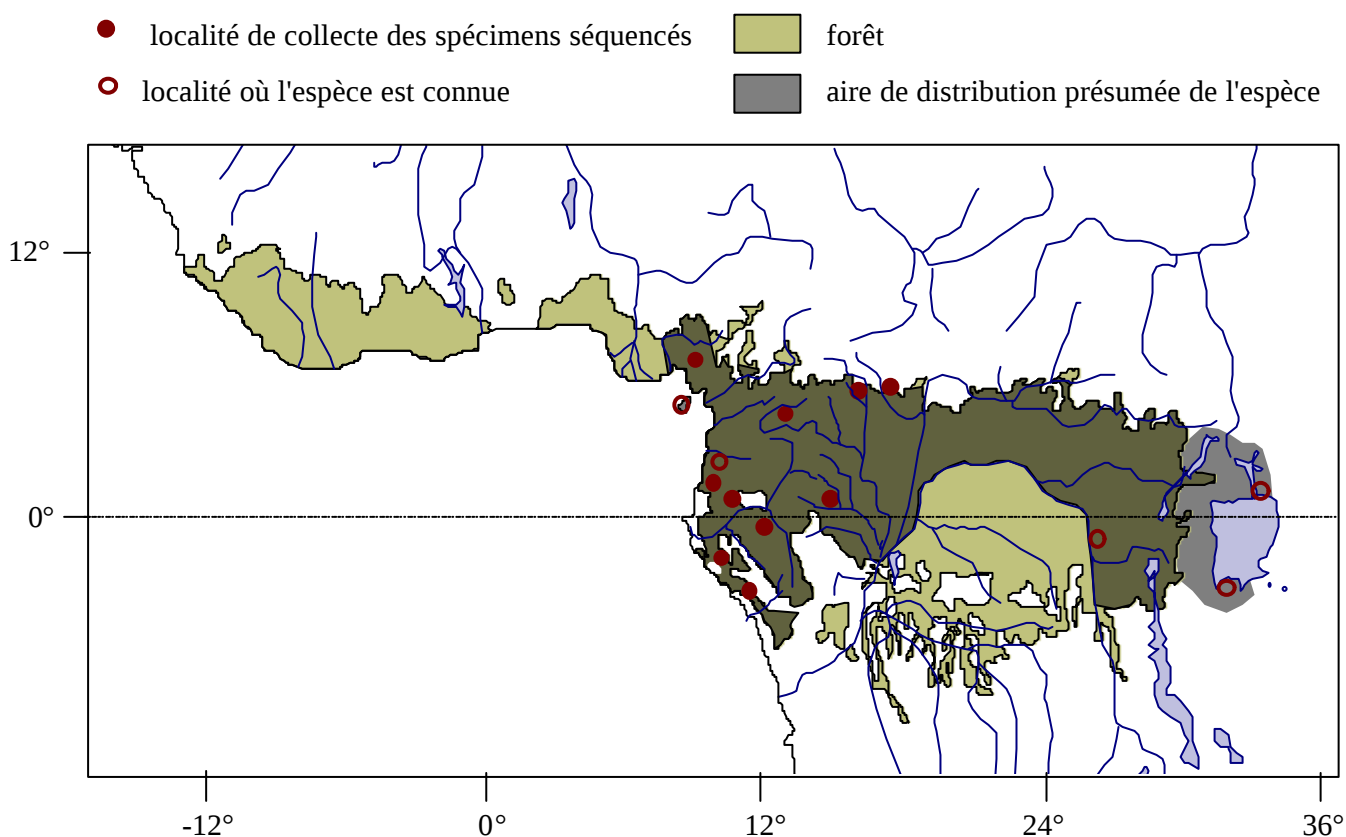
Chacune des espèces ou super-espèces retenues pour les analyses phylogéographiques est présentée dans les pages qui suivent (**doc. 2.1 à 2.5**).

Doc. 2.1

Sylvisorex johnstoni - (Dobson, 1888)

- **Description** : Dobson (1888), *Proc. Zool. Soc. Lond.*, 1887: 577.
- **Localité type** : Rio del Rey, Cameroun
- **Synonymie** : *S. dieterleni*
- **Distribution** : au Nord du fleuve Congo, du Sud-Ouest du Cameroun à l'Ouganda et la Tanzanie, ainsi que sur l'île de Bioko
- **Ecologie** : Espèce typiquement forestière vivant dans la litière.
Présente aussi bien en forêt dense que près de l'écotone, mais rarement en milieu dégradé.
Régime alimentaire ?
- **Biologie** : Très petite espèce (poids : 2-3 g).
Environ deux jeunes par portée. Longévité ?
- **Informations complémentaires** :
Révision par Hutterer (1986b)
Caryotype : $2n = 30$, $NF = 38$, $Nfa = 36$ - spécimen du Cameroun (Schlitter *et al.*, 1999)

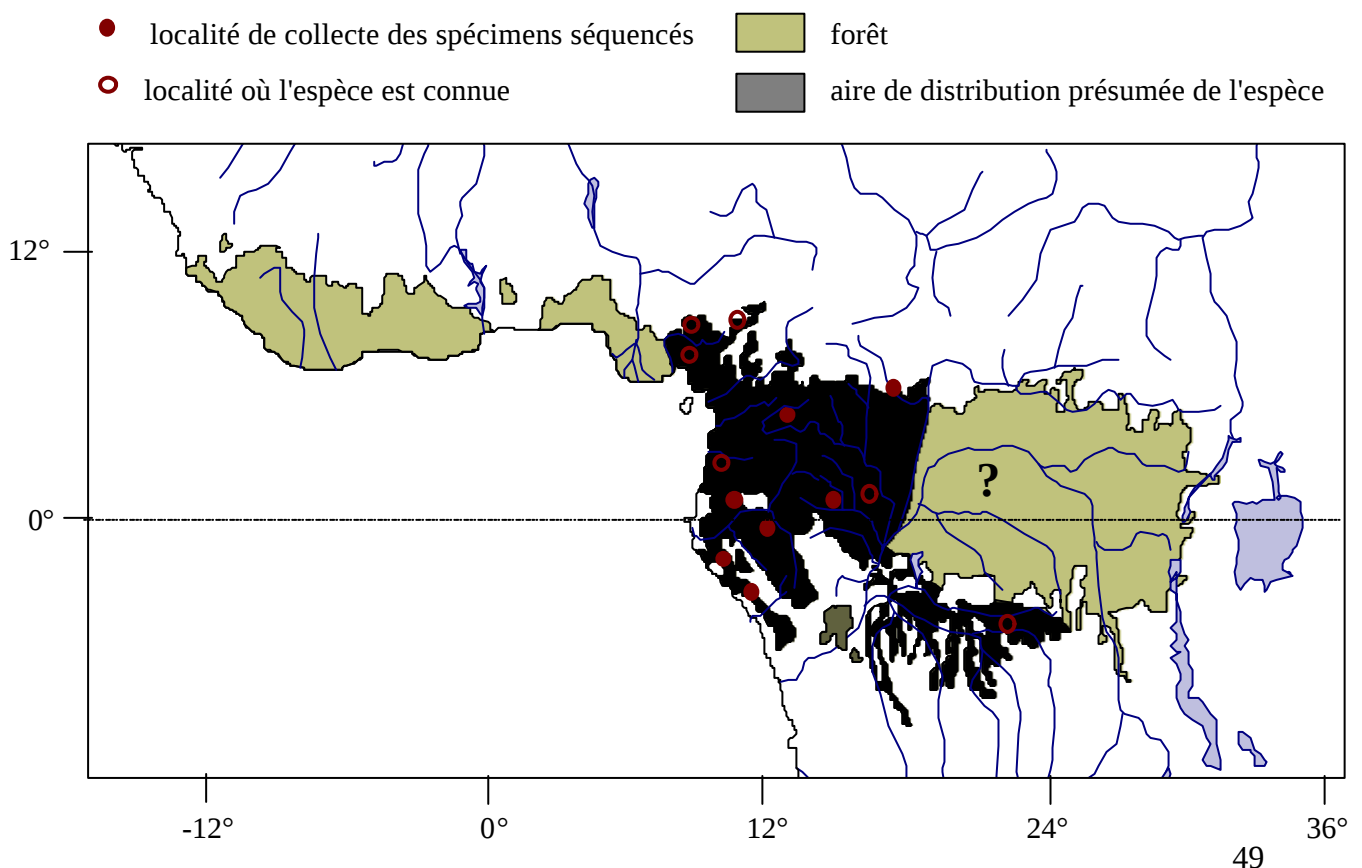
Distribution présumée de l'espèce et provenance des échantillons séquencés



Sylvisorex ollula - Thomas, 1913

- **Description** : Thomas (1913), *Ann. Mag. Nat. Hist.*, ser. 8, 11:321.
- **Localité type** : "Bitye, Ja River, Cameroons"
- **Distribution** : Sud Cameroun, Centrafrique, Gabon, Guinée Equatoriale, Sud-Ouest Nigeria, Sud R.D.C.
- **Ecologie** : Espèce typiquement forestière vivant dans la litière.
Présente aussi bien en forêt dense que près de l'écotone, mais rarement en milieu dégradé.
Régime alimentaire ?
- **Biologie** : Plus grande espèce du genre (poids : 15-20 g).
2 à 3 jeunes par portée. Longévité ?
- **Informations complémentaires** :
cf. Dieterlen et Heim de Balsac (1979)
Caryotype : $2n = 38$, $NF = 64$, $Nfa = 62$ - spécimen du Cameroun (Schlitter *et al.*, 1999)

Distribution présumée de l'espèce et provenance des échantillons séquencés



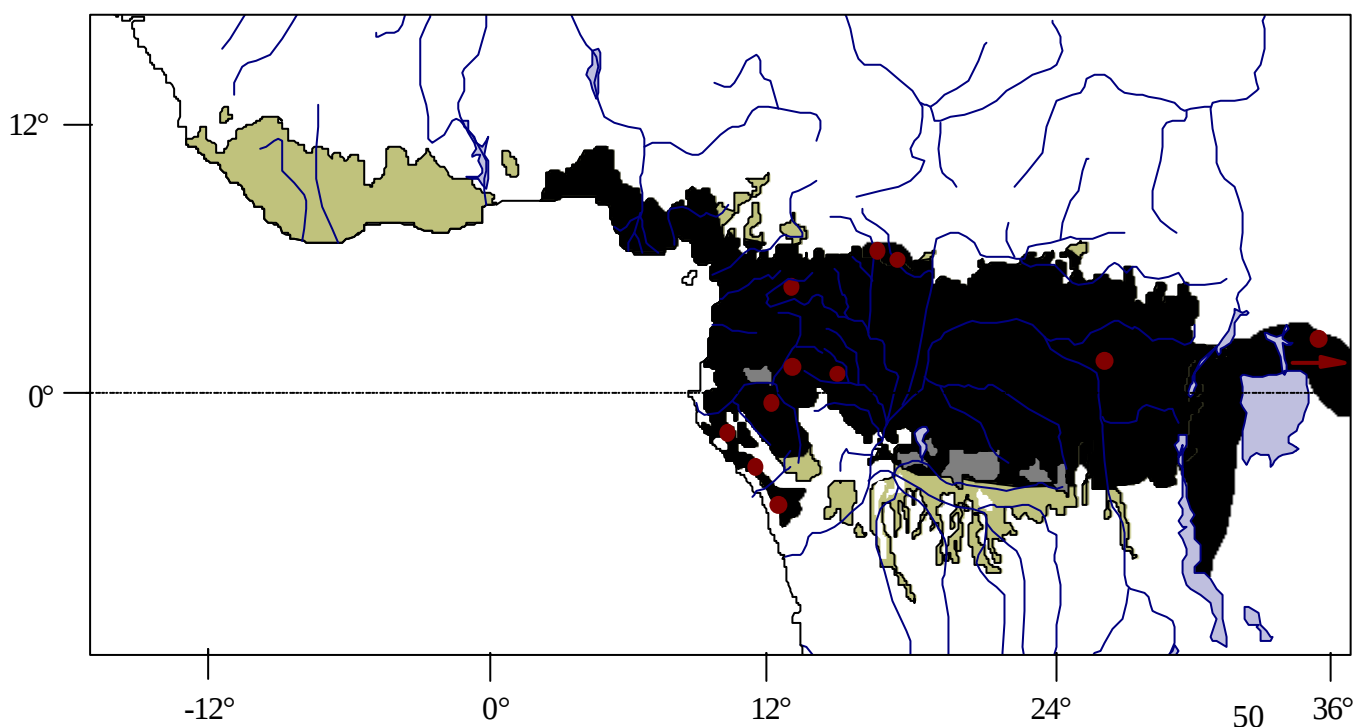
Doc. 2.3

Hylomyscus stella - (Thomas, 1911)

- **Description** : Thomas (1911), *Ann. Mag. Nat. Hist.*, ser. 8, 7:590.
- **Localité type** : forêt d'Ituri, NE R.D.C.
- **Synonymie** : confusion possible avec *alleni* et *simus* (Rosevear, 1969)
- **Distribution** : du Sud du Nigeria à l'Ouest du Kenya et de la Tanzanie, du Sud de la Centrafrique et du Soudan au Nord de l'Angola
- **Ecologie** : Espèce forestière, présente dans les zones boisées humides. (Rosevear, 1969 ; Kingdon, 1971)
Espèce nocturne, à régime omnivore à tendance végétarienne (fruits, graines et insectes).
- **Biologie** : Espèce de petite taille pour la famille (poids : 15-24 g, moyenne : 18 g).
1 à 5 jeunes par portée. Longévité sans doute > 1an. (Rosevear, 1969 ; Kingdon, 1971)
- **Informations complémentaires** :
Caryotype : $2n = 46$, $NF = 70$ - spécimens du Sud Cameroun et NE Gabon (Robbins *et al.*, 1980 ; Iskandar *et al.*, 1988)
 $2n = 46$, $NF = 72$ - spécimen de R.D.C. (Matthey, 1963)
 $2n = 46$, $NF = 74$ - spécimen de Centrafrique (Viegas-Pequignot *et al.*, 1983)
 $2n = 48$, $NF = 86$ - spécimen du Burundi (Maddalena *et al.*, 1989)

Distribution de l'espèce et provenance des échantillons séquencés

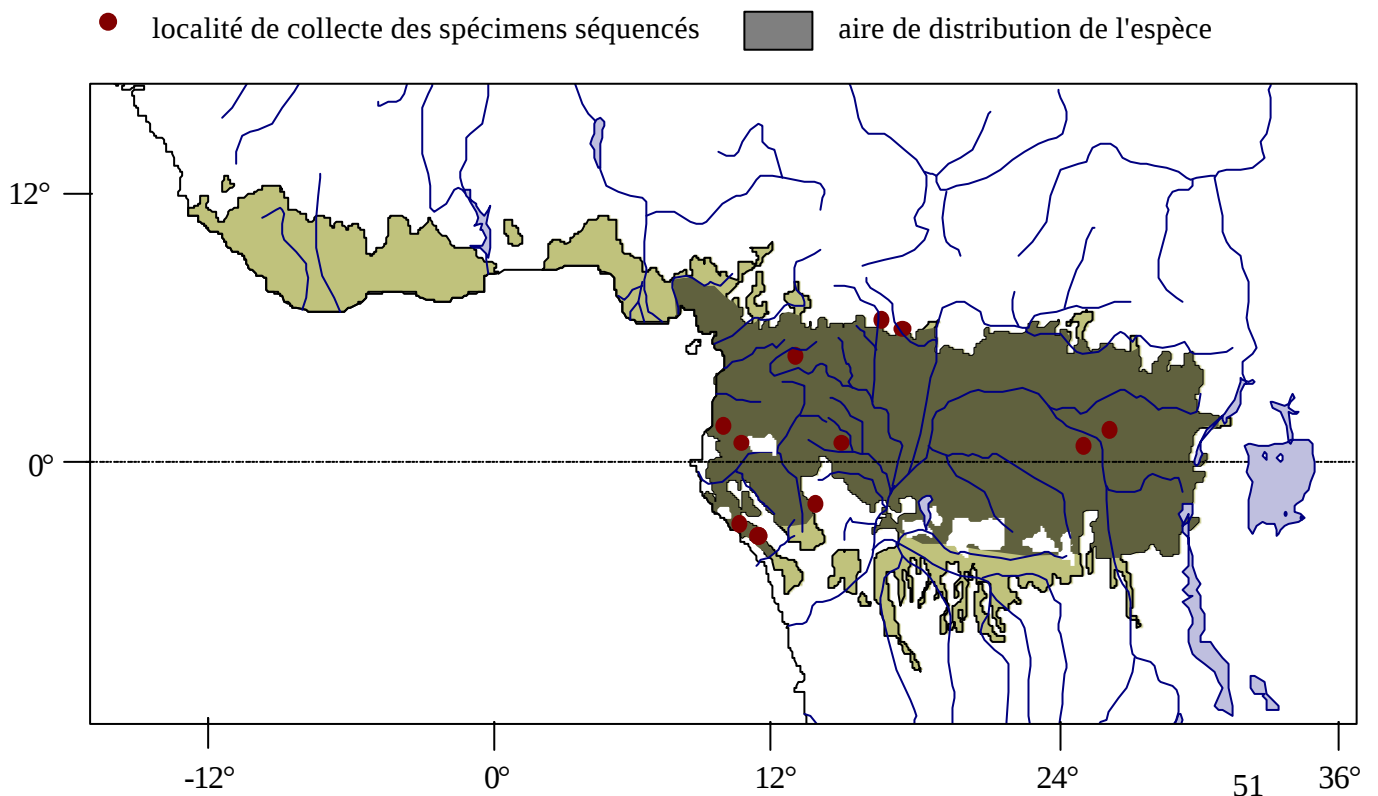
● localité de collecte des spécimens séquencés ■ aire de distribution de l'espèce



Stochomys longicaudatus - (Tullberg, 1893)

- **Description** : Tullberg (1893), *Nova Acta Reg. Soc. Sci. Upsala*, ser. 3, 16:36
- **Localité type** : Cameroun!
- **Synonymie** : *hypoleucus*, *ituricus*, *sebastianus*. Espèce successivement classée parmi les genres *Aethomys*, *Dasymys*, *Epimys*, *Mus*, *Rattus*, et *Stochomys* (Rosevear, 1969). Deux sous-espèces: *S. l. longicaudatus* et *S. l. ituricus* (Van der Straeten, 1984).
- **Distribution** : du Sud-Est du Nigeria à l'Ouest de l'Ouganda.
- **Ecologie** : Espèce inféodée aux forêts de plaine, présente aux abords des rivières et zones marécageuses. Nocturne. Régime alimentaire essentiellement végétarien (fruits et feuilles), mais consomme aussi quelques insectes. (Rosevear, 1969 ; Kingdon, 1971)
- **Biologie** : Espèce assez grande pour la famille (poid : 60-104 g, moyenne: 80 g). 2 à 3 jeunes par portée. Longévité ? (Rosevear, 1969 ; Kingdon, 1971)

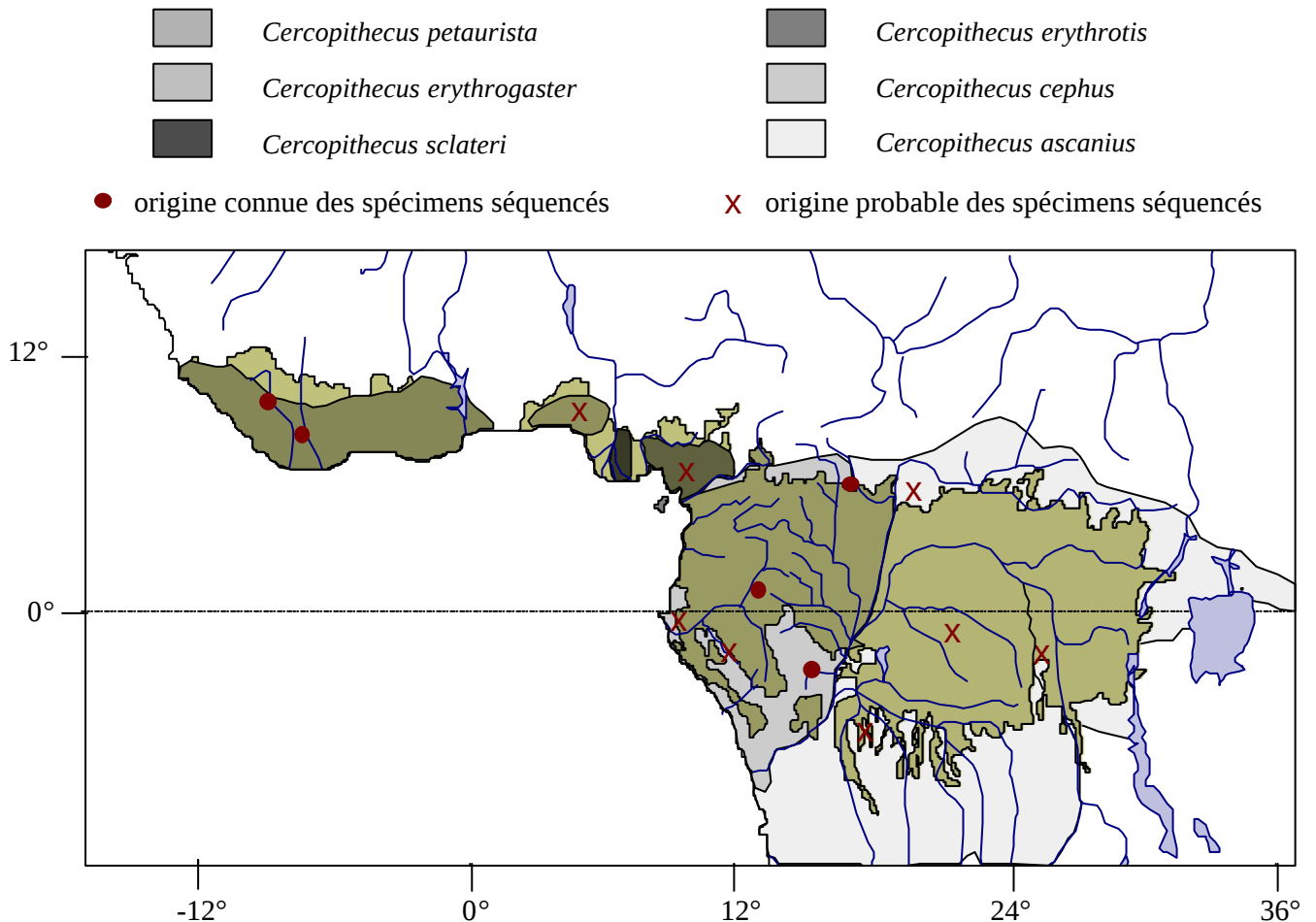
Distribution de l'espèce et provenance des échantillons séquencés



super-espèce *Cercopithecus cephus*
Linnaeus, 1758

- **Description** : Linnaeus (1758), *Syst. Nat.*, 10th ed., 1:27.
- **Localité type** : Afrique!
- **Espèces concernées** : *ascanius*, *cephus*, *erythrogaster*, *erythrotis* et *petaurista* (selon Napier, 1981), ainsi que *sclateri* (parfois considéré comme une sous-espèce de *erythrotis*)
- **Distribution** : Ensemble des forêts d'Afrique centrale et de l'Ouest, et zones de mosaïque forêt-savane qui les bordent, depuis la Guinée jusqu'au Kenya.
- **Ecologie** : Singes forestiers occupant préférentiellement les strates intermédiaires et basses des forêts secondaires au sous-bois dense. Régime alimentaire essentiellement végétarien incluant jusqu'à 80% de fruits, complété par des Invertébrés. Vivent en troupes d'une dizaine ou une vingtaine d'individus, comprenant un seul mâle adulte reproducteur. Ce sont essentiellement les mâles qui migrent (Gautier-Hion et al., 1999).
- **Biologie** : Espèces assez petite pour le genre (gamme de poids : 1,8 à 4 kg pour les femelles et 2,5 à 6 kg pour les mâles). 1 jeune par portée. Longévité : 15-25 ans en captivité (Gautier-Hion et al., 1999).

Distribution de la super espèce *C. cephus* et provenance des échantillons



Doc. 2.5 (suite)

• Informations complémentaires

Caryotype :

$2n = 66$, $NF = 110$, pour toutes les espèces du groupe (Dutrillaux *et al.*, 1980)

Pas de grandes différences anatomiques entre espèces et sous-espèces (Verheyen, 1962 ; Gebo et Sargis, 1994).

Grande variabilité des colorations de la face et du pelage (Kingdon, 1980, 1988).

Etude électrophorétique de cinq protéines (Walker *et al.*, 1988).

Analyse des cris (Gautier, 1988).

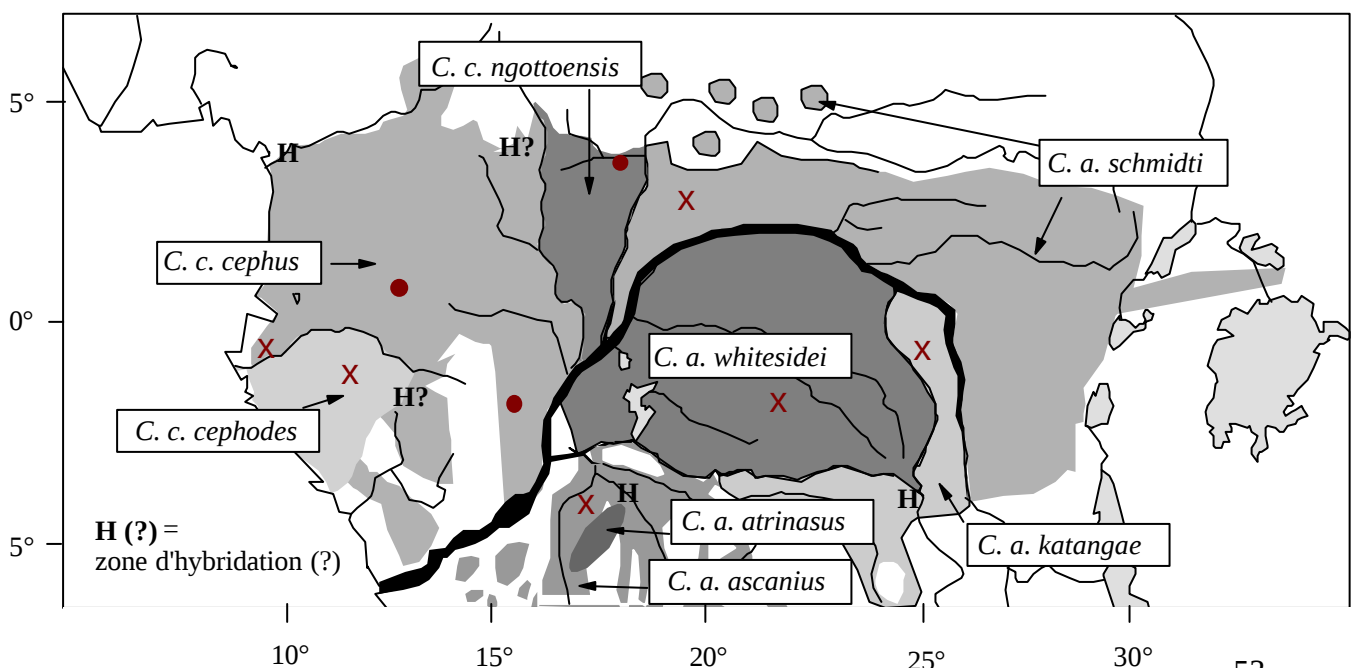
• Les différentes sous-espèces :

- *C. cephus* Linnaeus, 1758
 - C. c. cephus* Linnaeus, 1758 (synonymes *buccalis*, *inobservatus* et *pulcher*)
 - C. c. cephodes* Pocock, 1907 (synonyme *gabonensis*)
 - C. c. ngottoensis* Colyn, 1999
- *C. ascanius* (Audebert, 1799)
 - C. a. ascanius* (Audebert, 1799) (synonymes *melanogerys*, *histrio* et *picturatus*)
 - C. a. schmidt* Matschie, 1892 (synonymes *kaimosae*, *mpangae*, *sassae*, *enkamer*, *montanus*, *ituriensis*, *rutschuricus* et *orientalis*)
 - C. a. whitesidei* Thomas, 1909 (synonymes *omissus*, *cirrhorrhinus*, *kassaicus* et *pelorhinus*)
 - C. a. katangae* Lönneberg, 1919
 - C. a. atrinasus* Machado, 1965
- *C. erythrogaster* Gray, 1866
 - C. e. erythrogaster* Gray, 1866
 - C. e. pokocki* Grubb *et al.*, 1999
- *C. erythrotis* Waterhouse, 1838
 - C. e. erythrotis* Waterhouse, 1838
 - C. e. camerounensis* Hayman, 1940
- *C. petaurista* (Schreber, 1774) (synonymes *albinasus*, *büttikoferi* *pygrius* et *fantiensis*)
 - C. p. petaurista* (Schreber, 1774)
 - C. p. buettikoferi* Jentink, 1886
- *C. sclateri* Pocock, 1904

Distribution des sous-espèces de *C. cephus* et *C. ascanius*

● origine connue des spécimens séquencés

✗ origine probable des spécimens séquencés



2) Matériel biologique utilisé

Nous avons tenté d'analyser 458 spécimens, mais nous n'avons pu obtenir de résultats que pour 407 d'entre eux. Sur ces 407 spécimens, 275 ont été utilisés pour les analyses phylogénétiques (soit 165 Soricidae, 83 Muridae et 27 Cercopithecidae), mais seuls 139 sont présentés dans ce travail. Les spécimens qui n'y figurent pas sont tous des Soricidae. Certains ont servi à vérifier la monophylie des genres et des espèces, mais n'ont pas été présentés pour des raisons pratiques (cf. chapitre 3.I), d'autres ont été analysés dans la perspective de clarifier la taxinomie et les relations inter-spécifiques au sein du genre *Crociodura*, mais n'ont pu être exploités en raison de nombreuses difficultés taxinomiques encore irrésolues. Des travaux de révision sur base morphologique et craniologique sont en cours (R. Hutterer et P. Barrière, comm. pers.). Trente-six spécimens du genre *Lophuromys* (Muridae), analysés dans la perspective d'une étude phylogéographique, n'ont pas été retenus faute d'un échantillonnage suffisant. Parmi les 157 spécimens utilisés pour les analyses phylogéographiques, on dénombre 42 *Sylvisorex johnstoni* (dont un spécimen utilisé pour la phylogénie des musaraignes africaines), 40 *S. ollula* (idem), 39 *Hylomyscus stella* (tous utilisés pour la phylogénie du genre *Hylomyscus*), 18 *Stochomys longicaudatus* et 18 spécimens de la super-espèce *Cercopithecus cephus* (tous inclus dans l'analyse phylogénétique des Cercopithecini).

Après révision des musaraignes africaines, des espèces bien représentées dans la région d'étude, clairement identifiables et présentant une variabilité moléculaire relativement élevée ont été choisies pour une étude plus fine : *Sylvisorex ollula* et *Sylvisorex johnstoni*.

Parmi les Muridae, le genre *Hylomyscus*, typiquement forestier, semblait prometteur mais présentait un problème de détermination des espèces. Il a donc fait l'objet d'une révision phylogénétique, et l'espèce la plus commune, *Hylomyscus stella*, a été retenue pour les analyses intra-spécifiques. Un second modèle, appartenant à un genre monotypique, a été choisi sans révision préalable : l'espèce *Stochomys longicaudatus*.

Pour les Primates, une phylogénie moléculaire de la tribu des Cercopithecini a été réalisée, et un groupe d'espèces aux distributions allopatriques, la super-espèce *C. cephus*, a été étudié plus en détail.

III - ACQUISITION DES DONNEES MOLECULAIRES

Afin de déterminer les relations phylogénétiques inter- et intra-spécifiques, nous avons choisi l'approche moléculaire qui présente l'avantage de générer un grand nombre de données considérées comme indépendantes et non soumises à la subjectivité de l'observateur. Elle n'est cependant pas dénuée d'inconvénients (Patterson *et al.*, 1993) et ne dispense pas d'une approche morphologique. La biologie moléculaire recouvre un large éventail de techniques que nous allons présenter brièvement, de façon à justifier notre choix.

1) Choix de la technique moléculaire

a) Les différentes techniques moléculaires

Il existe un grand nombre de méthodes moléculaires présentant divers avantages et inconvénients. Certaines méthodes, notamment celles qui sont basées sur les propriétés physico-chimiques des protéines (mobilité électrophorétique des protéines ou allozymes, techniques immunologiques), se prêtent mal à la reconstruction de phylogénies. En revanche, il existe de nombreuses méthodes portant sur le support de l'information génétique (l'ADN) qui permettent d'estimer les relations phylétiques entre taxons plus ou moins proches. Le choix de la technique dépend essentiellement du niveau d'analyse désiré.

Plusieurs techniques reposent sur le principe d'hybridation de fragments d'ADN simple brin. Elles peuvent être appliquées soit à l'ADN simple copie soit à l'ADN génomique hautement répété. Elles reposent sur la comparaison de courbes d'élution d'ADN intra- et inter-spécifique. La courbe d'élution est obtenue en mesurant la quantité d'ADN apparié qui se dissocie en fonction de la température. Les méthodes d'hybridation permettent de distinguer entre elles des espèces plus ou moins éloignées (Wilson, 1995) et de construire des arbres phénétiques.

D'autres méthodes consistent à analyser le polymorphisme de taille de fragments d'ADN générés par amplification ou par découpage de l'ADN par des enzymes de restriction. Ces méthodes ne génèrent qu'un petit nombre de données et ne permettent pas forcément de reconstruire une phylogénie. Les plus employées sont :

- Analyse du polymorphisme des fragments de restriction (= "Restriction Fragment Length Polymorphism" ou RFLP) : le principe de cette analyse est d'amplifier une portion d'ADN, de la découper par des enzymes de restriction, puis de faire migrer les fragments sur un gel, de façon à obtenir

des informations sur la nature et l'ampleur des différences entre fragments d'ADN. Par cette technique, il est possible de différencier des lignées (Wilson, 1995).

- Finger-printing : cette méthode consiste à comparer les bandes résultant de l'amplification de régions comprises entre deux amorces universelles (par exemple situées au niveau d'ARNt agencés en tandem). Elle peut permettre de différencier les espèces et les genres chez les bactéries (Welsh et McClelland, 1991), mais n'a jusqu'à présent pas donné de résultat concluant chez les Vertébrés.

- Taxon-printing : cette méthode repose sur l'analyse de courts fragments de restriction d'ADN génomique hautement répété (20 à 300 pb) visualisés après marquage radioactif. Elle permet par exemple de distinguer les espèces les unes des autres chez les Lézards et les Primates (Fedorov *et al.*, 1999), mais n'est pas utilisable à l'échelle populationnelle.

- Marqueurs microsatellites : ces marqueurs présentent un taux de mutation très élevé qui les rend propre à l'étude de populations géographiquement proches (Jarne et Lagoda, 1996). Ils ne peuvent être utilisés qu'après réalisation d'amorces spécifiques et ne sont pas particulièrement adaptés à la reconstruction de phylogénies. Seuls des arbres phénétiques peuvent être construits, à condition de disposer d'un nombre de loci suffisant.

Une autre approche consiste à lire la séquence d'une portion d'ADN, autrement dit l'ordre d'agencement des quatre nucléotides qui constituent l'ADN dans une région ciblée du génome. On génère ainsi un grand nombre de caractères permettant d'inférer des relations phylogénétiques à des niveaux taxinomiques très variés, selon la séquence ciblée.

Nous avons choisi la technique qui génère le plus grand nombre de caractères et permet de construire des phylogénies à des niveaux taxinomiques variés : celle du séquençage de l'ADN. Il a ensuite fallu déterminer quelles seraient les portions d'ADN les plus pertinentes pour travailler aux niveaux taxinomiques choisis (comparaisons intra- et inter-spécifiques).

b) ADN mitochondrial ou ADN nucléaire ?

Qualités de l'ADN mitochondrial par rapport à l'ADN nucléaire

L'ADN mitochondrial évolue globalement plus vite que l'ADN nucléaire (5 à 10 fois chez les Mammifères, d'après Brown *et al.*, 1979, 1982). Il est haploïde et transmis uniquement par la mère ; il ne pose donc pas de problème de recombinaison ni d'hétérozygotie. Pour se placer dans des conditions similaires avec l'ADN nucléaire, il faut utiliser des gènes portés uniquement par le chromosome Y. Le

matériel biologique exploitable est alors réduit de moitié. De plus, chaque cellule renferme plusieurs molécules d'ADN mitochondrial, ce qui facilite son amplification.

Défauts de l'ADN mitochondrial

Il peut y avoir hétéroplasmie de séquence et de longueur, mais dans la pratique, le polymorphisme individuel est à peu près nul (Stoneking et Soodyall, 1996). En revanche, il existe des copies nucléaires de fragments d'ADN mitochondrial qui peuvent nuire à la lecture des séquences ciblées, voire être amplifiées à sa place et donner lieu à la comparaison de gènes non homologues (Arctander, 1995).

L'ADN mitochondrial est déconseillé dans le cas de dispersion totalement asymétrique des individus mâles et femelles, car il ne reflète que la distribution des individus femelles. Pour les espèces où ce sont les mâles qui se dispersent, comme chez les macaques, la transmission de l'ADN mitochondrial se fait essentiellement au sein d'une même population et les échanges entre populations sont limités, alors que l'ADN nucléaire est dispersé de façon assez homogène (Melnick et Hoelzer, 1993).

Nécessité d'analyser plusieurs séquences

L'utilisation de plusieurs gènes indépendants, c'est-à-dire non groupés sur une même molécule d'ADN (ce qui n'est pas le cas pour les gènes mitochondriaux), permet de diminuer les risques de présenter un arbre erroné. En effet, un arbre obtenu sur base d'une séquence moléculaire n'est pas l'équivalent de l'arbre des espèces. Non seulement le nombre de caractères peut être insuffisant pour obtenir une bonne résolution, mais en plus, certains gènes peuvent avoir une histoire particulière, différente de celle du taxon lui-même (Nei, 1987 ; Hey, 1994 ; Moore, 1995; Maddison, 1997). L'arbre des espèces dépend de l'histoire des populations ancestrales (isolement historique et isolement reproductif entre populations) tandis que l'arbre des gènes dépend des mutations qui se produisent au sein des populations.

Un arbre des gènes diffère de celui des espèces notamment lorsqu'il y a hybridation introgressive ou transfert horizontal de gènes via un vecteur. Ces phénomènes peuvent être décelés par l'analyse conjointe de portion d'ADN mitochondrial et d'ADN nucléaire, ou d'ADN et de caractères morphologiques (Moore, 1995).

Les arbres diffèrent aussi lorsque le polymorphisme ancestral et le phénomène de sélection des lignées sont importants. Ces phénomènes sont favorisés par un temps court entre deux événements de spéciation, une population initiale de grande taille et un processus de coalescence lent (Maddison, 1997). Le temps de coalescence est indépendant du taux de mutations pour les gènes sélectivement

neutres, mais il est 4 fois plus grand pour l'ADN mitochondrial qui est haploïde que pour les autosomes qui sont diploïdes (Moore, 1995).

L'ADN mitochondrial est généralement préféré à l'ADN nucléaire parce qu'il est haploïde et évolue assez rapidement. Il est employé dans 70 % des études phylogéographiques (Avisé, 2000). C'est également le choix que nous avons fait.

Sachant qu'il est important de séquencer plusieurs gènes indépendants (mitochondriaux et nucléaires) pour obtenir une phylogénie des organismes et non une phylogénie du gène, nous avons également tenté d'amplifier de l'ADN nucléaire, pour les Primates.

c) Choix des séquences mitochondriales

Organisation et variabilité du génome mitochondrial

Le génome mitochondrial est constitué par une molécule d'ADN circulaire double brin et haploïde (Bibb, 1981). Chez pratiquement tous les animaux, cette molécule comprend deux séquences codant pour des ARN-r et 22 pour des ARN-t, cinq gènes codant pour des protéines connues (cytochrome *b*, sous-unités I à III de la cytochrome oxydase, 6^e sous-unité de l'ATPase) et huit pour des protéines non identifiées (ARF1 à ARF8), ainsi qu'une partie non codante contenant l'origine de réplication du brin H et portant le nom de D-loop (**fig. 2.7**).

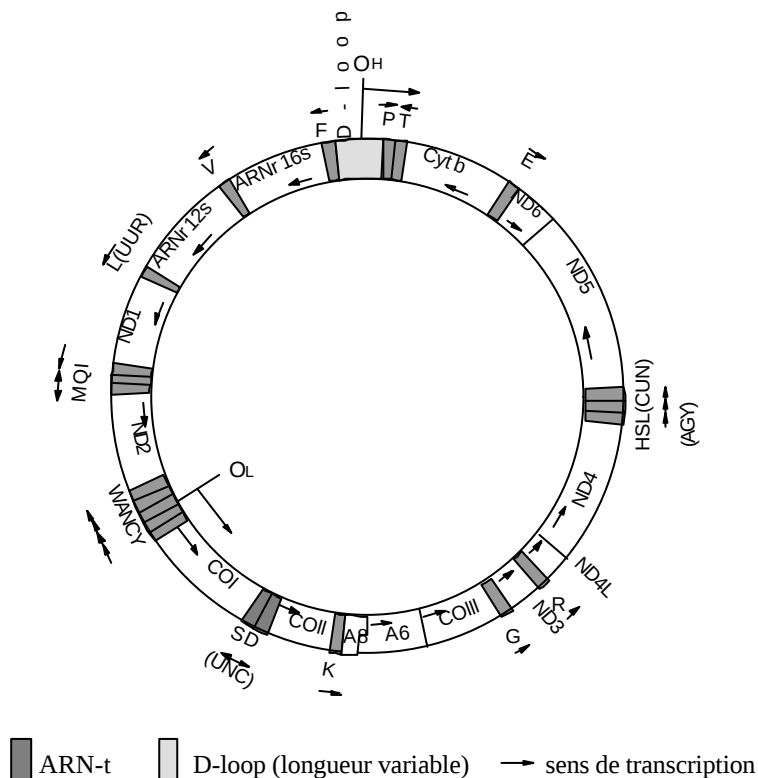


Figure 2.7. Organisation de la molécule d'ADN mitochondrial chez les Vertébrés (Simon *et al.*, 1994).

Le point fondamental lors du choix de la séquence à utiliser est de sélectionner une séquence évoluant à une vitesse adaptée au niveau taxinomique étudié. Le nombre de substitutions doit être suffisant pour résoudre la phylogénie ; mais le nombre de substitutions multiples doit être limité. En effet, lorsque plusieurs changements ont lieu successivement à une même position, il se produit un phénomène de saturation qui brouille l'information phylogénétique. Il est également important que le nombre d'insertions et de délétions (aussi appelés 'indels') soit faible, de façon à pouvoir aligner les séquences sans ambiguïté. D'après Hillis et Dixon (1991), il convient de choisir des séquences pour lesquelles le taux de substitution est inférieur à 30% entre les taxons étudiés.

Nous avons réalisé une comparaison de la variabilité des différentes parties de l'ADN mitochondrial entre le rat (*Rattus norvegicus*) et la souris (*Mus musculus*). Cette comparaison révèle des pourcentages de substitution croissants en fonction des contraintes structurales qui portent sur chaque classe de séquences :

ARN-r 12s < protéines connues < ARN-r 16s < protéines non identifiées < D-loop

8,4% < 16,7% (+/- 0,9) < 18,9% < 21,2% (+/- 2,4) < 26,8%

Pour les ARN-t, les pourcentages de substitution sont compris entre 2,7 et 15,6% (en moyenne 10,5% +/-3,6).

Avantages et inconvénients des différents types de séquences mitochondriales

Les séquences codant pour une **protéine** sont à privilégier pour des études portant sur des taxons proches ou assez proches (cf. Simon *et al.*, 1994). Du fait de l'organisation en codons de la séquence et de la rareté des insertions et délétions, l'alignement est généralement aisé. La séquence peut d'ailleurs être analysée par codon ou traduite en acides aminés. Les trois positions subissent des contraintes plus ou moins marquées et n'évoluent pas à la même vitesse. La deuxième position est la plus contrainte et la troisième est la plus libre de varier. La troisième position est donc informative pour une radiation récente, mais doit être éliminée pour une radiation ancienne en raison de la saturation des substitutions.

La séquence codant pour le Cytochrome *b* est la plus utilisée et la mieux connue (Irwin *et al.*, 1991), mais les sous-unités I et II de la Cytochrome Oxydase sont également employées.

Les séquences codant pour des **ARN-r** sont utiles pour des comparaisons entre taxons éloignés, et sont très fréquemment utilisées en phylogénétique (cf. Hillis et Dixon, 1991). L'ARNr 16S est globalement plus variable que l'ARN 12S et peut être employé pour des comparaisons entre genres,

voire entre espèces (cf. synthèses de Ferris *et al.*, 1983 ; Mindell et Honeycut, 1990 ; et exemples dans Horovitz et Meyer, 1995 ; Juste *et al.*, 1999 ; Clough et Summers, 2000).

Le taux de substitutions varie d'une région à l'autre, notamment en fonction de la structure secondaire de l'ARN-r qui présente des régions appariées, appelées tiges et des régions non appariées, appelées boucles (cf. structure secondaire de la seconde moitié de l'ARNr 16s chez l'homme ; **fig. 2.8**). Le choix d'une zone à cheval sur une région stable et une région variable permet d'évaluer simultanément différents niveaux taxinomiques. L'alignement peut être problématique, mais la connaissance de la structure secondaire aide à le retrouver. Certains auteurs proposent de donner aux positions impliquées dans la formation de tiges un poids de moitié inférieur à celui des boucles, à cause des substitutions compensatoires (Wheeler et Honeycutt, 1988), mais d'autres estiment que ce sont les positions appariées qui donnent le meilleur signal phylogénétique (Smith, 1989 ; Hillis et Dixon, 1991).

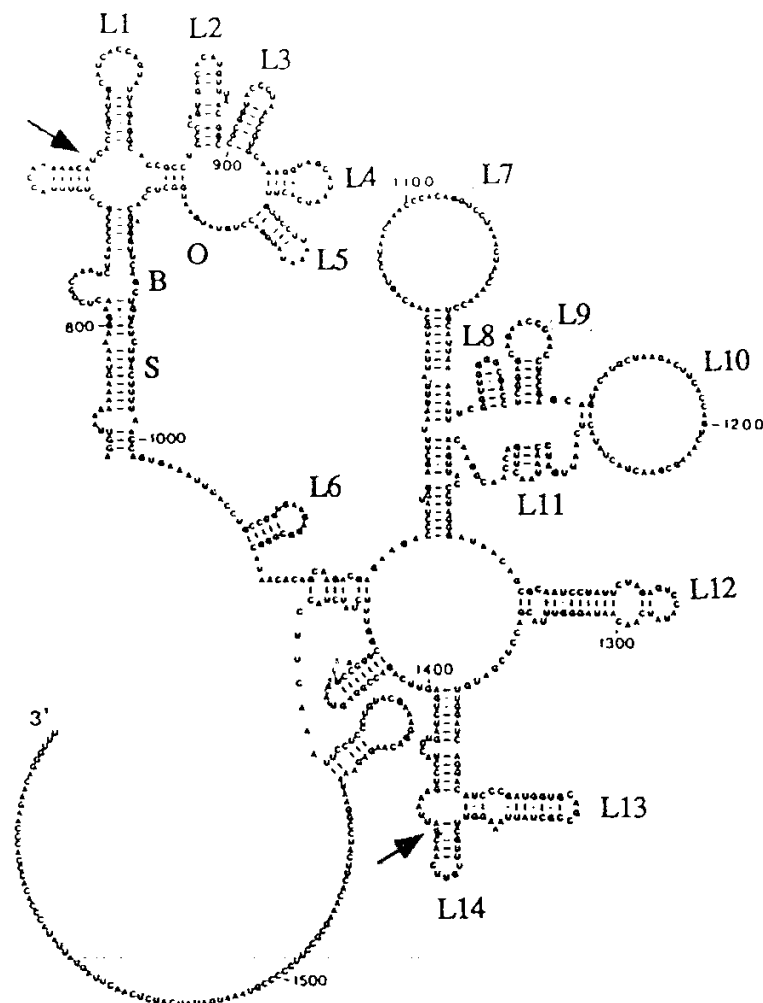


Figure 2.8. Structure secondaire de la seconde moitié de l'ARN-r 16s chez l'Homme, d'après Gutell et Fox (1988).

L = boucles ; les flèches délimitent la partie séquencée.

Les séquences codant pour des **ARN-t** sont peu utilisées car elles sont courtes et dispersées.

La région non codante ou **D-loop** contient des parties extrêmement variables, ce qui permet une étude assez fine au niveau spécifique ou populationnel. En contrepartie, l'alignement peut être délicat (cf. Sbisa *et al.*, 1997).

Nous avons utilisé principalement une partie de la séquence codant pour l'ARN 16s, ainsi qu'une fraction du gène du Cytochrome *b*. Nous avons également tenté de travailler avec la D-loop pour les analyses populationnelles.

Nous avons vérifié que les phylogénies obtenues avec l'ARN 16s et le cytochrome *b* coïncidaient pour certains taxons et avons ensuite privilégié l'ARN 16s qui a déjà été employé pour les rongeurs (e.g. Frye et Hedges, 1995) et les primates (e.g. Horovitz et Meyer, 1995).

2) Protocoles employés

Tous les protocoles détaillés sont fournis en annexe (Annexe 3).

a) Extraction

La plupart des extractions ont été réalisées selon la méthode Chelex (Walsh *et al.*, 1991), qui est la méthode la plus rapide et la moins coûteuse. En cas d'insuccès, une autre méthode a été appliquée. La méthode CTAB (Doyle et Sicks, 1987) offre un taux de réussite assez élevé et permet une bonne conservation de l'ADN extrait. Pour les tissus conservés dans l'azote liquide ou particulièrement récalcitrants, nous avons procédé à une dissolution par la Protéinase K, puis à une extraction au Phénol-Chloroforme (Sambrook *et al.*, 1989). Cette dernière méthode demande un plus grand investissement financier et temporel. Pour certains échantillons, plusieurs méthodes d'extraction ont dû être essayées. A partir de la seconde tentative infructueuse, nous avons conclu à un échec, sauf pour quelques échantillons particulièrement désirés.

Pour quelques espèces rares pour lesquelles nous ne disposons d'aucune biopsie, des extractions à partir de squelette et de fragments de peaux prélevés dans des collections muséologiques ont été tentées. De l'ADN de bonne qualité a ainsi été extrait à partir d'un squelette de *Scutisorex somereni* au Service de Systématique Moléculaire du Museum National d'Histoire Naturelle de Paris, en suivant le protocole d'Hassanin et collaborateurs (1998b). Par contre, les tentatives d'extraction à partir de peau

menées pour quelques spécimens de Primates sont demeurées infructueuses. Il est important de noter que l'infrastructure du laboratoire de Taxinomie et Systématique Biochimique de l'Institut Royal des Sciences Naturelles de Belgique, où presque toutes les analyses moléculaires ont été réalisées, n'a pas été conçue pour l'exploitation d'ADN en quantité limitée.

b) Amplification

Principe de la PCR ("Polymerase Chain Reaction")

La PCR permet d'obtenir rapidement un très grand nombre de copies de la portion d'ADN ciblée (jusqu'à un millier de paires de bases environ). Le fragment d'ADN est copié de multiples fois au cours de cycles successifs qui comprennent trois phases :

- 1) séparation des brins de l'ADN source par dénaturation thermique (94°C) ;
- 2) hybridation avec des amorces spécifiques flanquant la région ciblée (la température d'hybridation varie selon les amorces utilisées et doit être proche de la température dite "de fusion") ;
- 3) synthèse d'ADN au moyen d'une polymérase active à haute température (72°C dans le cas de la Taq-polymerase).

A chaque cycle, le nombre de copies est multiplié par deux.

Protocole de PCR

Les PCR ont été réalisées dans les conditions standards suggérées par le fournisseur de la Polymerase, en utilisant 1 à 10 µl de produit d'extraction pour un volume final de 25µl. Les cycles ont été programmés comme suit :

(60 sec à 94°C, 60 sec à ...°C et 90 sec à 72°C), 30 fois.

Les températures utilisées pour la phase d'hybridation sont précisées pour chaque amorce en annexe (Annexe 4). Après chaque PCR, le succès de l'amplification a été vérifié par visualisation de l'ADN aux UV après migration sur gel d'agarose 1,4% pendant 30 min. En cas de succès, une seconde amplification à partir du même extrait a été faite afin d'obtenir une plus grande quantité d'ADN. Dans le cas contraire, les conditions expérimentales ont été révisées (diminution de la température d'hybridation, touch-down PCR, test de différentes polymérases, modification de la salinité...).

Amorces utilisées et qualité des séquences

Pour l'ARNr 16s, le couple d'amorces 16Sar et 16SHm (Annexe 4) a donné un très bon produit d'amplification pour l'ensemble des taxons. De ce fait, cette région de l'ADN a été séquencée en

premier lieu, pour tous les échantillons analysés. Lorsque, malgré plusieurs tentatives, cette séquence n'a pas pu être amplifiée, l'échantillon a été mis de côté (11% des cas, n = 458).

En revanche, alors que de nombreux couples d'amorces ont été utilisés pour amplifier le cytochrome *b* et que diverses conditions ont été testées pour l'amplification, le taux d'échec est demeuré très élevé, tant au niveau de l'amplification que de la qualité des séquences (de 52% d'échec pour n = 99 à 98% d'échec pour n = 44, selon les amorces).

L'amplification de la D-loop a également été tentée en raison de la grande variabilité de cette région. Seules quelques séquences d'*Hylomyscus* ont pu être lues avec les amorces L15926 et H16498. Toutes les autres tentatives ont donné des résultats illisibles du fait de pics superposés (84% d'échec, n = 31).

Pour les Primates, il semblait pertinent de compléter les séquences d'ADN mitochondrial par une portion d'ADN nucléaire, à cause du biais important entre dispersion des mâles et des femelles (voir plus haut). Les amorces non spécifiques dont nous disposions au laboratoire pour l'ARN-18S et la région ITS n'ont pas permis d'obtenir de séquences. Dans le premier cas, le fragment d'ADN a pu être amplifié pour 5 échantillons sur 17, mais le séquençage réalisé pour deux de ces échantillons a donné des résultats illisibles. Dans le second cas, aucun produit d'amplification n'a été obtenu sur les 4 échantillons de Primates testés. Comme l'ARNr 12s mitochondrial a été utilisé par van der Kuyl et collaborateurs (1995, 2000a) pour reconstruire la phylogénie des Primates de l'ancien monde, nous avons aussi séquencé une portion de cette molécule avec les amorces universelles 12Sai et 12Sbi (0% d'échec, n = 7).

c) Purification

But et principe de la purification

Afin d'éliminer les amorces inutilisées, les protéines indésirables et l'excès de sels, une purification de l'ADN est nécessaire. L'ADN est d'abord fixé à une résine affine pour l'ADN dans des conditions de salinité élevée. Il est ensuite rincé par une solution à base d'éthanol, puis élué à l'eau distillée.

Protocole

Deux séries de kits de purification Pharmacia ont été utilisés selon le protocole fourni par le fabricant. L'élimination de la fraction soluble est obtenue par une surpression pour le premier kit, et une centrifugation pour le second.

d) Séquençage

Le principe du séquençage

Le principe du séquençage selon la méthode de Sanger repose sur la synthèse partielle du fragment d'ADN étudié. Pour chacune des bases azotées une réaction de type PCR employant une seule amorce (donc sans amplification exponentielle du nombre de copies) est réalisée séparément. Dans chaque tube, un di-désoxynucléotide est ajouté en concentration limitée (environ 1% par rapport au désoxynucléotide normal). L'insertion aléatoire de ce nucléotide modifié au cours de la polymérisation arrête aléatoirement la synthèse du brin complémentaire d'ADN. Ceci permet d'obtenir dans chaque tube des morceaux de différentes longueurs se terminant par la base azotée concernée.

Le produit des quatre réactions est mis à migrer sur gel de poly-acrylamide dans quatre puits successifs, libellés A, C, G et T. La lecture de la séquence se fait à l'aide d'un rayon laser qui détecte le passage des molécules marquées par une substance fluorescente. Comme les fragments migrent à des vitesses différentes selon leur longueur, il est possible de reconstituer la séquence nucléotidique en fonction de l'ordre de passage des fragments. Le plus court des fragments migre le plus vite et passe devant le rayon laser en premier. Le puits dans lequel il se trouve révèle le nucléotide auquel il correspond. Il en va de même pour tous les fragments qui suivent. A la fin, la liste ordonnée des fragments A, C, G ou T qui sont passés successivement devant le rayon laser révèle la séquence nucléotidique.

Protocole de séquençage (réactions, amorces, cycles)

Deux types de kits fabriqués par les firmes Amersham et Pharmacia ont été employés, selon les instructions du fabricant. Les cycles ont été programmés comme suit :

(30 sec à 94°C, 30 sec à ...°C et 80 sec à 72°C), 30 fois.

Les amorces utilisées sont des amorces fluorescentes de même séquence que celles utilisées pour la PCR, mais les températures d'hybridation sont augmentées de façon à améliorer la précision (Annexe 4).

Lecture des résultats

La plupart des séquences ont été lues à l'aide d'un séquenceur "Alf" de chez Pharmacia à 40 puits, permettant de lire 10 brins d'ADN à la fois. Un changement d'appareil a dû être opéré en cours de travail en raison d'une défaillance technique du premier. Une partie des séquences a donc été lue sur un séquenceur "Alf-express" (Pharmacia) muni d'un rayon laser de longueur d'onde différente de celle du

premier. Il en a résulté un changement du marquage des amorces qui ont dû être re-synthétisées. Le mode de fabrication des gels de poly-acrylamides à 6% a également été modifiée à ce moment-là. Les gels fabriqués par l'utilisateur à partir d'un mélange appelé "Ready-mix" et d'urée ont été remplacés par un mélange FastGel-LongRead (Pharmacia) qui polymérise par exposition prolongée aux UV. Les conditions d'électrophorèse ont été légèrement modifiées : la température de migration a subi une augmentation de 45°C à 55°C dans le but d'améliorer la résolution des séquences.

La lecture des séquences est réalisée automatiquement par le logiciel Alfinwin, mais nécessite une vérification minutieuse. Les résultats apparaissent sous forme de courbes montrant l'évolution de la quantité d'ADN provenant de chaque puits au cours de la migration. Toutes les séquences ont été relevées manuellement sur papier et vérifiées et/ou complétées par lecture du second brin. Dans quelques cas où la qualité des séquences était très bonne et où nous disposions de plusieurs individus de la même localité présentant le même haplotype, la lecture n'a été faite que sur un seul brin.

Risques d'erreur et solutions apportées

Afin de vérifier que les séquences obtenues correspondaient bien aux espèces et aux spécimens choisis, nous avons procédé à une double lecture des séquences (lecture des deux brins), et au moins deux échantillons par unité taxinomique (espèce ou localité) ont été utilisés dans la mesure du possible. Dans les cas où les séquences se groupaient de façon surprenante, nous avons procédé à une ré-extraction et à un nouveau séquençage de l'échantillon douteux. Deux permutations d'échantillons ont ainsi été détectées.

Afin de se prémunir contre les contaminations par de l'ADN extérieur, nous avons vérifié la cohérence des données et utilisé des amorces spécifiques aux taxons étudiés. Des contaminations par des copies nucléaires ne peuvent être exclues (Arctander, 1995), mais l'ADN mitochondrial étant en excès par rapport à l'ADN nucléaire simple copie, on peut espérer que la séquence lue est bien celle de la fraction mitochondriale. Dans certains cas où l'amplification a réussi mais le séquençage a donné des résultats illisibles avec plusieurs pics superposés, par exemple pour la D-loop, il est probable que des copies nucléaires aient masqué la séquence mitochondriale. Pour y remédier, il aurait vraisemblablement fallu cloner l'ADN amplifié par PCR et séquencer plusieurs clones indépendamment, ce qui est très coûteux.

Toutes les portions d'ADN souhaitées n'ont pas pu être amplifiées ou séquencées lisiblement. Pour la plupart des échantillons (n = 407), seule la séquence de l'ARN-16s a été obtenue. Pour 28 échantillons, nous avons séquencé également une partie du cytochrome *b* et pour 5 autres, une portion d'ARN 12s.

IV - RECONSTRUCTIONS PHYLOGENETIQUES

1) Alignement

Pour toute reconstruction phylogénétique, il est nécessaire de postuler que les caractères utilisés sont homologues et indépendants. Il est donc nécessaire d'aligner les séquences au préalable, de façon à mettre en correspondance les caractères homologues. Il est préférable d'exclure de l'analyse les régions pour lesquelles l'alignement est douteux (Gonzales *et al.*, 1990 ; Wheeler et Honeycutt, 1988).

Les séquences peuvent être alignées selon le principe de l'algorithme de Needleman et Wunsch (1970) qui, après avoir attribué un coût aux insertions et délétions (indels) et aux substitutions, positionne des insertions dans les séquences de telle façon que le coût global de l'alignement soit minimum. Les indels sont pénalisés par rapport aux substitutions, et les transversions sont pénalisées par rapport aux transitions.

Certains programmes fonctionnent par addition progressive des taxons les plus semblables et créent une séquence consensus à chaque addition (ex : Clustal - Thompson *et al.*, 1994). L'alignement final n'est pas forcément optimal, et il est nécessaire d'opérer une correction manuelle en révisant l'alignement en fonction de la ressemblance globale entre séquences (Collins *et al.*, 1994) ou de la structure secondaire de la molécule.

D'autres méthodes privilégient l'alignement qui maximise le signal phylogénétique. Elles se basent sur la topologie d'un arbre phylogénétique inféré de connaissances antérieures ou obtenu par un pré-alignement des données. L'alignement sélectionné est celui qui minimise le coût de l'arbre phylogénétique associé. Ce principe est utilisé par TreeAlign (Hein, 1989) qui ne prend en compte qu'une seule topologie de départ, et par Malign (Wheeler et Gladstein, 1994) qui fonctionne de façon itérative. Ce dernier programme est un des plus performants ; il permet en outre de régler le problème de la variabilité des résultats selon l'ordre d'entrée des séquences.

Hormis pour l'alignement de l'ensemble des Crocidurinae et des groupes externes pour les Muridés, cette opération a nécessité l'insertion de peu d'indels et n'a pas été considérée comme une difficulté majeure. La plupart des alignements ont été réalisés à l'aide du logiciel ClustalW qui est d'une utilisation plus aisée et plus rapide que Malign. Ils ont ensuite été corrigés manuellement, en tenant compte de la structure secondaire de la molécule.

2) Les différentes méthodes employées en phylogénie

Il existe deux types d'approches : d'une part les méthodes basées sur la ressemblance globale (méthodes phénétiques) et d'autre part les méthodes basées sur l'évolution des caractères (méthodes cladistiques et méthodes probabilistes).

a) Méthodes phénétiques (ou "de distances")

Le principe de ces méthodes est de relier successivement entre eux les taxons qui sont les plus proches, selon le critère d'évolution minimale défini par Saitou et Nei (1987). La reconstruction se fait à partir d'une matrice des distances entre taxons. Les distances correspondent au nombre de caractères qui diffèrent entre les taxons pris deux à deux. Elles reflètent la ressemblance globale entre taxons. Une première étape consiste à choisir les taxons les plus proches et à les lier entre eux par un nœud représentant l'ancêtre commun. La seconde étape consiste à recalculer les distances entre le nœud obtenu et tous les autres taxons. L'arborescence est construite par itération de ces deux étapes.

Les méthodes de distances sont les plus rapides, mais présentent des inconvénients. Leur utilisation entraîne une perte d'information lors de la conversion des données en une matrice de distance (Penny, 1982 ; Felsenstein, 1988). De plus, elles sont sensibles aux variations de vitesse, c'est-à-dire au non respect du principe de l'Horloge Moléculaire.

UPGMA (Unweighted Pair Group Method Using Arithmetic Averages)

Cette méthode peut être employée lorsque l'évolution se fait à vitesse constante et symétriquement le long des différentes branches. C'est la méthode la plus sensible aux violations du principe d'horloge moléculaire. D'après des simulations effectuées par Hillis (1996), elle conduit à un arbre erroné dans plus de la moitié des cas, même avec un nombre d'espèces et de caractères extrêmement important. Cette méthode n'est plus guère utilisée.

Neighbor Joining (NJ ; Saitou et Nei, 1987)

Cette méthode peut être employée lorsque l'évolution se produit à des vitesses différentes le long des différentes branches. Elle a longtemps été reconnue comme la plus fiable des méthodes de distance (e.g. Sourdis et Nei, 1988 ; Hillis, 1996) et elle peut être préconisée dans les cas où le nombre de caractères informatifs est limité (Cornish-Bowden, 1983).

Méthodes améliorées :

- BioNJ (Gascuel, 1997)

Ce programme utilise un algorithme inspiré de celui du NJ, mais dont la seconde étape est modifiée. Pour le calcul de la nouvelle matrice, la variance des distances moyennes calculées est réduite au moyen d'une correction basée sur la variance espérée.

-Weighted Neighbor Joining (WNJ) in Weighbor (Bruno et al., 2000)

Ce programme modifie les deux étapes de l'algorithme. Pour la première étape, le critère d'appariement utilisé est un critère de vraisemblance. Les distances sont considérées comme des variables gaussiennes et une formule permet de tenir compte de leurs variances et co-variances. Pour la seconde étape, le principe est le même que pour BioNJ.

Cette méthode, rapide et presque aussi performante que le maximum de vraisemblance, est à préconiser dans les cas où il y a présence de longues branches et non respect de l'horloge moléculaire (Bruno et al., 2000).

b) Méthodes basées sur les caractères

Méthodes cladistiques et méthodes probabilistes

La cladistique a été fondée par Hennig (1966) et repose sur l'idée que ce sont les caractères dérivés communs qui doivent être utilisés pour construire une phylogénie. Elle a été complétée par le principe de parcimonie qui consiste à retenir l'arbre qui nécessite le moins de changements, ce qui permet de limiter au maximum les hypothèses.

La construction d'un arbre dichotomique implique l'acceptation du principe de "spéciation divergente" (elle serait convergente s'il y avait hybridation généralisée ou transfert horizontal de caractères). Cela suppose que la similarité est le résultat d'un héritage en ligne directe, avec permanence de l'expression du caractère (Farris, 1989). En principe, seules les modifications héritées devraient être utilisées pour définir les groupements. Les autres cas de similarités sont théoriquement exclus. Mais, en pratique, il est difficile de distinguer les apomorphies (traits partagés dérivés) des symplesiomorphies

(traits partagés ancestraux) et des homoplasies (convergences). De fait, la topographie de l'arbre ne pourra être reconstituée correctement que si les homoplasies sont rares proportionnellement aux apomorphies.

Parmi les méthodes qui utilisent le principe de parcimonie, nous ne retiendrons que celle du maximum de parcimonie. Une autre méthode, dite de compatibilité ou "clique", tend à préférer l'arbre soutenu par le plus grand nombre de caractères, ce qui revient à ne conserver que le plus grand ensemble (ou clique) de caractères compatibles, donc à éliminer de l'analyse les caractères fortement homoplasiques (Le Quesne, 1969).

Les méthodes probabilistes évaluent la probabilité d'obtenir chaque arbre, en fonction des données et d'un modèle d'évolution. Il existe une méthode basée sur le principe des moindres carrés qui consiste à minimiser la somme des carrés des écarts relativement aux distances estimées, mais elle est très peu utilisée et nous ne la détaillerons pas. La méthode du maximum de vraisemblance, en revanche, est couramment employée.

Les méthodes cladistiques reposant sur le principe de parcimonie sont généralement opposées aux méthodes dites probabilistes. Mais en réalité, la parcimonie est un cas particulier du principe du Maximum de Vraisemblance, où le modèle serait le moins contraignant possible (Farris, 1999). Ce modèle a été mis en équations par Tuffley et Steel (1997), sous le nom de modèle NCM ("No Common Mechanism"), qui suppose une indépendance complète entre les processus évolutifs affectant les caractères sur les différentes branches du cladogramme.

Maximum de Parcimonie (MP)

La parcimonie recherche les arbres de longueur L minimale, où L est la somme sur les B branches (k) et N caractères (j) des $\omega_j \text{ diff}(x_{k,j}, x_{k',j})$. Cette longueur L est appelée "nombre de pas".

La longueur de l'arbre peut être calculée selon différentes méthodes, mais une seule est adaptée à l'analyse de séquences nucléotidiques : celle de Fitch (1971) qui suppose que tous les changements sont possibles, dans les deux sens. Les autres méthodes supposent soit que les caractères sont ordonnés mais non polarisés (méthode de Wagner, formalisée par Kluge et Farris, 1969 et Farris, 1970), soit que les changements sont ordonnés et polarisés, le gain d'un caractère n'étant possible qu'une seule fois, mais les réversions multiples étant autorisées (méthode Dollo ; Farris, 1977), soit que les changements sont irréversibles, ce qui n'est pas raisonnable (méthode de Camin et Sokal, 1965).

La recherche de l'arbre le plus parcimonieux nécessite d'analyser tous les arbres possibles afin de trouver l'arbre optimal. Le temps d'évaluation de tous les arbres possibles augmente donc exponentiellement avec le nombre de taxons. Au-delà d'une dizaine de taxons, il devient difficile d'entreprendre une recherche exhaustive, et au-delà d'une vingtaine, même une recherche de type branch-and-bound prend également beaucoup de temps (cette technique évalue pour chaque branche à construire les différentes topologies possibles et ne retient que celles menant à un arbre de taille inférieure à la taille limite fixée au départ). On a alors recours à une recherche heuristique qui ne permet pas d'examiner tous les cas de figure. L'heuristique peut être répétée plusieurs fois, en partant à chaque fois d'un arbre aléatoire de façon à éviter l'écueil de l'optimum local (cf. Stewart, 1993). Il existe 3 méthodes heuristiques permettant de rechercher l'arbre le plus parcimonieux :

- NNI (nearest neighbor interchange) : interversion de la position de séquences proches prises deux à deux, ou regroupement de ces séquences ;
- SPR (subtree pruning and regrafting) : déplacement de clades par coupure puis connexion à une autre position dans l'arbre ;
- TBR (tree bisection and reconnection) : réorganisation de clades par coupure puis re-connexion au même endroit, avec un agencement différent. C'est la méthode la plus communément employée.

Pour savoir si l'arbre obtenu a des chances d'être effectivement le plus court, certains comptent le nombre de pas qui séparent cet arbre des suivants. Cette estimation n'est pas fiable, car les chances de trouver un arbre unique beaucoup plus court que les autres sont élevées, surtout si le nombre de caractères est très supérieur au nombre de taxons (Hillis et Huelsenbeck, 1992). Une autre méthode consiste à évaluer le biais ("skewness") dans la distribution de la longueur des arbres au moyen de la statistique g_1 (Fitch, 1984 ; Hillis, 1991). Plus g_1 est négatif plus la résolution a des chances d'être bonne, mais cette condition nécessaire n'est pas suffisante. En effet, le biais peut résulter d'un petit nombre de taxons dont la position n'est pas ambiguë et ne pas être représentatif de l'ensemble.

La parcimonie ne donne de bons résultats que si les données sont bien structurées et si l'homoplasie est rare ou répartie de façon homogène au sein de l'arbre. Des simulations ont révélé que la parcimonie peut parfois créer des groupements fictifs bien supportés là où il devrait en réalité y avoir une polytomie (Slowinski, 2001). Lorsque le nombre de taxons est trop important et/ou le nombre de caractères trop réduit, la parcimonie peut ne pas parvenir à trouver l'arbre de longueur minimale ou

conduire à un très grand nombre d'arbres. Il est donc impératif que le nombre de caractères informatifs soit supérieur au nombre de taxons.

Maximum de vraisemblance (ML ; Felsenstein, 1981a)

Cette méthode recherche l'arbre qui maximise la vraisemblance de l'échantillon conditionnellement au modèle. En d'autres termes, elle tend à maximiser la probabilité qu'un modèle évolutif et une topologie donnés génèrent les données observées. Cela suppose de calculer le maximum de vraisemblance de chacun des modèles de transition possibles entre caractères, pour chacun des arbres possibles, puis de les comparer deux à deux. Le temps de calcul augmente donc exponentiellement avec le nombre d'échantillons.

La vraisemblance totale d'un arbre est le produit des vraisemblances obtenues pour chacun des sites. En pratique, on calcule généralement le log-likelihood qui est la somme des log-likelihoods de chacun des sites. Pour un certain site (j), la vraisemblance est égale à la somme des probabilités de toutes les reconstructions possibles de la forme ancestrale à chacun des nœuds, sous un modèle évolutif donné.

Le modèle prend en compte :

- une matrice de passage entre les différents états. C'est souvent le "general time reversible model" qui est préféré car il permet d'affecter une probabilité différente à chaque type de substitution.
- un modèle de variation différentielle selon les sites de type distribution gamma (facultatif). La distribution gamma est caractérisée par le paramètre alpha : plus alpha est grand plus la probabilité que chaque site change au moins une fois est élevée.

Les paramètres peuvent être fixés *a priori*, notamment si le jeu de données est limité, ou bien évalués à partir des données, ce qui est généralement préféré. Dans ce dernier cas, les paramètres sont calculés sur la base d'un arbre préliminaire qui peut être construit par une méthode quelconque car la topologie de l'arbre influence assez peu la valeur des paramètres (Yang *et al.*, 1994).

Cette méthode donne de bons résultats lors de tests de simulation (lorsque le modèle est adapté aux données) et permet d'estimer la cohérence entre modèle et données. Elle nécessite une certaine régularité de l'évolution puisqu'elle suppose que la probabilité d'un type de changement au cours du temps, c'est-à-dire entre les différentes branches, est fixe. La méthode ne teste pas la validité du modèle évolutif utilisé, ni celle des paramètres employés. Il n'existe pas de critère permettant de choisir une

pondération optimale (Barker et Lanyon, 2000), par contre la pertinence du modèle peut être évaluée grâce à des tests qui mesurent l'adéquation entre les données et différents modèles. Le logiciel Modeltest (Posada et Crandall, 1998) propose deux types de tests : le "likelihood-ratio" test qui compare le log-likelihood de modèles de plus en plus complexes, et le critère d'information d'Akaike qui récompense les modèles les plus adaptés aux données tout en pénalisant l'ajout de paramètres inutiles. La réalité biologique ne peut être abordée que par des modèles très complexes, mais plus le modèle est complexe et moins la méthode est cohérente. Enfin, le ML ne peut être employé que pour un petit nombre de taxons du fait du temps de calcul nécessaire.

c) Comment choisir ?

En général, on admet que les méthodes MP et ML donnent les meilleurs résultats (e.g. Kishino et Hasegawa, 1989 ; Yang, 1996). Si on considère un modèle évolutif très simple, MP et ML donnent des résultats comparables, ce qui est normal puisque la parcimonie est un cas particulier de ML où le modèle serait le moins contraignant possible (Farris, 1999). Mais si le modèle est complexe, ML a plus de chances de trouver la vraie topologie (Yang, 1996). Ainsi, dans le but de minimiser les hypothèses évolutives, on aura recours à la parcimonie qui suppose qu'il n'y a pas de mécanisme commun d'évolution plutôt qu'au ML qui suppose au moins que l'évolution est similaire pour toutes les séquences. Mais, dans le cas de séquences d'ADN, l'évolution est soumise à certaines contraintes qui méritent d'être prises en compte. En effet, la non intégration d'un modèle évolutif basé sur une connaissance préalable des mécanismes évolutifs impliqués (qui ne sont malheureusement pas encore bien connus) peut conduire à des résultats allant à l'encontre des phénomènes naturels (Swofford et Olsen, 1990). On pourra réaliser une analyse MP en affectant des poids inégaux aux caractères ou aux différents types de changements ou une analyse ML en choisissant un modèle d'évolution.

Les méthodes de distances sont moins fiables, et présentent un risque de création de groupements aberrants basés sur des symplésiomorphies. Par contre, elles sont beaucoup plus rapides, donc intéressantes lorsque le jeu de données est important. Il est possible de compenser les variations des taux de substitution par l'application de corrections sur les distances. La méthode WNJ allie la rapidité des méthodes phénétiques à la prise en compte de la variabilité des distances estimées. Elle est peu sensible à la présence de longues branches et presque aussi performante que le ML (Bruno *et al.*, 2000).

Selon Kim (1993), un moyen de ne pas biaiser les résultats par le choix de la méthode est d'opérer plusieurs approches. D'après les simulations effectuées par cet auteur, lorsque UPGMA, NJ et MP donnent des résultats similaires, il y a de fortes chances que la topologie de l'arbre obtenu soit juste, et inversement. Ce type d'approche œcuménique permet de palier à notre incapacité de choisir le modèle qui reflète au mieux le mode d'évolution des données. Cependant, l'utilisation d'une méthode erronée ne peut renforcer ni contredire un résultat "juste" obtenu avec une méthode adéquate. Dans tous les cas, le choix de la méthode doit pouvoir être justifié, en toute connaissance des postulats sous-jacents. Ce choix devrait être justifié *a priori*, en fonction des propriétés des différentes méthodes, et non pas *a posteriori*, au vu des résultats jugés plus ou moins satisfaisants (P. Deleporte, com. pers.).

Pour les reconstructions phylogénétiques, plusieurs méthodes ont été comparées de façon à retenir les topologies les mieux supportées. Parmi les méthodes de distances, nous avons privilégié le Weighted Neighbor Joining, qui est la méthode la plus robuste (logiciel Weighbor – Bruno *et al.*, 2000). La parcimonie a été utilisée à titre de comparaison, en pénalisant les transitions par rapport aux transversions lorsque les transitions paraissaient saturées (logiciel utilisé : PAUP 4b3a - Swofford, 2000). Un plus grand crédit a été accordé au maximum de vraisemblance car cette méthode repose sur un modèle intrinsèque d'évolution et est donc la plus à même de prendre en compte les contraintes portant sur l'évolution des séquences d'ADN (logiciel utilisé : Phylip 3.5c - Felsenstein, 1995, sauf mention contraire). Dans le cas de l'ARN 16s, notamment, on sait que les positions n'évoluent pas forcément indépendamment les unes des autres et que l'évolution n'est pas régulière. Le modèle d'évolution reflétant le mieux les données a été choisi au moyen du logiciel Modeltest (Posada et Crandall, 1998).

3) Amélioration de la qualité de la reconstruction

La reconstruction peut être délicate dans plusieurs cas : lorsque le niveau d'homoplasie est important, lorsque l'arbre a une topologie déséquilibrée (c'est-à-dire des vitesses d'évolution variables d'un taxon à l'autre), ou un faible facteur de "stemminess" (c'est-à-dire de courtes branches internes et de longues branches terminales). Il est important d'attacher une attention particulière au choix des séquences et des taxons à inclure dans l'analyse, en termes de qualité et de quantité.

a) Le problème de l'homoplasie et la saturation des substitutions

Afin de ne conserver que l'information pertinente, il faut repérer et pénaliser les homoplasies. Ceci ne peut être fait directement, mais, l'homoplasie étant liée à un phénomène de saturation résultant de substitutions multiples, on peut repérer les positions saturées et leur donner un poids faible par rapport aux autres positions.

Pour estimer le degré de saturation, on peut représenter graphiquement le taux de substitutions entre séquences prises 2 à 2 en fonction du temps (estimé) écoulé depuis leur divergence ou bien en fonction du nombre estimé de substitutions dans l'arbre le plus parcimonieux (e.g. Hassanin *et al.*, 1998). Lorsque plusieurs régions ou plusieurs classes de positions sont identifiées, la valeur de la pente de la régression pour chacune de ces catégories est utilisée pour localiser les positions saturées. On peut aussi tracer la courbe du pourcentage de transitions en fonction du temps. Cette courbe décroît exponentiellement et atteint un plateau à 35-45% lorsqu'il y a saturation (la valeur théorique est 33% car un tiers des substitutions possibles sont des transitions). Cependant un taux de transitions élevé n'est pas synonyme de non saturation, du fait des contraintes portant sur l'évolution de l'ADN.

Une façon de corriger l'homoplasie est d'appliquer des pondérations visant à compenser les biais affectant les taux de transitions (TS) et de transversions (TV), les fréquences nucléotidiques et les probabilités de substitutions entre sites. Ces corrections peuvent être déterminées *a priori*, en fonction des caractéristiques générales de la portion d'ADN étudiée (pondération en fonction de la position au sein des codons, de la structure secondaire...), ou *a posteriori*, en fonction des caractéristiques propres de la séquence (variabilité observée, mécanismes évolutifs impliqués). Pour les méthodes phénétiques, les distances sont recalculées à l'aide de formules qui prennent en compte tout ou partie des facteurs de variabilité. Pour les méthodes cladistiques, il faut affecter des poids aux changements d'état entre caractères de façon à pénaliser ceux qui sont les plus variables et définir une matrice attribuant un coût aux différents types de substitutions. Les modèles probabilistes effectuent des corrections de manière intrinsèque.

Différents modèles de correction des distances :

Une première catégorie de corrections considère que le taux de substitution est homogène le long de la séquence. La probabilité de substitutions multiples est estimée selon une loi de Poisson, après définition d'hypothèses sur les taux de substitution. Il existe une série de formules impliquant différentes

contraintes :

- Jukes et Cantor (1969) : $TS = TV$ et les fréquences des nucléotides sont identiques ;
- Kimura (1980) et Jin et Nei (1990) : $TS \neq TV$ et les fréquences des nucléotides sont identiques ;
- Felsenstein (1981b) et Tajima et Nei (1984) : $TS = TV$ et les fréquences des nucléotides sont variables ;
- Hasegawa *et al.* (1985) et Tamura (1992) : $TS \neq TV$ et les fréquences des nucléotides sont variables ;
- Kishino et Hasegawa (1989) et Tamura et Nei (1993) : $TS \neq TV$, les fréquences des nucléotides sont variables et le taux de transition peut être différent pour les purines et les pyrimidines.

Une seconde catégorie de corrections suppose que les taux de substitution varient d'un site à l'autre.

La formule la plus couramment employée est celle de Shannon (1948) qui permet de déduire le schéma de variabilité, appelé "distribution gamma", des données elles-mêmes. Les sites sont regroupés par classes de variabilité caractérisées par leur taux moyen de substitutions. Cette formule peut être appliquée aux méthodes de correction listées ci-dessus.

La méthode de Tamura et Nei (1993) suppose que les taux de substitutions par site suivent une loi binomiale négative régie par le paramètre α (Uzell et Corbin, 1971) qui peut être calculé à partir des données selon la méthode des moments (e.g. Tamura et Nei, 1993). Ce paramètre représente l'inverse du coefficient de variation du taux de mutation, donc, plus il est faible, plus le biais est important. Notons que le modèle de Tamura et Nei suppose que TS est différent de TV et que les fréquences nucléotidiques sont variables. Les positions sont pondérées en fonction de la fréquence des substitutions.

Le problème principal de ces corrections est que plus le modèle est compliqué, plus le nombre de paramètres entrant en ligne de compte est grand et plus l'adéquation est bonne, mais plus la variabilité est grande. Il faut faire un compromis entre meilleure adéquation et moindre variabilité. En pratique, si deux méthodes donnent les mêmes résultats, mieux vaut choisir la plus simple (Kumar *et al.*, 1993). Les pondérations affectent peu la topologie de l'arbre car il en résulte une augmentation des distances plus ou moins proportionnelle à ces distances. En revanche, elles sont très importantes lorsqu'on cherche à estimer le temps écoulé depuis un événement de divergence.

Correction des données pour les méthodes autres que celles de distances

Il existe des méthodes estimant la variabilité des données préalablement à la construction de l'arbre. La distorsion est réduite grâce à l'application d'une formule de correction de la dissemblance qui limite les différences de vitesse de substitution entre taxons (cf. Swofford et Olsen, 1990). Une correction plus sophistiquée consiste à pondérer indépendamment les positions en fonction du taux réel de substitutions par rapport au taux attendu, c'est le principe EOR ("expected-to-observed ratio") utilisé par la méthode de Collins et collaborateurs (1994).

Farris (1969) a proposé de procéder par pondération successive en se basant sur la topologie de l'arbre résultant pour calculer les poids utilisés à chaque étape. Hassanin et collaborateurs (1998a) suggèrent d'appliquer un poids inversement proportionnel au taux d'homoplasie, c'est-à-dire proportionnel au soutien qu'il apporte à la phylogénie. Ce poids peut être basé sur l'indice de cohérence normé (RC) ou sur un indice traduisant le degré de saturation mutationnelle (S). Après pondération, un nouvel arbre est construit, et cette opération est répétée jusqu'à obtention d'une solution stable.

Ce type de correction a l'avantage de découler des propriétés des données et non pas de conditions fixées *a priori*. Par contre, il est fortement influencé par l'arbre utilisé au départ pour le calcul des pondérations. Ce problème peut être résolu en appliquant la pondération successive à tous les arbres possibles, ce qui est très coûteux en temps (Golobov, 1993).

En fait, comme pour la correction des distances, aucune méthode vraiment satisfaisante n'existe (Simon *et al.*, 1994). En outre, les positions homoplasiques sont celles qui contiennent potentiellement le plus de signal, et ce signal peut être cohérent (Philippe *et al.*, 1996). Dans certains cas, les transitions peuvent se révéler plus fiables que les transversions (Broughton *et al.*, 2000).

b) Le problème de l'évolution mosaïque

C'est ce qui se produit lorsque certaines positions évoluent rapidement dans certaines régions de l'arbre alors qu'elles restent invariables dans d'autres. Dans ce cas, c'est l'approche MP qui sera la plus adaptée puisque cette méthode autorise n'importe quel type de changement. En revanche, l'utilisation de pondérations (successives ou non) ira à l'encontre de cette logique car elle implique qu'il existe un mécanisme commun d'évolution dans toutes les lignées. En effet, si un caractère ne se révèle fortement homoplasique que dans une partie du cladogramme, la pondération uniforme considérera ce caractère comme homoplasique pour toutes les lignées, même là où il présente une apparente stabilité locale

(P. Deleporte, com. pers.). Selon Farris (1999), aucun modèle ML ne prévoit la possibilité que les taux d'évolution diffèrent d'une partie de l'arbre à l'autre, alors que les cas d'évolution mosaïque ne sont pas rares. En réalité, un modèle d'évolution dit "Covarion Model" (Fitch, 1971) a été défini pour prendre en compte l'évolution mosaïque, mais il n'est pas proposé par les logiciels de traitement des données.

c) Le problème des longues branches

Les longues branches posent deux types de problèmes :

- attraction des longues branches ("Long Branch Attraction" ou LBA) : quand certains taxons évoluent plus vite que d'autres, ils ont tendance à se grouper de façon erronée à la base de l'arbre phylogénétique (Felsenstein, 1978). Pour limiter ce phénomène, il convient de choisir des groupes externes aussi proches que possible des taxons étudiés, de façon à limiter l'effet de saturation. On peut appliquer une correction prenant en compte la variabilité des taux de substitution entre sites, et même supprimer les sites invariants. On peut aussi reconstruire l'arbre en éliminant les espèces à évolution rapide, ou bien ajouter d'autres espèces afin de réduire les longues branches. Le problème LBA affecte principalement les méthodes de parcimonie, et, dans une moindre mesure, les méthodes phénétiques standard. A l'opposé, lorsque la proximité des longues branches est réelle, la parcimonie donne de meilleurs résultats que le maximum de vraisemblance qui aura tendance à les éloigner (Farris, 1999). Dans ce cas, on parlera de répulsion des longues branches ("Long Branch Repulsion", Siddal, 1998)

- perturbation par les longues branches ("Long Branch Distraction" ou LBD) : la présence de longues branches ne pose pas forcément de problème pour le positionnement de la branche elle-même, mais peut entraîner des erreurs dans le positionnement d'autres taxons (Pollock et Goldstein, 1995). C'est le cas lorsque la méthode utilisée ne peut intégrer la faible fiabilité des grandes distances, c'est-à-dire surtout pour le NJ et dans une moindre mesure pour la parcimonie (Bruno *et al.*, 2000).

d) Le codage des indels

La présence d'insertions et de délétions dans une séquence a une valeur phylogénétique qui est parfois considérée comme supérieure à celle des substitutions car moins susceptible d'homoplasie (e.g. Lloyd et Calder, 1991). Malgré cela, les positions présentant des indels sont souvent éliminées, soit pour tous les taxons, soit seulement pour les comparaisons deux à deux où l'un des taxons présente une délétion. Cette simplification est adoptée pour deux raisons : d'une part parce que l'alignement est souvent douteux dans les régions contenant des indels, et d'autre part parce que la prise en compte d'indels de plus d'une paire de base nécessite un codage particulier.

Comme les indels ne sont pas des données observées mais un état inféré lors de l'alignement, Wheeler (1996) a proposé de réaliser simultanément l'alignement et le codage des indels, qui sont alors considérés comme des événements et non comme des états du caractère. Cet auteur considère à tort l'alignement (qui recherche des homologies primaires) et la reconstruction de l'arbre (qui teste cette homologie primaire et propose des hypothèses d'homologie dérivée ou secondaire) comme des processus interdépendants (Simmons et Ochoterena, 2000).

Le codage des indels en vue de les comptabiliser dans l'analyse peut être effectué en traitant les délétions comme s'il s'agissait d'un cinquième état du caractère (cf. Simmons et Ochoterena, 2000). Les délétions affectant plusieurs positions adjacentes sont alors codées pour une seule de ces positions, choisie arbitrairement, tandis que les autres positions affectées sont considérées comme des données manquantes (Bena *et al.*, 1998) ou bien sont exclues de l'analyse. Ce codage n'est actuellement réalisable que pour des analyses effectuées selon le principe de parcimonie. En fait, la méthode qui semble la plus pertinente consiste à coder séparément les indels en termes de présence ou absence. La difficulté est alors de déterminer l'homologie entre indels partiellement chevauchants. Simmons et Ochoterena (2000) proposent deux méthodes pour effectuer ce codage, l'une simple et rapide, l'autre complexe mais plus précise. La plus simple consiste à coder toutes les délétions qui ne sont pas parfaitement chevauchantes comme des caractères distincts à trois états : présent, absent ou indéterminé. L'état "indéterminé" correspond à une situation où il n'est pas possible de dire si une délétion est présente, parce que celle-ci est totalement incluse dans une autre qui, elle, est présente dans la séquence concernée. A titre d'exemple, les délétions qui figurent dans les séquences ci-après seront codées comme indiqué dans la matrice qui suit, où chaque colonne représente une délétion :

ACGTTACCT		000
AC---CCT (délétion 1)	→	10?
A---TACCT (délétion 2)		01?
AC--TACCT (délétion 3 $\subset$ 1 et 2)		001

e) Le choix des groupes externes

Il est important de choisir plusieurs groupes externes qui soient assez proches des groupes internes. Il est même possible d'utiliser des gènes paralogues pour enraciner l'arbre.

Smith (1994) propose la stratégie suivante : commencer par deux groupes externes éloignés l'un de l'autre de façon à parfaire l'alignement et à choisir les sites à pondérer ; puis sélectionner pour l'analyse finale au moins deux groupes externes apparentés les plus proches possibles des groupes internes. Ceci permet de réduire les longueurs de branches et d'équilibrer l'arbre.

f) Plus de taxons ou plus de nucléotides ?

D'un côté, il semble nécessaire de travailler avec un grand nombre de caractères pour obtenir un arbre fiable. Givnish et Sytsma (1997) montrent que lorsqu'on augmente le nombre de caractères informatifs, la probabilité de trouver l'arbre juste par la parcimonie augmente. Nous avons également montré qu'il est important d'utiliser plusieurs gènes indépendants (cf. paragraphe III.1.b).

D'un autre côté, l'ajout de taxons permet d'améliorer la représentativité des données, de limiter les longues branches et de réduire le bruit de fond par dispersion de l'homoplasie (Hillis, 1996). D'après Lecointre et collaborateurs (1993), il est fondamental d'employer un grand nombre de taxons. Il faut choisir plusieurs espèces par groupe présumé monophylétique et utiliser plusieurs groupes externes, car chaque taxon peut avoir un impact sur les 2 ou 3 nœuds environnants.

4) Combinaison de plusieurs jeux de données

Lorsque l'on souhaite combiner l'information phylogénétique contenue dans différents jeux de données, les données peuvent être analysées en une analyse combinée ou bien séparément et être réunies ensuite.

a) Analyses séparées et "super-arbres"

Il peut être préférable de traiter les séquences indépendamment les unes des autres afin de ne pas perdre d'information et de ne pas introduire de biais (Harris et Disotell, 1998). L'analyse conjointe de plusieurs séquences risque de masquer des incohérences et de ne refléter que les séquences les plus informatives qui ne correspondent pas nécessairement à la phylogénie réelle.

Les arbres obtenus indépendamment peuvent être confrontés et leur cohérence estimée par un test ILD ou "incongruence length difference test" (Farris *et al.*, 1994). Les arbres cohérents peuvent ensuite être combinés de deux façons :

- Selon le principe de "congruence taxinomique" (Kluge, 1989), on peut réaliser un arbre consensus qui donne la même importance à tous les jeux de données. Seuls les taxons figurant dans tous ou presque tous les arbres peuvent être utilisés. Le choix de la méthode d'obtention du consensus est arbitraire (e.g. consensus strict, consensus semi-strict, consensus à la majorité, consensus par élimination des informations discordantes - Chippendale et Wiens, 1994).

- Les arbres peuvent être combinés directement en un "super-arbre" (Baum, 1992; Ragan, 1992). Il faut alors encoder chaque nœud de chacun des arbres dans une même matrice, puis analyser la matrice résultante selon le principe de parcimonie. Les taxons constituent les lignes de la matrice et les nœuds les

colonnes. Pour chaque nœud, on attribue la valeur "1" à l'ensemble des taxons qui rejoignent le nœud et la valeur "0" aux autres taxons. Cette méthode permet de combiner des arbres dont les branches terminales diffèrent, d'obtenir un signal même si les arbres de départ semblent incompatibles et de compenser le fait que les arbres les plus gros et les mieux résolus contribuent plus au résultat final en donnant un poids relatif aux arbres.

b) Analyse combinée

L'analyse combinée repose sur le principe de "congruence des caractères". Elle permet de combiner les séquences elles-mêmes en les mettant bout à bout et donc d'augmenter le nombre de caractères informatifs. Il est également possible de combiner des jeux de données de différents types (par exemple des données anatomiques, morphologiques, comportementales, génétiques...) et de les traiter par une analyse simultanée (Nixon et Carpenter, 1996 ; G. Lecointre et P. Deleporte, comm. pers.), selon le principe "d'évidence totale" (Kluge, 1989). Dans tous les cas, et surtout si l'échantillonnage est déséquilibré, l'analyse devra impérativement être faite par la parcimonie afin de ne pas biaiser les résultats.

Le problème de l'analyse combinée est que, en cas de discordance entre jeux de données, le cladogramme final sera peu résolu, peu supporté et ne reflétera pas nécessairement l'information phylogénétique contenue dans chacun des cladogrammes. Lecointre et Deleporte (comm. pers.) proposent une solution permettant de conserver à la fois les avantages des analyses séparées et de l'analyse combinée. Ils conseillent de commencer par traiter les données par des analyses séparées de façon à identifier l'information apportée par chaque jeu de données et à repérer les discordances, puis de tester l'incongruence des caractères pour tous les arbres pris deux à deux, de façon à éliminer de la matrice finale les données significativement discordantes. De cette façon, seule l'information provenant de certains jeux de données et concernant certains taxons est éliminée, et non pas un jeu de données ou un taxon dans leur totalité.

Nous avons tenté d'augmenter la fiabilité de la reconstruction phylogénétique en incluant toujours un grand nombre de spécimens dans l'analyse, provenant d'un panel de taxons le plus large possible. Ceci a été fait au détriment du nombre de nucléotides. Pour certains modèles cependant, les résultats obtenus avec un premier marqueur moléculaire ont été comparés avec ceux obtenus pour un second marqueur, de façon à s'assurer que la phylogénie obtenue n'était pas propre au gène utilisé. Les groupes externes ont été choisis de façon à être le plus proche possible des taxons étudiés.

Pour les analyses phénétiques, la non régularité de l'évolution a généralement été prise en compte par un modèle de correction des distances de type Tamura-Nei qui est un des plus complets. Pour les autres analyses, différentes pondérations ont été appliquées, en utilisant le plus souvent des paramètres estimés à partir des données elles-mêmes.

5) Test qualitatif de l'arbre obtenu

a) Tests de la fiabilité des branches

Quelques tests permettent d'estimer la fiabilité des branches d'un arbre. Ce ne sont pas des tests statistiques et un support élevé ne signifie pas que la topologie est juste, mais seulement qu'elle est bien soutenue par les données.

- Bootstrap

Ce test reflète le degré de support apporté par le jeu de caractères et peut donc être utilisé pour savoir si l'échantillonnage est suffisant. Il consiste en un ré-échantillonnage aléatoire avec remise des caractères jusqu'à obtention d'une série de données de la même taille que l'originale (Lecointre *et al.*, 1993). Le jeu de données ainsi généré est analysé et l'opération est répétée un nombre *n* de fois défini par l'utilisateur. Le pourcentage de fois où une branche apparaît est pris comme indice de fiabilité de cette branche.

- Jackknife

Le jackknife ressemble au bootstrap. Il consiste à enlever systématiquement, les uns après les autres, des sous-ensembles de caractères, à concurrence d'un pourcentage défini par l'utilisateur.

- Jackboot et Jackmono

Ces tests fonctionnent selon le même principe que le bootstrap, mais ré-échantillonnent les taxons au lieu des caractères. Le jackboot est souvent délaissé au profit du jackmono, car ce dernier fait varier moins de paramètres (Lecointre *et al.*, 1993). Un nœud est considéré comme fiable s'il est supporté à au moins 90% quel que soit l'échantillonnage des espèces, à partir d'un nombre minimum d'espèces considéré comme valide.

- Indice de Bremer (1994)

Cet indice de parcimonie, également appelé "autodecay index", estime la fiabilité des nœuds d'un arbre donné en mesurant la différence de longueur entre les arbres les plus courts contenant et ne contenant pas chacun des groupements. Plus l'indice est élevé plus la branche est fiable, mais on ne dispose pas d'échelle sur laquelle comparer les valeurs.

- test du taux de vraisemblance

Ce test permet d'estimer la probabilité qu'une branche soit de longueur non nulle sous un modèle évolutif donné et selon le principe de vraisemblance. Il permet de déceler la présence de groupements apparemment résolus, mais sans valeur statistique. Il est préconisé de l'employer systématiquement pour les analyses ML, d'autant plus que sa rapidité d'exécution le rend très concurrentiel par rapport au bootstrap (cf. Slowinski, 2001). Comme la plupart des logiciels considèrent que le taux de vraisemblance suit une distribution de Chi2 à 1 degré de liberté, alors qu'il suit un Chi2 à tantôt 0 degré de liberté et tantôt 1, les valeurs seuils de significativité doivent être corrigées en suivant la table proposée par Goldman et Whelan (2000).

- analyse de vraisemblance par quartet

Les analyses ML peuvent être réalisées en évaluant successivement la vraisemblance de tous les ensembles de quatre taxons possibles, selon la méthode de "Quartet Puzzling". Cette méthode, proposée par le programme Puzzle 4.0 (Strimmer et von Haeseler, 1996), permet d'obtenir simultanément la topologie de l'arbre le plus vraisemblable et le support des branches. Sa fiabilité tient au fait que les relations entre chaque groupe de trois taxons sont évaluées successivement par rapport à tous les autres taxons. Selon le même principe, il est possible d'évaluer la topologie la plus vraisemblable pour quatre sous-ensembles de taxons pris isolément en employant la méthode de "Likelihood mapping" (Strimmer et von Haeseler, 1997) qui effectue des permutations autour des trois topologies possibles.

b) Mesure de la qualité globale de l'arbre

La qualité de l'arbre dépend en réalité de la qualité du jeu de données. Elle peut être évaluée par le calcul d'une série d'indices :

- *Indice de cohérence* ou "consistency index" (CI) : cet indice reflète la fraction totale des changements attribuables à la variation observée (Kluge et Farris, 1969). Il est égal au nombre minimum de pas nécessaires pour relier les données, divisé par le nombre de pas de l'arbre le plus court. Il décroît exponentiellement de 1 vers 0 (jamais atteint) et est inversement corrélé au nombre de taxons. Plus cet indice est proche de 1, plus la cohérence des données est bonne.

- *Indice d'homoplasie* ou "homoplasy index" (HI) : cet indice reflète la proportion des changements attribuables à l'homoplasie (Kluge et Farris, 1969). Il vaut (1-CI), et est redondant avec l'indice CI.

- *Indice de cohérence normé* ou "rescaled consistency index" (RC ; Farris, 1989) : cet indice est calibré de façon à varier entre 0 (homoplasie maximale) et 1 (pas d'homoplasie).

- *Indice de rétention* ou "retention index" (RI) : cet indice représente le taux de non saturation des positions. C'est le complément du coefficient de distorsion d (Farris, 1973) qui représente le rapport entre homoplasie réelle et homoplasie maximale. Il se calcule en divisant l'écart entre longueur maximale et longueur observée par l'écart entre longueur maximale et longueur minimale de l'arbre. Il décroît linéairement de 1 (non saturation) vers 0 (saturation).

Ces indices permettent de comparer la qualité d'arbres construits à partir de jeux de données identiques ou voisins, mais ont peu d'intérêt pratique car on ne peut pas définir de valeur seuil acceptable.

Test de permutation aléatoire de la matrice

Il est aussi possible de tester la qualité des données en comparant le signal phylogénétique qu'elles contiennent à un signal aléatoire, par un test de permutation aléatoire.

L'hypothèse nulle est que les données ne sont pas plus structurées que si elles ne contenaient aucune information phylogénétique (absence de corrélation entre caractères). On permute aléatoirement les éléments de la matrice de façon à générer des matrices ayant la même composition en bases que la matrice de départ ; puis on évalue la proportion de ces matrices ayant un meilleur indice de qualité que la matrice de départ. L'hypothèse nulle est rejetée si cette proportion, appelée PTP ("permutation tail probability"), est inférieure à 0,05. Les indices de qualité utilisés peuvent être le nombre de pas de l'arbre le plus parcimonieux, l'indice de distorsion ou le nombre d'incompatibilités entre paires de caractères.

Ce test n'est pas très puissant. Il permet seulement d'identifier les très mauvais jeux de données, et ne localise pas les régions déviantes.

c) Tests d'arbres alternatifs

Il est possible de savoir si une phylogénie donnée est significativement plus parcimonieuse (nombre de pas plus petit) ou plus vraisemblable ($-\ln L$ plus grand) que des phylogénies alternatives. Le test paramétrique de Kishino-Hasegawa (1989) utilise la moyenne et la variance de la différence de longueur (ou de $-\ln L$) entre arbres et entre sites, pour déterminer si les arbres en compétition sont significativement différents. L'équivalent non-paramétrique de ce test est le test de Templeton (1983) qui compare le nombre de caractères soutenant chacun des arbres à une distribution binomiale.

Presque tous ces types de tests ont été réalisés, au moyen du logiciel PAUP 4b3a (Swofford, 2000). Des bootstraps ont été effectués pour les analyses de Neighbor Joining et de Maximum de

Parcimonie. Les différents indices qualitatifs (CI, RC et RI) ont été calculés, et les arbres obtenus sous différentes conditions ou par différentes méthodes ont été comparés par des tests de Kishino-Hasegawa (1989).

6) Datation

a) Principe d'horloge moléculaire

On pourrait supposer que l'évolution des caractères moléculaires se fait à vitesse constante et identique pour toutes les espèces. Ceci permettrait le calibrage d'une horloge universelle pour la datation des événements de spéciation. Malheureusement, la vitesse d'évolution n'est constante ni entre taxons éloignés, ni entre taxons proches (pas même entre deux espèces d'un même genre d'après Fieldhouse *et al.*, 1997), ni entre molécules ou portions d'ADN. D'un gène à l'autre, les contraintes qui pèsent sur la molécule qui en résulte peuvent être très variables et au sein d'un gène donné, certaines régions sont plus contraintes que d'autres. L'horloge moléculaire et la neutralité des substitutions ne sont pas plus respectés par l'ADN mitochondrial que par l'ADN nucléaire.

D'un taxon à l'autre, différents facteurs peuvent influencer sur le taux d'évolution moléculaire. Chez les Mammifères, il semble que le taux de substitution pour les positions quatre fois dégénérées de gènes nucléaires et mitochondriaux diminue avec la taille corporelle et le temps de génération (Bromhan *et al.*, 1996). D'autres auteurs ont mis en évidence une corrélation entre taux de métabolisme et taux de substitution qui les a conduits à proposer l'existence d'une horloge physiologique. Cette corrélation pourrait être liée à l'abondance de radicaux libres d'oxygène qui sont mutagènes et à la grande vitesse de renouvellement des mitochondries chez les espèces à métabolisme élevé (Martin et Palumbi, 1993). Le nombre de réplifications dans les cellules germinales par unité de temps, la température corporelle et le taux de GC pourraient aussi intervenir en agissant sur le taux de métabolisme (Li *et al.*, 1990).

La différence de taux de substitution entre ADN nucléaire et ADN mitochondrial s'explique par le fait que l'ADN-polymérase mitochondriale corrige moins bien les erreurs, qu'il n'y a pas de système de réparation de l'ADN dans les mitochondries et que des radicaux libres y sont présents (Reyes *et al.*, 1998). Les différences entre gènes mitochondriaux pourraient s'expliquer par le temps d'exposition du brin non recopié et donc non apparié lors de la répllication (le brin H), du fait qu'à l'état simple brin, l'ADN est plus sensible aux radicaux libres. On note en effet une corrélation entre distance par rapport à l'origine de répllication et taux de substitution, modulée par les contraintes portant sur les différents gènes (Reyes *et al.*, 1998).

Toutefois, il est possible de calibrer une horloge moléculaire locale (pour un taxon et une molécule donnés) afin de dater des radiations impliquées dans des scénarios évolutifs (e.g. Voelker, 1999). Cette approche peut être abordée selon différentes méthodes qui contribuent à repérer et corriger les biais éventuels (Martin *et al.*, 1992). Il est notamment nécessaire de corriger les distances moléculaires en fonction des mécanismes évolutifs impliqués.

b) Linéarisation des arbres

Pour toute datation d'événement de spéciation, il convient de se baser sur un arbre linéarisé, c'est-à-dire un arbre duquel auront été supprimées les taxons n'évoluant pas à la même vitesse que les autres. La détection des espèces ou groupes d'espèces ne respectant pas l'horloge moléculaire locale peut se faire selon plusieurs méthodes.

Deux méthodes ont été proposées par Takezaki et collaborateurs (1995). Elles reposent sur la comparaison de longueurs de branches calculées selon la méthode des moindres carrés ("least-square method").

- Le "Relative Rate Test" compare les distances reliant deux espèces (ou groupes d'espèces) à un groupe externe et détermine si l'évolution respecte une horloge moléculaire pour les lignées concernées. Les lignées sont considérées deux par deux, en commençant par les nœuds terminaux de l'arbre de façon à éliminer successivement les groupements divergents. L'horloge est rejetée si la probabilité que les longueurs de branches soient identiques est inférieure à 0,05 ; mais, d'après Voelker (1999), il faut que les vitesses diffèrent fortement pour que ce test rejette l'hypothèse nulle.

- Le test des longueurs de branches compare la longueur de branche totale entre la racine et un taxon à la longueur moyenne pour l'ensemble des taxons.

D'autres méthodes prennent en compte simultanément l'ensemble des données :

- Le test de Kishino-Hasegawa (1989) permet de comparer les valeurs de likelihood sous l'hypothèse d'un taux d'évolution constant ou non. Comme la fluctuation statistique est importante, il est alors conseillé de faire un bootstrap, ce qui est très long (cf. Goldman, 1993).

- Un autre test compare les résidus du moindre carré de Rzhetsky et Nei (1993) en faisant ces mêmes hypothèses. C'est une version remaniée du test de Felsenstein (1984 et 1988) qui supposait indûment que les distances sont indépendantes.

c) Datation

La datation ne peut être opérée qu'après calibrage de l'horloge moléculaire, c'est-à-dire calcul du taux de substitution (r) équivalent au nombre de substitutions moyen par million d'années. Ce calcul nécessite la connaissance indépendante du temps écoulé depuis différents événements de spéciation, de préférence inféré de données fossiles. Le taux de substitution est calculé par la formule : $r = D / 2T$ où D est la distance génétique entre les lignées qui ont divergé et T est le temps écoulé depuis la divergence. La datation des nœuds de l'arbre se fait ensuite au moyen de la formule : $T = D / 2r$.

L'estimation de D se fait par un calcul de distances génétiques entre lignées, par exemple à l'aide du programme PhilTest 2.0 - Kumar, 1996). Ces distances doivent être corrigées par une formule de correction des distances, de façon à compenser la saturation des substitutions. Si les transitions apparaissent trop saturées, elles peuvent être éliminées. Il est également possible d'utiliser directement les longueurs de branches d'un arbre construit sous l'hypothèse de l'existence d'une horloge moléculaire, à condition que les données respectent cette hypothèse (Voelker, 1999).

En général, on utilise une seule calibration pour une phylogénie donnée, même si on a mis en évidence des différences de vitesse au sein de l'arbre. Ces variations de vitesse sont souvent limitées pour des taxons proches et peuvent résulter d'un biais dans l'échantillonnage plutôt que d'une différence effective (Voelker, 1999). Cependant, on peut aussi moduler la valeur du taux de substitution le long de l'arbre phylogénétique en définissant des horloges locales (Kishino et Hasegawa, 1990).

Une datation des principaux événements de divergence a été envisagée, après vérification de la linéarité de l'arbre de départ (le plus souvent à l'aide du logiciel RRTree - Robinson-Rechavi et Huchon, 2000). Lorsque les données fossiles le permettaient, au moins deux datations indépendantes ont été effectuées, et le taux moyen de substitution a été utilisé pour calibrer une horloge propre aux taxons inclus dans l'arbre.

7) Reconstruction de scénarios évolutifs

Lorsqu'on dispose d'un arbre phylogénétique bien résolu et intégrant la plupart des espèces d'un taxon, il est possible de retracer l'évolution d'un caractère sur cette phylogénie, c'est-à-dire de rechercher l'état ancestral de ce caractère à chaque nœud. Le caractère peut avoir été utilisé pour construire l'arbre phylogénétique sans que cela entraîne de circularité, puisque le principe de la

reconstruction phylogénétique est indépendant des hypothèses évolutives à tester par la suite (Deleporte, 1993). Il est d'ailleurs recommandé d'utiliser tous les caractères homologues et héréditaires disponibles pour obtenir une phylogénie aussi fiable que possible. En revanche, les caractères dont l'homologie est incertaine ou qui pourraient être découpés plus finement en une série de caractères héréditaires ne doivent pas être inclus dans l'analyse initiale (Grandcolas *et al.*, 2001). Certains traits écologiques, de même que les caractères définis par rapport à leur aspect fonctionnel, entrent dans cette catégorie.

La reconstitution des états d'un caractère peut être faite selon le principe de parcimonie, à l'aide du programme MacClade (Maddison et Maddison, 1996), ou selon le principe du maximum de vraisemblance, à l'aide du programme Discrete (Pagel, 1999a). La parcimonie tente de minimiser le nombre d'événements évolutifs, et considère l'état ancestral estimé à chaque nœud comme juste. Lorsque le degré d'incertitude sur l'état du caractère est important, elle ne donne aucun résultat. L'utilisation de cette méthode implique trois hypothèses : les changements sont équiprobables sur toutes les branches, les probabilités de gain et de perte d'états du caractère sont égales, et le taux d'évolution est relativement lent (Omland, 1999). La méthode de vraisemblance présente l'avantage de calculer une probabilité que le caractère soit dans un état ou un autre. Le calcul des probabilités se fait en fonction des états du caractère chez les taxons terminaux, du taux d'évolution du caractère, et des longueurs de branches (Cunningham *et al.*, 1998). Il repose sur trois hypothèses : la probabilité de changement à un moment donné dépend seulement de l'état du caractère à ce moment là, et non des états précédents ; les transitions le long de chaque branche sont indépendantes des changements se produisant en d'autres endroits de l'arbre ; les taux de changements sont constants au cours du temps et d'une branche à l'autre (Cunningham *et al.*, 1998).

Ainsi, la reconstitution de scénarios évolutifs par la méthode MP peut être trompeuse quand les changements évolutifs sont rapides et que les probabilités de changement et de réversion ne sont pas égales. La méthode ML présente l'avantage d'intégrer l'incertitude par le calcul d'une probabilité, mais ne doit pas être appliquée si les longueurs de branches ne sont pas proportionnelles au temps et si l'évolution n'est pas régulière. Ces deux dernières conditions peuvent être testées, au moyen du programme Discrete (Pagel, 1999a, b).

Nous avons reconstruit des scénarios évolutifs pour des caractères écologiques et morphologiques à l'aide du programme MacClade, selon le principe de parcimonie. Cette méthode a été choisie parce que nous doutions de la régularité des changements au cours du temps et que l'arbre initial ne pouvait être considéré comme linéaire.

V - ANALYSES PHYLOGEOGRAPHIQUES ET BIOGEOGRAPHIQUES

La phylogéographie recherche les principes et processus évolutifs qui gouvernent la distribution géographique des lignées, aux niveaux intra et inter-spécifiques (Avice *et al.*, 1987). Les patterns phylogéographiques obtenus indépendamment les uns des autres peuvent être comparés entre eux, ce qui relève du domaine de la phylogéographie comparée. Ils peuvent aussi contribuer à l'étude des relations entre aires biogéographiques, par l'utilisation de méthodes propres à la biogéographie.

La biogéographie historique se divise en trois courants principaux : la biogéographie cladistique, la panbiogéographie et l'analyse parcimonieuse de l'endémicité (PAE). La panbiogéographie met l'accent sur la dimension spatiale et géographique de la diversité biologique (Craw *et al.*, 1999). Elle repose à la base sur une cartographie descriptive, mais peut intégrer une dimension phylogénétique qui la rapproche alors de la biogéographie cladistique. La PAE a pour but d'identifier les aires d'endémicité et de définir les affinités hiérarchiques entre zones biogéographiques via l'analyse des similarités entre leurs peuplements (Rosen, 1985). Elle ne prend pas en compte la phylogenèse. Seules les méthodes de biogéographie cladistique utilisent directement des informations issues de la phylogéographie de différentes lignées, et seront donc employées dans ce travail.

1) Phylogéographie

a) Construction d'arbres des haplotypes

La phylogéographie combine des informations phylogénétiques et géographiques. Au niveau supra-spécifique, les phylogénies peuvent être construites par les méthodes classiques décrites dans le chapitre précédent, car les différences entre taxons sont généralement nombreuses et bien définies. Au niveau infra-spécifique, ou pour des espèces proches, par contre, les substitutions sont peu nombreuses et les méthodes traditionnelles ont alors un faible pouvoir de résolution (Huelsensbeck et Hillis, 1993). La théorie de la coalescence prédit que les haplotypes ancestraux sont encore présents et qu'ils sont à l'origine de nombreux haplotypes dérivés, ce qui se traduit par un arbre présentant des multifurcations et non par un arbre dichotomique. Il est alors judicieux de construire un "arbre des haplotypes", c'est-à-dire un arbre généalogique reliant les haplotypes les plus proches entre eux, plutôt qu'un arbre phylogénétique (Templeton *et al.*, 1992 ; Crandall, 1996). Les arbres des haplotypes peuvent être construits à partir des distances génétiques entre haplotypes, à l'aide du programme MspNet (Excoffier, 1993). Pour chaque arbre alternatif, la somme (S) de toutes les longueurs de branches est calculée à

partir des distances estimées entre paires de taxons. L'arbre retenu selon le critère d'évolution minimale est celui qui a la plus faible valeur de S ; il est appelé arbre d'étendue minimale ou "Minimum Spanning Tree" (Prim, 1957).

Les arbres peuvent aussi être construits selon la méthode de parcimonie statistique (Templeton *et al.*, 1992), à l'aide du programme TCS (Clement *et al.*, 2000). Le programme calcule le nombre maximum de mutations autorisé pour que la connexion entre deux séquences soit considérée comme parcimonieuse avec une probabilité supérieure à 95%. Selon qu'il existe ou non des haplotypes dont la connexion ne peut se faire sans excéder cette valeur seuil, il génère un ou plusieurs réseaux parcimonieux d'haplotypes. La valeur seuil peut aussi être définie par l'utilisateur afin de considérer des connexions non parcimonieuses. Cette méthode suppose que les positions sont indépendantes et autorise que les mutations ne soient pas distribuées de façon homogène (Crandall, 1996). Elle permet de limiter les problèmes d'homoplasie car elle prend en compte la ressemblance entre séquences (nombre de positions identiques) pour calculer la probabilité de mutations multiples aux positions qui diffèrent. Elle traite également le cas de la recombinaison, mais ce problème ne devrait pas se poser pour des séquences mitochondriales. L'haplotype ancestral est désigné au moyen d'un algorithme qui affecte à chaque haplotype une probabilité que celui-ci soit ancestral, en se basant sur trois hypothèses : 1) les haplotypes intérieurs à l'arbre ont plus de chance d'être ancestraux que les haplotypes situés en périphérie, 2) la fréquence des haplotypes est un indicateur de leur probabilité d'être ancestraux, 3) les haplotypes intérieurs à l'arbre ont une probabilité d'autant plus élevée d'être ancestraux qu'ils sont plus fréquents ou reliés à des haplotypes fréquents (Castelloe et Templeton, 1994). Cette méthode doit être adaptée lorsque l'arbre présente des ambiguïtés. Elle est relativement robuste, mais en cas de sélection directionnelle favorisant un haplotype, elle risque de désigner cet haplotype comme ancestral (Crandall, 1996).

Comme dans le cas de reconstructions phylogénétiques, la méthode de parcimonie présente des avantages par rapport à celle basée sur les distances. Lorsque les distances entre haplotypes sont relativement importantes, il y a peu de chances que les nœuds correspondent justement aux haplotypes échantillonnés, or, seule la méthode de parcimonie permet qu'il y ait bifurcation au niveau d'haplotypes intermédiaires non observés. De plus, lorsque la distance par rapport aux groupes externes est importante, les critères utilisés pour enraciner l'arbre dans le programme TCS semblent plus pertinents que celui qui consiste à rechercher l'haplotype dont la distance au groupe externe est minimale.

Après superposition des arbres des haplotypes sur une carte de distribution, il est possible de repérer des unités géographiques et des échanges entre localités, puis d'en inférer des événements de vicariance. A l'échelle intra-spécifique, Avise et collaborateurs (1987) ont défini cinq catégories phylogéographiques en fonction de l'adéquation entre diversité génétique et distribution géographique (**tab. 2.2**). A chaque catégorie correspond une ou plusieurs interprétations évolutives. La catégorie V est difficile à distinguer des catégories III et IV, car il s'agit en fait d'un continuum.

Tableau 2.2: Description des cinq catégories phylogéographiques, selon Avise et collaborateurs (1987).

Catégorie	Diversité génétique	Distribution géographique	Circonstances évolutives	Remarques
I	discontinue	séparées	- barrière extrinsèque à long terme - flux génique limité et extinction des génotypes intermédiaires	courant
II	discontinue	superposées	- zone de contact récente ou secondaire - barrière intrinsèque entre espèces sœurs sympatriques	rare
III	continue	disjointes	- flux génique limité et absence de barrière géographique à long terme	
IV	continue	totalement recouvrantes	- flux génique important et absence de barrière géographique à long terme	
V	continue	imbriquées	- flux génique intermédiaire et absence de barrière géographique à long terme	

b) Reconstitution de l'histoire évolutive d'une lignée

L'histoire évolutive d'une lignée peut être appréhendée par une analyse des groupements imbriqués ou "**Nested Clade Analysis**" (Templeton *et al.*, 1987 ; Templeton, 1998), qui utilise trois sources d'information : la distribution géographique, la structure phylogénétique et la fréquence des allèles, en intégrant les principes tirés de la théorie de la coalescence. L'analyse se décompose en trois étapes :

- construction d'un arbre ou d'un réseau d'haplotypes reliant les haplotypes les plus proches entre eux, de préférence par la méthode de parcimonie statistique ;

- conversion de cette généalogie des haplotypes en une série de groupements emboîtés en suivant un algorithme d'emboîtement (Templeton *et al.*, 1987) et les règles complémentaires qui ont été définies

par la suite pour résoudre les cas de groupements ambigus (Templeton et Sing, 1993, Crandall, 1996) ;

- analyse de contingence emboîtée grâce au programme Geodis (Posada *et al.*, 2000). Un test de permutation permet d'évaluer l'hypothèse nulle qu'il n'y a pas d'association entre les groupements et les localisations géographiques, définies en termes de distances entre localités. Ces distances peuvent être soit des distances calculées directement à partir des coordonnées géographiques, soit des distances physiques (nombre et qualité des barrières potentielles). Pour chacun des groupements (ou clades) englobant des haplotypes distincts provenant de localités distinctes, le programme teste s'il existe une association géographique significative entre les haplotypes qui le composent au moyen d'un test de Chi-2. Il calcule les distances géographiques entre clades constituant le clade analysé (Dn) et entre clades de rang inférieur eux-mêmes emboîtés dans les clades constituant le clade analysé (Dc), et il évalue leur significativité. Il calcule également les distances moyennes entre clades internes (I) et clades terminaux (T), et teste si la différence entre les deux est significative. Les résultats significatifs sont ensuite interprétés à l'aide d'une clef d'interprétation (Annexe 5) de façon à déceler d'éventuels événements de fragmentation ou d'expansion de l'aire de distribution, à estimer le flux génique, et à déterminer si l'échantillonnage est suffisant.

c) Autres méthodes

La biologie des populations fournit également des outils utiles pour analyser les mouvements de populations. Malheureusement, du fait de la grande taille de l'aire géographique considérée et du petit nombre de spécimens disponibles et analysés par espèce et par localité, les statistiques employées en biologie des populations n'ont pas pu être utilisées, à l'exception de l'analyse de la distribution des substitutions entre séquences (ou "mismatch distribution") qui sera détaillée en temps utiles.

2) Phylogéographie comparée

La phylogéographie comparée consiste à comparer les patterns phylogéographiques de taxons ayant des distributions recouvrantes (exemples : Avise, 1992 ; Riddle, 1995 ; Moritz et Faith, 1998 ; Taberlet *et al.*, 1998) et à déterminer dans quelle mesure ces taxons ont évolué de concert. Plusieurs méthodes permettent d'évaluer la concordance phylogénétique et biogéographique entre taxons, notamment le **test de Kishino-Hasegawa** (1989) et le **test de bootstrap paramétrique** (Huelsenbeck *et al.*, 1996). Ces deux tests comparent le score MP ou ML associé à l'arbre phylogénétique optimal obtenu pour l'un des deux taxons à celui d'un arbre construit sous l'hypothèse nulle qu'il n'y a pas plus d'incompatibilité entre les deux arbres que celle qui peut être attribuée à l'incertitude liée à la

reconstruction phylogénétique. Pour le test de Kishino-Hasegawa, la distribution nulle est générée à partir des données, sous l'hypothèse qu'il n'y a pas de différence entre ces deux arbres. Pour le test de bootstrap paramétrique, elle est générée par simulation, indépendamment des données, et sous cette même hypothèse (Sullivan *et al.*, 2000).

Le **test de randomisation** de "reconciled tree maps" (Page, 1994a) permet de tester le degré de significativité de l'histoire partagée entre taxons estimé à partir du logiciel TreeMap (Page, 1995) qui compte le nombre d'événements de co-vicariance impliqués par la réconciliation des deux arbres. L'hypothèse nulle qu'il n'y a pas plus d'événements de co-vicariance que si les deux taxons avaient évolué indépendamment est évaluée par comparaison de paires d'arbres générés aléatoirement.

3) Biogéographie cladistique

Cette approche combine les informations phylogénétiques issues de différents taxons, dans le but de retracer l'histoire dominante du peuplement et d'en inférer les relations entre aires géographiques. Elle requiert la construction préalable de cladogrammes dichotomiques définissant les relations phylogénétiques entre taxons, une connaissance précise de la distribution actuelle des différentes unités taxinomiques, et la définition d'aires d'endémicité.

a) Définition des relations entre aires

Méthodes permettant de construire un arbre global des aires

Certaines méthodes visent à combiner l'information génétique issue de reconstructions phylogénétiques effectuées pour des lignées indépendantes de façon à générer un arbre global des aires. La première étape consiste à associer un arbre des aires à chacune des phylogénies. Si chaque taxon est endémique à une localité ou région donnée, les noms des taxons sont simplement remplacés par ceux des localités de collectes ou des régions de répartition. Dans le cas contraire, il est nécessaire de définir des postulats pour traiter des problèmes de taxons manquants, de taxons à large distribution et de taxons à distributions redondantes (Nelson et Platnick, 1981 ; Wiley, 1987 ; Zandee et Roos, 1987). La seconde étape vise à extraire l'information commune portée par les cladogrammes des aires ainsi créés. Il existe deux méthodes qui diffèrent à la fois dans les postulats et dans la méthode de combinaison des cladogrammes.

La "**component analysis**" (Nelson et Platnick, 1981 ; Page, 1990, 1994b) exprime l'information partagée par les arbres sous forme d'un arbre consensus construit selon la règle de la majorité ou par élimination des informations discordantes. Cette méthode est dite *a priori* car elle autorise la

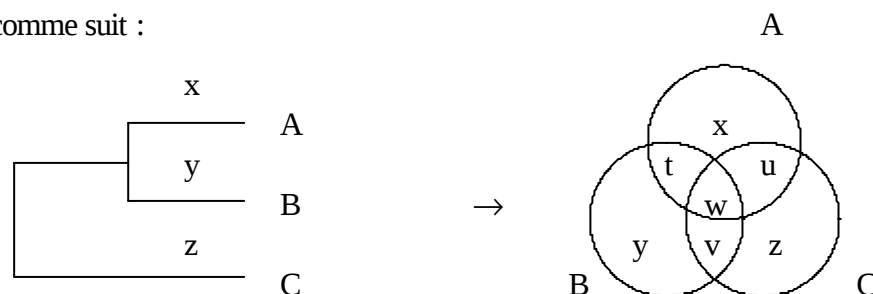
modification préalable de l'information contenue dans chaque cladogramme, de sorte que les relations historiques préétablies tiennent compte des processus supposés (Van Veller *et al.*, 2000). L'hypothèse de départ étant que tous les événements sont le produit de la vicariance, les méthodes *a priori* peuvent donner des résultats erronés si cette hypothèse est violée.

L'analyse de **parcimonie de Brooks** (Brooks, 1981 ; Wiley, 1988) utilise la parcimonie de Wagner pour construire un arbre global selon une méthode équivalente à celle du 'super-arbre'. Cette méthode est dite *a posteriori* car l'information contenue dans chaque cladogramme est incorporée sans altération dans l'analyse et les incompatibilités ne sont expliquées qu'après analyse (Van Veller *et al.*, 2000). Une version améliorée de la méthode (Brooks, 1990 ; Brooks *et al.*, 2001) autorise la violation de l'hypothèse nulle de pure vicariance.

Comme ces méthodes sont basées principalement sur des observations (des patterns phylogénétiques), la qualité et la quantité des modèles biologiques choisis comme représentants des peuplements locaux détermine la qualité du résultat de l'analyse. Dans le cadre de notre étude, le nombre de modèles et la concordance géographique dans la provenance des échantillons ont été jugés insuffisants pour pouvoir appliquer ces méthodes.

Diagrammes de Venn

Les méthodes exposées ci-dessus sont qualitatives, elles ne prennent pas en compte la diversité génétique contenue dans chaque phylogénie. Afin de prendre cette dimension en compte, Faith (1992) a proposé un indice de diversité phylogénétique (PD) qui mesure la diversité comprise dans des combinaison d'aires. Cet indice peut être sommé pour plusieurs phylogénies indépendantes (Oosterbroek, 1994), mais ne permet pas de distinguer la part relative de la diversité propre à chaque région et résultant du cumul de plusieurs régions. Aussi, Moritz et Faith (1998) ont-ils proposé d'affiner la méthode en construisant un diagramme de Venn. La diversité génétique intra- et inter-régionale est mesurée sur chaque arbre phylogénétique (par exemple un arbre construit selon la méthode NJ ; Moritz et Faith, 1998), à partir des longueurs des branches liant les haplotypes internes à chaque région et les haplotypes présents dans chaque combinaison de régions possible. Sa traduction en un diagramme de Venn s'effectue comme suit :



Les diagrammes de Venn correspondant à chacun des modèles sont ensuite sommés sous forme d'un diagramme global qui peut être traduit en un cladogramme des aires en appliquant la méthode MP. Selon la valeur des indices t , u , v et w , plusieurs cas de figure se présentent :

- $t \approx u \approx v$: pas d'histoire commune de vicariance ;
- $w \gg t, u$ et v : pas de relations hiérarchiques entre aires ;
- $v \gg t$ ou u : schémas congruents et existence de relations hiérarchiques.

Le principe de cette méthode est tel qu'une association entre régions apparaissant plusieurs fois dans un arbre phylogénétique pourra être aussi bien supportée qu'une association soutenue par un grand nombre de synapomorphies et n'apparaissant qu'une seule fois. Ceci tend à donner du poids à un événement qui se répète au cours du temps.

Un des inconvénients de cette méthode est qu'elle implique que les unités de changement aient la même valeur dans toutes les phylogénies. En théorie, seuls des modèles respectant la même horloge moléculaire et échantillonnés selon un plan d'échantillonnage équivalent devraient être combinés. De plus, les résultats sont influencés par la qualité et la quantité des modèles biologiques choisis. Cependant, la méthode est applicable même pour un petit nombre de modèles (e.g. Smith *et al.*, 2000), à condition de se référer aux phylogénies de base pour visualiser la contribution apportée par chacun d'eux (Moritz et Faith, 1998).

Méthode de Hovenkamp

Un des problèmes majeurs des méthodes biogéographiques est la définition des aires d'endémicité pour les analyses. Hovenkamp (1997) propose de circonvenir à ce problème en ne prédéfinissant pas les aires en fonction des espèces qui les occupent, mais en fonction de critères géographiques ou selon un quadrillage arbitraire. Les aires d'endémicité, si elles existent, émergeront de l'analyse sous forme de groupements de localités. L'auteur propose de rechercher les événements de vicariance communs à plusieurs phylogénies et d'en reconstituer la séquence temporelle. Sa méthode se décompose en quatre étapes : 1) identification des événements de vicariance sur des phylogénies indépendantes après remplacement des taxons par la liste des unités géographiques qu'ils occupent ; 2) extraction de la (ou des) séquence(s) temporelle(s) des événements pour chaque cladogramme, chaque séquence ne pouvant décrire que des événements se produisant sur des branches contiguës ; 3) recherche des événements de vicariance confirmés par plusieurs phylogénies, les aires séparées devant être concordantes mais pas nécessairement identiques (les autres sont éliminés) ; 4) reconstruction de la séquence temporelle globale pour les événements retenus. Cette méthode, contrairement aux

précédentes, ne traite pas les aires comme si elles étaient l'équivalent des taxons en phylogénie, ce qui est une démarche erronée. En effet, les aires n'évoluent vraisemblablement pas de façon purement divergente du fait que les barrières qui les séparent sont labiles. Les relations entre aires ne devraient donc pas être dichotomiques, mais réticulées.

b) Test de compatibilité entre un arbre théorique des aires et un arbre observé

Lorsqu'on dispose d'un arbre théorique des aires (obtenu par les analyses sus-mentionnées, issu d'une PAE, ou inféré d'une séquence d'événements géologiques ayant affecté la région), il est possible d'estimer la cohérence entre cet arbre et la phylogénie d'une lignée, dans le but d'expliquer l'histoire évolutive de cette lignée. Les tests de Kishino-Hasegawa (1989) et de bootstrap paramétrique (Huelsenbeck *et al.*, 1996) décrits précédemment peuvent être employés à cet effet. L'arbre phylogénétique obtenu pour la lignée est comparé à un arbre contraint, construit de façon à coïncider avec l'arbre des aires.

Il existe deux autres méthodes qui explicitent les différences entre arbres : **"reconciled trees"** et **"maximum co-speciation method"** (programme Component ; Page, 1993). Ces méthodes, qui font partie de la catégorie *"à priori"*, cherchent à maximiser la probabilité de co-divergence entre deux arbres de telle façon que le nombre de pertes, de duplications et de transferts horizontaux soit minimal. Elles ont été créées au départ pour l'étude des relations hôtes-parasites, mais peuvent être appliquées directement à l'étude de la vicariance en considérant que les hôtes sont les régions, les parasites les taxons étudiés, et en utilisant comme matrice de passage la matrice des distributions. Il est possible de tester la ressemblance entre les arbres, par génération aléatoire d'un grand nombre d'arbres phylogénétiques, puis évaluation de la probabilité que le nombre de changements observés soit inférieur au nombre de changements dans la situation aléatoire. L'emploi de ces méthodes est justifié dans le cas où on dispose d'un arbre théorique des aires qui décrit l'histoire réelle de la région. Par contre, si l'arbre théorique des aires est un arbre hypothétique basé sur l'histoire d'une partie du peuplement, il n'y a aucune raison de supposer *a priori* que l'histoire qu'il décrit devrait aussi être celle des taxons pris isolément. Dans ce cas, la comparaison des arbres théorique et observé n'est justifiée que si ceux-ci ne sont pas significativement différents.

c) Recherche des conditions géographiques ancestrales

D'autres méthodes visent à identifier les conditions géographiques ancestrales de spéciation pour une lignée donnée (aire ancestrale ou relation entre aires). Ces méthodes nécessitent la traduction de

chaque phylogénie et des distributions géographiques en cladogramme des aires, par remplacement de chacun des taxons du cladogramme original par toutes ses occurrences géographiques. Elles mettent en œuvre des modèles évolutifs explicites, donc les résultats dépendent de la qualité du modèle choisi (Ronquist, 1996).

La **méthode des aires ancestrales de Bremer** (1992) vise à identifier des centres d'origine, en se basant sur l'hypothèse que dans les centres d'origine, les espèces "ancestrales" étaient initialement présentes et les espèces nouvelles absentes. Les centres d'origine devraient donc correspondre aux régions où il y a plus de gain de nouvelles espèces que de pertes d'espèces ancestrales. Pour chaque aire, la probabilité que celle-ci ait appartenu à la distribution ancestrale est estimée par le rapport normalisé entre gain et perte d'espèces. Cette méthode, qui repose sur un principe douteux, est en outre sensible à la topologie de l'arbre de départ et à la définition de la distribution des espèces (cf. Voelker, 1999).

La **"dispersal-vicariance analysis"** (Ronquist, 1996), intégrée dans le programme DIVA, a pour but de reconstruire la distribution ancestrale d'un taxon. Le modèle évolutif utilisé privilégie la vicariance par rapport à la dispersion. Il attribue un coût neutre aux événements de vicariance, et une pénalité aux événements de colonisation et d'extinction. La méthode est inspirée de celle de Page (1990), mais n'implique aucun présupposé en ce qui concerne les patterns biogéographiques (pas de cladogramme général des aires prédéfini) et n'impose pas que la dispersion se traduise par un événement de spéciation.

La recherche des distributions ancestrales ne peut être réalisée que pour des taxons dont les distributions sont connues avec précision, elle n'a donc pas pu être opérée pour les petits mammifères étudiés dans ce travail. De plus, elle repose sur un postulat qui suppose que la vicariance prédomine sur la dispersion, ce qui n'est pas nécessairement le cas.

d) Test des modalités géographiques de spéciation

Chacune des approches mentionnées implique un modèle plus ou moins explicite d'évolution. La plupart sont basées sur le principe de vicariance donc considèrent que la spéciation se fait en allopatrie. Cependant, il peut aussi y avoir spéciation sympatrique.

Connaissant la phylogénie d'une lignée et l'aire de distribution des espèces qui la composent, il est possible d'aborder les modalités géographiques de spéciation. Le mode de spéciation prédominant peut être inféré de la courbe représentant le degré de sympatrie entre espèces en fonction de l'âge relatif des événements de spéciation (Barracough et Vogler, 2000). L'hypothèse de départ est que si la spéciation

se produit essentiellement en allopatrie, il ne devrait pas y avoir de chevauchement entre les aires de distribution des espèces ayant divergé récemment, mais le chevauchement devrait augmenter avec la durée de l'isolement des lignées, alors que si la spéciation se produit essentiellement en sympatrie, il devrait y avoir chevauchement total des aires de distribution entre espèces ayant divergé récemment, et le chevauchement devrait diminuer avec la durée de l'isolement des lignées.

Nous avons réalisé trois types d'approches :

1) Représentation d'arbres d'étendue minimale sur des cartes géographiques pour mettre en évidence d'éventuels mouvements de populations ou phénomènes de vicariance. Analyse de ces arbres par une analyse des groupements emboîtés. Test des scénarios évolutifs proposés à l'issue d'études biogéographiques par des analyses phylogéographiques indépendantes.

2) Recherche d'une histoire commune par la méthode de Hovenkamp (1997) et par construction d'un diagramme de Venn.

3) Test du mode géographique de spéciation.

CHAPITRE 3: ANALYSES PHYLOGENETIQUES

Dans le but de choisir des modèles adaptés aux analyses phylogéographiques, nous avons étudié la phylogénie des musaraignes africaines et d'un genre de muridés forestiers : *Hylomyscus*. Ces analyses préliminaires avaient pour principaux objectifs de vérifier la monophylie des espèces, d'évaluer la variabilité des marqueurs génétiques choisis et de calibrer une horloge moléculaire.

I - PHYLOGENIE DES MUSARAIGNES AFRICAINES (ARTICLE 1)

Article : Phylogénie et évolution des musaraignes africaines (Mammifères : Soricidae) inférées de séquences d'ARNr 16s.

Quéroutil, S., Hutterer, R., Barrière, P., Colyn, M., Kerbis Peterhans, J.C. et Verheyen, E. (2001) Phylogeny and evolution of African shrews (Mammalia: Soricidae) inferred from 16s rRNA sequences, *Mol. Phylogenet. Evol.*, **20** (2), 185-195.

Résumé : Parmi les musaraignes (Famille des Soricidae) on distingue deux Sous-Familles : celle des Soricinae, qui est surtout représentée dans l'hémisphère Nord, et celle des Crocidurinae, à laquelle appartiennent toutes les musaraignes africaines. Les hypothèses actuelles sur les relations phylétiques entre genres et espèces de musaraignes africaines étaient jusqu'à présent basées sur des études morpho-anatomiques, caryologiques, et allozymiques, et aucune phylogénie globale n'avait été proposée. Nous présentons ici une reconstruction phylogénétique basée sur des séquences mitochondriales partielles d'ARNr 16s, prenant en compte huit des neuf genres de musaraignes africaines. Pour les deux genres dont la distribution s'étend hors de l'Afrique, *Crocidura* et *Suncus*, des espèces Eurasiatiques ont été incluses en complément des espèces africaines de façon à étudier les échanges entre continents.

Les différentes reconstructions phylogénétiques effectuées ne convergent pas vers la reconnaissance d'un arbre unique, mais un certain nombre de groupements sont supportés quelle que soit la méthode utilisée (Weighted Neighbour Joining, maximum de vraisemblance par quartets, MP et NJ avec évaluation du support des branches par bootstrap). En accord avec la plupart des travaux antérieurs, il apparaît deux lignées au sein des Crocidurinae. La première inclut les genres *Myosorex* et *Congosorex*, et correspondrait aux Myosoricinés, auxquels il convient de donner soit le rang de tribu soit celui de Sous-Famille. La seconde lignée inclut les six genres restants. Au sein de cette lignée, au moins un des genres,

Sylvisorex, serait polyphylétique, ce qui confirme les conclusions d'autres auteurs. A l'inverse, le genre *Crocidura*, dont la monophylie a souvent été contestée, apparaît monophylétique, même lorsqu'un grand nombre d'espèces (non présentées dans cet article) sont ajoutées à l'analyse. Il semble que le genre *Suncus*, que certains auteurs supposent d'origine asiatique, serait en fait d'origine africaine. Les trois genres restants, *Paracrocidura*, *Ruwenzorisorex* et *Scutisorex*, seraient dérivés respectivement des genres *Crocidura*, *Suncus* et *Sylvisorex*, auxquels ils confèrent un statut paraphylétique. Leur reconnaissance en tant que genre peut donc être contestée du point de vue du critère de stricte monophylie. Elle peut cependant être acceptée en raison de caractères morpho-anatomiques nettement déviants (cf. "classification phylogénétique", dans le glossaire).

Les relations phylétiques entre espèces de différents continents révèlent des échanges multiples entre les faunes asiatiques et africaines. Les datations obtenues, après calibrage minutieux d'une horloge moléculaire pour la séquence et le taxon étudiés, coïncident avec celles qui ont été proposées pour des événements de migration similaires au sein d'autres taxons (bovidés, chiroptères, rongeurs). Ainsi, il y aurait une certaine synchronie entre les événements tectoniques déterminant la connectivité entre Afrique et Eurasie, la continuité supposée de l'habitat forestier ou savanicole, et la dispersion des taxons de mêmes caractéristiques écologiques. Des ancêtres du genre *Sylvisorex*, actuellement forestier, auraient migré en période de transgression forestière, tandis que des membres du genre *Suncus*, actuellement savanicole, auraient migré en période de forte extension savanicole. Ce résultat laisse supposer que les préférences écologiques de ces genres n'auraient pas changé.

Cette étude nous a permis de choisir deux espèces pouvant se prêter à une analyse phylogéographique : *Sylvisorex johnstoni* et *S. ollula*. En effet, bien qu'un seul spécimen par espèce soit présenté dans l'article, plusieurs spécimens d'origines différentes ont été analysés afin de vérifier la monophylie des espèces, de préciser leur abondance et leur distribution géographique, et d'estimer la diversité génétique intra-spécifique. En outre, l'étude a permis de calibrer une horloge moléculaire pour les musaraignes africaines et de choisir des groupes externes pour les analyses.

Phylogeny and Evolution of African Shrews (Mammalia: Soricidae) Inferred from 16s rRNA Sequences

Sophie Quéroutil,* Rainer Hutterer,† Patrick Barrière,* Marc Colyn,*
Julian C. Kerbis Peterhans,‡§ and Erik Verheyen¶

*Laboratoire Ethologie, Evolution, et Ecologie, Station Biologique, CNRS-U.M.R. 6552, Université de Rennes I, F-35380 Paimpont, France; †Section of Mammals, Zoologisches Forschungsinstitut und Museum Alexander Koenig, Adenauerallee 160, D-53113 Bonn, Germany; ‡University College, Roosevelt University, 430 South Michigan Avenue, Chicago, Illinois 60605; §Field Museum of Natural History, Chicago, Illinois 60605-2496; and ¶Section of Biochemical Systematics and Taxonomy, Royal Institute for Natural Sciences, Vautier Street 29, B-1000 Brussels, Belgium

Received August 16, 2000; revised January 25, 2001

Current phylogenetic hypotheses on the African Crocidurinae (Soricidae) are based upon morpho-anatomical, karyological, and allozyme studies. The present study attempts to resolve the interrelationships among African Crocidurinae and their relationships to Eurasian Crocidurinae and to the subfamily Soricinae, on the basis of partial mitochondrial 16s rRNA sequences (549 bp). This is the first molecular study to include all but one of the nine currently recognized African shrew genera. In agreement with current views, two major lineages emerge. The first lineage includes *Myosorex* and *Congosorex* and supports the existence of a mysoricine taxon. The second lineage includes the six remaining genera. The genus *Sylvisorex* appears to be polyphyletic, whereas species of the controversial genus *Crociodura* are monophyletic. The genus *Suncus* presumably originated in Africa. The monospecific genera *Ruwenzorisorex* and *Scutisorex* and the two representatives of *Paracrociodura* cluster with species of other genera. Grouping patterns of species from different continents suggest that there have been multiple exchanges between Africa and Eurasia. The time estimates of these exchanges, inferred from two independent fossil-based calibrations of a molecular clock, coincide with the time estimates for migration events in other mammalian taxa. © 2001 Academic Press

Key Words: phylogeny; phylogeography; shrews; Crocidurinae; Mysoricinae; Africa; Eurasia.

INTRODUCTION

Shrews constitute a very speciose and ecologically diversified taxon, with 335 described species (Wolsan and Hutterer, 1998). They are cosmopolitan except for the Antarctic, Australia, and most of the Arctic and Oceania. The extant species are usually classified into

two subfamilies: Soricinae and Crocidurinae (Repenning, 1967; Reumer, 1998). Despite a recent suggestion to define the Soricini and Crocidurini as two tribes within a soricine subfamily (Hutterer, 1993), we will follow the currently used classification (Reumer, 1998). The Soricinae mostly inhabit the Northern hemisphere, whereas the majority of the Crocidurinae occur in Africa, where all species belong to this subfamily. The fossil record suggests that the Crocidurinae have evolved from a Eurasian soricine ancestor and colonized the tropical forests of Africa where they subsequently diversified (Butler, 1998).

To date, about 150 species of African shrews, allocated to nine crocidurine genera, have been described (Wolsan and Hutterer, 1998). Shrew species are notoriously difficult to identify. The morphological and cranial characters that are traditionally used, such as body size, coat color, vibrissae abundance and length, skull shape, and tooth number and shape, are often difficult to interpret. Consequently, there are taxonomic uncertainties at the genus and species levels, especially within the genus *Crociodura* (Hutterer, 1993).

Meester (1953) and Heim de Balsac and Lamotte (1956, 1957) have tried to resolve the relationships among the African genera on the basis of a set of external, cranial, and dental characters. They suggested the existence of two lineages that differed primarily in their tooth formulae. The first lineage presently includes *Myosorex* ($n = 13$ species), *Congosorex* ($n = 2$), and *Surdisorex* ($n = 2$). The second lineage includes *Crociodura* ($n = 107$), *Sylvisorex* ($n = 14$), *Suncus* ($n = 6$), *Paracrociodura* ($n = 3$), *Scutisorex* ($n = 1$), and *Ruwenzorisorex* ($n = 1$) (updated number of species per genus from Wolsan and Hutterer (1998)). Although the grouping of African shrews in two lineages is now generally accepted (e.g., Maddalena and Bronner, 1992), their taxonomic status remains de-



bated (Hutterer, 1993). The *Myosorex* lineage may constitute the subfamily Myosoricinae (or Crocidosoricinae for some authors) or a myosoricine tribe. Accordingly, the second lineage would be the subfamily Crocidurinae or a crocidurine tribe. To avoid confusion, we will call the whole subfamily the Crocidurinae (s.l.) and the second lineage, potentially raised to the subfamily level, the Crocidurinae (s.s.).

It is rather difficult to reconstruct the phylogeny of shrews. As the fossil record is very poor, reconstruction depends largely on the comparison of the extant species (Butler, 1998). A first synthesis of African shrew phyletic relationships was proposed by Heim de Balsac and Lamotte (1957; Fig. 1). A recent attempt to infer the evolutionary history of African shrews from the fossil record resulted in a hypothetical phylogenetic inference (Butler, 1998; Fig. 2). The study of mitochondrial DNA (mtDNA) sequences has improved our understanding of the relationships among shrews by resolving the phylogeny of *Sorex* (Fumagalli *et al.*, 1999) and some Asian *Crocidura* species (Ruedi *et al.*, 1998). This approach seems appropriate to test the morphology-based phylogenies of the African Crocidurinae. Although alternative methods such as enzyme electrophoresis and cytogenetics have been applied to African shrews (Maddalena, 1990a,b; Maddalena and Ruedi, 1994; Schlitter *et al.*, 1999), our study is the first that analyses mtDNA sequences to infer the phylogenetic relationships among all the extant African shrew genera (except *Surdisorex*).

The first objective of this study was to test the current evolutionary scenarios (Figs. 1 and 2) proposed for the African Crocidurinae (s.l.) by means of mitochondrial sequences. In particular, the existence of a myosoricine taxon and the monophyly of the genera were tested. The second purpose of this study was to trace the evolutionary history of the subfamily Crocidurinae in space and time. As two of the genera occurring in Africa also occur in Europe and Asia (i.e., *Crocidura* and *Suncus*), we have tested the monophyly of these genera at a worldwide scale in an attempt to track the

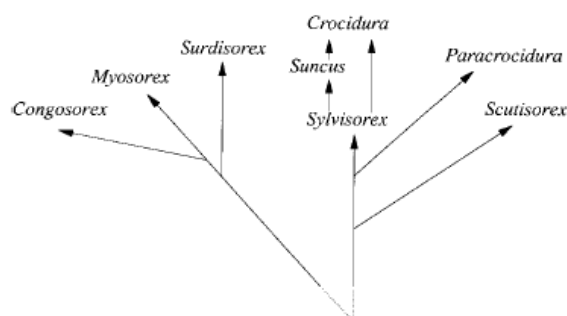


FIG. 1. Phyletic relationships of African shrew genera resulting from a detailed cranio-dental study of the species known by that time (from Heim de Balsac and Lamotte, 1957).

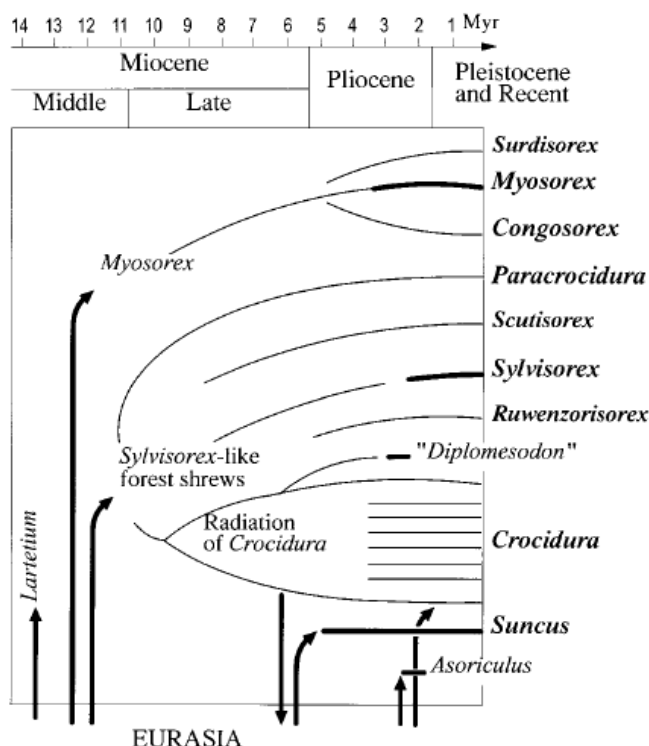


FIG. 2. Hypothesized history of African shrews based on anatomical studies of extant and fossil species (from Butler, 1998). Extant genera are in boldface. Heavy horizontal lines indicate the period for which there is a fossil record in Africa for a given genus. Arrows show entry into and exit from Africa.

migration events that may explain their current distribution ranges. These questions have already been addressed for the genus *Crocidura* on the basis of karyotypes (Maddalena and Ruedi, 1994) and allozyme data (Maddalena, 1990b). We tested their findings using different species and a different method. The use of mtDNA sequences allowed us to propose time estimates for the migration events between Africa and Eurasia by carefully calibrating a molecular clock with the available fossil record. Finally, the inferred timing and direction of transcontinental dispersion events were compared to migrations recorded in other mammalian orders (e.g., bovids: Thomas, 1984; murine rodents: Chevret, 1994; bats: Juste *et al.*, 1999).

MATERIALS AND METHODS

Biological Material

Over 100 specimens representing 47 species were analyzed, including two sequences taken from GenBank. To ascertain that each specimen was correctly assigned, a minimum of 2 specimens per species was used, whenever possible. After neighbor-joining phylogenetic analyses were conducted to ensure that the

TABLE 1
List of Specimens Analyzed

Genus	Species	Distribution	Ref. number	Country	Collector
Crocidurinae s.l.					
<i>Congosorex</i>	sp.	Central Africa	R22903	Congo	M. Colyn <i>et al.</i>
<i>Crocidura</i>	<i>fuliginosa</i>	Southeast Asia	IZEA 3747	Malaysia	M. Ruedi
<i>Crocidura</i>	<i>goliath</i>	Central Africa	GA70	Gabon	M. Colyn <i>et al.</i>
<i>Crocidura</i>	<i>lamottei</i>	West Africa	T-619	Senegal	L. Granjon*
<i>Crocidura</i>	<i>leucodon</i>	Europe and West Asia	IZEA 3923	Greece	P. Vogel
<i>Crocidura</i>	<i>malayana</i>	South Asia	IZEA 3550	Malaysia	M. Ruedi
<i>Crocidura</i>	<i>obscurior</i>	West Africa	R24051	Ivory Coast	M. Colyn <i>et al.</i>
<i>Crocidura</i>	<i>roosevelti</i>	Central and East Africa	NC147	C.A.R.	M. Colyn <i>et al.</i>
<i>Crocidura</i>	<i>russula</i>	Europe and North Africa	Cru 8.4	Switzerland	M. Ruedi
<i>Myosorex</i>	<i>geata</i>	East Africa	6315	Tanzania	W. Verheyen
<i>Paracrocidura</i>	<i>maxima</i>	East Africa	FMNH 157828	Uganda	J. C. Kerbls-Peterhans
<i>Paracrocidura</i>	<i>schoutedeni</i>	Central Africa	R16829	Congo	M. Colyn <i>et al.</i>
<i>Ruwenzorisorex</i>	<i>suncoides</i>	East Africa	FMNH 148939	Burundi	J. C. Kerbls-Peterhans
<i>Scutisorex</i>	<i>somereni</i>	Central Africa	817	D.R.C.	M. Colyn <i>et al.</i>
<i>Suncus</i>	<i>dayi</i>	South Asia	V-563	India	F. Catzefflis
<i>Suncus</i>	<i>etruscus</i>	Europe, Asia, and North Africa		France	R. Fons
<i>Suncus</i>	<i>infinitesimus</i>	Africa (except West)	NC450	C.A.R.	M. Colyn <i>et al.</i>
<i>Suncus</i>	<i>montanus</i>	South Asia	V-559	India	F. Catzefflis
<i>Suncus</i>	<i>murinus</i>	Asia + introduced in East Africa	IZEA 4433	Indonesia	M. Ruedi
<i>Suncus</i>	<i>remyi</i>	Central Africa	R22636	Congo	M. Colyn <i>et al.</i>
<i>Sylvisorex</i>	<i>johnstoni</i>	Central and East Africa	R16716	Congo	M. Colyn <i>et al.</i>
<i>Sylvisorex</i>	<i>cf. konganensis</i>	Central Africa	R16749	Congo	M. Colyn <i>et al.</i>
<i>Sylvisorex</i>	<i>megalura</i>	Tropical and Southern Africa	R22350	Congo	M. Colyn <i>et al.</i>
<i>Sylvisorex</i>	<i>ollula</i>	Central Africa	R12029	C.A.R.	M. Colyn <i>et al.</i>
Soricinae					
<i>Sorex</i>	<i>araneus</i>	Europe and Asia	T-1500	Switzerland	C. Vaucher*
<i>Sorex</i>	<i>palustris</i>	North America	Access No. SPU 97343		From Springer <i>et al.</i> (1997)
<i>Soriculus</i>	<i>fumidus</i>	East Asia	T-274	Taiwan	M. Harada*
Outgroups					
<i>Scalopus</i>	<i>aquaticus</i>	North America	Access No. AF069539		From Stanhope <i>et al.</i> (1998)
<i>Talpa</i>	<i>europa</i>	Europe	ZFMK 96285	Germany	R. Hutterer

Note. Country: Congo, Republic of the Congo; D.R.C., Democratic Republic of Congo; C.A.R., Central African Republic. Collector: Colyn *et al.*, field work directed by M. Colyn, collectors are either P. Barrière, M. Colyn, G. Ossessa, or O. Perpète; *, received from F. Catzefflis.

specimens clustered, 1 specimen was kept for each species.

All African genera except *Surdisorex* were represented. Four species each of the European and Asian representatives of the transcontinental genera *Crocidura* and *Suncus* were analyzed. Three representatives of the subfamily Soricinae (*Sorex palustris*, *S. araneus*, and *Soriculus fumidus*) were included. Two moles (*Scalopus aquaticus* and *Talpa europa*: Talpidae; Insectivora) were chosen as outgroups, for their suggested proximity to Soricidae (Stanhope *et al.*, 1998; Asher, 1999).

Finally, 29 species were selected for this study (Table 1). For reasons of clarity and homogeneity, only 8 of the 26 *Crocidura* species analyzed are considered here. We retained (1) the type species of the genus: *C. leucodon*; (2) three Eurasian species: *C. fuliginosa*, *C. malayana*, and *C. russula*; (3) three African species which had

been originally described as representative of different genera or subgenera: *C. goliath* (*Praesorex*), *C. lamottei* (*Afrosorex*), and *C. roosevelti* (*Heliosorex*); and (4) an African species with plesiomorphic characters: *C. obscurior* (Maddalena and Ruedi, 1994; where *C. obscurior* is referred to as *C. cf. bottegii*).

All original sequences and the alignment were submitted to GenBank, under Accession Nos. AF252518, AF253953, and AF274538–AF274562.

Laboratory Methods

We used the second half of the 16s rRNA gene, situated between the positions 1948 and 2489 of the mouse mitochondrial DNA (Bibb, 1981). This gene fragment was chosen because it seems to evolve at an appropriate rate to be informative at the intergeneric level (see Ferris *et al.*, 1983; Mindell and Honeycut, 1990). The cytochrome *b* gene could have been a good

marker also, but despite our efforts to sequence it following the protocol elaborated for *Sorex* shrews by Fumagalli *et al.* (1999), only 1 of 10 specimens could be reliably sequenced.

DNA was isolated from ethanol-preserved muscle or liver tissues by standard proteinase K dissolution, followed by phenol–chloroform extraction and ethanol precipitation (Sambrook *et al.*, 1989), or by the Chelex method (Walsh *et al.*, 1991). As no tissue was available for the genus *Scutisorex*, DNA was extracted from the skeleton of a desiccated specimen. This extraction was performed following the proteinase K and phenol–chloroform protocol adapted by Hassanin and Douzery (1999).

The primers used for amplification and sequencing were the 16S_r L primer, 5'-CGCCTGTTTAAACAAAACCAT-3' (Palumbi *et al.*, 1991) and a new primer named 16S-Hm, 5'-AGATCACGTAGGACTTTAAT-3'. The PCR consisted of 30 cycles: 60 s at 94°C, 60 s at 49°C, and 90 s at 72°C. The double-stranded PCR products were purified with an EasyPrep kit from Pharmacia. The samples were sequenced in both directions with an automated Alf-sequencer from Pharmacia and an Amersham cycle-sequencing kit, following the manufacturer's protocol. The sequencing reaction consisted of 30 cycles: 36 s at 94°C, 36 s at 52°C for the L primer or 55°C for the H primer, and 80 s at 72°C. When the sequences yielded unexpected results, fresh extractions were made to reanalyze the samples. In all cases, these controls confirmed the initial sequencing results.

Data Analyses

Sequence alignment. The sequences were aligned by CLUSTAL W version 1.5 (Thompson *et al.*, 1994) with the default settings. The obtained alignment was corrected by eye, on the basis of the secondary structure of the 16s gene fragment for *Mus* (De Rijk, 1995). Most of the ambiguities were situated in the large hypervariable loop. Following Wheeler and Honeycutt (1988), we excluded the 24 positions for which nucleotide homology was doubtful from the analyses (from positions 249 to 263 and 278 to 286 of the alignment). We also deleted the first six and last nine bases, which were missing in several sequences. The final alignment was 510 bp long.

Phylogenetic analyses. Phylogenetic inferences were made by neighbor-joining (NJ) and maximum-parsimony (MP) with PAUP version 4.02d (Swofford, 1997). The genetic distances (d) were corrected by the Tamura and Nei (1993) formula which allows differential substitution rates for transitions (TS) and transversions (TV) between purines and pyrimidines and substitution rate heterogeneity along the sequences. Heuristic searches for the most parsimonious trees were conducted with 10 replicates of random addition of taxa and a uniform weighting scheme. A stepmatrix was used to weight transversions

twice as much as transitions because transitions appeared to be partially saturated. Gaps were treated as missing.

The robustness of internal nodes was estimated by bootstrapping: 10,000 neighbor-joining and 1000 maximum-parsimony replicates (each with a single replication of random addition of taxa) were carried out.

Data were also analyzed with Weighbor (Bruno *et al.*, 2000). The weighted neighbor-joining (WNJ) algorithm reconstructs trees in such a way that the variance and covariance of the distances are taken into account for the choice of the sequences to be linked and for the calculation of the new distance matrix. Simulations show that it performs nearly as good as maximum-likelihood even in instances when long branches are present (Bruno *et al.*, 2000). Two options were defined: the number of variable sites (L) was set to 180 and the number of bases corrected for the nucleotide bias (b) was set to 3.85.

Maximum-likelihood (ML) analyses were performed with DNAML in PHYLIP 3.57c (Felsenstein, 1995) and the quartet puzzling (QP) method implemented in PUZZLE 4.0 (Strimmer and von Haeseler, 1996). The HKY general time-reversible model was chosen, because it allowed for transition bias and did not assume equal base frequencies (Hasegawa *et al.*, 1985). For DNAML analyses, four classes of substitutions were defined, with the number of changes estimated to 0, 1, 3, and 5 for, respectively, 80, 6, 8, and 6% of the positions. The transition/transversion ratio was set to 2. Ten replications with random addition of taxa were carried out. For the quartet puzzling method, the model of substitution used was HKY and the rate heterogeneity was modeled by a discrete gamma distribution (with four mutation rate classes) and by allowing invariable sites. Parameters were inferred from the data set.

We also performed a series of four-cluster likelihood analyses to evaluate the supports for distinct internal branches critical for the interpretation of the evolutionary pathways by use of the likelihood mapping approach performed by PUZZLE 4.0 (Strimmer and von Haeseler, 1996). This approach is a permutation method evaluating the likelihood supports for the three alternative topologies of multiple sets of four taxa to assess the overall phylogenetic signal in the data set. The test was carried out to clarify all possible quartet relationships for the polytomies of the QP tree and to evaluate the relationships among the *Crocidura* species. All settings were defined as for the QP method.

Monophyly of the genera. Kishino–Hasegawa tests (1989) were performed to evaluate the hypothesis of monophyly of the polyspecific genera, by use of PAUP 4.02d (Swofford, 1997). The length difference between constrained and unconstrained trees was tested under

the maximum-parsimony criterion. Settings were as previously described.

Divergence time estimates. To minimize the effect of saturation, we chose to rely on transversions only and to apply a Kimura two-parameter correction (Kimura, 1980). First, we tested whether the data were consistent with the molecular clock hypothesis. Two-cluster substitution rate homogeneity tests were performed for all the nodes to be dated, by use of Phyltest (Kumar, 1996). For each of the selected ingroups in the WNJ tree, the most closely related taxa were defined as outgroups. Differences were declared significant for $P < 0.05$. Second, we attempted to calibrate a shrew molecular clock. The limited fossil record provided only two values for dating the divergence events. First, the split between Soricinae and Crocidurinae (s.l.) is estimated to have occurred 20 million years (Myr) ago (Reumer, 1989, 1994). This calibration point has already been used to date divergences within the *Sorex* lineage (Fumagalli *et al.*, 1999). Second, although the oldest fossils of *Crocidura* date back to 3 Myr, the radiation of this genus is believed to have started in the late Miocene, at least 5–6 Myr ago (Maddalena, 1990a; Butler, 1998). These dates were used to bracket the substitution rate (r), by means of the formula $r = D/2T$, where D is the genetic distance between groups of taxa which have diverged and T is the time elapsed since the divergence event. Genetic distances between clusters of sequences were computed with Phyltest (Kumar, 1996).

RESULTS

DNA Sequence Variation

Total length after alignment is 549 bp. The addition of the outgroups in the dataset required the inclusion of four gaps in the ingroup sequences. Of the 510 nucleotide positions analyzed, 357 are constant and 120 are informative. The mean transition/transversion ratio is 1.35 for the whole of the ingroups and 1.55 for closely related species such as the *Crocidura* species. These low values suggest that transitions are partially saturated.

Tree Comparison

The shortest MP tree (803 steps) is not significantly shorter than the NJ and WNJ trees (both 808 steps; $P = 0.622$). The homoplasy index for the MP tree is 0.592. It is indicative of a high level of reversal and uncertainty which has to be compensated for by differential weighting and cautious data analysis, through different reconstruction methods and branch support evaluation. The WNJ tree reflects most of the information embedded in the NJ and the MP trees. It is less influenced by the violation of the molecular clock assumption than the NJ tree and is not dependent on arbitrary weighting nor on a heuristic search as is the

MP tree. The QP tree is less resolved than the other trees. It is also more reliable as it discards doubtful groupings. The WNJ and QP tree comparison illustrates the relative consistency between tree topologies obtained by different methods (Fig. 3). The clades supported by small P values by WNJ and ML ($P < 0.05$) are also supported by high NJ and MP bootstrap values and receive a high quartet puzzling support (Fig. 3).

Quartet likelihood mapping results are shown only in cases in which there is a high support for one of the three possible topologies (Fig. 4). This approach did not allow us to resolve the five-cluster basal polytomy found in the QP tree (Fig. 3).

Phylogenetic Analyses

All the reconstruction methods except the quartet puzzling algorithm detect two clades among the Crocidurinae (s.l.) (Fig. 3). The first clade includes the genera *Myosorex* and *Congosorex*. The second clade consists of the species of the remaining genera, i.e., *Crocidura*, *Paracrocidura*, *Ruwenzorisorex*, *Scutisorex*, *Suncus*, and *Sylvisorex*. The lack of basal resolution of the QP tree is compensated by the high likelihood mapping value obtained for the grouping of *Sylvisorex johnstoni* with the remaining members of the second clade (Fig. 4A).

Within the first clade (Fig. 3), the *Myosorex* and *Congosorex* species that we analyzed are closely related (Tamura–Nei distance = 4.7%). The two genera are associated with the three Soricinae, *Sorex palustris*, *S. araneus*, and *Soriculus fumidus*. This association is supported at 87% by quartet likelihood mapping (Fig. 4A), whereas it is not supported by the QP tree and receives a low bootstrap support, which falls below 50%, under the parsimony criterion.

Within the second clade (Fig. 3), there are high genetic distances among the representatives of the genus *Sylvisorex* ($\leq 15.1\%$ divergence). This genus appears to be polyphyletic (Fig. 3). *S. ollula* and *S. cf. konganensis* cluster ($d = 7.2\%$) and may be related to *S. johnstoni*, as the independent branching of this latter species is poorly supported (bootstrap values $< 50\%$ and QP support of 58%; Fig. 3). *S. megalura* stands apart from the other *Sylvisorex* species and clusters with *Suncus* species. Its closest relative among the studied species is *Suncus dayi* ($d = 8.3\%$).

The monospecific genus *Scutisorex* is closely related to *Sylvisorex ollula* and *Sylvisorex cf. konganensis* (mean distance: 7.6%). This association is supported by all reconstruction methods except MP.

For the genus *Suncus*, all the species analyzed have an intermediate position between representatives of the genus *Sylvisorex* (*S. megalura* excepted) and *Crocidura*. With regard to their geographic origin, the two African *Suncus* species cluster ($d = 6.6\%$), and two of the three Asian species (*S. murinus* and *S. montanus*) cluster ($d = 3.0\%$). The affiliations of the widely distributed species *S. etruscus* remain unresolved. Accord-

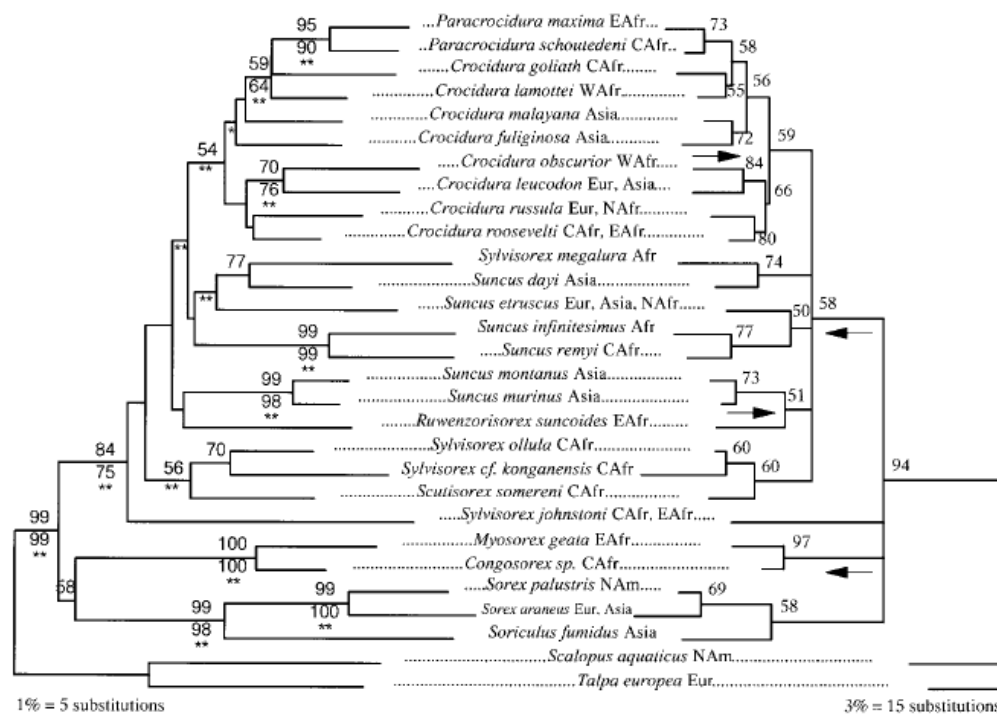


FIG. 3. (Left) Weighted neighbor-joining tree. Bootstrap supports resulting from 10,000 neighbor-joining and 1000 maximum-parsimony replicates are shown, respectively, above and below the branches. Maximum-likelihood support is indicated below the branches (* $P < 0.05$ and ** $P < 0.01$). (Right) Quartet puzzling tree with branch support. Distributions of the taxa: CAfr, EAfr, and NAfr, Central, East, and North Africa; NAm, North America; Eur, Europe; Asia, Asia. The region where the specimen was collected is listed first. Arrows pointing, respectively, to the left and to the right indicate potential migration events to and from Africa.

ing to the different tree-building methods it could be closer either to the Asian *S. dayi* (WNJ and ML trees) or to the two African species (QP tree).

The monospecific genus *Ruwenzorisorex* shows affinities with the Asian *Suncus* species (mean distance:

8.5%), but this relationship is not supported by bootstrapping nor by ML.

All *Crocidura* species tested are closely related ($\leq 8.9\%$ divergence). The Indian and European species do not form a separate clade relative to the African species. Actually, quartet likelihood mapping suggests that the Asian species are more closely related to the African species *C. goliath* and *C. lamottei* than to the mostly European species *C. leucodon* and *C. russula* (Fig. 4B). Furthermore, the European *C. leucodon* clusters with the West African *C. obscurior* with high bootstrap support (Fig. 3).

Whatever the method used, the two *Paracrocidura* species cluster ($d = 3.7\%$) within the *Crocidura* clade.

Monophyly of the Genera

The shortest MP tree (802 steps) was obtained under the constraint that the genus *Suncus* is monophyletic. This tree is not significantly shorter than the tree obtained by unconstrained parsimony analyses (803 steps; $P = 0.925$), but shows that the heuristic search failed in finding the shortest MP tree. When phylogenetic analyses are carried out under the constraint of monophyly of all the genera (821 steps) or monophyly of *Sylvisorex* (819 steps), the constrained trees are sig-

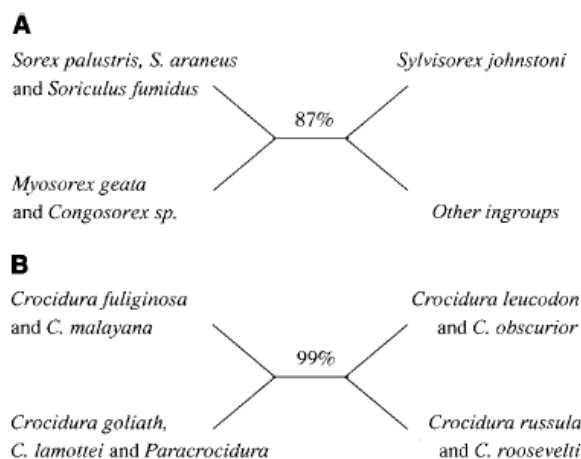


FIG. 4. Quartet likelihood mapping showing the most likely pairings.

TABLE 2
Substitution Rate Estimates

Taxon 1	Taxon 2	Mean distance (%)	SE (distance) (%)	Divergence (Myr)	<i>r</i> (%)	SE (<i>r</i>) (%)
Soricinae	Crocidurinae s.s.	7.53	1.09	20	0.188	0.027
First clade of <i>Crocidura</i> *	Other <i>Crocidura</i> + <i>Paracrocidura</i>	2.11	0.42	5	0.211	0.042
				6	0.176	0.035

* = *C. obscurior*, *C. russula*, *C. leucodon*, and *C. roosevelti*.

nificantly longer than the shortest tree ($P = 0.032$ and $P = 0.004$, respectively).

The monophyly of the genera *Crocidura* and *Paracrocidura* was not tested because, whatever the method of alignment and analysis used, all the specimens clustered with a high bootstrap and ML support for *Paracrocidura* and a lower support for *Crocidura*.

Divergence Time Estimates

No significant differences were found for the two cluster tests. The highest Z value, which was well below 1.96, was observed between the Soricinae and the Myosoricinae ($Z = 1.295$).

As the molecular clock hypothesis was not rejected, we computed between-cluster distances and estimated divergence times. The substitution rate calculated with the Soricinae/Crocidurinae (s.l.) split was intermediate between the rates based on the *Crocidura* radiation (Table 2). Therefore, we used the rate based on a 6-Myr-old radiation of the genus minus the standard error ($r = 0.141$) as the lower limit and the rate based on a 5-Myr-old radiation plus the standard error ($r = 0.253$) as the higher limit. The mean r value ($r = 0.197$) together with the minimum and the maximum r values were used to estimate the Soricinae/Myosoricinae split (about 19.1 Myr ago) and the appearance of each genus relative to its nearest sister species (Table

3). The most basal *Sylvisorex* species in our tree is *S. johnstoni*. We estimated that it would have diverged from the common crocidurine ancestor about 12.9 Myr ago. The first *Suncus* species would have evolved between 7.5 and 9.3 Myr ago and the genus *Paracrocidura* about 2.1 Myr ago.

DISCUSSION

Parts of the branching patterns that we obtained are poorly supported by bootstrapping and likelihood methods (Fig. 3). This can reflect either very short branches due to a polytomic radiation or a series of rapid consecutive cladogenetic events that cannot be resolved by the information embedded in the 16s gene fragment. The analysis of longer sequences could help to rule out the second possibility, but we preferred to obtain double-checked sequences for a greater number of specimens rather than multiple or longer sequences for fewer specimens. Actually, it has been emphasized that use of a great number of sequences improves the reliability of the phylogenetic trees by increasing the representativity of the data (Lecointre *et al.*, 1993) and reducing long-branch attraction (Hillis, 1996).

Phylogeny

Our first result is the occurrence of two major lineages in African Crocidurinae (s.l.).

TABLE 3
Divergence Time Estimates

Clade 1	Clade 2	Mean dist. (%)	SE (dist.) (%)	Divergence time (Myr)		
				Min	Mean	Max
Soricinae	"Myosoricinae"	7.52	1.17	14.9	19.1	26.7
<i>Sylvisorex johnstoni</i>	Other Crocidurinae s.s.	5.09	0.90	10.1	12.9	18.0
<i>Sylvisorex cf. konganensis</i> , <i>S. ollula</i> , and <i>Scutisorex somereni</i>	Other Crocidurinae except <i>S. johnstoni</i>	3.58	0.58	7.1	9.1	12.7
<i>Suncus murinus</i> , <i>S. montanus</i> , and <i>Ruwenzorisorex suncoides</i>	Other <i>Suncus</i> species, <i>Sylvisorex megalura</i> , <i>Crocidura</i> , and <i>Paracrocidura</i>	2.96	0.49	5.8	7.5	10.5
Other <i>Suncus</i> and <i>Sylvisorex megalura</i>	<i>Crocidura</i> and <i>Paracrocidura</i>	3.67	0.56	7.3	9.3	13.0
<i>Crocidura lamottel</i> and <i>C. goliath</i>	<i>Paracrocidura</i>	0.82	0.33	1.6	2.1	2.9

The first clade contains *Myosorex* and *Congosorex* which are in most instances grouped with the soricine genera *Sorex* and *Soriculus*. This clustering is in agreement with the observations of Heim de Balsac and Lamotte (1956) who reported that *Myosorex* shared characters with the Soricinae.

The apparent split within Crocidurinae (s.l.) is supported by morpho-anatomical characters (Figs. 1 and 2) and behavioral traits (Meester, 1953; Heim de Balsac and Lamotte, 1957). *Myosorex* and *Congosorex* possess third lower premolars that are lacking in Crocidurinae (s.s.) and share morphological adaptations to a fossorial life style.

Maddalena and Bronner (1992) suggested that *Myosorex* should be classified in a separate subfamily called Crocidosoricinae. As this name is restricted to extinct taxa, we prefer the name Myosoricinae, which is based on the extant *Myosorex* and has formal priority (cf. Wolsan and Hutterer, 1998). According to the study of Heim de Balsac and Lamotte (1957), the genus *Surdisorex*, which we were unable to analyze, may also belong to this taxon.

The second clade includes all the remaining genera, i.e., *Crocidura*, *Paracrocidura*, *Ruwenzorisorex*, *Scutisorex*, *Suncus*, and *Sylvisorex*. Within this group, which should be termed the Crocidurinae (s.s.), our results are mostly consistent with the morpho-anatomical phyletic reconstructions of Heim de Balsac and Lamotte (1957; Fig. 1) and Butler (1998; Fig. 2).

The main disagreement between our data and current opinions concerns the genus *Paracrocidura* (Fig. 3 as opposed to Figs. 1 and 2). The two studied species appear to be each other's closest relatives but occupy an unexpected position within the genus *Crocidura*, from which they seem to have evolved recently. The divergence time estimate of 2.1 Myr may be underestimated because it is based on transversions only and the TS/TV ratio is high among closely related taxa. An independent calculation based on all substitutions suggests that *Paracrocidura* actually diverged from *Crocidura* about 4 Myr ago. In any case, the relatively recent origin of *Paracrocidura* is in agreement with the fact that no fossils have been found for this genus (Butler, 1998), even if the probability of finding fossils is very low for forest species. The close relationship between *Crocidura* and *Paracrocidura* is only weakly supported anatomically: *Paracrocidura* has the same dental formula as *Crocidura*, but it differs from all the *Crocidura* species by a set of unique characters (Hutterer, 1986a).

Another striking difference between the obtained phylogeny and the current views concerns the origin of the genus *Suncus*. All the *Suncus* species examined are positioned between representatives of the *Crocidura* and *Sylvisorex* genera, as suggested by Heim de Balsac and Lamotte (1957) (Fig. 1). None of them clusters among the *Crocidura* species, as observed for *S. muri-*

nus in a morphological comparison (McLellan, 1994) and in a molecular study based on cytochrome *b* sequences (Motokawa *et al.*, 2000). Our results are inconsistent with those of Butler (1998) and Jenkins *et al.* (1998) who suggested that *Suncus* differentiated in Asia and ultimately migrated to Africa (Fig. 2). On the contrary, results indicate that *Suncus* derived from a *Sylvisorex*-like ancestor. Given that *Sylvisorex* is endemic to Africa, it is probable that *Suncus* originated in Africa. This interpretation is supported by previous studies (Meester, 1953; McLellan, 1994) and by the fossil record. No unmistakable *Suncus* fossil older than the Holocene has been discovered in Europe (Rzebik-Kowalska, 1998) nor in Asia (Storch *et al.*, 1998); whereas some 3 to 5 Myr old fossils have been found in East and South Africa (Butler, 1998). Our results suggest that the genus is much older, maybe as old as 9.3 Myr. The comparison between the obtained phylogenies (Fig. 3) and the divergence time estimates (Table 3) reveals a discrepancy among *Suncus* species. *Suncus murinus* and *S. montanus* have an ancestral position relative to the other *Suncus* species in the WNJ tree (but not in the QP tree) but have presumably diverged later (7.5 Myr ago versus 9.3 Myr ago).

Results for the remaining polyspecific genera agree with earlier syntheses of African shrew genera relationships, but shed new light on interspecific relationships.

This study suggests that *Crocidura* is monophyletic. The omitted species also cluster with their relatives (data not shown). Hence, our results do not support the hypothesis that *Crocidura* is not a natural taxonomic unit (Butler *et al.*, 1989; Motokawa *et al.*, 2000) nor that it includes several independent lineages (Ruedi, 1998) which could be paraphyletically derived from *Sylvisorex* and *Suncus* (Heim de Balsac and Lamotte, 1957; Fig. 1). We do not distinguish a mostly Palaearctic from an African clade as suggested by an allozyme study (Maddalena, 1990b). The European species do not cluster with the Asian species, and the African species are related both to European and to Asian species (Figs. 3 and 4B). The poor resolution found within the genus and the short basal branches reveal a rapid radiation of the genus. It is presumed to have occurred in Africa during the late Miocene, and despite the fact that the oldest uncontroversial fossils of *Crocidura*, dated from the Early Pliocene, were found in Russia (Rzebik-Kowalska, 1998). Our results support a late Miocene diversification, because the molecular clock calibrated with a 5 to 6 Myr estimate for the radiation of *Crocidura* is congruent with the clock calibrated with the Soricinae/Myosoricinae split.

Regarding the genus *Sylvisorex*, our study once again casts doubts on its monophyly (e.g., Butler and Greenwood, 1979; Butler *et al.*, 1989; Hutterer, 1993). *Sylvisorex cf. konganensis*, *S. ollula*, and *S. johnstoni* form an ancestral group, and *S. megalura* is more

closely related to *Suncus*. Our results agree with those of Schlitter *et al.* (1999) who distinguished two groups on the basis of karyotypes. *S. ollula* and *S. johnstoni* are part of a group with low diploid and fundamental numbers, whereas *S. megalura* exhibits high diploid and fundamental numbers. In addition, the ecology of the latter species differs from that of the other species of the genus, as it is not a typical forest dweller. We suggest that its *Sylvisorex* affiliation should be reconsidered.

The positions of the two monospecific genera do not sustain their generic status as they do not appear to be independent lineages. The monospecific genus *Scutisorex* clusters with *Sylvisorex ollula* and *S. cf. konganensis*. Its dental characters are close to those of *Sylvisorex*, but it is classified in a distinct genus due to its unique spine (Thomas, 1913; Allen, 1917; Cullinane and Aleper, 1998; Cullinane *et al.*, 1998). The monospecific genus *Ruwenzorisorex* is positioned among the *Suncus* species. It bears cranio-dental resemblance to *Suncus* but also to *Sylvisorex ollula* (Hutterer, 1986b). The genus is also remarkable for its semiaquatic behavioral adaptations (J. C. Kerbis Peterhans, unpublished data) reflected by its unique brain anatomy (Stephan *et al.*, 1991).

On the whole, our results support the findings of most previous studies. Our data confirm the existence of two subfamilies (or tribes) within the African Crocidurinae (s.l.). They indicate that *Sylvisorex* is certainly a polyphyletic genus poorly differentiated from *Suncus* and that both genera are in need of revision. Our results also shed new light on the origin of *Paracrocidura*, *Scutisorex*, and *Ruwenzorisorex*. They suggest that the three genera may have derived, respectively, from *Crocidura*, *Sylvisorex*, and *Suncus* and raise the question of whether they should be considered independent genera. The present study is a starting point for further molecular and morphological work. It would benefit from a comparative approach with other DNA sequences, especially from the nuclear genome.

Evolutionary History of African Shrews

Shrews are thought to have originated in Eurasia where fossils dating back to the Oligocene have been found (Rzebik-Kowalska, 1998; Storch *et al.*, 1998). They may have entered Africa after the closing of the Tethys sea at the beginning of the Miocene (Butler, 1998). It is presumed that by that time, around 19 Myr ago, forested corridors linked the Asian and African forests (Cox and Moore, 1993), which allowed several forest mammals to disperse from Asia into Africa where some underwent successful radiation (e.g., bovids: Thomas, 1985; bats: Juste *et al.*, 1999). Our data suggest that this has also been the case for the Crocidurinae (s.l.). There would have been at least two migration events leading to the two present African lineages (Fig. 3). If the ancestors of the extant *Sylvisorex*

and *Myosorex* species were—like the extant taxa—primarily forest dwellers, they could have reached Africa through the forested corridors.

During the subsequent period of isolation of the African and Asian forests (Cox and Moore, 1993; Hallam, 1994), the two African lineages diversified separately. Later on, about 7.5 Myr ago, during the Upper Miocene, some early *Suncus* ancestors must have migrated back to Asia (Fig. 3). At that time, the climate was dry (Poumot, 1989; Maley, 1996), and the species that migrated belong to the genus *Suncus*, which is adapted primarily to savanna. During the same period, other savanna-dwelling mammals migrated between Africa and Asia, such as bovids (Thomas, 1984) and the ancestor of the murine genus *Arvicanthis* (Chevret, 1994).

The *Crocidura* radiation seems to have started in Africa at the end of the Miocene (Butler, 1998), just after a major extinction phase (Maglio, 1978). The availability of a wide range of vacant ecological niches may have facilitated the ecological and morphological diversification of this genus. It is very likely that one or more *Crocidura* ancestors migrated from Africa to Europe and Asia at an early stage of the radiation (Fig. 3). At that time, about 5 Myr ago, the climate was dry (Maley, 1996) and presumably not favorable to forest species expansion. This suggests that the species that migrated were adapted to dry habitats. Dispersal events were possible due to the existence of a land connection between Africa and Western Europe during the later stage of the Miocene (Azzaroli and Guazzone, 1979). Consistently, faunal exchanges between Africa and Eurasia dating from the Miocene–Pliocene transition have been evidenced for several small mammals (Thomas *et al.*, 1982; Chevret, 1994).

More recent migration events cannot be inferred from the present study because it includes too few European and Asian species. The close relationship found between *Crocidura leucodon* and *C. obscurior* can be explained in two ways. Either *C. obscurior* is an African descendant of the migrant ancestor and both species retained plesiomorphic characters or some European *Crocidura* species traveled back to Africa during the Pleistocene. The first hypothesis seems more likely because the mammalian fossil record reveals only few migration events between the two continents during the Pleistocene (Hallam, 1994).

Finally, our observation that multiple exchanges must have occurred between Africa and Eurasia follows a general trend that has been documented in other taxa. Simultaneous faunal exchanges have occurred in different taxa during periods of land bridge connection between Africa and Asia or Europe. This, in addition to lending support to the validity of our molecular clock and divergence time estimates, indicates a tight relationship between the paleoecological changes and the faunal migration events.

ACKNOWLEDGMENTS

The authors are indebted to F. Catzeffis, R. Fons, M. Ruedi, W. Verheyen, and P. Vogel, for providing valuable tissue samples. We thank the field collectors O. Perpète, for the fruitful years he spent collecting in the field, and L. Granjon, M. Harada, T. Maddalena, and C. Vaucher, for having provided useful specimens. We address special thanks to A. Hassanin, S. Tillier, and the Museum National d'Histoire Naturelle, Paris, for allowing the extraction of DNA from the *Scutisorex somereni* skeleton. We also thank A. Gautier-Hion for helpful comments on the manuscript. The project, field work in Central and West Africa, and molecular analysis were funded by (1) EU-DGVIII—Ecofac program "Conservation et Utilisation Rationnelle des Écosystèmes Forestiers en Afrique Centrale" (managed by AGRECO, GEIE, BDPA, SCETAGRI, SECA, CIRAD-FORET); (2) EU-DGVIII-PDRN "Programme de Développement de la Région Nord" (NORCADEV); (3) EU-DGVIII—Biofac program "Origine et maintien de la biodiversité en Afrique Centrale" (University of Rennes 1/CNRS, UMR 6552); (4) WWF Dzanga-Sangha Reserve Project; and (5) project Ebola—Forêt de Taï, OMS Abidjan. Work at the Royal Institute for Natural Sciences was supported by DWTC project 31.64 entitled "Evaluating the Effect of Pleistocene Climate Changes on Speciation Patterns in African Vertebrates." Field work in Burundi and Uganda was funded by grants from the Ellen Thorne Smith Fund and Marshall Field III Fund of the Field Museum, the John D. and Catherine T. MacArthur Foundation (Chicago), and the Chicago Zoological Society. We thank the Uganda Department of Forestry, Uganda National Parks, the Institute of Tropical Forest Conservation, the Peace Corps Biodiversity Program in Burundi, and the "Institut National pour la Conservation de la Nature."

REFERENCES

- Allen, J. A. (1917). The skeletal characters of *Scutisorex* Thomas. *Bull. Am. Mus. Nat. Hist.* **37**: 769–784.
- Asher, R. J. (1999). A morphological basis for assessing the phylogeny of the "Tenrecoidea" (Mammalia, Lipotyphla). *Cladistics* **15**: 231–252.
- Azzaroli, A., and Guazzone, G. (1979). Terrestrial mammals and land connections in the Mediterranean before and during the Messinian. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **29**: 155–167.
- Bibb, M. (1981). Sequence and gene organization of mouse mitochondrial DNA. *Cell* **26**: 167–180.
- Bruno, W. J., Socci, N. D., and Halpern, A. L. (2000). Weighted Neighbor Joining: a likelihood approach to distance-based phylogeny reconstruction. *Mol. Biol. Evol.* **17**: 189–197.
- Butler, P. M. (1998). Fossil history of shrews in Africa. In "Evolution of Shrews" (J. M. Wojcik and M. Wolsan, Eds.), pp. 121–132. Mammal Research Institute, Polish Academy of Science, Białowieża.
- Butler, P. M., and Greenwood, M. (1979). Soricidae (Mammalia) from the Early Pleistocene of Olduvai Gorge, Tanzania. *Zool. J. Linn. Soc.* **67**: 329–379.
- Butler, P. M., Thorpe, R. S., and Greenwood, M. (1989). Interspecific relations of African crocidurine shrews (Mammalia: Soricidae) based on multivariate analysis of mandibular data. *Zool. J. Linn. Soc.* **96**: 373–412.
- Chevret, P. (1994). "Étude Évolutive des Murinae (Rongeurs: Mammifères) Africains par Hybridation ADN/ADN: Comparaison avec les Approches Morphologiques et Paléontologiques," Ph.D. dissertation, University of Montpellier II, Montpellier.
- Cox, C. B., and Moore, P. D. (1993). "Biogeography: An Ecological and Evolutionary Approach," Blackwell Sci., Oxford.
- Cullinane, D. M., and Aleper, D. (1998). The functional and biomechanical modifications of the spine of *Scutisorex somereni*, the hero shrew: Spinal musculature. *J. Zool. Lond.* **244**: 453–458.
- Cullinane, D. M., Aleper, D., and Bertram, J. E. A. (1998). The functional and biomechanical modifications of the spine of *Scutisorex somereni*, the hero shrew: Skeletal scaling relationship. *J. Zool., Lond.* **244**: 447–452.
- De Rijk, P. (1995). "Optimisation of a Database for Ribosomal RNA Structure and Application in Structural and Evolutionary Research." Ph.D. dissertation, University of Antwerp, Antwerp.
- Felsenstein, J. (1995). "PHYLIP (Phylogeny Inference Package) version 3.57c." Computer program and manual available at <http://evolution.genetics.washington.edu/phylip.html>.
- Ferris, S. D., Sage, R. D., Prager, E. M., Ritte, U., and Wilson, A. C. (1983). Mitochondrial DNA evolution in mice. *Genetics* **105**: 681–721.
- Fumagalli, L., Taberlet, P., Stewart, D. T., Gelly, L., Hausser, J., and Vogel, P. (1999). Molecular phylogeny and evolution of *Sorex* shrews (Soricidae: Insectivora) inferred from mitochondrial DNA sequence data. *Mol. Phylogenet. Evol.* **11**: 222–235.
- Hallam, A. (1994). "An Outline of Phanerozoic Biogeography," Oxford Biogeography Series No. 10, Oxford Univ. Press, Oxford.
- Hasegawa, M., Kishino, H., and Yano, T. (1985). Dating of the human–ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* **22**: 160–174.
- Hassanin, A., and Douzery, E. J. P. (1999). Evolutionary affinities of the enigmatic saola (*Pseudoryx nghetinhensis*) in the context of the molecular phylogeny of Bovidae. *Proc. R. Soc. Lond. B* **266**: 893–900.
- Heim de Balsac, H., and Lamotte, M. (1956). Evolution et phylogénie des Soricidés africains-I. *Mammalia* **20**: 140–167.
- Heim de Balsac, H., and Lamotte, M. (1957). Evolution et phylogénie des Soricidés africains-II. *Mammalia* **21**: 16–46.
- Hillis, D.M. (1996). Inferring complex phylogenies. *Nature* **383**: 130–131.
- Hutterer, R. (1986a). Synopsis der Gattung *Paracrocidura* (Mammalia: Soricidae), mit Beschreibung einer neuen Art. *Bonn. Zool. Beitr.* **37**: 73–90.
- Hutterer, R. (1986b). Eine neue Soricidengattung aus Zentralafrika (Mammalia: Soricidae). *Z. Säugetierk.* **51**: 257–266.
- Hutterer, R. M. (1993). Order Insectivora. In "Mammal Species of the World: A Taxonomic and Geographic Reference" (D.E. Wilson and D.M. Reeder, Eds.), 2nd ed., pp. 69–130. Smithsonian Institution Press, Washington, DC.
- Jenkins, P., Ruedi, M., and Catzeffis, F. M. (1998). A biochemical and morphological investigation of *Suncus dayi* (Dobson, 1888) and discussion of relationship in *Suncus* Hemprich & Ehrenberg, 1833, *Crociodura* Wagler, 1832, and *Sylvisorex* Thomas, 1904 (Insectivora: Soricidae). *Bonn. Zool. Beitr.* **47**: 257–276.
- Juste, J., Alvarez, Y., Tabares, E., Garrido Perterra, A., and Ibañez, C. (1999). Phylogeography of African fruitbats (Megachiroptera). *Mol. Phylogenet. Evol.* **13**: 596–604.
- Kimura, M. (1980). A simple method for estimating evolutionary rate of base substitution through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**: 111–120.
- Kishino, H., and Hasegawa, M. (1989). Evaluation of the maximum likelihood estimates of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. *J. Mol. Evol.* **29**: 170–179.
- Kumar, S. (1996). PHYLTEST version 2.0: A program for testing phylogenetic hypothesis version 2.0, distributed by the author, Institute of Molecular Evolutionary Genetics and department of Biology. The Pennsylvania State University from <ftp.bio.indiana.edu> in directory molbio/ibmpc.
- Lecointre, G., Philippe, H., Van Le, H. L., and Le Guyader, H. (1993).

- Species sampling has a major impact on phylogenetic inference. *Mol. Phylogenet. Evol.* **2**: 205–224.
- Maddalena, T. (1990a). "Systématique, Évolution et Biogéographie des Musaraignes Afro-Tropicales et Paléartiques de la Sous-Famille des Crocidurinae: Une Approche Génétique," Ph.D. dissertation, University of Lausanne, Lausanne.
- Maddalena, T. (1990b). Systematics and biogeography of Afrotropical and Palearctic shrews of the genus *Crocidura* (Insectivora: Soricidae): An electrophoretic approach. In "Vertebrates in the Tropics" (G. Peters and R. Hutterer, Eds.), pp. 297–308. Museum Alexander Koenig, Bonn.
- Maddalena, T., and Bronner, N. (1992). Biological systematics of the endemic African genus *Myosorex* Gray, 1838 (Mammalia: Soricidae). *Israel J. Zool.* **38**: 245–252.
- Maddalena, T., and Ruedi, M. (1994). Chromosomal evolution in the genus *Crocidura* (Soricidae, Insectivora). In "Advances in the Biology of Shrews" (J.F. Merritt, G.L. Kirkland, Jr, and R.K. Rose, Eds.), *Carnegie Mus. Nat. Hist. Spec. Publ.* **18**: 335–344.
- Maglio, V. J. (1978). Patterns of faunal evolution. In "Evolution of African Mammals" (V.J. Maglio and H.B.S. Cooke, Eds.), pp. 603–619. Harvard Univ. Press, Cambridge, MA.
- Maley, J. (1996). The African rain forest—Main characteristics of changes in vegetation and climate from the Upper Cretaceous to the Quaternary. *Proc. R. Soc. Edinburgh* **104B**: 31–73.
- McLellan, L. J. (1994). Evolution and phylogenetic affinities of the African species of *Crocidura*, *Suncus* and *Sylvisorex* (Insectivora: Soricidae). In "Advances in the Biology of Shrews" (J.F. Merritt, G.L. Kirkland, Jr, and R.K. Rose, Eds.), *Carnegie Mus. Nat. Hist. Spec. Publ.* **18**: 379–391.
- Meester, J. (1953). The genera of African shrews. *Ann. Transvaal Mus.* **22**: 205–214.
- Mindell, D. P., and Honeycutt, R. L. (1990). Ribosomal RNA in vertebrates—Evolution and phylogenetic applications. *Annu. Rev. Ecol. Syst.* **21**: 541–566.
- Motokawa, M., Suzuki, H., Harada, M., Lin, L. K., Koyasu, K., and Oda, S. (2000). Phylogenetic relationships among East Asian species of *Crocidura* (Mammalia, Insectivora) inferred from mitochondrial cytochrome *b* gene sequences. *Zool. Sci.* **17**: 497–504.
- Palumbi, S., Martin, A., Romano, S., McMillan, W. O., Stice, L., and Grabowski, G. (1991). "The Simple Fool's Guide to PCR, version 2.0," Univ. of Hawaii Press, Honolulu.
- Poumot, C. (1989). Palynological evidence for eustatic events in the tropical Neogene. *Bull. Centre Rech. Explor. Prod. Aquit.* **13**: 437–453.
- Repenning, C. A. (1967). Subfamilies and genera of the Soricidae. *U. S. Geol. Surv. Prof. Paper* **565**: 1–74.
- Ruedi, M. (1998). Protein evolution in shrews. In "Evolution of Shrews" (J.M. Wojcik and M. Wolsan, Eds.), pp. 269–294. Mammal Research Institute, Polish Academy of Science, Białowieża.
- Ruedi, M., Auberson, M., and Savolainen, V. (1998). Biogeography of sulawesian shrews: Testing for their origin with a parametric bootstrap on molecular data. *Mol. Phylogenet. Evol.* **9**: 567–571.
- Reumer, J. W. F. (1989). Speciation and evolution in the Soricidae (Mammalia: Insectivora) in relation with the paleoclimate. *Rev. Suisse Zool.* **96**: 81–90.
- Reumer, J. W. F. (1994). Phylogeny and distribution of the Crocidorhinae (Mammalia: Soricidae). In "Advances in the Biology of Shrews" (J.F. Merritt, G.L. Kirkland, Jr, and R.K. Rose, Eds.), *Carnegie Mus. Nat. Hist. Spec. Publ.* **18**: 345–356.
- Reumer, J. W. F. (1998). A classification of the fossil and recent shrews. In "Evolution of Shrews" (J.M. Wojcik and M. Wolsan, Eds.), pp. 5–22. Mammal Research Institute, Polish Academy of Science, Białowieża.
- Rzebiak-Kowalska, B. (1998). Fossil history of shrews in Europe. In "Evolution of Shrews" (J.M. Wojcik and M. Wolsan, Eds.), pp. 23–92. Mammal Research Institute, Polish Academy of Science, Białowieża.
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). "Molecular Cloning: A Laboratory Manual," 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Schlitter, D. A., Hutterer, R., Maddalena, T., and Robbins, L. W. (1999). New karyotypes of shrews (Mammalia: Soricidae) from Cameroon and Somalia. *Ann. Carnegie Mus.* **68**: 1–14.
- Springer, M. S., Cleven, G. C., Madsen, O., de Jong, W. W., Waddell, V. G., Amrine, H. M., and Stanhope, M. J. (1997). Endemic African mammals shake the phylogenetic tree. *Nature* **388**: 61–64.
- Stanhope, M. J., Waddell, V. G., Madsen, O., de Jong, W., Blair Hedges, S., Cleven, G. C., Kao, D., and Springer, M. (1998). Molecular evidence for multiple origins of Insectivora and for a new order of endemic African insectivore mammals. *Proc. Natl. Acad. Sci. USA* **95**: 9967–9972.
- Stephan, H., Baron, G., and Frahm, H. D. (1991). "Comparative Brain Research in Mammals. Vol. 1: 'Insectivora, with a Stereotaxic Atlas of the Hedgehog Brain,'" Springer-Verlag, New York.
- Storch, G., Qiu, Zh., and Zazhigin V. S. (1998). Fossil history of shrews in Asia. In "Evolution of Shrews" (J.M. Wojcik and M. Wolsan, Eds.), pp. 93–120. Mammal Research Institute, Polish Academy of Science, Białowieża.
- Strimmer, K., and von Haeseler, A. (1996). Quartet puzzling: A quartet maximum likelihood method for reconstructing tree topologies. *Mol. Biol. Evol.* **13**: 964–969.
- Swofford, D. L. (1997). "PAUP*: Phylogenetic Analysis Using Parsimony (and other methods)," Version 4.0, Sinauer, Sunderland, MA.
- Tamura, K., and Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* **10**: 512–526.
- Thomas, H. (1984). Les Bovidae (Artiodactyla: Mammalia) du Miocène du sous-continent Indien, de la péninsule Arabique et de l'Afrique: biostratigraphie, biogéographie et écologie. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **45**: 251–299.
- Thomas, H. (1985). The Early and Middle Miocene land connection of the Afro-Arabian plate and Asia: A major event for hominoid dispersal. In "Ancestors: The Hard Evidence" (E. Delson, Ed.), pp. 42–50. A. R. Liss, New York.
- Thomas, H., Bernor R., and Jaeger J. J. (1982). Origine du peuplement mammalien en Afrique du Nord durant le Miocène terminal. *Geobios* **15**: 283–397.
- Thomas, O. (1913). On African bats and shrews. *Ann. Mag. Nat. Hist.* **8**: 314–321.
- Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**: 4673–4680.
- Walsh, P. S., Metzger, D. A., and Higuchi, R. (1991). Chelex 100 as a medium for simple extraction of DNA for PCR based typing from forensic material. *Biotechniques* **10**: 506–513.
- Wheeler, W., and Honeycutt, R. (1988). Paired sequence difference in ribosomal RNAs: Evolutionary and phylogenetic implications. *Mol. Biol. Evol.* **5**: 90–96.
- Wolsan, M., and Hutterer, R. (1998). Appendix: A list of the living species of shrews. In "Evolution of Shrews" (J.M. Wojcik and M. Wolsan, Eds.), pp. 425–448. Mammal Research Institute, Polish Academy of Science, Białowieża.

II - PHYLOGENIE DU GENRE *HYLOMYSCUS*, MURIDES FORESTIERS AFRICAINS

Le genre *Hylomyscus* Thomas, 1926, est largement répandu en Afrique tropicale. Son affinité pour la forêt et sa grande abondance sur la plupart des sites forestiers (17 à 33% des captures de muridae sur les sites échantillonnés par notre équipe) font qu'il pourrait fournir de bons modèles pour les études phylogéographiques. Au sein de ce genre, la détermination des espèces est délicate et ne peut être réalisée que grâce à une analyse minutieuse des crânes. De ce fait, la plupart des spécimens capturés n'ont pas été déterminés, et une révision systématique s'est avérée nécessaire.

A l'instar de *Mastomys* et de *Myomyscus* (ou *Myomys* cf. Rosevear, 1969 et Van der Straeten et Verheyen, 1978), *Hylomyscus* est soit considéré comme un sous-genre de *Praomys* (Misonne, 1969 ; Meester et Setzer, 1971), soit comme un genre à part (Rosevear, 1969 ; Robbins *et al.*, 1980). Nous adopterons ici la seconde alternative. La proximité entre les quatre genres est confirmée par une phylogénie moléculaire basée sur des hybridations ADN / ADN (Chevret *et al.*, 1994). La limite entre *Hylomyscus* et *Praomys* est difficile à préciser, et certaines espèces, comme *H. denniae*, sont alternativement classées dans l'un ou l'autre genre (cf. Kingdon, 1971). En 1993, Musser et Carleton reconnaissaient 7 espèces appartenant au genre *Hylomyscus* : *H. aeta* (Thomas, 1911), *H. alleni* (Waterhouse, 1838), *H. baeri* Heim de Balsac et Allen, 1965, *H. carillus* (Thomas, 1904), *H. denniae* (Thomas, 1906), *H. parvus* Brosset, Dubost et Heim de Balsac, 1965 et *H. stella* (Thomas, 1911). *H. baeri* est cantonnée à l'Afrique de l'Ouest, *H. carillus* au Nord de l'Angola, *H. denniae* à l'Afrique de l'Est, et les quatre autres espèces sont présentes en Afrique centrale, où elles sont parfois sympatriques (au Gabon : Brosset *et al.*, 1965 ; dans le Sud du Cameroun : Robbins *et al.*, 1980). Brosset et collaborateurs reconnaissent aussi l'espèce *H. simus* Allen et Coolidge, 1930, espèce décrite de l'Afrique de l'Ouest dont la validité et la présence en Afrique centrale ont été remises en question (Rosevear, 1969 ; Musser et Carleton, 1993). Ils décrivent également une cinquième espèce sympatrique, *H. fumosus*, mais celle-ci appartiendrait à un genre distinct : *Heimyscus* (Misonne, 1969). Trois complexes d'espèces ont été définis au sein des *Hylomyscus* (Robbins *et al.*, 1980). Le premier de ces complexes est constitué de quatre espèces qui se remplacent géographiquement *H. aeta*, *H. baeri*, *H. carillus* et *H. denniae*, le second de *H. alleni* et *H. stella*, et le troisième de *H. parvus*.

Les espèces du genre *Hylomyscus* sont souvent difficiles à distinguer les unes des autres (e.g. Rosevear, 1969 ; Robbins *et al.*, 1980). De nombreuses espèces, telles *H. kaimosae*, *H. simus* ou *H. vulcanorum*, ont été décrites, puis mises en synonymies avec d'autres ; mais pourraient en réalité représenter des taxons valides, au sein de complexes d'espèces (Musser et Carleton, 1993). Nous avons

tenté de distinguer des groupements phylétiques et de déterminer les relations phylogénétiques entre espèces à l'aide d'une analyse moléculaire, basée sur une portion de la séquence de l'ARN 16s et du cytochrome *b*. Dans un premier temps, nous avons analysé un grand nombre de spécimens dans le but de faciliter la révision craniologique nécessaire à leur détermination spécifique. Cette révision, indispensable à la détermination des espèces du genre *Hylomyscus* et réalisée par Walter Verheyen, spécialiste des rongeurs africains, est encore en cours. Dans un second temps, nous avons essayé d'obtenir une phylogénie du genre à partir d'un nombre plus restreint de spécimens, ceci de façon à équilibrer l'échantillonnage entre espèces et à pouvoir comparer les résultats produits par différentes méthodes de reconstruction phylogénétique. Nous avons aussi tenté de calibrer une horloge moléculaire propre aux muridés et de procéder à une datation des principaux événements de divergence entre espèces. Pour finir, comme la formule chromosomique de la majorité des espèces a été déterminée (Mathey, 1963 ; Robbins *et al.*, 1980 ; Viegas-Pequignot *et al.*, 1983 ; Iskandar *et al.*, 1988 ; Maddalena *et al.*, 1989), nous nous sommes intéressés à l'évolution des caryotypes.

1) Matériel et méthodes

a) détermination spécifique

Après analyse moléculaire d'une première série d'échantillons choisis en aveugle, les premiers résultats ont permis de procéder à un tri des crânes disponibles, qui a permis en retour de proposer une liste additionnelle de spécimens à analyser. Tous les spécimens analysés ont été déterminés sur la base de caractères craniologiques. Au total, 81 échantillons d'*Hylomyscus* ont été séquencés pour l'ARNr 16S et 28 d'entre eux l'ont été également pour le cytochrome *b* (tab. 3.1). Nous avons choisi comme groupes externes un spécimen de *Praomys delectorum* (2732, Mbete, Tanzanie), et un spécimen de *P. tullbergi* (R24476, Taï, Côte d'Ivoire). Ce dernier a été séquencé pour l'ARN 16s uniquement.

Pour l'ARN 16s, l'alignement a été réalisé à l'aide de ClustalW, puis corrigé manuellement. Les 26 positions flanquantes non disponibles pour la plupart des spécimens ont été éliminées. La longueur des séquences après alignement est de 523 paires de bases (pb). Pour le cytochrome *b*, aucun indel n'a été détecté. La portion d'ARNt amplifiée conjointement au cytochrome *b* a été éliminée, donnant un segment de 492 pb.

Etant donné le grand nombre d'échantillons, nous avons construit des arbres phylogénétiques par la méthode du Neighbor Joining uniquement. Cette méthode est suffisante pour regrouper des séquences proches entre elles. Par contre, elle est peu fiable pour déterminer les relations entre groupements, ce qui

fait que nous ne discuterons pas de la valeur phylogénétique des arbres obtenus par cette méthode. Nous avons corrigé les distances par la formule de Tamura-Nei et considéré les délétions comme des informations manquantes. Nous avons procédé à un bootstrap avec 10.000 répliques.

Tableau 3.1 : Liste des spécimens analysés par localité, avec nom du collecteur pour les spécimens non collectés par notre équipe.

Code	Localité, pays	Echantillons séquencés	Collecteur
Bab	Bamingui-Brendja, RCA	NC438	
Bur	Kibira NP, Burundi	T-0477	T. Maddalena
Cfp	Four-Place, Gabon	GA66, GA87, GA120	
Dja	Dja, Cameroun	R14014 , R14028, R14220 , R14476*, R14477 , R14510, R14659* , R14703, R14771	
Fab	Forêt des abeilles, Gabon	G10003 , G10006, G10022, G10035*, R16004* , R16014 , R16038 , R16132, R16137	
Geo	Escarpements d'Odzala, Congo	R16841, R16844	
Kak	Kakamega, Kenya	SP 5032, 5058	D. Schlitter
Kik	Kikwit, RDC	915, 2727, 2359	H. Leirs
Kil	Kili, Gabon	GA217, GA225, GA256	
KiN	Kisangani Nord, RDC	F2559 , F2654* , F2739	
KiS	Kisangani Sud, RDC	F2679, F2745*	
Kor	Korup, Cameroun	SP 10502, 10504, 10507, 10514	
Kou	Kouilou, Congo	T-0793*, T-0799*	L. Granjon
Mal	Malounga, Gabon	GA166, GA131, GA199	
Mbe	Mbete, Tanzanie	4011	W. Verheyen
Mbi	Mbizi, Tanzanie	13253	W. Verheyen
Mak	Makokou, Gabon	T-0415	V. Nancé
Muf	Mufindi, Tanzanie	10871, 11036*	W. Verheyen
Mvo	Mvoum, Gabon	GA13	
Ngo	Ngotto, RCA	R12381, R12586, R12626, R13001, R13061, R13800, R18835, R18723, R19304	
Odz	Odzala, Congo	R22014 , R22018 , R22062, R22226 , R22265* , R22302, R22353, R22354, R22966* , R23116 , R23146* , R23181*	
Sag	Salonga, RDC	S002	
Sal	Salo ² , RCA	R13187, R13222 , R13657 , R13714	
Taï	Taï, Côte d'Ivoire	R24225* , R24278*	

Les échantillons indiqués en gras ont été séquencés pour l'ARN 16s et le cytochrome *b*, les autres ne l'ont été que pour l'ARN 16s. Les échantillons repérés par un astérisque ont été employés pour construire la phylogénie du genre.

b) Phylogénie du genre *Hylomyscus*

Dans la mesure du possible, nous avons retenu deux échantillons par espèce, ayant été séquencés à la fois pour l'ARN 16s et le cytochrome *b* (cf. tab. 3.1 et **fig. 3.1**, où les séquences sont repérées par des astérisques). Pour quatre de ces échantillons, seule la séquence de l'ARN 16s était disponible. Dans le cas de *H. stella*, nous avons conservé deux spécimens d'Afrique centrale et un spécimen d'Afrique de l'Est. En ce qui concerne le complexe d'espèces *H. kaimosae*, nous disposions d'une seule séquence de *H. kaimosae* pour le cytochrome *b*. En conséquence, nous avons conservé une spécimen appartenant à une espèce non décrite que nous appellerons *H. nsp1*. Nous avons retenu *Praomys delectorum* en tant que groupe externe proche, et nous avons ajouté quatre groupes externes plus éloignés de façon à mieux enraciner l'arbre et à procéder à une datation. Les séquences ont été extraites de GenBank : *Gerbillus nigeriae* (numéros d'accès : AF141257 pour l'ARN 16s et AF141226 pour le cytochrome *b*), *Rattus norvegicus* (X14848, pour les deux séquences), *Mus musculus* (J01420, idem), et *Mastomys natalensis* (Z83905, ARN 16s seulement).

Nous avons analysé les données obtenues pour les deux portions d'ADN mitochondrial d'abord séparément, puis conjointement. Avant de passer à cette seconde étape, nous avons vérifié la compatibilité entre les arbres. Pour chaque portion d'ADN, nous avons estimé graphiquement la saturation des transitions et des transversions. Nous avons effectué des analyses NJ (après correction des distances par la formule de Tamura-Nei), WNJ, MP (10 recherches heuristiques avec addition aléatoire des séquences), et ML (10 recherches heuristiques avec addition aléatoire des séquences, 4 classes de variabilité, paramètres estimés à partir des données). Dans tous les cas, *Gerbillus nigeriae* a été désignée comme groupe externe. Nous avons aussi évalué le support des branches par des tests de bootstrap avec 10.000 réplifications pour le NJ et 1000 réplifications avec une seule addition de séquences pour le MP. Seules les valeurs supérieures à 50% seront reportées sur les arbres correspondant.

Pour l'**ARN 16s**, nous avons mesuré la variabilité des positions appariées et non appariées (déterminées par analogie avec la structure II de l'ARN 16s chez *Mus*, De Rijk, 1995), mais nous n'avons pas pondéré ces positions différenciellement pour les raisons invoquées dans le chapitre 2.III.1 Par contre, nous avons essayé d'évaluer l'impact du codage des indels sur l'analyse. Dans un premier temps, nous avons considéré les délétions comme des informations manquantes. Dans un second temps, nous avons procédé au plus simple des deux codages proposés par Simmons et Ochoterena (2000). Ce codage n'a pu être effectué que pour les analyses NJ et MP, car les méthodes WNJ et ML ne peuvent traiter que des caractères nucléotidiques.

Pour le **cytochrome b**, nous avons évalué l'effet de la pondération différentielle des positions des codons sur les analyses MP. Nous avons calculé le taux de substitution pour chacune des positions, et analysé les données d'abord sans pondération, puis en attribuant un poids de 4 à la seconde position et de 2 à la première position.

Pour la **combinaison des deux** séquences, nous avons choisi de ne pas coder les indels et de ne pas pondérer les différentes positions des codons de façon à ne pas donner plus de poids au cytochrome *b* qu'à l'ARN 16s. Les spécimens pour lesquels nous ne disposions pas de la séquence du cytochrome *b* ont été inclus dans les analyses MP et ML uniquement, et nous avons vérifié que leur inclusion n'entraînait pas un biais. Pour finir, nous avons construit un arbre consensus reprenant les groupements confortés par la plupart des analyses. En nous servant des données caryologiques existantes, nous avons pisté l'évolution des formules chromosomiques sur cet arbre.

Afin d'estimer le **temps écoulé depuis certains événements de divergence**, nous avons vérifié que les données respectaient l'horloge moléculaire par un "relative rate test". Seul l'ARN 16s a été comptabilisé, de façon à pouvoir exploiter les résultats dans la partie "Phylogéographie" (chapitre 4.II) qui n'est basée que sur ce gène. Nous nous sommes basés sur les groupements mis en évidence sur l'arbre NJ qui implique une certaine régularité de l'évolution. Nous avons calibré une horloge moléculaire à partir de données fossiles déjà utilisées pour procéder à des datations moléculaires chez les murinés (Ducroz, 1998 ; Michaux et Catzeflis, 2000) : la divergence entre *Mus* et *Gerbillus*, estimée à 16 MA (Tong et Jaeger, 1993) et celle entre *Mus* et *Rattus*, estimée à 12 MA (10-15 MA - Jaeger *et al.*, 1986 ; Jacobs et Downs, 1994). Nous avons pris en compte les transversions et les transitions de façon à conserver une précision acceptable pour dater les événements récents. Nous avons corrigé les distances par la formule de Tamura-Nei de manière à tenir compte de la saturation des substitutions. De ce fait, nous n'avons pu comptabiliser les indels.

2) Résultats

a) Détermination spécifique

Pour l'ARN 16s, parmi les 523 caractères analysés pour les groupes internes, 410 sont constants, 113 sont variables et 78 sont informatifs pour la parcimonie. Lorsqu'on inclut les groupes externes, 392 caractères sont constants, 131 sont variables et 81 sont informatifs pour la parcimonie. La longueur de l'arbre NJ est de 389 pas (**fig. 3.1**). L'indice de cohérence (CI) vaut 0,470 et l'indice de cohérence normée (RC) vaut 0,365, ce qui signifie que plus de la moitié des substitutions sont homoplasiques. L'indice de rétention (RI) vaut 0.776, ce qui indique que les données ne sont pas trop saturées.

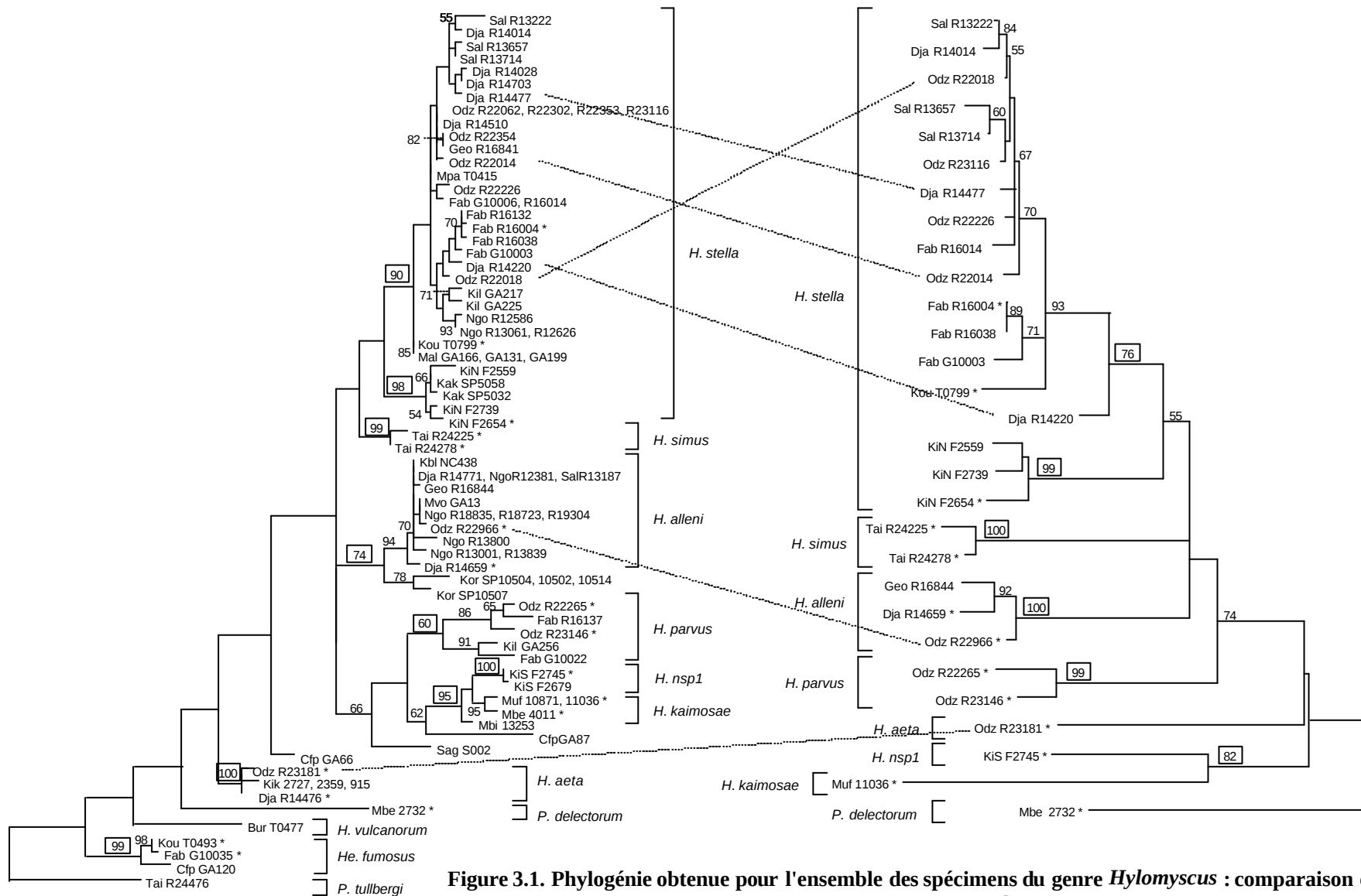


Figure 3.1. Phylogénie obtenue pour l'ensemble des spécimens du genre *Hylomyscus* : comparaison entre arbres NJ basés sur l'ARNr 16s (à gauche) et le cytochrome *b* (à droite).

Les traits hachurés relient les séquences d'un même spécimen qui ne peuvent être mises en correspondance sans modification de la topologie des arbres. Les valeurs de bootstrap sont indiquées au-dessus des branches, et entourées pour chacune des espèces.

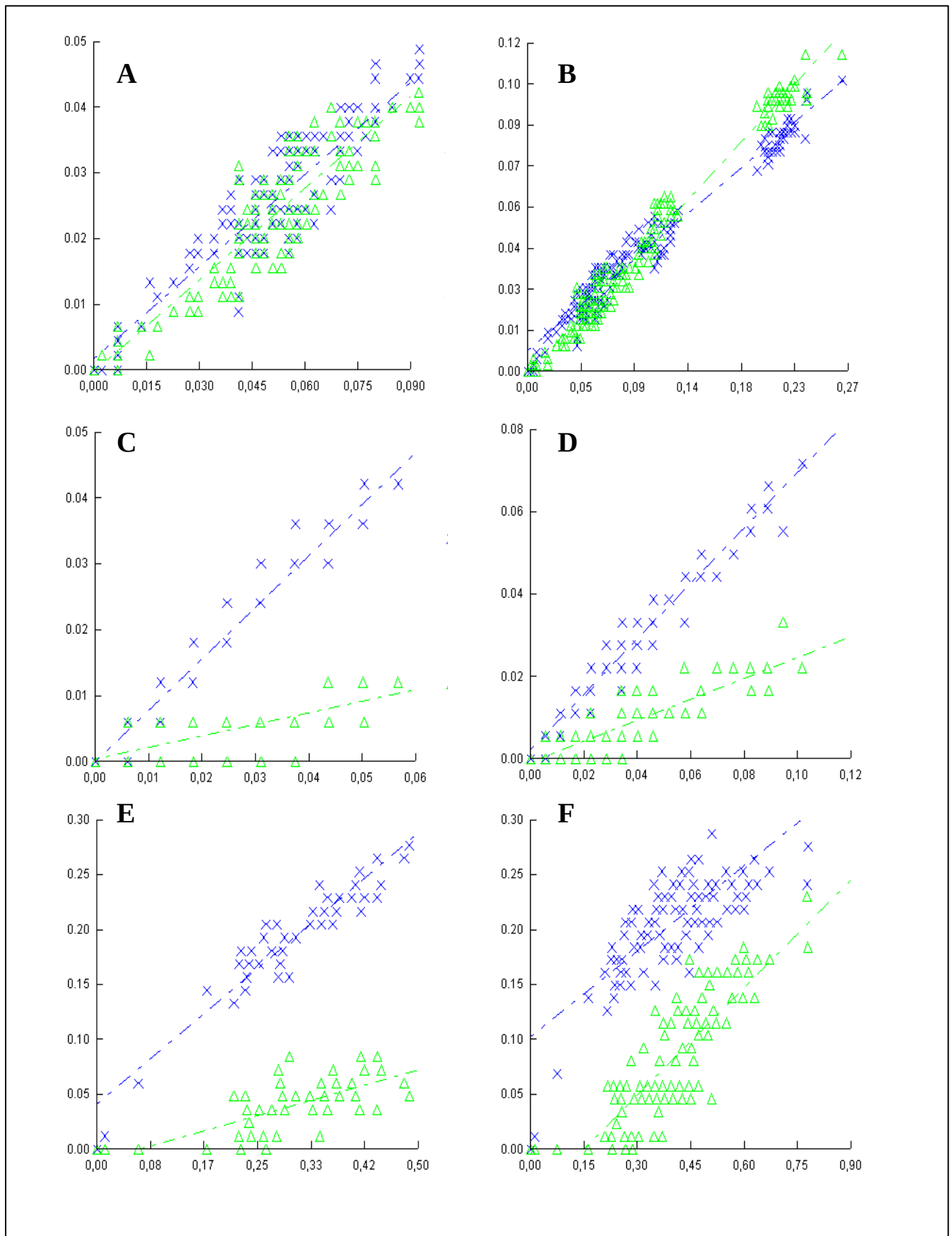


Figure 3.2. Courbes de saturation de l'ARN 16s (graphes A et B) et du cytochrome b (graphes C et E : 1ère et 2ème positions des codons, graphes D et F: 3ème position).

Colonne de gauche : groupes internes uniquement, colonne de droite : en incluant les groupes externes.

x : TS = f(% substitutions K2p), Δ : TV = f(% substitutions K2p).

Pour le cytochrome *b*, parmi les 492 caractères analysés, 372 sont constants, 120 sont variables et 89 sont informatifs pour la parcimonie (ou respectivement 365, 127 et 91, en incluant les groupes externes). L'arbre NJ est caractérisé par une longueur de 266 pas (**fig. 3.1**), et des indices de cohérence et de rétention assez proches des précédents (CI = 0,553, RC = 0,393 et RI = 0,710).

Les arbres obtenus pour l'ARN 16s et le cytochrome *b* ne sont pas parfaitement compatibles (**fig. 3.1**). La position de certains spécimens diffère d'un arbre à l'autre alors même qu'elle est supportée par des valeurs de bootstrap élevées dans les deux cas. Cependant, les groupements entre échantillons appartenant à une même espèce coïncident pour les deux portions d'ADNmt séquencées. Ils sont supportés par des valeurs de bootstrap élevées, comprises entre 60 et 100%.

Les groupements entre échantillons étant en accord avec les déterminations craniologiques, nous avons pu donner un nom à presque tous les spécimens et choisir deux spécimens par espèce pour aborder la phylogénie du genre.

b) Phylogénie du genre *Hylomyscus*

ARN 16s

Après alignement, la longueur de la séquences est de 529 caractères, dont 463 caractères constants, 68 variables, et 59 informatifs pour la parcimonie (ou respectivement 347, 182 et 88, en incluant les groupes externes). Le codage binaire des indels a entraîné l'ajout de 22 caractères supplémentaires, tous variables, dont 13 sont informatifs pour la parcimonie. Sur les 302 positions incluses dans des tiges, 84 sont variables (27,8%), tandis que sur les 229 positions non appariées, 98 sont variables (42,8%). Les tiges apparaissent donc moins variables que les positions non appariées. La variabilité de ces dernières est due pour plus d'un tiers à une grande boucle, comprise entre les positions 237 et 283, qui comprend 36 positions variables sur 47.

Les transitions (TS), tout comme les transversions (TV) ne semblent pas saturées, puisque la relation entre taux de TS ou TV et taux de substitutions est linéaire (**fig. 3.2 A**). Lorsqu'on inclut les groupes externes, le taux de TV devient supérieur au taux de TS, sans pour autant atteindre un rapport TV sur TS égal à 2, qui indiquerait que les transitions sont saturées (**fig. 3.2 B**).

Les méthodes WNJ et NJ, avec ou sans codage des indels, ont donné des arbres de topologie identique (**fig. 3.3 A**). L'analyse MP avec codage des indels a mis en évidence un seul arbre plus parcimonieux (**fig. 3.3 B**). Cet arbre diffère du précédent par la position de *H. alleni*, *H. simus* et *H. stella*, mais les groupements supportés par une forte valeur de bootstrap sont les mêmes pour les quatre analyses susmentionnées. En revanche, l'arbre MP est presque identique à l'arbre ML (**fig. 3.3 D**), si ce n'est pour la position de *Praomys delectorum*.

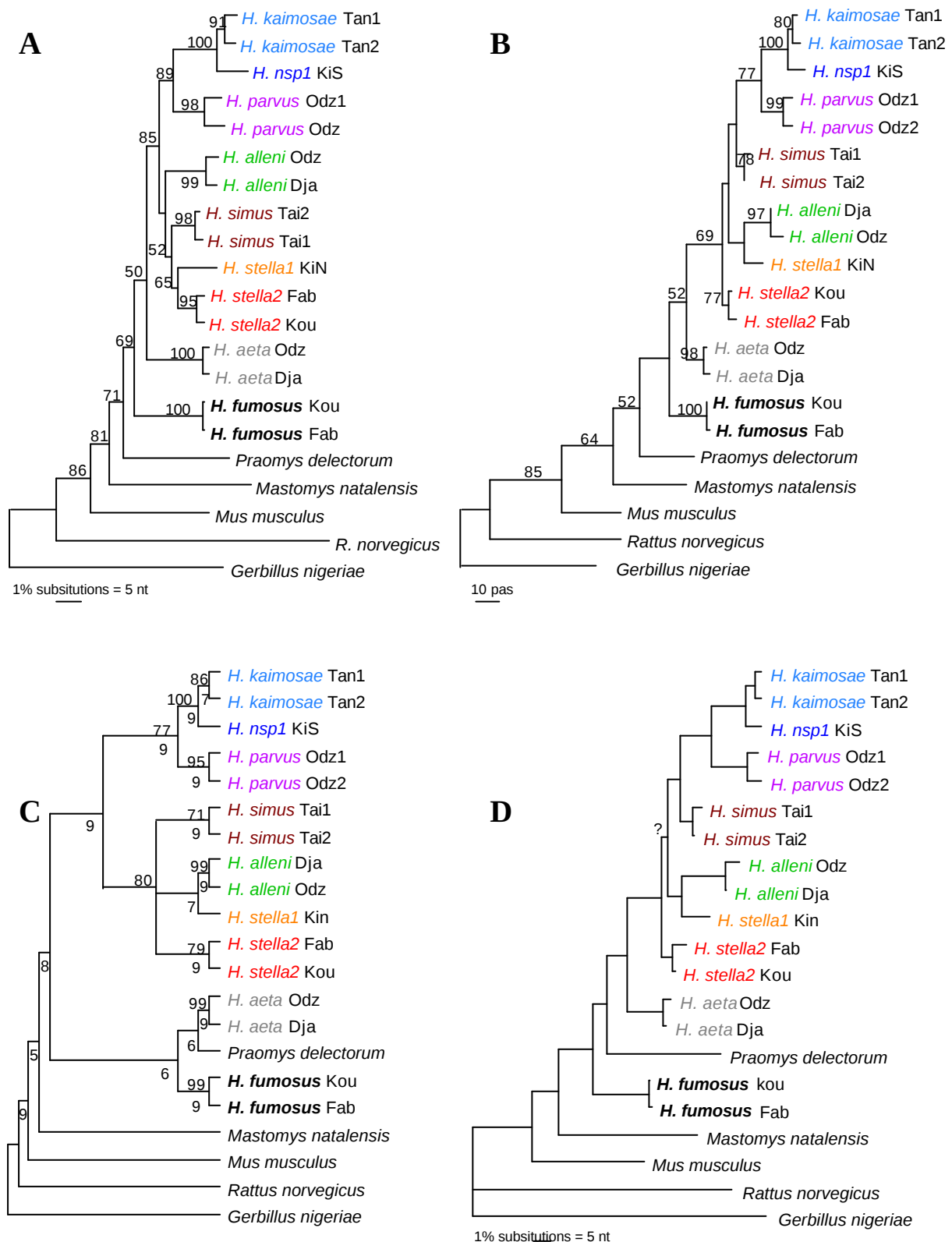


Figure 3.3. Phylogénie du genre *Hylomyscus* basée sur l'ARN 16s - arbres phylogénétiques obtenus par différentes méthodes, avec support bootstrap.

A - arbre NJ avec codage des indels (même arbre sans codage et par WNJ) ; B - arbre MP avec codage des indels ; C - consensus des 9 arbres MP sans codage des indels, avec nombre d'arbres présentant chaque noeud indiqué sous les branches ; D - arbre ML (? = branche non significativement différente de zéro).

L'analyse MP sans codage des indels a conduit à 9 arbres, dont le consensus (**fig. 3.3 C**) fait apparaître un regroupement entre *H. aeta*, *H. fumosus* et *P. delectorum*, présent dans 6 arbres mais qui n'est pas conforté par les autres méthodes d'analyse.

Pour chacune des conditions d'analyse, les arbres sont de longueurs comparables ($p > 0,05$) et se caractérisent par des indices RC, RI et HI sensiblement identiques (tab. 3.2). Les indices de cohérence normé sont peu élevés (valeur moyenne 0,408), et les indices de rétention révèlent une certaine saturation des substitutions (valeur moyenne 0,611).

Tableau 3.2 : Longueur des arbres obtenus par différentes méthodes pour les trois portions d'ADN analysées, et valeur des indices RI, RC et HI.

Arbre	Longueur	p	CI	RC	RI
ARN 16s - 21 taxons - sans codage des indels					
NJ et WNJ	385	0,638	0,665	0,403	0,607
MP et ML	383	-	0,668	0,410	0,613
ARN 16s - 21 taxons - avec codage des indels					
NJ	420	0,527	0,662	0,406	0,603
MP	418	-	0,665	0,412	0,620
cytochrome b - 16 taxons					
NJ et WNJ	401	0,336	0,569	0,271	0,477
MP sans poids	396	-	0,576	0,284	0,492
MP avec poids	(502)	-	0,604	0,292	0,484
ML	405	0,170	0,563	0,262	0,465
ARN 16s + cyt b - 16 taxons					
NJ et WNJ	732	0,197	0,630	0,320	0,508
MP	727	-	0,634	0,328	0,517
ARN 16s + cyt b - 21 taxons					
MP	791	-	0,612	0,328	0,536
ML	803	0,134	0,603	0,312	0,518

Cytochrome b

La longueur de la séquence est de 492 caractères, dont 379 caractères constants, 113 variables, et 82 informatifs pour la parcimonie (ou respectivement 321, 171 et 109, en incluant les groupes externes). Chacune des positions représente 164 caractères. La troisième position est la plus variable avec 90 caractères variables dont 70 informatifs pour la parcimonie. La seconde position est la moins variable avec 7 caractères variables dont un seul informatif. La première position est intermédiaire avec 16 caractères variables dont 11 informatifs. En incluant les groupes externes, ces valeurs deviennent respectivement 36 et 16 pour la première position, 11 et 1 pour la seconde, et 124 et 92 pour la troisième.

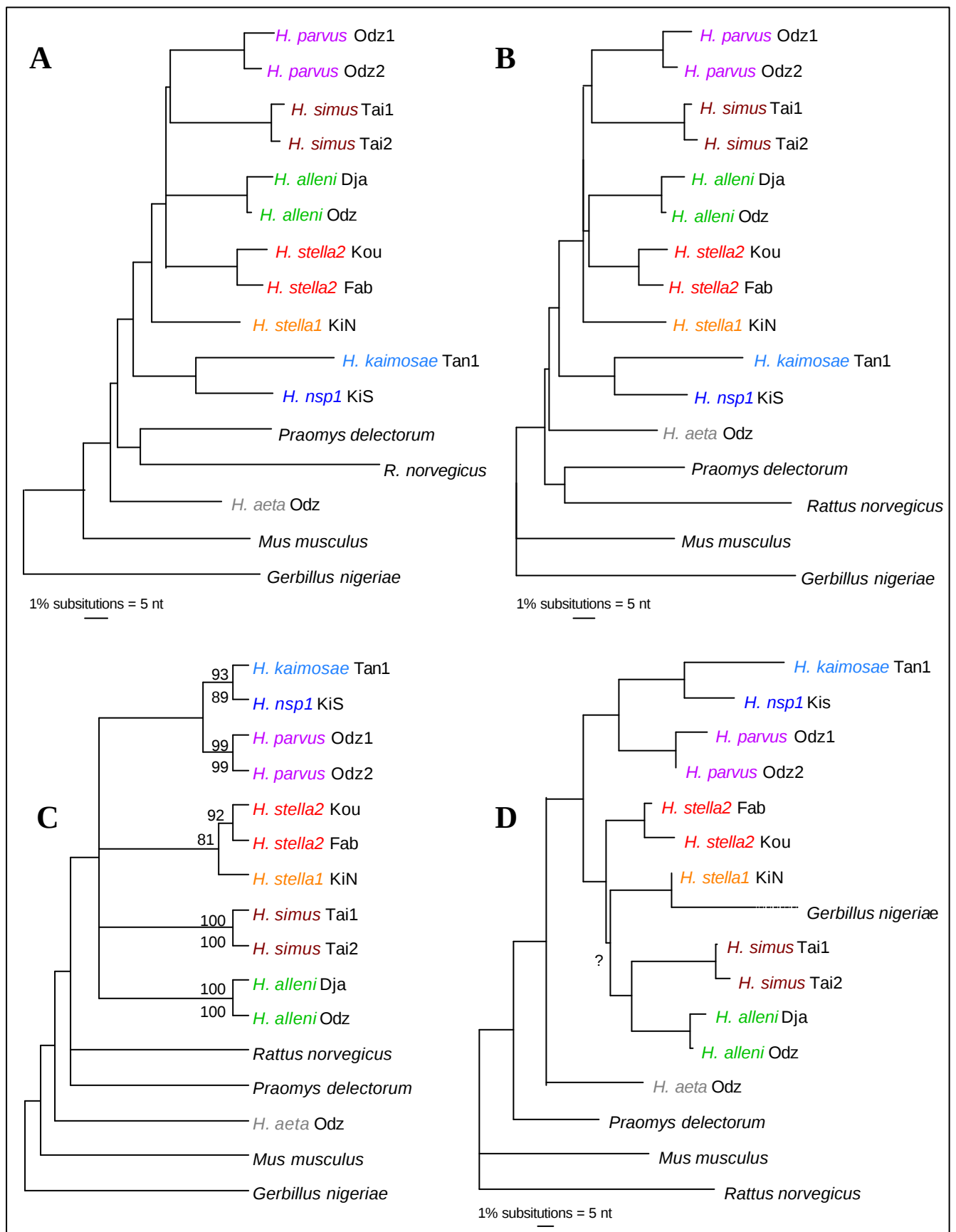


Figure 3.4. Phylogénie du genre *Hylomyscus* basée sur le cytochrome *b* - arbres obtenus par différentes méthodes, avec support bootstrap.

A - arbre NJ ; B - arbre WNJ ; C - strict consensus des 2 arbres MP sans pondération = consensus des 4 arbres MP avec pondération (bootstrap avec / sans pondération, respectivement au-dessus et au-dessous des branches) ; D - arbre ML (? = branche non significativement différente de zéro ; ... = branche réduite de 2/3).

Etant donné la très faible variabilité de la seconde position, nous l'avons groupée avec la première pour représenter les courbes de saturation. Ces courbes indiquent que les premières et secondes positions ne sont pas saturées (**fig. 3.2 C et D**). Par contre, pour la troisième position, l'obtention de nuages de points et de droites de régressions dont l'ordonnée à l'origine diffère de zéro sont des indices de saturation (**fig. 3.2 E et F**). Celle-ci est plus marquée lorsqu'on ajoute les groupes externes.

Les arbres obtenus par les méthodes NJ et WNJ comptent le même nombre de pas, mais ne sont pas identiques (tab. 3.2, **fig. 3.4 A et B**). Ils diffèrent par la position de *H. stella 1* et *H. aeta*. Les consensus à la majorité des 2 arbres MP obtenus sans pondération et des 4 arbres MP obtenus avec pondération sont identiques (**fig. 3.4 C**). Ils présentent peu de résolution, mais regroupent *H. parvus* avec *H. kaimosae* et *H. nsp1*, et *H. stella 1* avec *H. stella 2*, ce qui n'est pas le cas dans les deux arbres phénétiques. L'arbre ML présenté a été obtenu en relâchant la contrainte stipulant que *Gerbillus nigeriae* est un groupe externe (**fig. 3.4 D**), car celle-ci entraînait le placement des trois autres groupes externes au milieu des groupes internes. L'arbre ML regroupe *H. parvus* avec *H. kaimosae* et *H. nsp1*, ainsi que *H. simus* avec *H. allenii*. Les arbres NJ, WNJ et ML sont moins parcimonieux que l'arbre MP, mais la différence n'est pas significative (tab. 3.2). Les indices de cohérence et de rétention sont plus bas que pour l'ARN 16s (tab. 3.2).

Combinaison des deux séquences

La longueur totale de la séquence combinée est de 1023 caractères, dont 843 caractères constants, 181 variables, et 141 informatifs pour la parcimonie (ou respectivement 681, 342 et 185, en incluant les groupes externes).

Les méthodes NJ et WNJ ont donné des arbres de topologie identique (**fig. 3.5 A**). La reconstruction MP avec 16 taxons a donné deux arbres différents de l'unique arbre obtenu pour les 21 taxons (**fig. 3.5 B**) par la position de *H. allenii* et *H. simus*, mais dont le consensus est compatible avec l'arbre obtenu pour les 21 taxons. Les groupements d'espèces supportés par des valeurs de bootstrap élevées sont les mêmes que dans l'arbre NJ, mais la position du groupe *H. stella 1* et 2 est différente. L'arbre ML présenté a été obtenu en utilisant les 21 taxons et en relâchant la contrainte forçant *Gerbillus nigeriae* comme groupe externe (**fig. 3.5 C**). Dans cet arbre, les affinités de *H. stella 1* et 2 sont plus proches de celles révélées par l'arbre NJ que par l'arbre MP. Pour le reste, et à l'exception de la position de *Gerbillus nigeriae*, la topologie de l'arbre est voisine de celle de l'arbre MP. Notons que les arbres ont été présentés sous forme de phénogrammes, mais que les longueurs de branches ne sont pas fiables puisque seule la séquence de l'ARN 16s est prise en compte pour 5 des taxons.

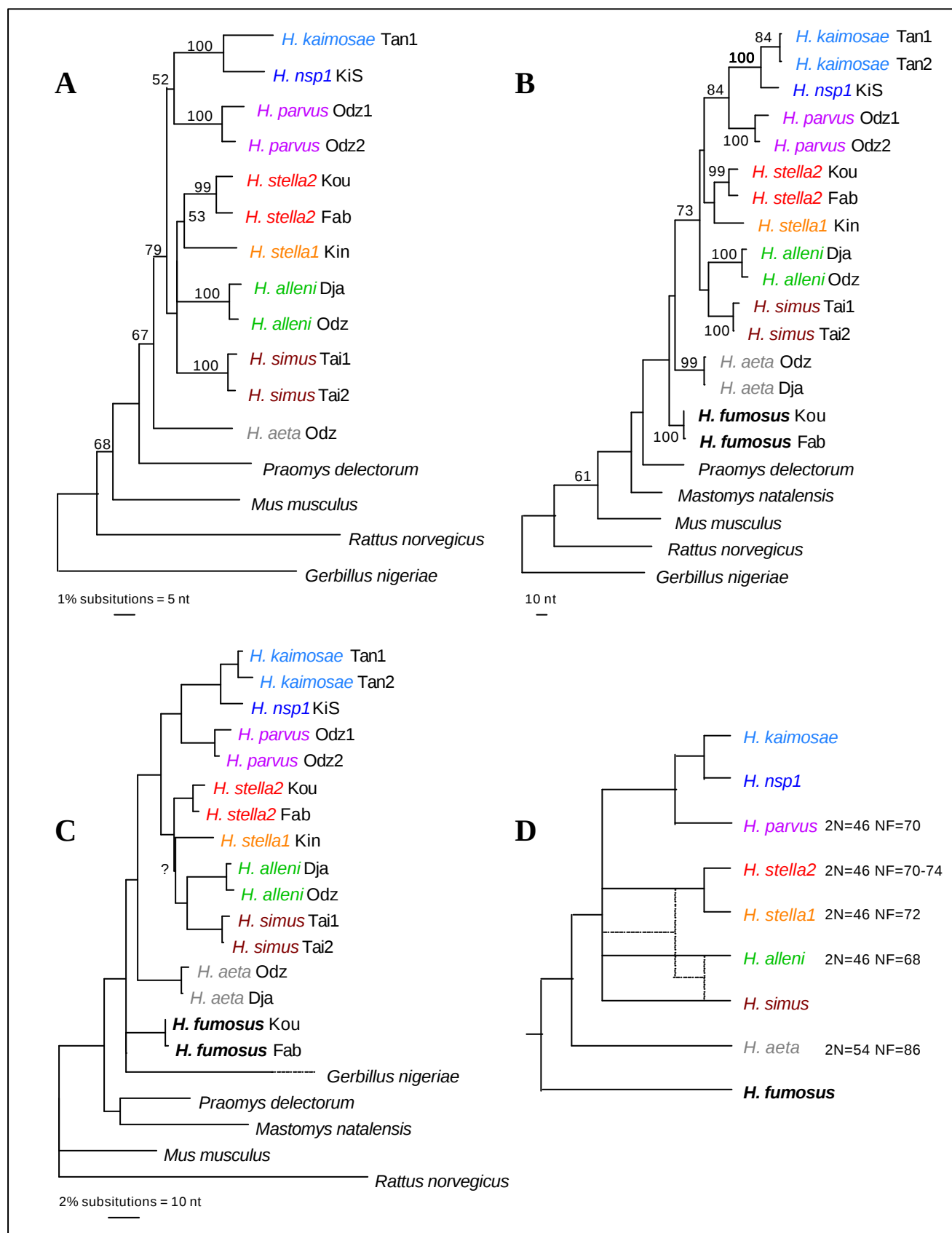


Figure 3.5. Phylogénie du genre *Hylomyscus* basée sur l'ARN 16s + le cytochrome *b* - arbres obtenus par différentes méthodes, avec support bootstrap.

A - arbre NJ pour les 16 taxons séquencés pour les deux portions d'ADN (identique arbre WNJ) ; B - arbre MP pour l'ensemble des taxons ; C - arbre ML pour l'ensemble des taxons (? = branche non significativement différente de zéro; = branche réduite de moitié) ; D - relations probables entre les espèces inférées de l'ensemble des analyses, les relations les plus hypothétiques sont représentées en pointillés.

Que les analyses soient réalisées avec 16 ou 21 taxons, l'arbre MP est toujours le plus parcimonieux, mais les différences de longueur ne sont pas significatives (tab. 3.2). Les indices de cohérence et de rétention sont intermédiaires entre ceux trouvés pour l'ARN 16s et le cytochrome *b*.

Datation

Les résultats du "relative rate test" permettent d'accepter l'hypothèse de l'horloge moléculaire puisqu'aucune valeur de *p* n'est inférieure à 0,05. La valeur minimale de *p*, obtenue entre *H. simus* et *H. aeta*, est de 0,229.

En se basant sur la divergence *Mus-Rattus*, le taux de substitution estimé est de 0,825 substitutions par millions d'années. En se basant sur la divergence *Mus-Gerbillus*, il est de 0,718 substitutions par millions d'années. La valeur moyenne est 0,772. Nous avons daté les événements de divergence mis en évidence sur l'arbre NJ (**fig. 3.3 A**) pour ces trois valeurs de *r* (tab. 3.3).

Tableau 3.3 : Estimation du temps de divergence (en millions d'années) pour les principales lignées d'*Hylomyscus*.

		Calibration <i>Mus-Rattus</i>	Calibration <i>Mus-Gerbillus</i>	Moyenne
Taxon 1	Taxon 2	temps +/- σ	temps +/- σ	temps +/- σ
<i>Mus</i>	<i>Rattus</i>	(12)	13,8	12,8
<i>Mus</i>	<i>Gerbillus</i>	13,9	(16)	14,9
<i>H. fumosus</i>	autres *	4,1 +/- 0,4	4,7 +/- 0,5	4,3 +/- 0,5
<i>H. aeta</i>	autres *	3,7 +/- 0,7	4,2 +/- 0,8	3,9 +/- 0,8
<i>H. parvus</i> , <i>H. nsp1</i> <i>H. kaimosae</i>	autres *	3,4 +/- 0,4	3,9 +/- 0,5	3,6 +/- 0,5
<i>H. parvus</i>	<i>H. nsp1</i> <i>H. kaimosae</i>	3,0 +/- 0,3	3,5 +/- 0,4	3,2 +/- 0,3
<i>H. allenii</i>	<i>H. stella1/2</i> <i>H. simus</i>	2,6 +/- 0,3	3,0 +/- 0,3	2,8 +/- 0,3
<i>H. simus</i>	<i>H. stella1/2</i>	1,9 +/- 0,3	2,1 +/- 0,3	2,0 +/- 0,3
<i>H. stella 1</i>	<i>H. stella 2</i>	1,8 +/- 0,3	2,0 +/- 0,3	1,9 +/- 0,3

* = autres groupes internes par rapport au taxon 1.

3) Discussion

Détermination spécifique

Les groupements entre échantillons sont en accord avec les déterminations craniologiques. Un nom d'espèce a pu être proposé pour la plupart des échantillons, mais plusieurs spécimens (Kor SP 10502, 10504, 10507, 10514, Sag S002, Cfp GA87 et GA66, KiS F2745 et F2672)

pourraient correspondre à des espèces nouvelles. La détermination de *H. vulcanorum*, appartenant au complexe d'espèces *H. denniae*, n'est pas certaine, car elle est basée sur un seul crâne, ce pourquoi nous n'avons pas gardé le spécimen concerné pour retracer la phylogénie du genre. Deux noms d'espèces ont été (provisoirement) réhabilités : *H. simus* et *H. kaimosae*, qui avaient été mis en synonymie respectivement avec *H. alleni* et *H. stella* (Musser et Carleton, 1993). Notons que *H. kaimosae* et *H. nsp1* correspondent à un complexe d'espèces qui nécessiterait une révision approfondie. Il est probable que les échantillons d'*Hylomyscus stella* analysés correspondent en réalité à deux espèces distinctes. Comme l'holotype de l'espèce provient de l'Est du Zaïre, il convient de conserver ce nom pour la forme orientale. Par la suite, nous parlerons de *H. stella 1* pour désigner la forme orientale et de *H. stella 2* pour celle d'Afrique centrale.

Ces déterminations nous permettent d'établir la représentation minimale des espèces dans les différentes localités de collecte (tab. 3.4, cf. fig. 2.6 pour la position géographique des localités). Nous avons identifié au maximum quatre espèces sympatriques, en comptant *H. fumosus*. Ce maximum a été obtenu uniquement au Parc National d'Odzala, soit dans une des deux localités qui ont fait l'objet des programmes de piégeage les plus importants en terme de zone géographique couverte et de durée de l'étude. L'espèce présente dans le plus grand nombre de localités, *H. stella*, est également la mieux représentée en nombre d'individus. Les espèces présentes dans une seule localité sont des espèces d'Afrique de l'Est ou de l'Ouest, provenant de régions peu ou pas échantillonnées par notre équipe. Leur occurrence unique n'est donc pas un indice de rareté. Au vu des résultats, il est possible d'étendre la distribution de certaines espèces par rapport à la synthèse de Musser et Carleton (1993). Ainsi, *H. alleni*, *H. parvus*, *H. stella* et *H. fumosus* sont présentes au Congo, et *H. alleni* est également présente en République Centrafricaine.

Tableau 3.4 : Localités de présence des espèces d'*Hylomyscus* analysées. Les abréviations font référence aux localités du tableau 3.1.

Espèce	Ba b	Cf p	Dj a	Fa b	Ge o	K ak	Ki k	Ki l	Ki N	Ki S	K or	K ou	M ak	M al	M be	M bi	M uf	M vo	N go	O dz	Sa l	Ta i
<i>H. aeta</i>		x	x				x							x						x		
<i>H. alleni</i>	x	x	x		x						x							x	x	x		
<i>H. kaimosae</i>															x	x	x					
<i>H. parvus</i>				x				x												x		
<i>H. simus</i>																						x
<i>H. stella 1</i>						x			x													
<i>H. stella 2</i>			x	x	x			x				x	x	x					x	x	x	
<i>H. fumosus</i>		x		x								x										

Comparaison des différentes méthodes d'analyse

Les arbres obtenus par les méthodes NJ et WNJ, toutes deux basées sur la ressemblance globale, sont semblables dans deux cas sur trois (fig. 3.3 A et fig. 3.5 A). Dans le cas du cytochrome *b*, ils diffèrent pour des branches très courtes et peu supportées (fig. 3.4 A et B). Les analyses de parcimonie identifient le plus souvent plusieurs arbres plus parcimonieux, dont le consensus à la majorité est peu résolu (fig. 3.3 C et fig. 3.4 C). Ce manque apparent de résolution pourrait traduire un manque d'informativité des données. D'ailleurs, l'analyse simultanée des deux séquences décèle un seul arbre plus parcimonieux (fig. 3.5 B). En ce qui concerne les analyses ML, l'agencement des groupes externes va à l'encontre de ce qui est connu de la phylogénie des muridae dans deux cas sur trois (fig. 3.4 D et 3.5 C). Ceci pourrait être dû à un mauvais positionnement des longues branches inhérent à la méthode, et aurait pu être évité en n'incluant pas dans l'analyse les groupes externes utilisés pour la datation. Pour chacun des jeux de données, les arbres NJ et WNJ diffèrent des arbres MP, qui sont eux-mêmes différents des arbres ML. Malgré les disparités entre méthodes, les groupes bien supportés sont les mêmes quelle que soit la méthode d'analyse, à de rares exceptions près.

La prise en compte des indels pour l'ARN 16s n'a pas eu d'influence sur la reconstruction de l'arbre par la méthode des distances, mais a entraîné la réunion de *H. stella*, *H. simus* et *H. allenii* par la méthode MP. Ce groupement pourrait avoir une valeur phylogénétique puisqu'on le retrouve dans les arbres ML construits pour le cytochrome *b* et pour la combinaison des deux séquences (fig. 3.4 D 3.5 C). Les indels sont une source d'information supplémentaire qu'il semble avantageux d'inclure dans les analyses. Malheureusement, seule la méthode MP permet de les prendre en compte sans codage particulier. Pour la méthode NJ, leur codage empêche d'appliquer une correction aux distances, et les programmes conçus pour employer la méthode ML ne traitent pas les indels pour l'instant.

La pondération des positions des codons pour le cytochrome *b* n'a pas eu d'impact dans le cadre de cette étude (fig. 3.4 C), mais nos résultats ne permettent pas de conclure sur l'effet de la pondération du fait du peu de résolution de l'arbre consensus obtenu. En revanche, la comparaison entre les arbres MP et ML permet d'appréhender l'effet d'une pondération différentielle des positions selon quatre classes de substitution. D'une manière générale, l'emploi d'une pondération en parcimonie est un procédé arbitraire car il s'effectue par classe de caractères (Siddall, 1997). Dans le cas du cytochrome *b*, il a été montré que l'attribution de poids aux différentes positions des codons est à proscrire, car les positions les plus variables ne sont pas forcément les plus saturées (Jansa *et al.*, 1999).

Valeur phylogénétique des séquences

L'ARN 16s apparaît moins saturé que le cytochrome *b* (fig. 3.2). L'agencement des groupes externes trouvé pour ce dernier, non conforme à ce que l'on connaît de la phylogénie des muridés, pourrait résulter de la saturation des substitutions.

Dans la mesure où le traitement séparé des deux séquences d'ADN mitochondrial a donné lieu à des arbres globalement compatibles (du moins pour les branches supportées), il est vraisemblable que l'analyse combinée conduise à des arbres plus proches de la phylogénie réelle des espèces que les analyses séparées, puisque le nombre de caractères est alors plus grand. Dans ce cas, les espèces *H. stella* 1 et 2 pourraient être des espèces sœurs (fig. 3.5 A et B, non contredit par fig. 3.5 C), de même que *H. allenii* et *H. simus* (fig. 3.5 B et C, non contredit par fig. 3.5 A).

Phylogénie et datation

Deux arbres consensus peuvent être proposés, l'un reprenant les relations phylogénétiques mises en évidence dans la majorité des analyses, et l'autre, plus résolu, prenant en compte les relations additionnelles discutées dans le paragraphe qui précède (fig. 3.5 D, respectivement en traits pleins et en pointillés). Les deux arbres confirment la proximité entre les espèces *H. allenii*, *H. simus* et *H. stella* (Rosevear, 1969), et rapprochent ces espèces de *H. parvus*. Si l'arbre le plus résolu est valide, alors nos résultats sont compatibles avec les trois groupes d'espèces définis par Robbins et collaborateurs (1980). *H. aeta*, qui est le seul représentant de son groupe dans notre étude, se distingue de toutes les autres espèces. Le groupe de *H. allenii* et *H. stella* comprendrait aussi *H. simus*. Le groupe de *H. parvus* ne serait pas monospécifique, mais inclurait un complexe d'espèces.

L'ensemble des analyses conforte le maintien de l'espèce *fumosus* dans un genre distinct. En effet, même si sa position n'est pas soutenue par des valeurs de bootstrap élevées, cette espèce est toujours la plus basale du groupe. De plus, dans certaines analyses, *Praomys delectormum* s'insère entre *H. fumosus* et les autres *Hylomyscus* (fig. 3.1 et 3.3 C et D). Nos résultats ne permettraient pas à eux seuls d'affecter cette espèce à un genre distinct, mais la reconnaissance du genre *Heimyscus*, qui a été proposée sur la base de caractères morphologiques (Misonne, 1969), est aussi supportée par des caractères chromosomiques (Robbins *et al.*, 1980).

Les datations proposées sont vraisemblablement surestimées en raison de la grande différence des taux de substitutions entre les taxons utilisés pour calibrer l'horloge moléculaire et ceux dont on cherche à estimer les temps de divergence (tab. 3.3). En effet, la saturation des substitutions, qui n'est que partiellement corrigée par la formule de Tamura-Nei, est d'autant plus importante que la divergence est ancienne, ce qui implique une sous-estimation des taux de substitution. Nous proposons donc de

donner un plus grand crédit aux datations basées sur la divergence entre *Mus* et *Rattus*, qui engendrent les estimations les plus faibles. L'âge du genre *Hylomyscus* peut être approché par la divergence entre *Heimyscus fumosus* et les différentes espèces d'*Hylomyscus*. Il serait d'environ 4,1 MA (+/- 0,4), ce qui est un peu plus ancien que les 3,3 MA qui ont été proposé sur la base d'hybridation entre molécules d'ADN (Chevret *et al.*, 1994). Le complexe d'espèces *H. stella* 1 et 2 daterait d'environ 1,9 MA (+/- 0,3) et la divergence entre ces deux taxons de 1,8 MA (+/- 0,3). Ces divergences seraient beaucoup plus anciennes que les divergences entre espèces sœurs appartenant aux genres *Praomys* et *Mastomys*, qui se situeraient entre 0,3 et 1,0 MA (Chevret *et al.*, 1994).

Evolution des caryotypes

Les caryotypes des espèces pour lesquelles ils ont été déterminés sont indiqués sur l'arbre consensus (fig. 3.5 D). Trois des espèces regroupées par nos analyses, *H. alleni*, *H. parvus* et *H. stella*, ont des caryotypes voisins (2N = 46 et NF = 68, 70, 72 ou 74 ; Robbins *et al.*, 1980). L'espèce *H. aeta*, qui se distingue des autres dans notre analyse, a un caryotype bien distinct (2N = 54 et NF = 86). Les analyses génétiques et caryologiques sont donc cohérentes.

Il nous paraît important de noter que l'échantillon T0477, pour lequel le nom *H. vulcanorum* est proposé, nous a été envoyé comme spécimen de *H. stella*, récolté au Burundi par des membres de l'IZEA sous la référence 2703. Au vu de cette erreur de détermination, il est envisageable que le spécimen IZEA 2743 utilisé pour caryotyper l'espèce *H. stella* au Burundi (Maddalena *et al.*, 1989) n'appartienne pas non plus à l'espèce *stella*. Ceci expliquerait que le caryotype trouvé (2N = 48 et NF = 86) soit fort différent de celui des représentants centrafricains et camerounais de l'espèce.

Cette étude a permis d'améliorer la détermination spécifique au sein du genre *Hylomyscus* et d'éclaircir les relations phylogénétiques entre espèces. Elle a également permis de choisir comme modèle pour les analyses phylogéographiques une espèce bien représentée sur l'ensemble des sites échantillonnés : *H. stella*. Cependant, nous émettons une réserve quant à la conspécificité des formes présentes dans l'Est et dans l'Ouest de la zone échantillonnée, aussi nous les appellerons respectivement *H. stella* 1 et 2.

CHAPITRE 4: PHYLOGEOGRAPHIE

Dans ce chapitre, nous abordons l'histoire évolutive de deux modèles de musaraignes, deux modèles de rongeurs et une super-espèce de primates. Nous avons confronté les schémas phylogéographiques obtenus dans le but de rechercher d'éventuels événements communs et d'identifier les différences qui les opposent.

I - PHYLOGEOGRAPHIE COMPAREE DE DEUX MUSARAIGNES

1) Phylogéographie comparée de *Sylvisorex johnstoni* et *S. ollula* (Article 2)

Titre : Diversification génétique de deux musaraignes africaines forestières, *Sylvisorex johnstoni* et *S. ollula* (Soricidae, Insectivora) en relation avec les changements paléo-environnementaux.

Quérrouil, S., Verheyen, E., Dillen, M. et Colyn, M. (soumis à *J. Evol. Biol.*) Patterns of diversification in two African forest shrews : *Sylvisorex johnstoni* et *S. ollula* (Soricidae, Insectivora) in relation to paleo-environmental changes.

Résumé : Nous avons cherché à évaluer l'impact du choix des espèces sur le résultat des tests de scénarios biogéographiques, en prenant comme modèles deux espèces de musaraignes forestières d'Afrique centrale, syntopiques et congénériques, mais de taille différente. Nous avons séquencé une portion d'ARNr 16s pour 42 spécimens de *Sylvisorex johnstoni* (**doc. 2.1**) et 40 spécimens de *S. ollula* (**doc. 2.2**), en provenance de 11 localités d'Afrique centrale. Nous avons construit des arbres d'étendue minimale, et procédé à une datation des principaux événements de divergence, en adaptant l'horloge moléculaire calibrée pour les musaraignes africaines, à des taxons ayant divergé récemment. Nous avons comparé les schémas phylogéographiques des deux espèces entre eux et avec un cladogramme théorique des aires établi sur la base de scénarios biogéographiques obtenus par notre équipe.

Les datations réalisées indiquent que les deux espèces seraient d'origine Plio-Pléistocène, mais les temps de divergence estimés pour *S. johnstoni* remontent plus loin dans le temps que pour *S. ollula* (environ 3,2 contre 1,3 MA). Certains événements de divergence coïncident avec des périodes de sécheresse du Pléistocène, ce qui laisse supposer un rôle de la fragmentation forestière dans la différenciation génétique des populations. Pour chacune des espèces, la comparaison des arbres

phylogénétiques à des arbres contraints de façon à être compatibles avec le cladogramme théorique des aires, indique que leur histoire diffère de l'histoire générale des aires. Cependant, les résultats confirment l'existence probable d'un refuge forestier dans chacune des principales régions biogéographiques, étayant ainsi l'hypothèse du rôle primordial joué par les refuges et les barrières fluviales.

Les schémas phylogéographiques obtenus pour les deux espèces diffèrent fortement. Bien que *Sylvisorex johnstoni* soit présente dans tout le bloc forestier, on observe une forte ségrégation géographique des haplotypes, qui pourrait traduire l'existence d'un complexe d'espèces allopatriques. À l'inverse, *S. ollula* aurait colonisé récemment la majeure partie de la zone d'étude. La distribution géographique des haplotypes indique soit que les deux espèces n'ont pas évolué dans le même cadre géographique et temporel, soit qu'elles ont répondu différemment aux mêmes changements environnementaux. Dans cette seconde alternative, les différences observées pourraient être imputables aux capacités de survie et de dispersion de chaque espèce. Comme ces capacités sont liées aux besoins écologiques et aux aptitudes physiques de l'espèce, elles pourraient être influencées par la taille moyenne des individus. *S. johnstoni* étant beaucoup plus petite que *S. ollula*, ses capacités de dispersion pourraient être moindres, et ses chances de survie dans des îlots forestiers de petite taille pourraient être supérieures à celles de *S. ollula*. Des études éco-éthologiques et une comparaison entre séquences nucléaires et mitochondriales permettraient d'approfondir la question.

Les différences phylogéographiques observées entre deux espèces de musaraignes syntopiques et congénériques (mais phylogénétiquement assez éloignées, cf. chapitre 3.I) incitent à ne pas tirer hâtivement des conclusions biogéographiques sur la base d'études phylogéographiques portant sur un petit nombre de taxons. Elles montrent qu'il convient de choisir comme modèle des espèces dont les caractéristiques écologiques, le potentiel de dispersion, et les paramètres démographiques sont relativement bien connus et comparables.

Patterns of Diversification in two African Forest Shrews: *Sylvisorex johnstoni* and *S. ollula* (Soricidae, Insectivora) in Relation to Paleo-environmental Changes

Sophie Quérrouil¹, Erik Verheyen², Martine Dillen² and Marc Colyn¹

¹ Laboratoire Ethologie, Evolution, et Ecologie, Station Biologique, CNRS-U.M.R. 6552, Université de Rennes, F-35380 Paimpont, France;

² Royal Belgian Institute of Natural Sciences, Section of Biochemical Systematics and Taxonomy, Vautierstraat 29, B-1000 Brussels, Belgium

Running title: Phylogeography of African forest shrews

Abstract

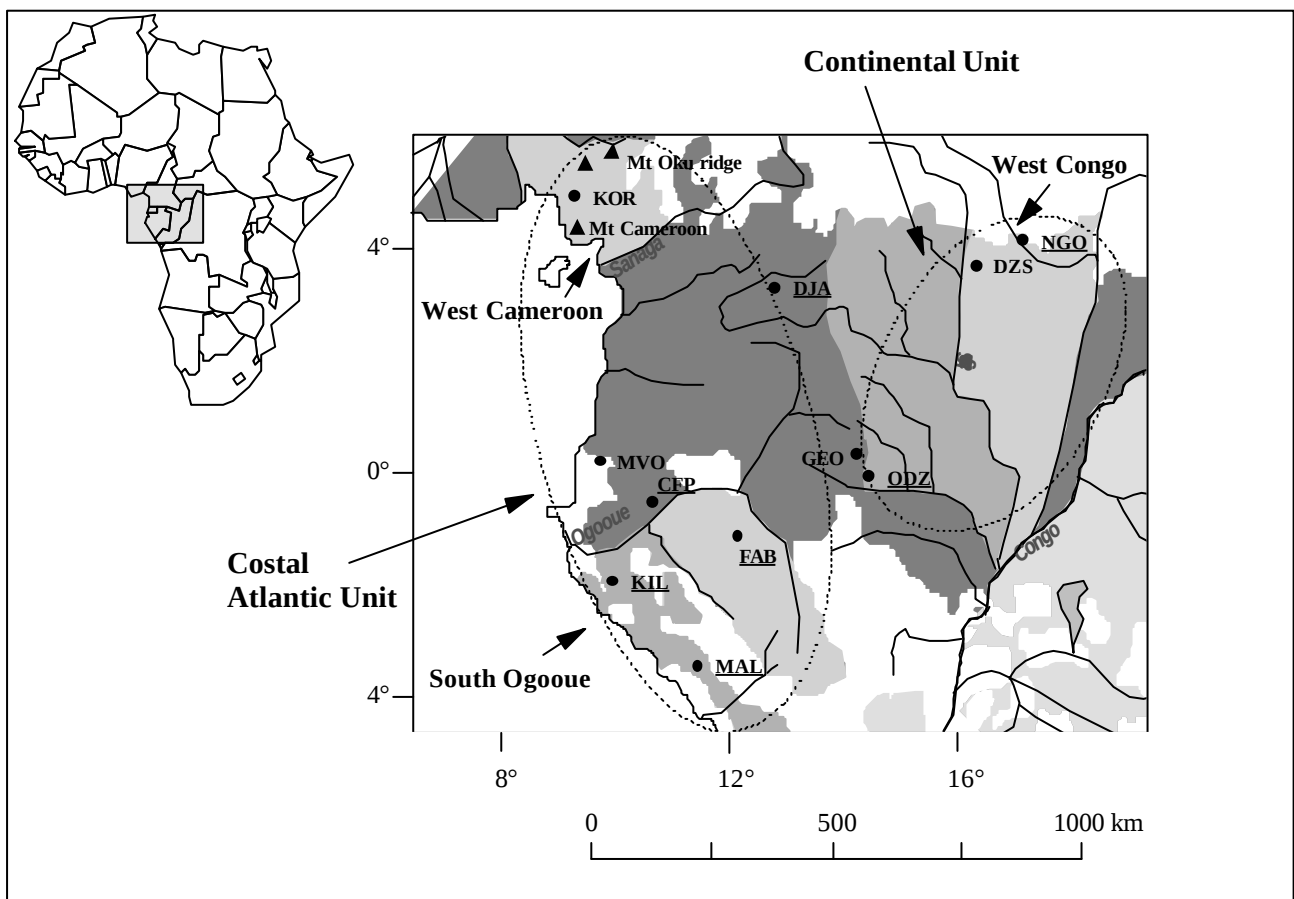
We evaluated the impact of species choice on the outcome of tests of biogeographical scenarios proposed for the Central African tropical forest. We compared the phylogeographic patterns of two congeneric and syntopic species of forest shrews, inferred from a total of 82 partial 16s rRNA mitochondrial sequences. Our data suggest that both species must have originated in the Plio-Pleistocene and their haplotype distributions show traces of the effect of forest fragmentation and expansion associated with the paleoclimatic fluctuations of the Pleistocene. However, it seems that the two species responded very differently to similar environmental changes. While *S. johnstoni* populations exhibit ancient haplotype segregation that may even represent currently unrecognized allopatric species, *S. ollula* haplotypes are much less differentiated and suggest that this taxon has undergone a recent range expansion. The observed differences between these taxa may be explained by their presumably different ecological requirements and colonisation abilities, which in turn may be the result of a significant difference in body size. In conclusion, it appears necessary to incorporate well-documented ecologically divergent species in studies that attempt to infer evolutionary processes from phylogeographic patterns.

Keywords: phylogeography; refuges; Central Africa; Forest shrews; *Sylvisorex*; mitochondrial sequences; macro-evolution

Introduction

The available information on global faunal diversity demonstrates that biodiversity is heterogeneously distributed over the world, both among and within the continents (e.g. Brown, 1988). Even within an apparently continuous stretch of habitat, such as a forest, species ranges are often restricted. Overlap of species range boundaries have led to the recognition of faunal regions where the fauna is relatively homogenous (Wallace, 1876). For tropical Africa, early studies, based on the distribution ranges of butterflies and birds, distinguished only two faunal regions centred around mountains (Carcasson, 1964; Moreau, 1969). Later studies, on the Central African lowland forest fauna, distinguished a higher number of faunal regions (e.g. Bigalke, 1978; Prigogine, 1988). More recent analyses based on the distribution ranges of mammals (primarily primates) and insects revealed a considerably more complex pattern of small regions separated by buffer zones (i.e. intergradation areas) where faunas from adjacent regions meet (Colyn, 1999; Deleporte & Colyn, 1999). For the Western part of Central Africa, the latter studies define five faunal sub-regions, grouped into three faunal regions and two faunal units (Fig. 1).

Fig. 1 Map of Central Africa showing the two faunal units (circled in dashed lines) and the embedded three faunal regions (in grey) and five sub-regions (two levels of grey), together with the intergradation area (oblique bars) and forest cover limit (dots). Collecting localities are abbreviated as in table 1. Localities where both species were collected are underlined, whereas localities where only *S. johnstoni* was collected are not.



The distribution ranges are determined by ecological and historical factors. The complex river system is presumed to be the major ecological factor influencing the distribution of forest taxa (Grubb, 1990; Happold, 1996; Deleporte & Colyn, 1999). Actually, except for the presence of extensive river systems, the Central African tropical forest is rather homogenous. The topography of this region is only interrupted by Mount Cameroon and the Mount Oku ridge which presumably do not act as barriers for the lowland forest fauna given that they are separated by a wide lowland forest corridor. Climatic variations in this region are characterised by local differences in rainfall abundance and seasonal fluctuations. Finally, the forest cover is presently uninterrupted, but differs in composition between coastal and continental areas (White, 1983).

The evolution of the forest faunas has been shaped by geomorphological events. Among these are the fact that a water body that flooded the Congo basin found its way to the sea about 400.000 years ago (Colyn, 1991), and the successive phases of forest fragmentation and expansion related to the climatic oscillations of the Pleistocene (e.g. Maley, 1996; Jahns et al., 1998). The refuge theory (Haffer, 1969) assumes that, in tropical areas, during very arid phases, some conserved lowland forest patches may have acted as «refuges» for the forest fauna which had the opportunity to diverge and eventually speciate in isolation (Haffer, 1969, 1982; Mayr & O'Hara, 1986). In these refuges, the cessation of gene flow may have led to genetic drift resulting in genetic divergence between allopatric populations (Mayr, 1963). During wet periods, forest expansion gave way to colonisation events possibly associated with hybridisation or switches to vacant niches (Grubb, 1978). This theory was first formulated to explain the high species diversity in the tropics, but it can also apply to intra-specific differentiation, in which case reproductive isolation is not required. It has often been evoked to explain faunal differentiation in African birds (Diamond & Hamilton, 1980; Mayr & O'Hara, 1986; Louette, 1992) and mammals (Booth, 1958; Grubb, 1990; Colyn, 1991). A number of refuges has been postulated for the Central African forest (cf. synthesis in Maley, 1996) and a major fluvial refuge along the Congo river has been hypothesised (Colyn et al., 1991). However, the refuge theory remains controversial (cf. Moritz et al., 2000).

The actual role of evolutionary processes suggested by biogeographic scenarios can be investigated by a phylogeographic approach, through the analysis of the geographic distribution of polymorphism within separate lineages (Avice et al., 1987). The development of molecular techniques has provided the opportunity to compare intra-specific phylogenies of co-distributed species based on mitochondrial-DNA sequences (Avice, 1992). The strength of this approach has been illustrated by comparative phylogeographic studies which investigated the impact of past habitat fragmentation on South American, Australian and European faunas (e.g. Schneider et al., 1998; Taberlet et al., 1998;

Ditchfield, 2000; Patton et al., 2000). However, only a few studies have attempted to apply this approach to tropical African faunas. These studies focused on mountain and lowland forest birds (Roy, 1997; Beresford & Cracraft, 1999) and savanna-dwelling bovids (Arctander et al., 1999).

This study investigates the potential biogeographic scenarios proposed for the Central African lowland forest fauna by testing the congruence between phylogeographic patterns obtained for two species of forest shrews and the faunal regions previously recognised (Colyn, 1999; Deleporte & Colyn, 1999). As earlier studies have mainly focused on colourful birds and large mammals such as primates, we have chosen to study small mammals. Their diversity is less apparent but yet important. Actually, non-flying small mammals account for a third of the specific diversity of mammalian communities in the African tropical forest: 24-25% for rodents and 8-10% for insectivores (Happold, 1996). The drawback with small mammals is that their taxonomy is often uncertain. However, in the present case, a phylogenetic revision of African shrews has been conducted prior to the phylogeographic study (Quérrouil et al., 2001). Shrews were selected because they constitute a rather old taxon that was probably well diversified at the beginning of the Pleistocene, and because they are considered to be very sensitive to environmental fluctuations, primarily the temperature and humidity (Reumer, 1989). We selected two congeneric species: *Sylvisorex johnstoni* (Dobson, 1887) and *S. ollula* Thomas, 1913, which are both insectivorous, terrestrial and typical forest-dwellers, but differ in body size. We inferred phylogeographic patterns from partial 16s rRNA sequences, using the same gene fragment as for the phylogenetic study of African shrews (Quérrouil et al., 2001), and cautiously dated divergence events according to the molecular clock calibrated in that study. Comparison of the phylogeographic patterns obtained for the two model species may reveal discrepancies resulting from different evolutionary histories.

Materials and Methods

Study Area

The study area is situated between the 6th N and 5th S parallels and the 7th E and 19th E meridians (Fig. 1). It consists of a portion of the Central African forest where intricate biogeographical patterns have been described as a result of intensive field-work. Five sub-regions, grouped into three regions and two biogeographical units are recognised, in addition to an area of intergradation. Collections were made at eleven sites in the lowland forest (Table 1).

Table 1. Number of specimens sequenced per species and locality, with geographical coordinates of the localities.

Locality code	Locality	Country	Faunal region	Coordinates		<i>S. johnstoni</i>	<i>S. ollula</i>
				Latitude	Longitude		
Kor	Korup N.P.	Cameroon	W. Cameroon	5°16'N	09°08'E	2	0
Dzs	Dzanga-Sangha	CAR	West Congo	2°47'N	16°25'E	1	0
Ngo	Ngotto	CAR	West Congo	4°02'N	17°20'E	6	10
Odz	Odzala N.P.	Congo	West Congo	0°26'N	14°41'E	10	12
Fab	Forêt des Abeilles	Gabon	South Ogooue	0°47'S	11°56'E	2	2
Kil	Kili	Gabon	South Ogooue	1°35'S	10°10'E	3	2
Mal	Malounga	Gabon	South Ogooue	3°09'S	10°45'E	3	2
Cfp	Four-Place Quarry	Gabon	Intergradation	0°15'S	10°20'E	3	2
Dja	Dja reserve	Cameroon	Intergradation	3°20'N	13°02'E	6	10
Geo	Great Escarpment of Odzala	Congo	Intergradation	1°04'N	14°29'E	2	0
Mvo	Mvoum	Gabon	Intergradation	0°35'N	09°53'E	4	0
Total						42	40

Biological Material

We selected two congeneric species of shrews (Soricidae, Insectivora): the pygmy forest musk shrew, *Sylvisorex johnstoni* and the greater forest musk shrew, *S. ollula*. The two species are morphologically distinguishable (cf. Hutterer, 1986; Dieterlen & Heim de Balsac, 1979), widely distributed and abundant. Both species are typical forest-dwellers that live in the litter, from deep forest to the forest edge. They are rarely caught in degraded habitat. *S. johnstoni* is the smallest species of the genus that occurs in Central Africa (average weight: 2-3g), while *S. ollula* is the largest *Sylvisorex* species (average weight: 15-20g). The species co-occur in Cameroon, the Central African Republic, Equatorial Guinea, Gabon, and the Republic of Congo. In these countries, *S. johnstoni* accounts on average for 25% of the forest shrew community and *S. ollula* for 10% (P. Barrière, pers. comm.), and the distribution of the two species are presumably continuous. The distribution range of *S. johnstoni* extends east to Burundi, Tanzania and Uganda whereas the range of *S. ollula* extends west to Nigeria and south to the left bank of the Congo river, in the Democratic Republic of Congo.

A total of 42 specimens of *S. johnstoni* and 40 specimens of *S. ollula* were selected which covered as many trapping sites at each locality as possible. Depending on trapping success, between 0 and 12 specimens were analysed per trapping locality (Table 1). *S. johnstoni* was captured at all trapping sites in the main forest block, while *S. ollula* was not captured in localities where the trapping effort was lower (Fig. 1). The absence of capture in one locality does not mean actual absence of the

species from that locality. Sequences were deposited in GenBank, under accession numbers AF252505 to AF252534 and AY017312 to AY017323 for *S. johnstoni*, and AF253941 to AF253967 and AY017324 to AY017336 for *S. ollula*. Two outgroups were chosen among published sequences of crocidurine shrews in a way that the two model species were monophyletic relative to the outgroups: *Myosorex geata* AF274560 and *Congosorex nsp.* AF274559 (Quérrouil et al., 2001).

Laboratory Methods

DNA was isolated from ethanol-preserved muscle or liver tissues by standard proteinase K digestion, followed by phenol-chloroform extraction and ethanol precipitation (Sambrook et al., 1989) or by the Chelex method (Walsh et al., 1991).

The studied 16s rRNA gene fragment comprises much of the second half of the 16s rRNA gene. After alignment, it is 537 bp long for *S. johnstoni* and 534 bp long for *S. ollula*.

The primers used for amplification and sequencing were the 16Sar L-primer, 5'-CGCCTGTTTAACAAAAACAT-3' (Palumbi et al., 1991) and a new primer named 16S-Hm, 5'-AGATCACGTAGGACTTTAAT-3'. The PCR consisted of 30 cycles: 60 s at 94°C, 60 s at 49°C and 90 s at 72°C. The samples were sequenced with both primers using an automated Alf DNA Sequencer (Amersham-Pharmacia) and an Amersham sequencing kit, following the manufacturer's protocol. The cycle-sequencing reaction consisted of 30 cycles: 36 s at 94°C, 36 s at 52°C for the L-primer or 55°C for the H-primer and 80 s at 72°C.

Genetic Diversity

The alignment was performed with ClustalW version 1.5 (Thompson et al., 1994) and corrected by eye. It required only six point insertions or deletions for *S. johnstoni* and none for *S. ollula*. Fifteen nucleotides (the first six and last nine) were excluded prior to analyses because they were missing from several sequences: 522 nucleotides were analysed for *S. johnstoni* and 519 for *S. ollula*. For all distance measurements and analyses, the distances were corrected by the Kimura-2 parameter formula (Kimura, 1980) which allows for differential transition and transversion rates. More sophisticated corrections were deemed unnecessary given the relatively low level of substitutions and saturation. Intra- and inter-locality genetic distances were compared graphically by means of a box plot and analysed statistically by a Mann-Whitney U-test (Sokal & Rohlf, 1981). For each species, we determined the mean genetic distances among haplotypes within each locality and overall.

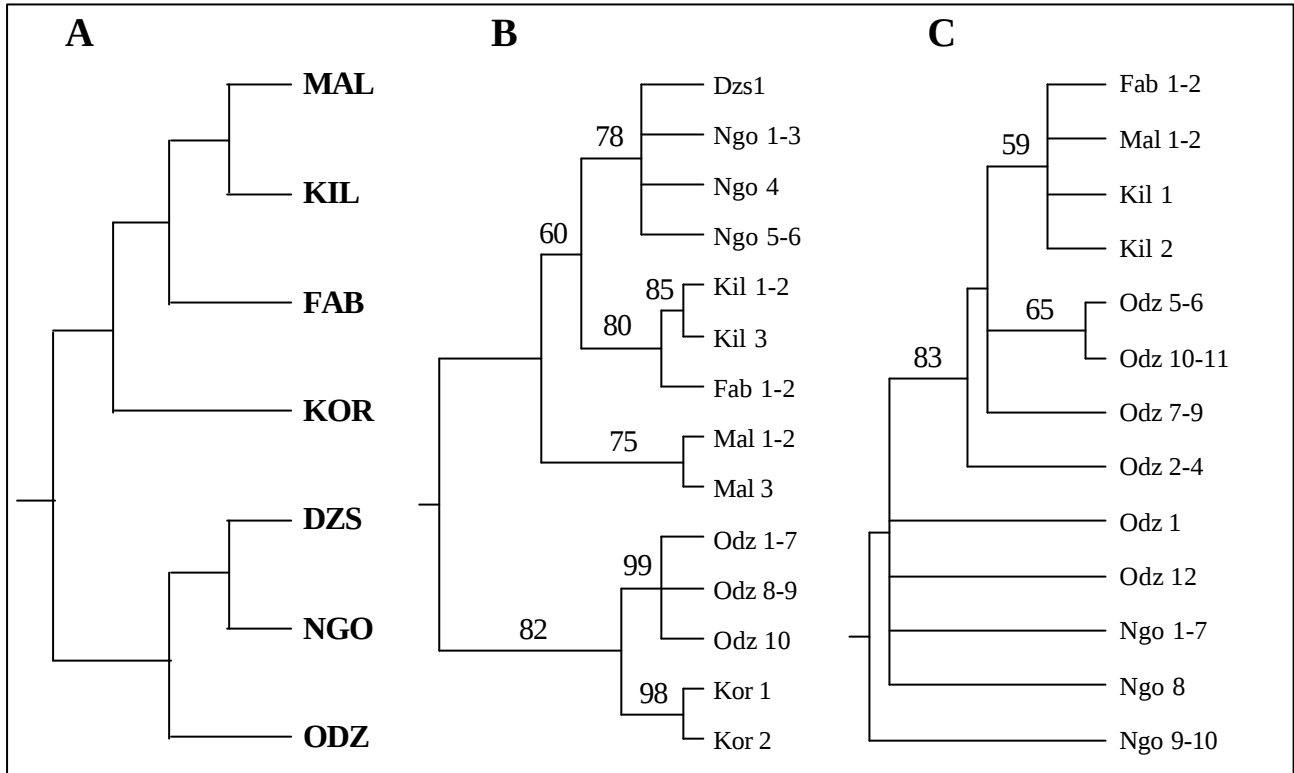
Phylogenetic Analyses

As phylogenetic relationships were evaluated below the species level, it seemed reasonable to assume that the common ancestors of the haplotype lineages were not necessarily extinct, and to expect a multifurcating tree (Crandall, 1996). In addition, when the number of informative characters is low, as it is the case for one of the two species, traditional phylogenetic methods perform poorly (cf. Crandall, 1996). Therefore, minimum spanning trees (MST) rather than phylogenetic trees were constructed with MinSpNet (Excoffier, 1993). In these trees, we call "intermediate haplotypes", the haplotypes differing by one substitution from one another and from sampled haplotypes, which were not uncovered but potentially exist either in sampled populations or in localities which were not sampled.

Test of Phylogeographical Hypotheses

If individual species were shaped by the same processes which have shaped faunal assemblages, genetic variation within a species could reflect the relationships between faunal areas. To test for congruence between the sequence data and the faunal regions (Fig. 1), we constructed a theoretical area cladogram based on the definition of the faunal units, regions and sub-regions and their presumed hierarchical relationships (Fig. 2A). Within sub-regions, the fauna is considered as rather homogenous. The faunas from two sub-regions within a region are presumably more alike than faunas from different regions, and the faunas from two regions within a faunal unit more alike than faunas from distinct units. Localities from the 'intergradation region' were excluded from the area cladogram given that this region is expected to be colonisable from any adjacent region. Therefore, the sequences from the 'intergradation region' were pruned from the data-sets prior to analysis. The theoretical area cladogram was then used to build a constrained maximum parsimony (MP) tree for each species using PAUP 4.0b1 (Swofford, 1997). An unconstrained MP tree was built for comparison. Heuristic searches for the most parsimonious trees were conducted using 10 replicates of random addition of taxa, with a uniform weighting scheme because substitutions were apparently not saturated. Gaps were treated as missing. The robustness of internal nodes was estimated by bootstrap analyses with 1000 MP replicates, each with a single replication of random addition of taxa. Finally, we compared the parsimony indices and the likelihood values of the constrained and unconstrained MP trees by a Kishino-Hasegawa test (1989), using PAUP 4.0b1 (Swofford, 1997). For the estimation of the likelihood values, we assumed that the data followed the general time reversible model that allows for transition bias and does not assume equal base frequencies (Hasegawa et al., 1985). The rate of variability of the nucleotide sites was assumed to follow a gamma-distribution, and a part of the sites was assumed to be invariable. All parameters were estimated from the data.

Fig. 2 (Left.) Theoretical area cladogram derived from the embedded relationships between faunal sub-regions, regions and units. (Centre.) Unconstrained MP tree for *S. johnstoni*, after pruning of localities from the ‘intergradation region’, with bootstrap support (1000 MP replicates). (Right.) Unconstrained MP tree for *S. ollula*, after pruning of localities from the ‘intergradation region’, with bootstrap support (1000 MP replicates).



Age Estimations

First, we tested if the data were consistent with the molecular clock hypothesis. Substitution rate heterogeneity tests were performed with RRTree (Robinson-Rechavi & Huchon, 2000). This program compares substitution rates in a series of lineages and evaluates the probability P that two lineages are evolving at the same rate. The null hypothesis is rejected if $P < 0.05$ divided by the number of lineages. Tests were carried out within each species and between the species.

Secondly, we attempted to date the main divergence events by recalibrating the molecular clock that was used in an earlier study (Qu  rouil et al., 2001). To calibrate our new molecular clock, we used all nucleotide positions and all kinds of substitutions, and the estimated date of radiation of the genus *Crocidura*, about 5-6 million years (Myr) ago (Butler, 1998). The average value for the substitution rate (r) was 0.65% per Myr (range 0.51-0.78%). Distances between clades of haplotypes identified on a Neighbor Joining tree were computed using Phyltest (Kumar, 1996). Identical haplotypes were counted only once.

For *S. ollula*, as the phylogenetic pattern seemed to reflect a recent population expansion (see results), we tried to date this event by means of the mismatch distribution method using Arlequin 2.000 (Schneider et al., 2000). In the case of an on-going expansion of a population, the curve is unimodal (Rogers & Harpending, 1992). The time elapsed since the demographic event is proportional to the mode (τ) of the curve and can be estimated by the formula: $T = \tau / 2\mu$, where μ is the substitution rate expressed in number of substitutions per million years. Goodness of fit between the expected mismatch distribution under a sudden expansion model (using estimated parameters) and the observed distribution was tested by the generalised least-square method. Confidence intervals ($\alpha = 0.05$) were estimated by a Monte-Carlo simulation with 5000 replicates.

Results

Genetic Diversity

Among the 522 sites analysed for *S. johnstoni*, 56 are variable and 48 are parsimony informative. Among the 519 sites analysed for *S. ollula*, 25 are variable and only 8 are parsimony informative.

The transition / transversion ratio amounts to 2.5 (+/- 1.9) for *S. johnstoni* and 3.8 (+/- 3.2) for *S. ollula*. The number of transitions plotted against the number of transversions does not level off (data not shown). Hence, transitions and transversions are presumably unsaturated.

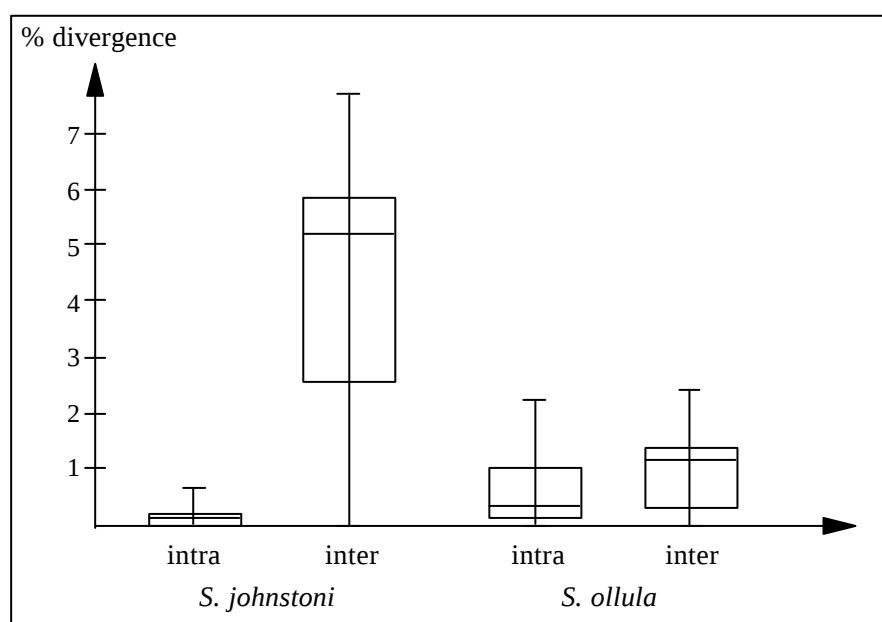
The percentage of distinct haplotypes is comparable for the two species (Table 2). There are 21 haplotypes out of 42 samples for *S. johnstoni* and 17 haplotypes out of 40 samples for *S. ollula*. Local mean genetic distances are low (around 0.1 - 0.3%), but can reach high values for *S. ollula* (1.1% in Odzala - Table 2). Maximum local distances are the highest in Ngotto for *S. johnstoni* (0.7%) and in Odzala for *S. ollula* (2.2% - Table 2).

Table 2. Number of specimens sequenced (N.seq), number of distinct haplotypes (N.hapl), Kimura 2-parameters mean genetic distance (D.mn (+/-s.e.)) and maximum genetic distance between haplotypes (D.max) per locality, for each of the two species.

Code	Faunal region	<i>S. johnstoni</i>				<i>S. ollula</i>			
		N.seq	N.hapl	D.mn (s.e.)	D.max	N.seq	N.hapl	D.mn (s.e.)	D.max
Kor	W. Cameroon	2	2	0.2 +/- 0.3					
Dzs	West Congo	1	1	--					
Ngo	West Congo	6	3	0.3 +/- 0.2	0.7	10	3	0.2 +/- 0.2	0.6
Odz	West Congo	10	3	0.2 +/- 0.2	0.2	12	6	1.1 +/- 0.6	2.2
Fab	South Ogooue	2	1	0		2	1	0	
Kil	South Ogooue	3	2	0.3 +/- 0.3		2	2	0.2 +/- 0.3	
Mal	South Ogooue	3	2	0.1 +/- 0.2		2	1	0	
Cfp	Intergradation	3	1	0		2	2	0.2 +/- 0.3	
Dja	Intergradation	6	3	0.1 +/- 0.1	0.2	10	6	0.3 +/- 0.2	0.8
Geo	Intergradation	2	2	0.2 +/- 0.3					
Mvo	Intergradation	4	2	0.1 +/- 0.1					

Inter-locality distances are comprised between a minimum of 0% and a maximum of 7.6% for *S. johnstoni* and 2.4% for *S. ollula* (Fig. 3). Inter-locality genetic distances reach much higher values than intra-locality distances for *S. johnstoni*, while the maximum values are comparable for *S. ollula* (Fig. 3). However, intra-locality distances are significantly smaller than inter-locality distance for both species ($P < 0.001$).

Fig. 3 Box plot showing intra-locality (left box) and inter-locality (right box) Kimura 2-parameters genetic distances for *S. johnstoni* and *S. ollula*. Boxes represent the 25% and 75% quartiles, horizontal lines figure the medians, and vertical bars delineate the whole ranges of distances.



The average genetic distance between all pairs of sequences is 3.6% (+/- 2.2) for *S. johnstoni* and 0.9% (+/- 0.6) for *S. ollula*. Genetic distances are significantly higher for *S. johnstoni* than for *S. ollula* ($P < 0.001$).

MST

Minimum spanning trees have been represented so that they fit the geographic location of haplotypes (Fig. 4 and 5). For *S. johnstoni*, the haplotypes from Ngotto and Dzanga could not be assigned their correct location due to their short genetic distance to haplotypes from the South Ogooue region.

The tree obtained for *S. johnstoni* has long branches with several missing haplotypes (Fig. 4). All the haplotypes from a given locality are monophyletic. If we arbitrarily distinguish clades on the basis of 1% divergence, 6 clades appear. The first clade corresponds to Odzala (West Congo region), the second one to the Dja, the Escarpment and Four-Place ('intergradation region'), the third one to Mvoum ('intergradation region'), the fourth one to Korup (West Cameroon region), the fifth one to the Forêt des Abeilles, Kili and Malounga (South Ogooue region), and the last one to Dzanga and Ngotto (West Congo region). The last two clades are geographically distant but genetically rather close.

In contrast, the tree obtained for *S. ollula* is star-shaped, and haplotypes from a given locality do not necessarily cluster (Fig. 5). Three clades can be distinguished on the basis of 1% divergence. The first clade joins all haplotypes from the Dja, Four-Place and Ngotto (intergradation and West Congo region) with haplotypes Odz 1 and Odz 12 from Odzala, the second one consists only of haplotype Odz 2-4 from Odzala, and the third one joins all haplotypes from the Forêt des Abeilles, Kili and Malounga (South Ogooue region) with the remaining haplotypes from Odzala (Odz 5-6, Odz 7-9, Odz 10-11). Samples from Odzala are found in the three clades. The first clade contains a widely-distributed dominant haplotype and several derived haplotypes which differ from it by 1 to 3 substitutions.

Fig. 4 Minimum spanning tree for *S. johnstoni*. Black dots represent original haplotypes and their size is proportional to the observed haplotype frequency. Open circles represent potential intermediate haplotypes that were not detected. Haplotypes that differ by less than 5 substitutions (1%) are grouped by square boxes.

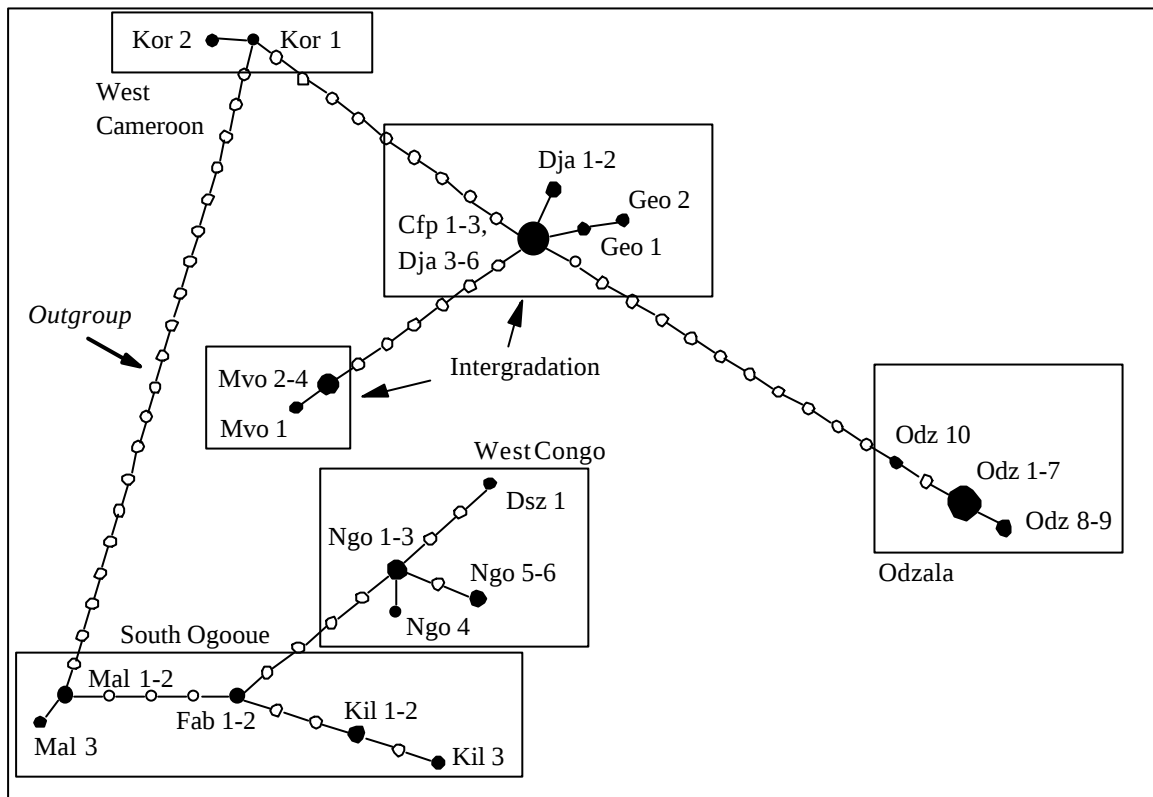
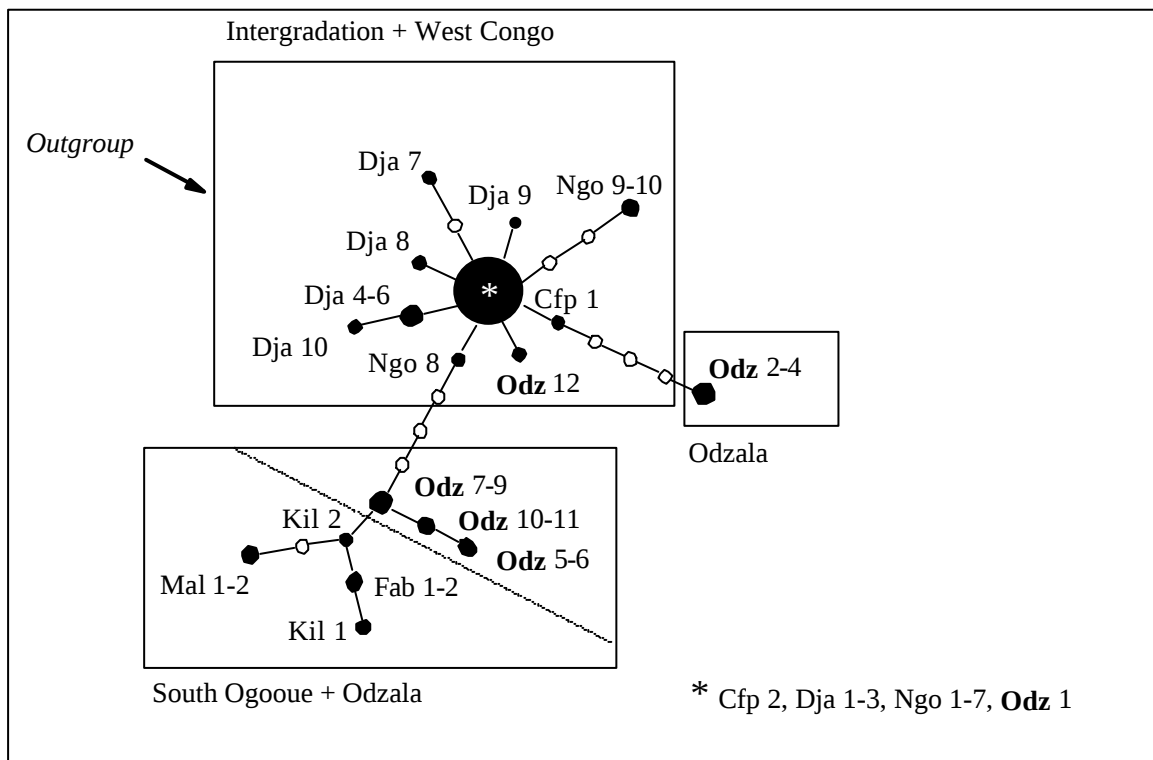


Fig. 5 Minimum spanning tree for *S. ollula*. Symbols and annotations are the same as for Fig. 4.



Test of Evolutionary Scenarios

After samples from the ‘intergradation region’ are removed from the data set, the shortest constrained MP tree for *S. johnstoni* is 149 steps long. It is significantly longer than the shortest unconstrained MP tree (Fig. 2B) which is 131 steps long ($P < 0.001$). The likelihood values also differ significantly ($-\ln L = 1512.3$ versus 1437.2 , $P < 0.001$). For *S. ollula*, the shortest constrained MP tree is 82 steps long. It is significantly longer than the shortest unconstrained MP tree (Fig. 2C) which is 78 steps long ($P = 0.045$). The likelihood values also significantly differ ($-\ln L = 1174.6$ versus 1146.8 , $P = 0.021$). Therefore, the unconstrained phylogenies for both species appear to differ significantly from phylogenies constrained to accommodate the faunal region model.

Age Estimations

For each species considered separately, the relative rate tests reveal no rate difference among branches (all P -values > 0.05). The test conducted to compare the two species reveals that the 16 rRNA fragment has evolved about 1.4 times faster for *S. johnstoni* ($K = 0.119$) than for *S. ollula* ($K = 0.087$). Consequently, we corrected the r -values in such a way that $r_{johnstoni} = 1.4 * r_{ollula}$. Accordingly, the derived r -values are $r = 0.76$ for *S. johnstoni* and $r = 0.54$ for *S. ollula*. Age estimates ranged between 1.1 and 3.2 Myr for *S. johnstoni* and between 1.2 and 1.3 Myr for *S. ollula* (Table 3).

Table 3. Estimated ages of divergence between clusters of haplotypes.

<i>S. johnstoni</i>			<i>S. ollula</i>		
Cluster 1	Cluster 2	Age (Myr +/- s.e.)	Cluster 1	Cluster 2	Age (Myr +/- s.e.)
W. Cameroon,	S. Ogooue,	3.2 +/- 0.6	W. Congo +	S. Ogooue	1.3 +/- 0.6
Intergradation,	W. Congo		Intergradation,	+ Odzala	
Odzala			Odzala		
W. Cameroon	Intergradation,	1.6 +/- 0.5	W. Congo +	Odzala	1.2 +/- 0.4
	Odzala		Intergradation		
Intergradation	Odzala	1.9 +/- 0.5			
Intergradation	Intergradation	1.1 +/- 0.4			
(Mvoum)	(other)				
S. Ogooue	W. Congo	1.1 +/- 0.3			

The mismatch distribution of the central group of haplotypes in *S. ollula* is characterised by $\tau = 1.34$ (0.337-2.297), $\theta_0 = 0$ (0-1.031) and $\theta_1 = 89.69$ (1.548-6725.938), for 11 polymorphic sites. The distribution fits the sudden expansion model ($P = 0.389$). The confidence intervals for θ_0 and θ_1 ,

which respectively represent the size of the population before and after the expansion, are not overlapping, suggesting that an expansion did occur. As $\tau = 1.34$ and $\mu = 2.8$, this expansion is estimated to have occurred about 0.24 Myr ago (0.06-0.41).

Discussion

*Genetic variation of haplotypes in the two *Sylvisorex* species*

The percentage of variable sites is higher (10.7 versus 4.8%) and the percentage of informative sites much higher (9.2 versus 1.5%) in *Sylvisorex johnstoni* than in *S. ollula*. The genetic patterns observed for the two species show an opposite trend. *S. johnstoni* haplotypes exhibit high inter-locality distances (up to 7.6% substitutions) and small intra-locality distances (maximum of 0.7% substitutions, Table 2 and Fig. 3). In each locality, haplotypes are monophyletic and unique to the locality, except for one haplotype which is shared between Four-Place and the Dja reserve (Fig. 4). The branching pattern suggests ancient divergence events between localities. In contrast, the branching pattern obtained for *S. ollula* reflects recent differentiation associated with small overall genetic distances between haplotypes (Fig. 5). Inter-locality distances between haplotypes are small (up to 2.4% substitutions, Fig. 3) and one haplotype is shared by four localities (Fig. 5). However, intra-locality distances can be high, such as in Odzala where the maximum distance between two haplotypes is 2.2% (Table 2 and Fig. 3). The high degree of divergence found in Odzala does not appear to result from a sampling bias in favour of this locality ($n = 12$) since diversity is low at the two localities where 10 samples were analysed (i.e. the Dja reserve and Ngotto).

Phylogeny and Geographic Distribution of Haplotypes

The distribution of genetic diversity is discontinuous for *S. johnstoni*, as shown by the MST where several intermediate haplotypes are missing (Fig. 4). As haplotypes also have a non-overlapping geographic distribution, the obtained phylogeographic pattern coincides with the first phylogeographic category defined by Avise et al. (1987). Such a distribution pattern reflects either a long-term extrinsic geographic barrier or, provided that gene flow is low for this species, the extinction of intermediate haplotypes.

For *S. ollula*, the distribution of genetic diversity is nearly continuous and the geographic distribution of haplotypes is overlapping (Fig. 5). This species meets the criteria of the fifth phylogeographic category (Avise et al., 1987). Its distribution pattern presumably results from a lack of long-term geographic barrier combined with an intermediate gene flow. This is compatible with the star-

shaped MST (Fig. 5) which is indicative of a recent radiation and (re)colonisation of the area. This finding is also corroborated by the analysis of the mismatch distribution.

Evolutionary Time-frame and the Origin of Intra-specific Lineages

The age estimates for intra-specific differentiation in the two investigated taxa encompass the Pleistocene for the two *Sylvisorex* species and date back to the end of the Pliocene for *S. johnstoni* (Table 3). As the species must have originated before the oldest intra-specific divergence, they must be of Plio-Pleistocene origin. This is in agreement with the findings of recent studies of neotropical taxa indicating that most taxa originated during the Pliocene rather than during the Pleistocene (e.g. Schneider et al., 1998; Clough & Summers, 2000). As it appears that the two species have co-existed over at least the last 1.3 Myr (Table 3), it is unlikely that the observed differences in phylogeographic structures can be attributed to differences in species ages. However, it remains possible that the two species did not have the same geographic distribution over this time-frame and consequently did not experience the same environmental changes.

Despite the uncertainty which is associated to the use of a molecular clock, it is noteworthy that some divergence events correspond to periods of great aridity. Most of the estimated divergence ages range between 1.1 and 1.3 Myr, which may correspond to a period of marked increase in glacial climate amplitude (deMenocal, 1995). The events dated at 1.6 and 1.9 Myr for *S. johnstoni* also coincide with a period of increased aridity in Africa (deMenocal, 1995). Although not much is known about the paleoclimatic conditions in Africa during the Pliocene, it seems that the oldest estimated age of 3.2 Myr corresponds to a forested period prior to the transition between a humid and warm climate to more arid and cooler conditions (deMenocal, 1995). Hence, the first divergence event affecting *S. johnstoni* cannot be explained by the refuge theory, but coincides with the flourishing of shrews after the Early Pliocene extinction of the sub-family Crocidossoricinae (Reumer, 1989).

The expansion of *S. ollula* in the intergradation and West Congo regions is estimated to be about 240,000 years old, based on the mismatch distribution. If this estimation is accurate, the expansion would have taken place just after an arid period (Jahns, 1996). The fact that it occurred long before the last glacial maximum (about 18,000 years ago) suggests that recent phases of forest fragmentation were possibly too short to provide sufficient isolation time for lineages to differentiate. If this assumption is false, then the recorded events are much more recent than estimated and the analysed 16s rRNA gene fragment has evolved at a very fast rate. The substitution rate required for this event to have followed the last phase of forest transgression (about 8,000 years ago), and for the most recent divergence events identified in the two species to coincide with the last glacial maximum, would be 30 to 60 folds faster

than our estimation. Such a variation in the molecular clock is unlikely and would certainly have been detected on the overall phylogeny of African shrews (Quérrouil et al., 2001).

Evaluation of the Faunal Regions and Assessment of Refuge Locations

As the age estimates for intra-specific differentiation in the two investigated taxa are ancient, it is relevant to search for the effects of recent climatic oscillations on the population distribution ranges. We evaluated the congruence between the biogeographic scenario suggested by mammal and insect distribution ranges (Deleporte & Colyn, 1999) and the phylogeographic pattern of each shrew species, by means of a comparison between a theoretical area cladogram and the area cladogram associated to each species phylogeny. The intra-specific phylogeographic patterns differ significantly from the general biogeographic pattern (Fig. 2). However, a detailed analysis of the shrew data sets shows that they are partly compatible with the faunal regions currently recognised for the Central African forest. The haplotype distribution of *S. johnstoni* suggests some differentiation between the regions South Ogooue, West Congo (Ngotto and Dzanga) and West Cameroon (Fig. 4). The pattern found for *S. ollula* also demonstrates some differentiation between the South Ogooue region and the other sampled regions (Fig. 5), but unfortunately, no specimens were obtained from the West Cameroon region for this species. Finally, the regions South Ogooue and West Cameroon may have served as refuges, at least for *S. johnstoni*.

The phylogeographic patterns depart from the area relationships predicted by the faunal regions (Fig. 2A) in two instances. First, all haplotypes from the ‘intergradation region’ cluster for both species (Fig. 4 and 5). This suggests the colonisation of this area from a single region for each species. The assumption that its fauna is of mixed origin (Deleporte & Colyn, 1999) holds for a series of taxa, but does not necessarily apply to a single species. Secondly, Odzala haplotypes, which were expected to group with other haplotypes from the West Congo region, cluster with haplotypes from the ‘intergradation region’ for the two species, and additionally with haplotypes from the South Ogooue region for *S. ollula* (Fig. 4 and 5). For *S. johnstoni*, the monophyly and endemism expressed by Odzala haplotypes suggest that a secondary refuge could have been located in the vicinity of this locality. For *S. ollula*, Odzala hosts a population with a high diversity of mitochondrial DNA haplotypes of polyphyletic origin, where some haplotypes are endemic to the locality while other ones are shared with the ‘intergradation region’. Consequently, Odzala, which is peripheral to the West Congo region and geographically very close to the ‘intergradation region’, must have been a place of faunal convergence for *S. ollula*.

The faunal region model (Fig. 1) indicates that rivers have acted as barriers to dispersal in Central Africa. The present study also shows that the main rivers, Ogooue and Sanaga, certainly contributed to the segregation of mitochondrial DNA haplotypes in shrews. The Ogooue river appears to have acted as a barrier between the South Ogooue region and the other localities for both *S. johnstoni* and *S. ollula*, and the Sanaga river seems to have separated the Korup population from the other populations for *S. johnstoni*. This is in agreement with earlier studies suggesting that rivers constitute the main barriers affecting the mammalian fauna in Central Africa (especially the Ogooue, Sanaga and Sangha rivers, e.g. Colyn et al., 1991; Happold, 1996) and affecting the primate and bird faunas in the Neotropics (e.g. Capparella, 1991; Ayres & Clutton-Brock, 1992).

Existence of Cryptic Sister Species in S. johnstoni?

There is a discrepancy between the small genetic distances and the great geographic distances found between the regions South Ogooue and West Congo (Ngotto and Dzanga) for *S. johnstoni*. As the two regions are separated by the 'intergradation region' and Odzala, the populations must have experienced successive phases of isolation and subsequent range expansion. The allopatric distribution of the two lineages suggests that what is presently recognised as a single species could actually be two mutually exclusive cryptic species. The very high genetic distances observed between the two clades of haplotypes also support this hypothesis, as differences of up to 7.6% within the same species are very rare for the molecular marker used (e.g. Baric & Sturmbauer, 1999; Maggioni et al., 2001), although not unprecedented (Thomaz et al., 1996). For a comparison, eight species of shrews belonging to the genus *Crocidura* differ on average by 7.1% (Quérrouil et al., 2001).

Even if reproductive isolation is not achieved, *S. johnstoni* populations could have differentiated into ecotypes whose distribution would depend on precipitation and vegetation, as suggested for African primate sub-species (Dandelot, 1965). Actually, there is a good fit between the distributions of the lineages and the forest types depicted by White (1983). For instance, the geographically distant populations of the regions South Ogooue and West Congo both occur in semi-deciduous forests. However, there is no morphological differentiation suggesting ecological adaptation along the whole range of the species (Hutterer, pers. comm.). In the future, it would be interesting to carry out a taxonomic revision of this species over its whole distribution range and to investigate whether secondary contact zones between these lineages can be detected.

Ecology of the Species and their Responses to Paleo-Environmental Changes

The geographic distribution of *S. johnstoni* haplotypes seems to be the result of ancient vicariance events (Fig. 4). Relict *S. johnstoni* populations must have survived in the vicinity of each sampled region and have not introgressed when their distribution ranges were no longer isolated. On the other hand, it seems that *S. ollula* recently (re)colonised a large part of the investigated area (Fig. 5).

The observed differences in phylogeographic structure could reflect different responses to paleo-environmental changes in relation to the ecological characteristics and the level of gene flow in the two species (cf. Arctander et al., 1999). The only known difference between the two *Sylvisorex* species is a difference in body size, but this allows us to make inferences on their persistence and dispersal abilities. Body size has been shown to influence the probability of population extinction in *Sorex* shrews: it is negatively correlated with population extinction rates and the impact of environmental stochasticity (Peltonen & Hanski, 1991). It could thus explain the extinction of intermediate haplotypes in *S. johnstoni*. Moreover, body size may be positively correlated with the dispersal abilities of these species. If this hypothesis applies to *S. johnstoni*, then this species is probably a poor disperser, which would explain the reduced gene flow observed among its populations. In addition, it is possible that its small size may have allowed *S. johnstoni* to survive even in small patches of forest, which did not act as refuges for bigger species such as primates (cf. Grubb, 1978). This is what has been observed in neotropical birds for which the probability of presence in small patches of forest decreases with increasing bird size (Cornelius et al., 2000). In that case, the observed population differentiation is likely to be the result of isolation by distance and it is to be expected that a finer grain sampling of the area investigated may reveal the presence of intermediate haplotypes (Templeton, 1998). Following the same reasoning, the large *S. ollula* could be a species with greater dispersal abilities, which may explain its low level of population differentiation.

Other biological and ecological features, such as demographic parameters, could certainly explain the differential responses to paleo-environmental changes. Preliminary results indicate that litter size may be greater and reproduction more seasonal in *S. ollula* than in *S. johnstoni* (V. Nicolas, pers. comm.). Unfortunately, our present knowledge on the ecology of tropical African shrews is too limited to address such alternative hypotheses.

Main Conclusions

Both *Sylvisorex* species must have originated before the oldest intra-specific divergences, therefore they must be of Plio-Pleistocene origin. This finding agrees with recent studies indicating that most neotropical taxa originated during the Pliocene rather than during the Pleistocene.

The haplotype distributions of the two species of shrews seem to accommodate the effect of forest fragmentation and expansion associated with the paleoclimatic fluctuations of the Pleistocene, in conjunction with ecological barriers such as rivers. Our results confirm that there must have been at least one forest refuge in each of the main biogeographical regions.

The two congeneric and syntopic species selected for this study exhibit very different phylogeographic patterns. Although they evolved under the same environmental conditions during the Quaternary, the geographic distribution of mitochondrial DNA haplotypes suggests that these taxa responded differently to similar environmental changes. It seems reasonable to assume that the main reason for this finding could be that these two shrew species do not have the same ecological requirements nor the same colonisation abilities, possibly due to a difference in body size. Our conclusions are similar to those of other studies which found that presently detectable patterns in different taxa will tell different stories depending on the age of the taxa involved and on the species dispersal or survival capacities (e.g. Sturmbauer & Meyer, 1992; Zink, 1996; Ditchfield, 2000).

Consequently, when phylogeographic studies are to be used to test biogeographic hypotheses, it appears necessary to incorporate several species in the analysis. It also seems important to pay attention to the choice of the model species and to acquire independent information on their taxonomy, ecology and sex-specific dispersal potential in order to investigate the evolutionary processes underlying their phylogeographic patterns.

Acknowledgements

The authors are grateful to P. Barrière and O. Perpète who collected the specimens and to P. Barrière and R. Hutterer for the species identifications and the sharing of unpublished results. We also wish to thank J. Ray and D. Schlitter for providing additional tissue samples respectively from the Dzanga Reserve and Korup National Park. We also acknowledge E. Petit for useful comments on the manuscript and M. I. Taylor for improving our use of the English language.

The field work and molecular analysis were funded by EU-DGVIII- Ecofac program 'Conservation et Utilisation Rationnelle des Ecosystèmes Forestiers en Afrique Centrale' (managed by AGRECO, GEIE, BDPA, SCETAGRI, SECA, CIRAD-FORET), EU-DGVIII- Biofac program 'Origine et maintien de la biodiversité en Afrique Centrale' (University of Rennes 1 / CNRS, UMR 6552). M. Dillen and E. Verheyen were supported by a grant of the Belgian Federal Services for Scientific, Cultural and Technical Affairs (DWTC project 31.64) entitled 'Evaluating the Effect of Pleistocene Climate Changes on Speciation Patterns in African Vertebrates.'

References

- Arctander, P., Johansen, C. and Coutellec-Vreto, M.A. 1999. Phylogeography of three closely related African bovids (tribe Alcelaphini). *Mol. Biol. Evol.* 16: 1724-1739.
- Avice, J.C. 1992. Molecular population structure and the biogeographic history of a regional fauna: a case study with lessons for conservation biology. *Oikos* 638: 62-76.
- Avice, J.C., Arnold, J., Ball, R.M., Bermingham, E., Lamb, T., Neigel, J.E., Reeb, C.A. and Saunders, N.C. 1987. Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. *Ann. Rev. Ecol. Syst.* 18: 489-522.
- Ayres, J.M. and Clutton-Brock, T.H. 1992. River boundaries and species range size in Amazonian primates. *Am. Nat.* 140: 531-537.
- Baric, S. and Sturmbauer, C. 1999. Ecological parallelism and cryptic species in the genus *Ophiothrix* derived from mitochondrial DNA sequences. *Mol. Phylogenet. Evol.* 11(1):157-162.
- Beresford, P. and Cracraft, J. 1999. Speciation in African forest robins (*Stiphornis*): species limits, phylogenetic relationships, and molecular biogeography. *Am. Mus. Novitates* 3270: 1-22.
- Bigalke, R.C. 1978. Present day mammals of Africa (Chapter 1), pp. 1-16. In J. Maglio and H.B.S. Cooke (Eds.), *Evolution of African Mammals*. Harvard University Press, Cambridge.
- Booth, A.H. 1958. The zoogeography of West African primates. A review. *Bull. Inst. Fond. Afr. Noire*. 20: 587-622.
- Brown, J.H. 1988. Species diversity, pp. 57-89. In A.A. Myers and P.S. Giller (Ed.), *Analytical biogeography: an integrated approach to the study of animal and plant distributions*. Chapman and Hall, London.
- Butler, P.M. 1998. Fossil history of shrews in Africa, pp. 121-132. In J.M. Wojcik and M. Wolsan (Eds), *Evolution of Shrews*. Mammal Research Institute, Polish Academy of Science, Bialowieza.
- Capparella, A.P. 1991. Neotropical avian diversity and riverine barriers. *Ibid.* 20: 307-316.
- Carcasson, R.H. 1964. A preliminary survey of the zoogeography of African butterflies. *East Afr. Wildl. J.* 2: 122-157.
- Clough, M. and Summers, K. 2000. Phylogenetic systematics and biogeography of the poison frogs: evidence from mitochondrial DNA sequences. *Biol. J. Linn. Soc.* 70: 515-540.
- Colyn, M. 1991. L'importance géographique du bassin du fleuve Zaïre pour la spéciation: le cas des Primates simiens. *Ann. Mus. R. Afr. Centr. Tervuren, Belgique* 264.
- Colyn, M. 1999. Etude populationnelle de la super-espèce *Cercopithecus cephus* habitant l'enclave forestière Sangha-Oubangui (République Centrafricaine). Description de *C. cephus ngottoensis* subsp. nov. *Mammalia* 63: 137-147.

- Colyn, M., Gautier-Hion, A. and Verheyen, W. 1991. A re-appraisal of paleoenvironmental history in Central Africa: evidence for a major fluvial refuge in the Zaire Basin. *J. Biogeogr.* 18: 403-407.
- Cornelius, C., Cofre, H. and Marquet, P.A. 2000. Effects of habitat fragmentation on bird species in a relict temperate forest in semiarid Chile. *Conserv. Biol.* 14: 534-543.
- Crandall, K.A. 1996. Multiple interspecies transmissions of human and simian T-cell leukemia / lymphoma virus type I sequences. *Mol. Biol. Evol.* 13: 115-131.
- Dandelot, P. 1965. Distribution de quelques espèces de Cercopithecidae en relation avec les zones de végétation de l'Afrique. *Zool. Afr.* 1: 167-176.
- deMenocal, P.B. 1995. Plio-Pleistocene African climate. *Science* 270: 53-59.
- Deleporte, P. and Colyn, M. 1999. Biogéographie et dynamique de la biodiversité: application de la PAE aux forêts planitiaies d'Afrique centrale. *Biosystema* 17: 37-43.
- Diamond, A.W. and Hamilton, A.C. 1980. The distribution of forest passerine birds and quaternary climatic change in tropical Africa. *J. Zool. Lond.* 191: 379-402.
- Dieterlen, F. and Heim de Balsac, H. 1979. Zur Ökologie und Taxonomie der Spitzmäuse (Soricidae) des Kivu-Gebietes. *Säugetierkundliche Mitteilungen* 27(4): 241-287.
- Ditchfield, A.D. 2000. The comparative phylogeography of Neotropical mammals: patterns of intraspecific mitochondrial DNA variation among bats contrasted to nonvolant small mammals. *Mol. Ecol.* 9: 1307-1318.
- Dobson, G.E. 1887. On the genus *Myosorex*, with description of a new species from the Rio Del Rey (Cameroons) district. *Proc. Zool. Soc. Lond.* 1887: 575-578.
- Excoffier, L. 1993. *MinSpNet*. Genetics and Biometry Laboratory, University of Geneva, Switzerland.
- Grubb, P. 1978. Patterns of speciation in African mammals. *Bull. Carnegie Mus. Nat. Hist.* 6: 152-167.
- Grubb, P. 1990. Primate geography in the Afro-Tropical forest biome, pp. 187-214. In G. Peters and R. Hutterer (Eds.), *Vertebrates in the Tropics*. Museum Alexander Koenig, Bonn.
- Haffer, J. 1969. Speciation in neotropical birds. *Science* 165: 131-137.
- Haffer, J. 1982. General aspects of the refuge theory, pp. 6-24. In G.T. Prance (Eds), *Biological Diversification in the Tropics*. Columbia University Press, New York.
- Happold, D.C.D. 1996. Mammals of the Guinea-Congo rain forest. *Proc. R. Soc. Edinburgh* 104B: 243-284.
- Hasegawa, M., Kishino, H., and Yano, T. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* 22: 160-174.
- Hutterer, R. 1986. Diagnosen neuer spitzmäuse aus Tansania (Mammalia: Soricidae). *Bonn. Zool.*

Beiträge 37: 23-33.

- Jahns, S. 1996. Vegetation history and climate changes in West Equatorial Africa during the Late Pleistocene and Holocene, based on a marine pollen diagram from the Congo fan. *Veget. Hist. Archaeobot.* 5: 207-213.
- Jahns, S., Hüls, M. and Sarntheim, M. 1998. Vegetation and climate history of West Equatorial Africa based on a marine pollen record off Liberia (site GIK 16776) covering the last 400,000 years. *Rev. Palaeobot. Palynol.* 102: 277-288.
- Kimura, M. 1980. A simple method for estimating evolutionary rate of base substitution through comparative studies of nucleotide sequences. *J. Mol. Evol.* 16: 111-120.
- Kishino, H. and Hasegawa, M. 1989. Evaluation of the maximum likelihood estimates of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. *J. Mol. Evol.* 29: 170-179.
- Kumar, S. 1996. PHYLTEST version 2.0: A program for testing phylogenetic hypothesis, distributed by the author, Institute of Molecular Evolutionary Genetics and department of Biology, The Pennsylvania State University from ftp.bio.indiana.edu in directory molbio/ibmpc.
- Louette, M. 1992. Barriers, contact zones and subspeciation in Central Equatorial Africa. *Bull. B.O.C. Centenary Suppl.* 112A: 209-216.
- Maggioni, R., Rogers, A.D., Maclean, N. and D'Incao, F. 2001. Molecular phylogeny of Western Atlantic *Farfantepenaeus* and *Litopenaeus* shrimp based on mitochondrial 16s partial sequences. *Mol. Phylogenet. Evol.* 18 (1): 66-73.
- Maley, J. 1996. The African rain forest - main characteristics of changes in vegetation and climate from the Upper Cretaceous to the Quaternary. *Proc. R. Soc. Edinburgh* 104B: 31-73.
- Mayr, E. 1963. *Animal Species and Evolution*, 1st edition. The Belknap Press of Harvard University Press, Cambridge, Massachusetts.
- Mayr, E. and O'Hara, R.J. 1986. The biogeographic evidence supporting the Pleistocene forest refuge hypothesis. *Evolution* 40(1): 55-67.
- Moreau, R.E. 1969. Climatic changes and the distribution of forest vertebrates in west Africa. *J. Zool. Lond.* 158: 39-61.
- Moritz, C., Patton, J.L., Schneider, C.J. and Smith, T.B. 2000. Diversification of rainforest faunas: An integrated molecular approach. *Annu. Rev. Ecol. Syst.* 31: 533-563.
- Palumbi, S.R., Martin, A., Romano, S., McMillan, W.O., Stice, L. and Grabowski, G. 1991. *The simple fool's guide to PCR*, version 2.0. University of Hawaii, Honolulu.
- Patton, J.L., Da Silva, M.N.F. and Malcolm, J.R. 2000. Mammals of the Rio Jurua and the

- evolutionary and ecological diversification of Amazonia. *Bull. Am. Mus. Nat. Hist.* (244): 13-306.
- Peltonen, A. and Hanski, I. 1991. Patterns of island occupancy explained by colonization and extinction rates in shrews. *Ecology* 72(5): 1698-1708.
- Prigogine, A. 1988. Speciation pattern of birds in the Central African forest refugia and their relationships to other refugia, pp. 144-157. In *Acta XIX Congressus Internationalis Ornithologici*, vol. II. University of Ottawa Press, Ottawa.
- Quérrouil, S., Hutterer, R., Barrière, P., Colyn, M., Peterhans, J.C.K. and Verheyen, E. (2001). Phylogeny and evolution of African shrews (Mammalia : Soricidae) inferred from 16s rRNA sequences. *Mol. Phylogenet. Evol.* 20(2): 185-195.
- Reumer, J.W.F. 1989. Speciation and evolution in the Soricidae (Mammalia: Insectivora) in relation with the paleoclimate. *Rev. Suisse Zool.* 96(1): 81-90.
- Robinson-Rechavi, R. and Hucon, D. 2000. RRTree: relative rate tests between groups of sequences on a phylogenetic tree. *Bioinformatics* 16(3): 296-297.
- Rogers, A.R. and Harpending, H. 1992. Population growth makes waves in the distribution of pairwise genetic differences. *Mol. Biol. Evol.* 9: 552-569.
- Roy, M.S. 1997. Recent diversification in African greenbul (Pycnonotidae: *Andropadus*) supports a montane speciation model. *Proc. R. Soc. Lond. B* 264: 1337-1344.
- Sambrook, J., Fritsch, E.F. and Maniatis, T. 1989. *Molecular cloning: A laboratory manual*, Second Edition. Chapter 14: In vitro amplification of DNA by the polymerase chain reaction, pp. 14.1-14.35. Cold Spring Harbour Laboratory Press.
- Schneider, C.J., Cunningham, M. and Moritz, C. 1998. Comparative phylogeography and the history of vertebrates endemic to the Wet Tropics rainforest of Australia. *Mol. Ecol.* 7: 487-498.
- Schneider, S., Roessli, D. and Excoffier, L. 2000. *Arlequin ver. 2.000: A Software for Population Genetics Data Analyses*. Genetics and Biometry Laboratory, University of Geneva, Switzerland.
- Sokal, R.R. and Rohlf, F.J. 1981. *Biometry*, 2nd edn. W.F. Freeman and Company, New York.
- Sturmbauer, C. and Meyer, A. 1992. Genetic divergence, speciation and morphological stasis in a lineage of African cichlid fishes. *Nature* 358(638): 578-581.
- Swofford, D.L. 1997. *PAUP*: Phylogenetic Analysis Using Parsimony (and other methods)*, version 4.0. Sinauer Associates, Sunderland, Massachusetts.
- Taberlet, P., Fumagalli, L., Wust Saucy, A.G. and Cosson, J.F. 1998. Comparative phylogeography and post-glacial colonization routes in Europe. *Mol. Ecol.* 7(4): 453-464.
- Templeton, A.R. 1998. Nested clade analyses of phylogeographic data: testing hypotheses about

- gene flow and population history. *Mol. Ecol.* 7: 381-397.
- Thomas, O. 1913. Four new shrews. *Ann. Mag. Nat. Hist., ser. 8*, 11: 214-218.
- Thomaz, D., Guiller, A. and Clarke, B. 1996. Extreme divergence of mitochondrial DNA within species of pulmonate land snails. *Proc. R. Soc. Lond. B* 263: 363-368.
- Thompson, J.D., Higgins, D.G. and Gibson, T.J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22: 4673-4680.
- Wallace, A.R. 1876. *The Geographical Distribution of Animals*. Hafner, New York.
- Walsh, P.S., Metzger, D.A. and Higuchi, R. 1991. Chelex 100 as a medium for simple extraction of DNA for PCR based typing from forensic material. *Biotechniques* 10: 506-513.
- White, F. 1983. The vegetation of Africa: a descriptive memoir to accompany the UNESCO/AETFAT/UNSO vegetation map of Africa. UNESCO, Paris, France.
- Zink, R.M. 1996. Comparative phylogeography in North American birds. *Evolution* 50: 308-317.

2) Analyse complémentaire : analyse des groupements emboîtés

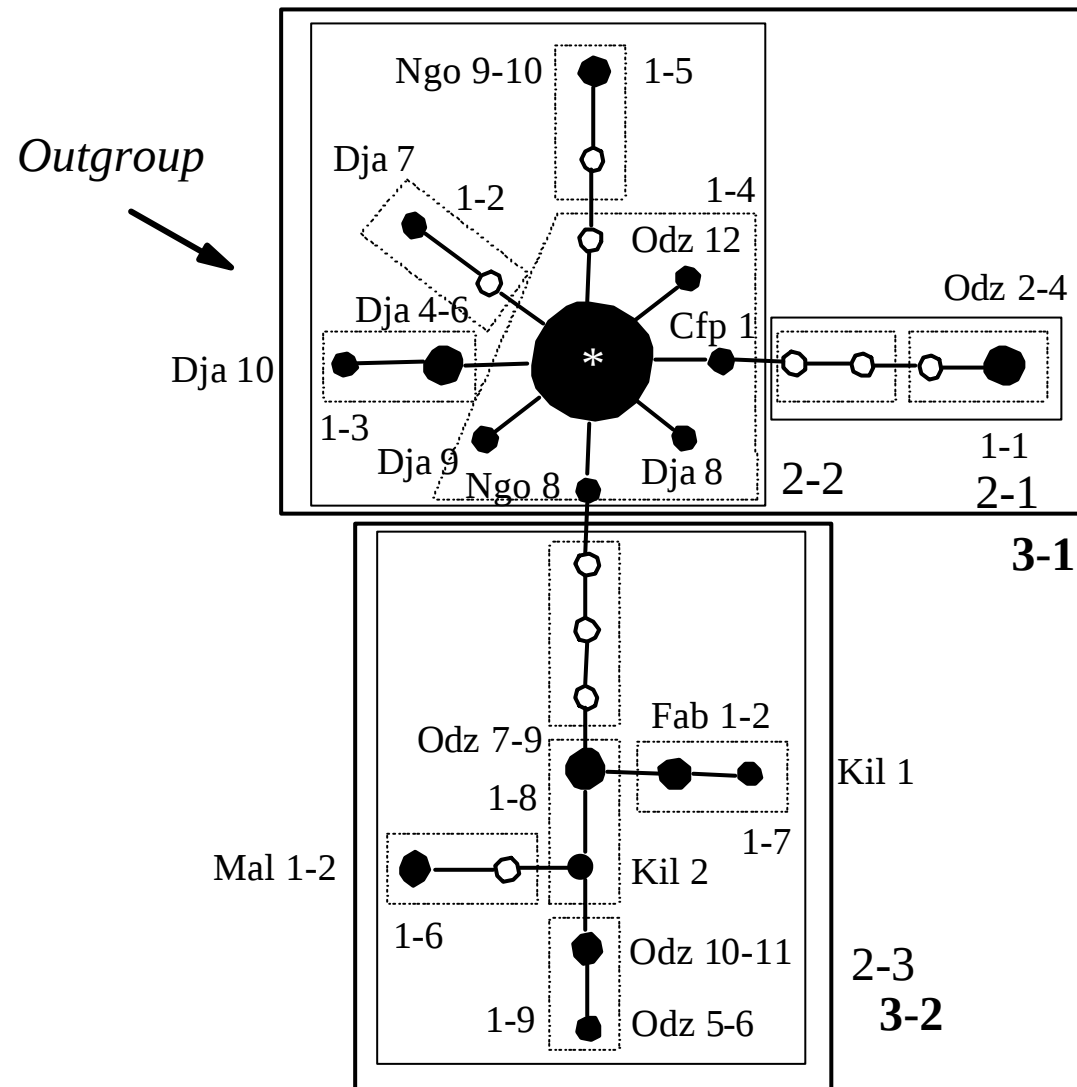
a) Matériel et méthodes

Nous avons réalisé une analyse de contingence emboîtée des groupements mis en évidence sur les arbres d'étendue minimale présentés dans l'article ci-dessus. Nous avons utilisé des arbres construits à partir de matrices des distances plutôt que des réseaux parcimonieux, parce que ces derniers présentaient des connexions multiples difficiles à traduire en groupements imbriqués (Annexes 6 et 7). En effet, bien que le programme TCS soit conçu pour produire des réseaux d'haplotypes, les règles d'emboîtement proposées sont insuffisantes pour définir les clades emboîtés en cas de connexions multiples. En outre, pour *S. johnstoni*, cinq cladogrammes indépendants ont été obtenus (Annexe 6). En imposant que les haplotypes soient connectés même si leur connexion n'est pas parcimonieuse à 95%, il a été possible de réunir ces cinq cladogrammes en un seul, mais le réseau présentait alors de très nombreuses connexions multiples.

Nous avons défini les groupements et leurs relations d'emboîtement comme décrit dans le chapitre 2.V (**fig. 4.1** et **4.2**), puis effectué l'analyse au moyen du programme Geodis. Pour chacun des groupements (ou clades) englobant des haplotypes distincts provenant de localités distinctes, le programme teste s'il existe une association géographique significative entre les haplotypes qui le composent au moyen d'un test de Chi-2. Il calcule les distances géographiques entre clades constituant le clade analysé (D_n) et entre clades de rang inférieur eux mêmes nichés dans les clades constituant le clade analysé (D_c), et il évalue leur significativité (tab. 4.1 et 4.2). Il calcule également les distances moyennes entre clades internes (I) et clades terminaux (T), et teste si la différence entre les deux est significative. Les résultats significatifs sont ensuite interprétés à l'aide d'une clef d'interprétation (Annexe 5). Dans la mesure où nous nous sommes basés sur un arbre d'étendue minimale alors que le programme est conçu pour interpréter des réseaux parcimonieux d'haplotypes, ces résultats sont à considérer avec précaution.

b) Résultats

Les résultats des tests sont présentés sous forme de tableau (tab. 4.1 et 4.2). Les indices D_c et D_n sont indiqués uniquement pour les clades pour lesquels il existe une association géographique significative entre haplotypes. L'interprétation relative à ces indices est rapportée, avec la liste des étapes suivies sur la clef de détermination, en dessous de chacun des tableaux.



* Cfp 2, Dja 1-3, Ngo 1-7, Odz 1

Figure 4.2. Définition des groupements emboîtés sur l'arbre des haplotypes obtenu pour *S. ollula*.

Tableau 4.1 : Résultat de l'analyse des groupements emboîtés réalisée pour *Sylvisorex johnstoni*. Les clades pour lesquels le test de contingence exact est significatif, ainsi que les distances significativement grandes (^L) ou petites (_s), sont en caractères gras.

Haplo.	Dc	Dn	Cl.1	Dc	Dn	Cl.2	Dc	Dn	Cl.3	Dc	Dn	Cl.4	Dc	Dn	Cl.5
Dzs1			1-1	0	86										
Ngo5			1-2	0	86	2-1			3-1	86_s	412				
Ngo1			1-3	0	86										
Ngo4												4-1	413	413	
			I-T	0	0										
Kil1			1-4			2-2	0	108							5-1
Kil3			1-5						3-2	108_s	413				
Fab1			1-6			2-3	0	108							
						I-T	0	0							
									I-T	22	0				
Mal1			1-7			2-4			3-3			4-2	0_s	472	
Mal3															
												I-T	-413_s	61	
Kor1			1-8			2-5			3-4			4-3	0	499^L	
Kor2															
Cfp1	0	249	1-9	249	267^L										
Dja1	239	249													
	I-T	239	0			2-6			3-5	248	254				
Geo1			1-10	0	210_s										
Geo2															5-2
			I-T	249	56^L							4-4	262_s	278_s	
Mvo1			1-11			2-7			3-6	0_s	293^L				
Mvo2															
Odz10			1-12												
Odz1			1-13			2-8			3-7	0_s	253				
Odz8															
									I-T	248^L	-10				
												I-T	-262	221^L	

Tableau 4.1 (suite)

Cl.5	Dc	Dn	Total
5-1	426^L	443^L	
5-2	315_s	318_s	Total
I-T	110^L	125^L	

Interprétation des résultats d'après la clef

Clade 2-6 : 1 2 11 17 4 9 → fragmentation passée

Clade 4-1 : 1 2 11 12 13 ? → colonisation sur longue distance

Clade 4-4 : 1 2 3 5 15 → fragmentation passée

Clade 5-1 : 1 2 11 12 → expansion contiguë

Clade 5-2 : 1 2 3 4 9 10 → plan d'échantillonnage inadapté pour trancher entre fragmentation et isolement par la distance

Cladogramme total : 1 2 3 4 9 ? → fragmentation passée

Tableau 4.2 : Résultat de l'analyse des groupements emboîtés réalisée pour *Sylvisorex ollula*. Les clades pour lesquels le test de contingence exact est significatif, ainsi que les distances significativement grandes (^L) ou petites (_s), sont en caractères gras.

Haplo.	Dc	Dn	Cl.1	Dc	Dn	Cl.2	Dc	Dn	Cl.3	Dc	Dn	Total
Odz2			1-1			2-1	344	348				
Dja7			1-2	0	135							
Dja4			1-3	0_s	135_s							
Dja10												
Odz1	381	393										
Odz12	0	164				2-2	0	218	3-1	338	400^L	
Dja8	0	164	1-4	378	378							
Dja9	0	421										
Ngo8	0	491										
Cfp1	0	214										
I-T	352	159										
Ngo9			1-5	0	455							Total
			I-T	378^L	151							
						I-T	-344_s	-130				
Mal1			1-6	0	207							
Kil1	0	144	1-7	95	109_s							
Fab1	0	72										
I-T	0	-72										
Kil2	0	183	1-8	244	239	2-3			3-2	191_s	293	
Odz7	0	366										
I-T												
Odz5			1-9	0	406							
Odz10												
			I-T	212^L	-24							
									I-T	147^L	107^L	

Interprétation des résultats

Clade 2-2 : 1 2 3 4 → flux génique limité plus isolement par la distance

Clade 2-3 : 1 2 3 4 → flux génique limité plus isolement par la distance

Clade 3-1 : 1 2 11 12 → expansion de range contiguë

Cladogramme total : 1 2 3 4 9 → fragmentation passée

Pour *S. johnstoni*, la structure phylogéographique des clades s'explique le plus souvent par une fragmentation passée (3 voire 4 cas sur 6). La relation entre les clades 3-1 et 3-2, formant le clade 4-1, est interprétée comme le résultat d'une colonisation sur grande distance, de la région Sud Ogooué vers Ngotto et Dzangha. Au sein du clade 5-1, constitué par les clades 4-1 et 4-2, il y aurait eu une expansion contiguë entre localités de la région Sud Ogooué.

Pour *S. ollula*, les clades 2-2 et 2-3 résulteraient d'un flux génique limité associé à un phénomène d'isolement par la distance. Le clade 3-1, composé d'un haplotype d'Odzala (clade 2-1) et du clade 2-2

qui contient entre autres des haplotypes d'Odzala, résulterait d'une expansion contiguë. La relation entre les deux groupements majeurs qui constituent l'arbre global (clades 3-1 et 3-2) serait le résultat d'une fragmentation passée.

c) Discussion

Les résultats de l'analyse des groupements emboîtés réalisée pour *S. johnstoni* sont conformes aux interprétations du MST que nous avons données dans l'article qui précède. Nous avons effectivement interprété la plupart des événements comme résultant d'une fragmentation passée. C'est aussi l'interprétation qui est donnée pour un schéma phylogéographique de première catégorie (Avisé *et al.*, 1987). En ce qui concerne les haplotypes de la région Sud Ogooué, notre première analyse les regroupait dans un même clade, ce qui ne contredit pas la possibilité d'une expansion contiguë entre Malounga et la Forêt des Abeilles. La faible distance génétique entre les haplotypes de Ngotto et de la région Sud Ogooué, pourtant séparés par une grande distance géographique, laisse effectivement supposer qu'il y a eu colonisation sur grande distance.

Pour *S. ollula*, l'analyse des groupements emboîtés explique la dichotomie entre les deux principaux clades par une fragmentation passée, ce qui était également le cas dans l'article qui précède. L'explication de la relation entre les clades 2-1 et 2-2 par une expansion contiguë semble pertinente. L'interprétation de la structure des clades 2-2 et 2-3 par un flux génique limité associé à un isolement par la distance est acceptable pour le clade 2-3, mais surprenante pour le clade 2-2. En effet, l'un des haplotypes de ce second clade est partagé par toutes les localités incluses dans le clade, et le clade présente par ailleurs une structure typique d'une expansion récente. Cependant, les haplotypes satellites de cet haplotype partagé sont différents d'une localité à l'autre, ce qui pourrait traduire un certain isolement par la distance.

L'analyse des groupements emboîtés présente l'avantage de mieux préciser les relations entre haplotypes, de distinguer des événements qui auraient été négligés et de proposer une explication aux événements mis en évidence. Elle nous a permis de confirmer certaines hypothèses, telles que la fragmentation passée et la colonisation sur grande distance dans le cas de *S. johnstoni*, et d'affiner nos interprétations dans le cas de *S. ollula*. En effet, pour *S. ollula*, l'analyse met en évidence des traces d'isolement par la distance et de dispersion au sein du groupe relativement homogène d'haplotypes que nous avons interprété comme résultant d'un événement d'expansion. En revanche, l'analyse des groupements emboîtés, qui ne prend pas directement en compte les distances entre haplotypes, n'intègre pas la dimension temporelle.

II - PHYLOGEOGRAPHIE COMPAREE DE DEUX RONGEURS

Nous avons répété le travail réalisé dans le chapitre précédent pour deux modèles de rongeurs typiquement forestiers : *Hylomyscus stella* (**doc. 2.3**) et *Stochomys longicaudatus* (**doc. 2.4**). Le premier de ces modèles a été retenu après révision du genre *Hylomyscus* (chapitre 3.II). Le second a été choisi parce qu'il ne présentait pas de difficulté taxinomique particulière, le genre *Stochomys* étant monophylétique, et qu'il a été collecté sur un grand nombre de sites. Ces espèces ont des aires de distribution étendues et largement recouvrantes, mais *H. stella* s'étend un peu plus loin à l'Ouest et à l'Est que *S. longicaudatus*. Les deux espèces sont supposées vivre dans le même type d'habitat, bien qu'il ait été suggéré que *S. longicaudatus* pourrait avoir une préférence pour les bords de rivières ou les marécages (Rosevear, 1969 ; Kingdon, 1971). Bon nombre de spécimens de cette espèce ont effectivement été collectés dans des zones humides, mais, d'une façon générale, les taux de captures sont plus élevés dans les zones humides qu'ailleurs, et *S. longicaudatus* peut aussi être capturée à distance des cours d'eau, en forêt primaire et secondaire ou dans des cultures (Duplantier, 1982 ; Granjon, 1991 ; P. Barrière et V. Nicolas, comm. pers.).

1) Matériel et méthodes

Les échantillons analysés proviennent de 6 localités de la région Sud Ogooué, 3 localités de la région Ouest Congo et 5 localités de la région d'intergradation (**fig. 4.3**). La région Cameroun n'est pas représentée, en revanche nous avons inclus des localités des régions Est et Sud Congo (Kisangani Nord et Sud, en R.D.C.) et d'Afrique de l'Est (Kakamega, au Kenya). Nous avons utilisé les 39 échantillons de *H. stella* 1 (forme orientale) et *H. stella* 2 (forme d'Afrique centrale) analysés dans le chapitre 3.II, et choisi les spécimens *H. alleni* R22966 et *H. parvus* R23146 comme groupes externes. Nous avons séquencé la seconde moitié de l'ARN 16s de 22 échantillons de *S. longicaudatus*, tous récoltés par des membres de notre équipe (tab. 4.3), ainsi que des spécimens *Arvicanthis neumanni* 9295 (Ruaha National Park, Tanzanie) et *Hybomys univittatus* G10023 (Forêt des Abeilles, Gabon), pris comme groupes externes.

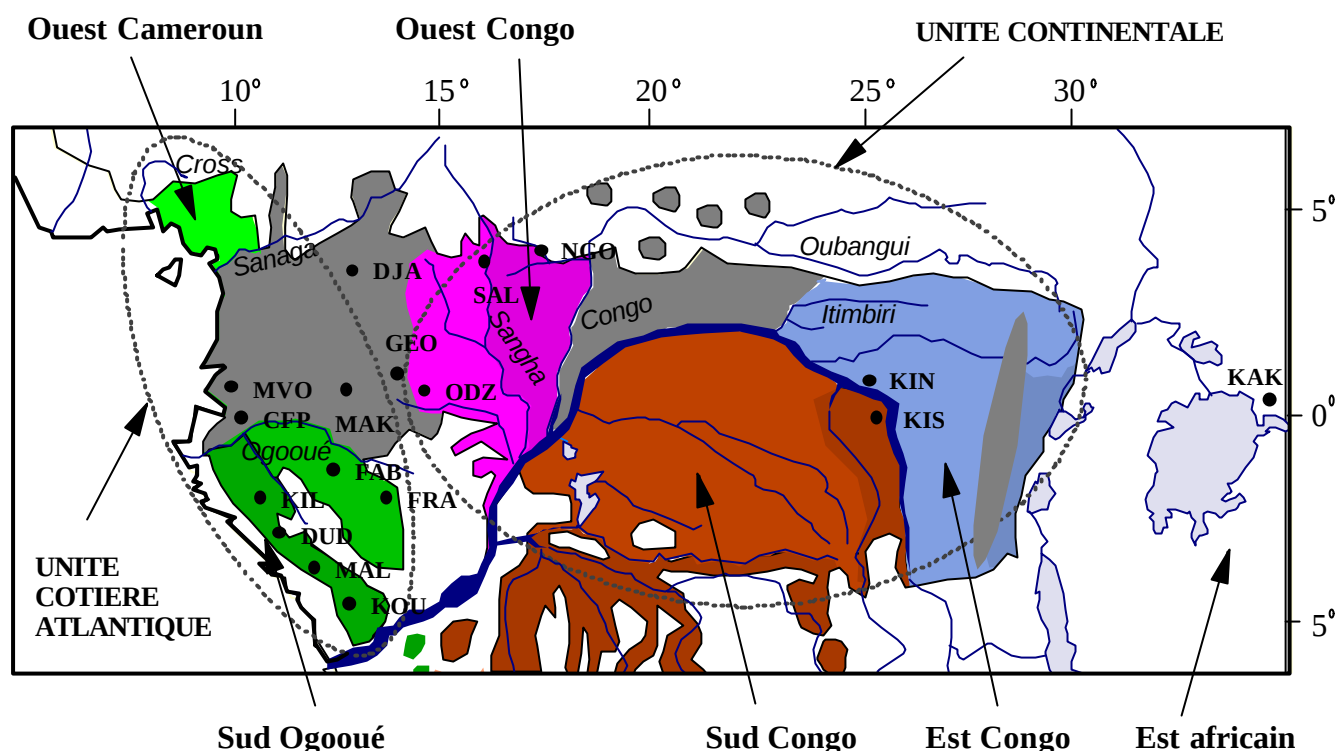


Figure 4.3. Localités de collecte des spécimens de *H. stella* et *S. longicaudatus*, et unités, régions et sous régions fauniques d'après M. Colyn (comm. pers.).

Les zones hachurées représentent les régions d'intergradation, et les pointillés caractérisent des sous-régions au sein des régions fauniques.

Tableau 4.3 : Liste des échantillons de *S. longicaudatus* séquencés

Code	Localité	Pays	Références
Cfp	Four-Place	Gabon	GA99
Dja	Dja	Cameroun	R14033, R14107, R14148
Dud	Monts Dudu	Gabon	GA545
Fra	Franceville	Gabon	G10124
KiN	Kisangani Nord	RDC	Z2707, Z2744
KiS	Kisangani Sud	RDC	Z2685, Z2693
Mal	Malounga	Gabon	GA145, GA200
Mvo	Mvoum	Gabon	GA04, GA32, GA36
Ngo	Ngotto	RCA	R12033, R18311, R18818
Odz	Odzala	Congo	R22320, R22321
Sal	Salo ²	RCA	R13450, R13659

Nous avons procédé aux mêmes analyses que dans le chapitre qui précède. Pour les datations, nous avons appliqué une correction de type Tamura-Nei au lieu de la correction de Kimura à 2 paramètres en raison de la grande distance entre les taxons utilisés pour calibrer l'horloge et les groupes internes. Pour *H. stella*, nous avons utilisé le taux de substitution par million d'années obtenu dans le chapitre 3.II le plus élevé, soit $r = 0,825$.

2) Résultats

Diversité génétique

Pour *H. stella*, l'alignement des groupes internes a nécessité la création de 5 indels dont un affectant deux positions. Sur les 511 sites analysés, 465 sont constants, 46 sont variables et 27 sont informatifs pour la parcimonie (en incluant les groupes externes, 447 sont constants, 64 sont variables et 35 sont informatifs). Pour *S. longicaudatus*, l'alignement des groupes internes a nécessité la création de 2 indels simples, et l'ajout des groupes externes a entraîné la création de 4 indels supplémentaires affectant 7 positions. Sur les 522 sites analysés, 486 sont constants, 36 sont variables et 30 sont informatifs pour la parcimonie (en incluant les groupes externes, 450 sont constants, 72 sont variables et 42 sont informatifs). Le rapport moyen entre transitions et transversions est de 2,0 (+/- 1,4) pour *H. stella* et de 4,2 (+/- 1,9) pour *S. longicaudatus*, ce qui laisse supposer que les substitutions ne sont pas saturées, au moins pour *S. longicaudatus*.

Tableau 4.4 : Nombre d'échantillons séquencés (N.seq), nombre d'haplotypes différents (N.hapl), distances génétiques moyennes (D.moy, avec écart type) et maximales (D.max) au sein de chacune des localités de collecte (Loc.), pour les deux modèles étudiés.

Loc.	Région faunique	<i>H. stella</i>				<i>S. longicaudatus</i>			
		N.seq	N.hapl	D.moy	D.max	N.seq	N.hapl	D.moy	D.max
Ngo	Ouest Congo	3	2	0,1 +/- 0,1	0,2	3	2	0,1 +/- 0,1	0,2
Odz	Ouest Congo	8	5	0,4 +/- 0,4	1,0	2	2	0,6	
Sal	Ouest Congo	3	3	1,0 +/- 0,7	1,5	2	2	0,2	
Dud	Sud Ogooué					1	1	---	
Fab	Sud Ogooué	6	5	0,6 +/- 0,3	1,0				
Fra	Sud Ogooué					1	1	---	
Kil	Sud Ogooué	2	2	0,8					
Kou	Sud Ogooué	1	1	---					
Mal	Sud Ogooué	3	1	0		2	1	0	
Cfp	Intergradation					1	1	---	
Dja	Intergradation	6	6	0,8 +/- 0,4	1,4	3	1	0	
Geo	Intergradation	1	1	---					
Mak	Intergradation	1	1	---					
Mvo	Intergradation					3	3	0,3 +/- 0,1	0,4
KiN	Est Congo	3	3	1,1 +/- 0,6	1,4	2	1	0	
KiS	Sud Congo					2	2	0,2	
Kak	Afrique Est	2	2	0,2					
	Total	39	32 / 29			22	17 / 13		

La proportion d'haplotypes différents est plus grande pour *H. stella* (74%, soit 29 haplotypes différents pour 39 échantillons en provenance de 12 localités) que pour *S. longicaudatus* (59%, soit 13

haplotypes différents pour 22 échantillons en provenance de 11 localités ; tab. 4.4). La variabilité intra-localité moyenne est maximale à Kisangani Nord pour *H. stella* (1,1 +/- 0,6, n= 3) et à Odzala pour *S. longicaudatus* (0,6, n= 2 spécimens). La distance maximale entre haplotypes d'une même localité atteint 1,5 à Salo², pour *H. stella* et 0,6 à Odzala pour *S. longicaudatus*.

La représentation graphique de la variabilité intra- et inter-localités sous forme de boxplot (fig. 4.4) fait apparaître un net contraste entre une faible variabilité intra-localité et une forte variabilité inter-localités pour les deux espèces, que l'on prenne en compte les localités de l'Est africain ou non. La variabilité intra-localité est plus forte pour *H. stella* (0,7 +/- 0,4) que pour *S. longicaudatus* (0,1 +/- 0,2), tandis que la variabilité inter-localités, globale ou limitée à l'Afrique centrale, est plus forte pour *S. longicaudatus* (2,2 +/- 1,6 et 2,5 +/- 1,6) que pour *H. stella* (1,5 +/- 1,1 et 0,9 +/- 0,4). La variabilité inter-localités augmente lorsqu'on ajoute les localités de l'Est africain pour *H. stella*, alors qu'elle aurait plutôt tendance à diminuer pour *S. longicaudatus* (p = 0,063). En dépit du faible nombre de spécimens analysés par localité (surtout pour *S. longicaudatus*), toutes les comparaisons exceptée la dernière sont significatives (p < 0,001).

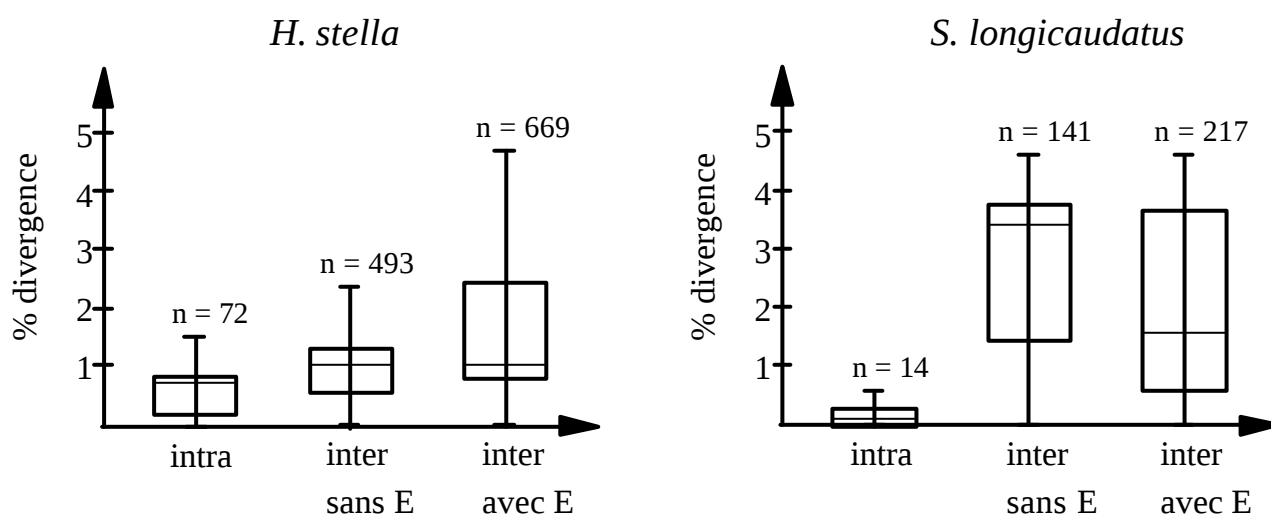


Figure 4.4. Boxplots représentant les distances intra- et inter- localités pour *H. stella* et *S. longicaudatus*, en incluant et en excluant les localités de l'Est africain ("E") pour les distances inter-localités.

Les boîtes représentent les valeurs comprises entre le premier et le troisième quartile, tandis que les barres horizontales indiquent la médiane, et les barres verticales relient les valeurs minimales et maximales.

Arbres d'étendue minimale

L'arbre obtenu pour *H. stella* présente deux groupements majeurs, l'un comprenant les haplotypes d'Afrique centrale, correspondant à *H. stella* 2, et l'autre ceux d'Afrique de l'Est, correspondant à *H. stella* 1 (**fig. 4.5**). La distribution en double étoile (c'est-à-dire centrée autour de deux haplotypes) des haplotypes d'Afrique centrale laisse supposer qu'il y a eu une expansion récente dans toute la région. Les haplotypes satellites de l'haplotype repéré par un astérisque ("Dja5, Odz3-6") proviennent du Dja, de Salo² et d'Odzala, tandis que les haplotypes satellites de l'haplotype 'Mak 1' proviennent de Kili, Malounga et la forêt des Abeilles, qui sont trois localités situées dans la région Sud Ogooué, ainsi que d'Odzala et de Ngotto. Malgré la présence d'haplotypes d'Odzala dans ces deux ensembles, la plupart des haplotypes se regroupent par localité ou région. Notons que l'un des haplotypes, en provenance de Salo², diffère des autres haplotypes d'Afrique centrale par six substitutions, ce qui est supérieur à la limite qui nous a semblé pertinente pour délimiter des ensembles distincts d'haplotypes.

L'arbre obtenu pour *S. longicaudatus* met en évidence quatre groupements (**fig. 4.6**). L'un d'eux s'étend sur une très grande zone géographique, depuis le Dja et Ngotto, situées respectivement dans les régions d'intergradation et Ouest Congo, jusqu'à Kisangani dans la région Est Congo. En revanche, les haplotypes des régions Sud Ogooué, intergradation et Odzala forment trois ensembles distincts. Les haplotypes d'Odzala et de Ngotto, qui sont deux localités situées dans la région Ouest Congo, diffèrent fortement. En revanche, les haplotypes de Salo², également située dans la région Ouest Congo, diffèrent par 0 ou 1 substitution de ceux de Ngotto.

Test de scénarios évolutifs

Pour *H. stella*, les 38 arbres les plus parcimonieux obtenus en l'absence de contrainte (l = 69 pas et -ln L = 1154,0 - **fig. 4.7 B**) sont plus courts que les arbres MP contraints selon l'arbre des relations théoriques entre localités, représenté sur la figure **4.7 A** (l = 72 pas et -ln L = 1162,8). La différence n'est pas significative, que le test de Kishino-Hasegawa soit effectué sous le critère de parcimonie (p = 0,083) ou sous celui de vraisemblance (p = 0,108).

Pour *S. longicaudatus*, l'unique arbre MP obtenu en l'absence de contrainte (l = 83 pas et -ln L = 1159,2 - **fig. 4.7 C**) est significativement plus court que les 12 arbres MP contraints (94 pas et -ln L = 1191,4), que le test soit effectué sous le critère de parcimonie (p < 0,001) ou sous celui de vraisemblance (p = 0,005). La différence entre les deux arbres réside dans le groupement des haplotypes d'Odzala avec ceux de la région Sud Ogooué dans l'arbre non contraint, et avec les autres haplotypes de la région Ouest Congo dans l'arbre contraint.

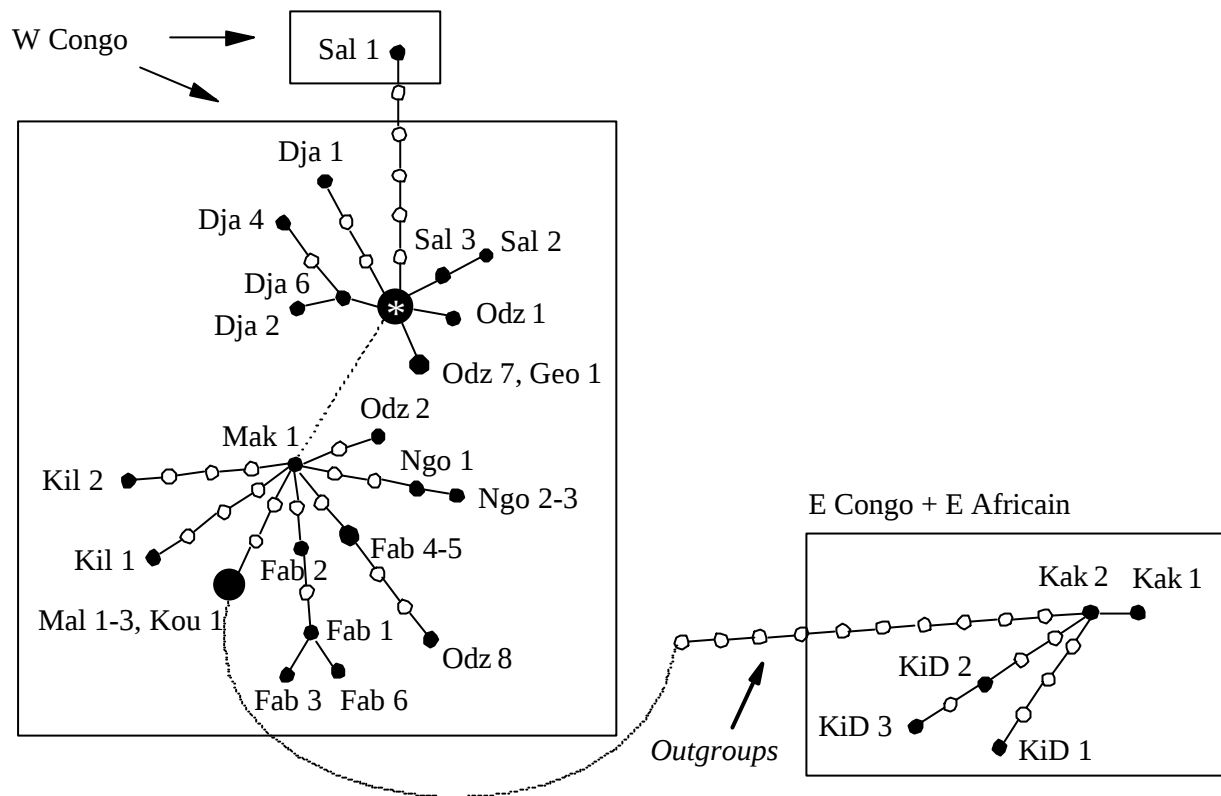


Figure 4.5. Arbre d'étendue minimale obtenu pour *H. stella*.

Les traits pointillés relient des haplotypes séparés par une grande distance géographique mais une faible distance génétique, de façon à respecter les localisations géographiques.

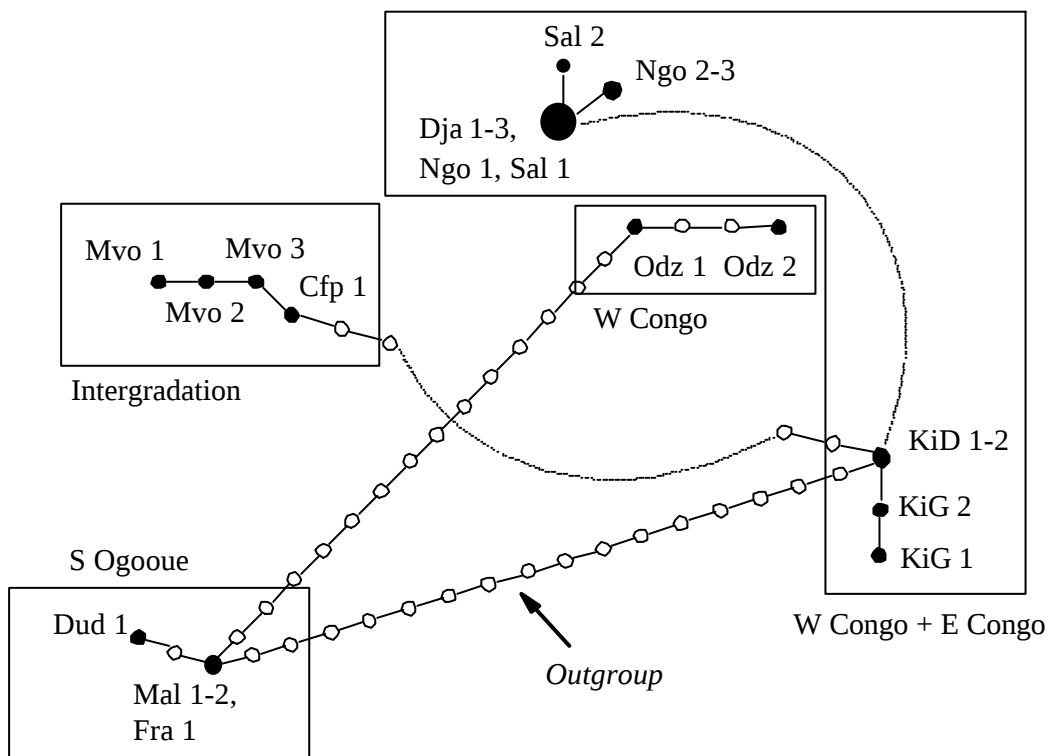


Figure 4.6. Arbre d'étendue minimale obtenu pour *S. longicaudatus*.

Même remarque que pour la figure 4.6.

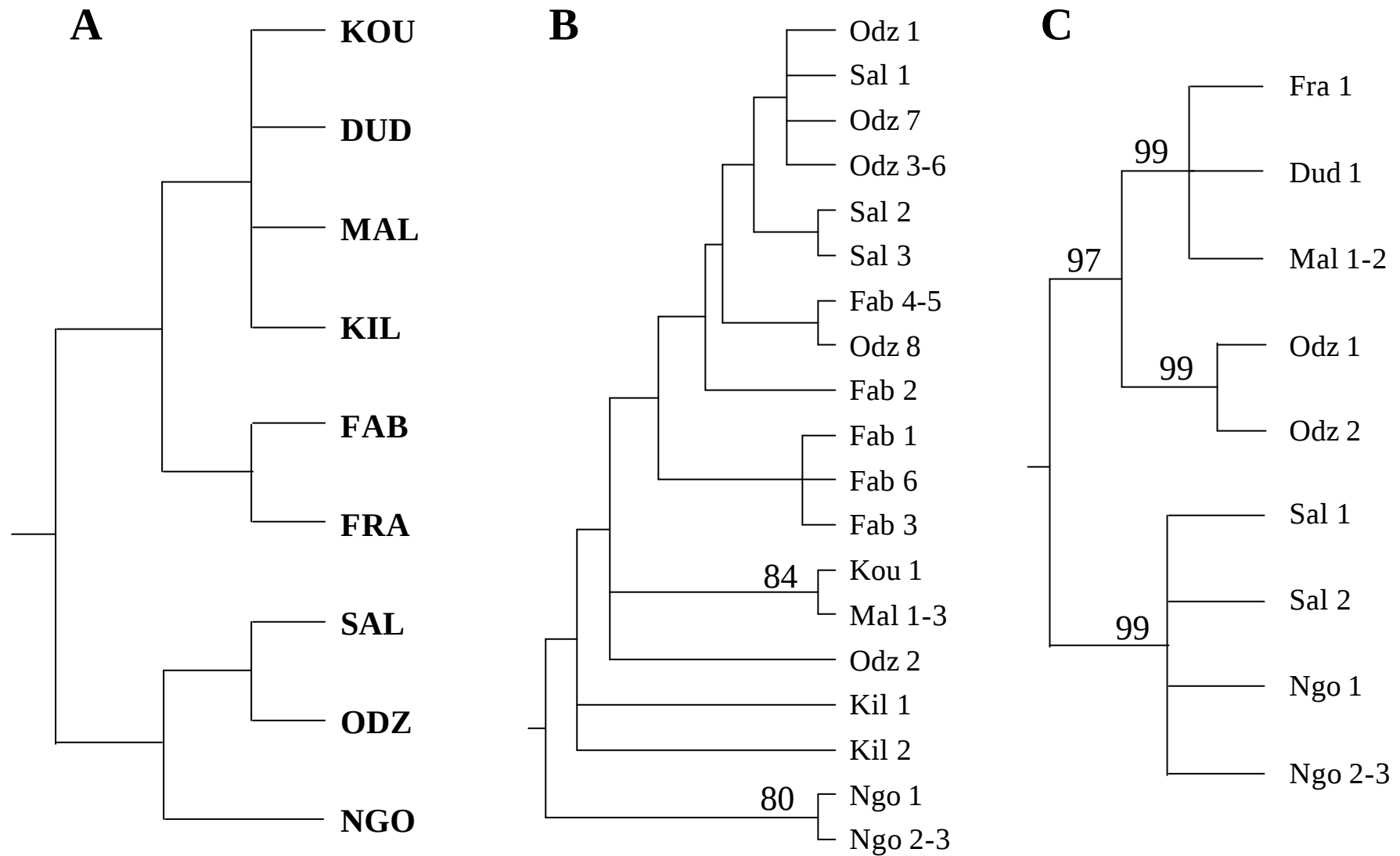


Figure 4.7. Arbre théorique des aires (A), et arbres MP non contraints avec valeurs de bootstrap pour *H. stella* (B) et *S. longicaudatus* (C). L'arbre B représente le consensus à la majorité des 38 arbres obtenus.

Datation

Pour *H. stella*, nous rappelons que la divergence entre *H. stella 1* et *H. stella 2* a été estimée à 1,9 MA +/- 0,3 dans le chapitre 3.II. D'autre part, l'arbre d'étendue minimale faisant présager une expansion récente en Afrique centrale, nous avons procédé à la datation de cette expansion à l'aide d'un test de 'mismatch distribution'. La distribution suit un modèle tel que le nombre moyen de différences entre séquences prises deux à deux est de 4,78 (3,67-6,16), $\tau = 4,97$ (2,98-6,27), $\theta_1 = 0$ (0-1,98) et $\theta_2 = 145,6$ (29,6-8230,6), pour 33 sites polymorphes. Les intervalles de confiance de θ_1 et θ_2 sont non chevauchants. Le modèle d'une expansion soudaine n'est pas rejeté ($p = 0,09$). L'expansion est estimée à 0,59 MA (0,35-0,74).

Pour *S. longicaudatus*, le "relative rate test" n'a pas mis en évidence de différence significative des taux de substitution. La valeur minimale a été obtenue entre les échantillons d'Odzala et des localités du Nord ($p = 0,109$). En se basant sur la divergence *Mus-Rattus*, le taux de substitutions estimé est de 0,875 substitutions par million d'années, et en se basant sur la divergence *Mus-Gerbillus*, il est de 0,723 substitutions par million d'années. Si on se réfère à la valeur moyenne de r , soit 0,799, les principaux événements de divergence auraient eu lieu autour de 0,9, 2,0 et 2,2 MA (tab. 4.5).

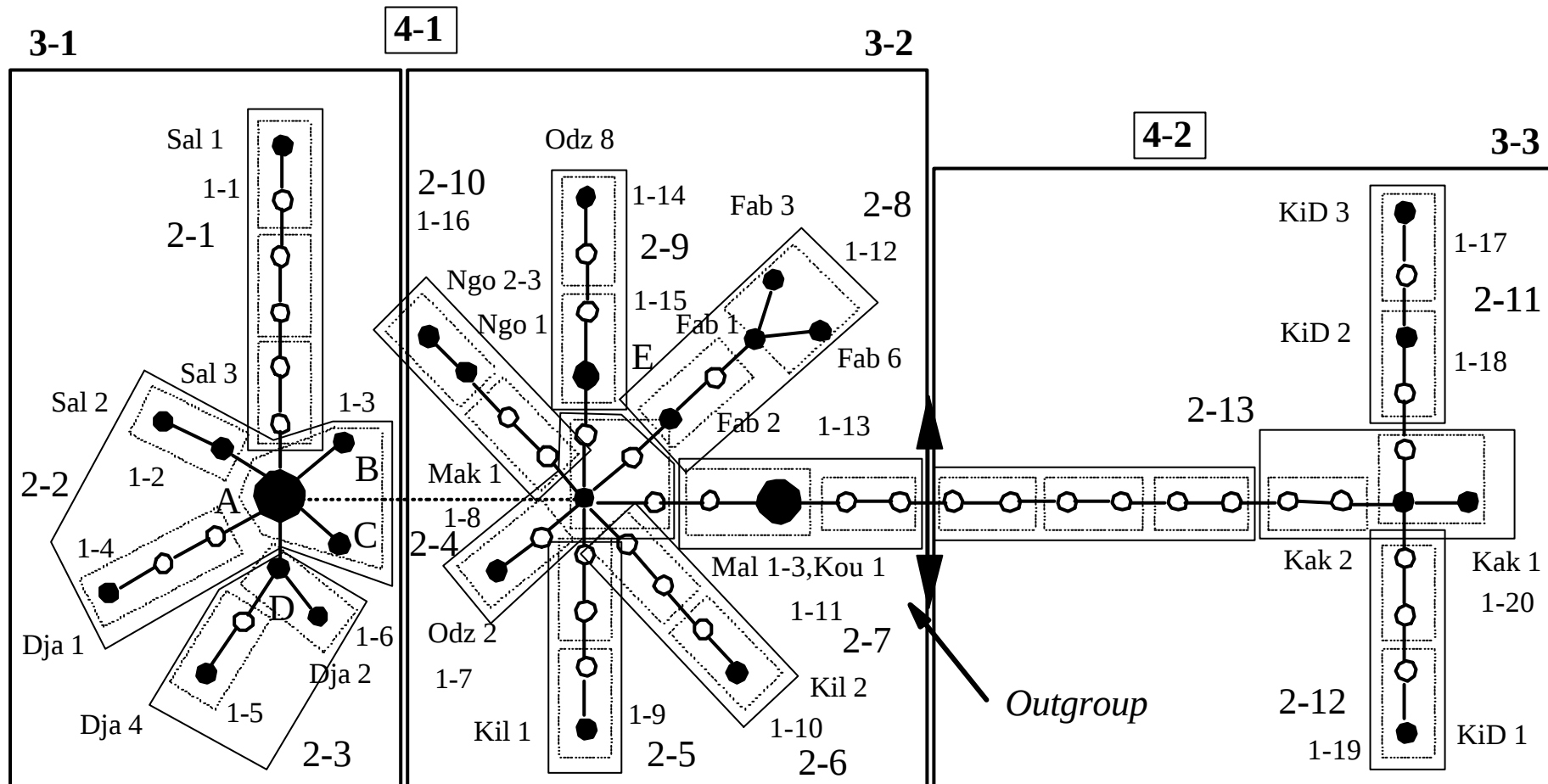
Tableau 4.5 : Estimation du temps de divergence (en millions d'années) pour les principales lignées de *S. longicaudatus*.

		calibration <i>Mus-Rattus</i>	calibration <i>Mus-Gerbillus</i>	moyenne
Taxon 1	Taxon 2	temps +/- σ	temps +/- σ	temps +/- σ
<i>Mus</i>	<i>Rattus</i>	(12)	14,5	13,1
<i>Mus</i>	<i>Gerbillus</i>	13,2	(16)	14,5
Mvo-Cfp	Nord	0,8 +/- 0,1	1,0 +/- 0,1	0,9 +/- 0,1
Mal-Fra-Dud	Odz	1,9 +/- 0,2	2,2 +/- 0,2	2,0 +/- 0,2
Mvo-Cfp-Nord	Mal-Fra-Dud-Odz	2,0 +/- 0,6	2,4 +/- 0,7	2,2 +/- 0,7

Analyses des groupements emboîtés

Des réseaux parcimonieux d'haplotypes ont été construits et figurent en annexe (Annexes 8 et 9). Ils consistent en deux réseaux distincts pour les deux espèces. Dans le cas d'*H. stella*, l'un de ces réseaux présente de nombreuses connexions multiples. Pour les raisons explicitées dans le chapitre qui précède, nous avons préféré nous baser sur les arbres d'étendue minimale pour définir les groupements emboîtés (**fig. 4.8 et 4.9**). Les résultats des tests et leur interprétation sont présentés ci-après (tab. 4.6 et 4.7).

Figure 4.8. Définition des groupements emboîtés sur l'arbre des haplotypes obtenu pour *H. stella*.
La double-flèche délimite les clades de rang 4.



A: Dja 5 -Odz 3-6 - B: Odz 1 - C: Odz 7 et Geo 1 - D: Dja 6 - E: Fab 4-5

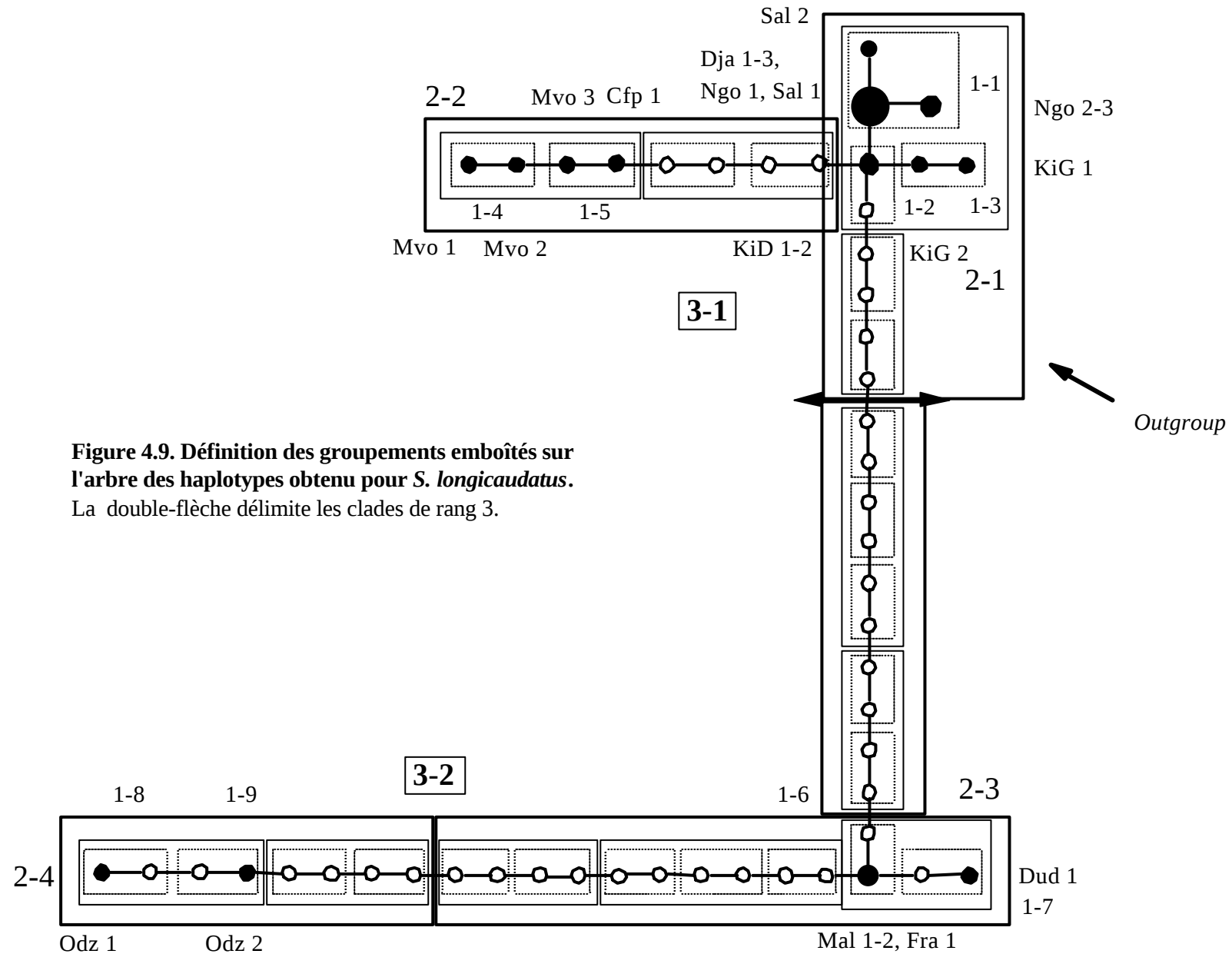


Figure 4.9. Définition des groupements emboîtés sur l'arbre des haplotypes obtenu pour *S. longicaudatus*.
La double-flèche délimite les clades de rang 3.

Tableau 4.6 : Résultat de l'analyse des groupements emboîtés réalisée pour *Hylomyscus stella*. Les clades pour lesquels le test de contingence exact est significatif, ainsi que les distances significativement grandes (^L) ou petites (_s), sont en caractères gras.

Haplo.	Dc	Dn	Cl.1	Dc	Dn	Cl.2	Dc	Dn	Cl.3	Dc	Dn	Cl.4	Dc	Dn	Total
Sal1			1-1			2-1	0	203							
Sal2			1-2	0	222										
Sal3															
Odz1	0	70													
Odz7	15	15	1-3	58_s	116_s	2-2	151	166							
Dja5	139	128							3-1	178_s	282				
I-T	129	95													
Dja1			1-4	0	264										
			I-T	58	-120_s										
Dja4			1-5												
Dja2			1-6			2-3	0	228							
Dja6												4-1	335_s	454_s	
						I-T	151	-56_s							
Odz2			1-7	0	176	2-4	39	159							
Mak1			1-8	0	22										
			I-T	0	-154										
Kil1			1-9			2-5	0	281							
Kil2			1-10			2-6	0	281							
Mal1			1-11			2-7	82_s	355							
Fab1															
Fab3			1-12			2-8	0_s	71_s	3-2	325	366				
Fab6															
Fab2			1-13												
Odz8			1-14	0	243	2-9	132	126_s							
Fab4			1-15	0	91										
			I-T	0	-152										
Ngo1			1-16			2-10	0_s	754^L							
Ngo2															Total
						I-T	35	-1							
									I-T	147	84				
KiD3			1-17			2-11	0	541							
KiD2			1-18												
KiD1			1-19			2-12	0	541	3-3			4-2	541	1541 ^L	
Ken1			1-20			2-13	0	541							
Ken2															
						I-T	0	0							
												I-T	205	1087 ^L	

Interprétation des résultats

Clade 2-2: 1 2 11 12 → expansion contiguë

Clade 3-1: 1 2 11 17 → impossible de conclure

Clade 3-2: 1 2 3 5 6 7 → flux génique limité plus dispersion sur longues distances (→ Ngotto)

Clade 4-1: 1 2 3 4 9 → fragmentation passée

Cladogramme total: 1 2 3 4 9 10 → plan d'échantillonnage inadapté pour trancher entre fragmentation et isolement par la distance

Tableau 4.7 : Résultat de l'analyse des groupements emboîtés réalisée pour *Stochomys longicaudatus*. Les clades pour lesquels le test de contingence exact est significatif, ainsi que les distances significativement grandes (^L) ou petites (_s), sont en caractères gras.

Haplo.	Dc	Dn	Cl.1	Dc	Dn	Cl.2	Dc	Dn	Cl.3	Dc	Dn	Total
Ngo1	197	208										
Ngo2	0	213	1-1	187_s	464_s							
Sal2	0	76										
I-T	197	40				2-1	551	562				
KiD1			1-2	0	659							
KiG1			1-3	0	705							
KiG2									3-1	617	649^L	
I-T				-149	147							
Mvo1			1-4	0	52							Total
Mvo2						2-2	52_s	756				
Mvo3	0	78	1-5	39	52							
Cfp1	0	26										
I-T				39	0							
I-T						499^L	-193					
Mal1			1-6	174	177							
Dud1			1-7	0	121	2-3	159	196				
I-T				174	56							
Odz1			1-8			2-4	0	348	3-2	234_s	394	
Odz2			1-9									
I-T						159	-151					
I-T									383^L	255^L		

Interprétation des résultats :

Clade 2-1: 1 2 3 4 9 10 → pas assez de localités pour trancher entre fragmentation et isolement par la distance

Clade 3-1: 1 2 3 9 → fragmentation passée

Cladogramme total : 1 2 3 4 9 → fragmentation passée

Pour *H. stella*, la structure du clade 2-2 est interprétée comme le fruit d'une expansion contiguë. Celle du clade 3-2 s'explique par un flux génique limité accompagné d'une dispersion sur longue distance. La relation entre les deux clades d'Afrique centrale composant le clade 4-1 résulterait d'une fragmentation passée. Les clades correspondant aux haplotypes d'Afrique centrale (4-1) et de l'Est (4-2) se distingueraient en raison d'une fragmentation passée ou de l'isolement par la distance.

Pour *S. longicaudatus*, au moins deux des trois clades pour lesquels il existe une association géographique significative des haplotypes résulteraient d'une fragmentation passée.

3) Discussion

Variabilité génétique

La variabilité intra-localité est inférieure à la variabilité inter-localités pour les deux espèces (fig. 4.4). La variabilité intra-localité est plus faible et la variabilité inter-localités plus grande pour *S. longicaudatus* que pour *H. stella*, ce qui révèle un groupement des haplotypes par localité plus marqué pour *S. longicaudatus*. Le fait que la variabilité inter-localités ait tendance à diminuer lorsqu'on ajoute les localités de l'Est africain pour *S. longicaudatus* traduit vraisemblablement une absence d'isolement par la distance pour cette espèce.

Distribution géographique des haplotypes

Pour *H. stella*, la représentation des relations entre haplotypes sous forme de MST met en évidence une expansion récente en Afrique centrale, avec ségrégation partielle des haplotypes par localités. Certains haplotypes sont présents dans plusieurs localités, tandis que les haplotypes d'une même localité sont parfois polyphylétiques. Ce schéma correspond à la catégorie phylogéographique V qui s'explique par un flux génique modéré et par une absence de barrières géographiques à long terme (Avisé *et al.*, 1987). Les haplotypes d'Afrique centrale se regroupent en deux ensembles. Le premier correspond aux haplotypes de la région Sud Ogooué, d'une localité de la région d'intergradation (Makokou) et de deux localités de la région Ouest Congo (Odzala et Ngotto). Le second correspond aux haplotypes d'une localité de la région d'intergradation (Dja) et de deux localités de la région Ouest Congo (Odzala et Salo²). Ces deux ensembles sont géographiquement chevauchants, mais ils seraient mutuellement exclusifs si des haplotypes d'Odzala n'étaient pas présents dans les deux.

Pour *S. longicaudatus*, en dépit de deux haplotypes partagés par plusieurs localités, il existe une ségrégation géographique nette des haplotypes. La divergence entre populations est nettement plus élevée que la divergence intra-populationnelle. Ce schéma correspond à la catégorie phylogéographique I, et serait la conséquence de barrières géographiques à long terme (Avisé *et al.*, 1987). Le nombre limité de spécimens séquencés par localité nous incite cependant à rester prudent quand à la monophylie des haplotypes de chaque localité. La faible distance génétique entre les haplotypes de Ngotto et Salo² et ceux de Kisangani Nord et Sud, qui s'oppose à la grande distance géographique entre ces localités, traduit vraisemblablement un phénomène de colonisation sur grande distance.

Mécanismes impliqués (Analyse des groupements emboîtés)

Pour *H. stella*, plusieurs mécanismes différents sont invoqués pour expliquer la relation entre structure génétique et localisation géographique des haplotypes. L'explication de la dichotomie entre Afrique centrale et Afrique de l'Est par une fragmentation passée ou un isolement par la distance nous indique que l'hypothèse de la vicariance n'est pas la seule qui puisse expliquer le schéma obtenu et qu'il conviendrait d'échantillonner des sites intermédiaires. Pour les clades 3.1 et 3.2 (fig. 4.8), alors que la radiation observée autour de l'haplotype "Dja 5, Odz 3-6" semble être le fruit d'une expansion contiguë, la radiation autour de l'haplotype "Mak 1" s'expliquerait plutôt par un flux génique limité et une dispersion sur longue distance. Dans notre interprétation de l'arbre des haplotypes avant découpage en groupements emboîtés (fig. 4.5), nous n'avions pas dissocié les deux clades, et nous avons conclu à une expansion pour l'ensemble. Le résultat obtenu ici pour le second clade s'explique par le fait que la plupart des haplotypes se groupent par localité, et que les localités représentées couvrent une grande aire géographique.

Pour *S. longicaudatus*, la fragmentation passée semble être le mécanisme prédominant. Ceci est en accord avec nos propres hypothèses dans deux cas sur trois. En effet, à notre avis, le clade 2-1 (fig. 4.9) correspond à une colonisation sur grande distance et non pas à une fragmentation passée ou à un isolement par la distance. Cette opposition entre les résultats obtenus par l'analyse des groupements emboîtés et nos interprétations pourrait venir du fait que le programme Geodis tient compte des relations phylétiques qui unissent les haplotypes mais pas des distances génétiques. De plus, il est possible que le découpage en clades imbriqués reflète mal les relations phylogénétiques entre haplotypes, surtout lorsque l'échantillonnage est faible.

Datation

La dichotomie entre *H. stella* 1 et *H. stella* 2 a été estimée à 1,9 MA +/- 0,3. Si *H. stella* 1 et *H. stella* 2 sont effectivement deux espèces distinctes, *S. longicaudatus* serait plus ancienne que l'espèce *H. stella* 2 puisque son apparition doit être antérieure au plus ancien événement de divergence mis en évidence, soit 2,0 MA +/- 0,6. Les principaux événements de divergence entre groupements de *S. longicaudatus* auraient eu lieu entre 0,8 MA +/- 0,1 et 2,0 MA +/- 0,6, si on se réfère à la divergence entre *Mus* et *Rattus*. Le plus récent de ces événements pourrait avoir fait suite à l'augmentation de la sévérité des glaciations qui se serait produite il y a 1 MA +/- 0,2 (deMenocal, 1995). En fait, en raison du polymorphisme intra-spécifique et du temps de coalescence, les temps de divergence estimés sont sous-estimés par rapport à l'âge des événements de vicariance (Taberlet *et al.*, 1998), donc la synchronie entre un événement présumé être à l'origine d'une divergence et un temps

estimé ne peut être qu'apparente, et vice-versa. L'expansion de *H. stella* 2 en Afrique centrale se serait produite assez récemment, puisqu'elle a été estimée à 0,6 MA +/- 0,15. Elle pourrait avoir masqué des événements de divergence antérieurs. Les oscillations climatiques ultérieures à cette expansion n'auraient pas entraîné de différenciation génétique sans doute parce que les phases d'allopatrie ont été trop courtes pour qu'il y ait accumulation de substitutions au niveau de la portion d'ADN étudiée et / ou parce qu'il y a eu mélange des populations et homogénéisation pendant les phases de sympatrie (cf. Bernatchez et Wilson, 1998).

Mise en correspondance avec les régions fauniques

L'hypothèse d'une ségrégation géographique des haplotypes selon les régions biogéographiques est rejetée pour *S. longicaudatus*, mais pas pour *H. stella*. Pour cette seconde espèce, les haplotypes diffèrent par peu de substitutions (fig. 4.5) et les relations qui les unissent sont compatibles avec les régions fauniques, en dépit d'un événement probable de colonisation vers Ngotto. Pour *S. longicaudatus*, les haplotypes se groupent par région (fig. 4.6), mais les relations entre régions sont différentes de celles prédites par l'arbre théorique des relations entre aires biogéographiques. Il est probable que les phases de fragmentation et d'expansion de la forêt aient eu des conséquences sur la structure génétique de cette espèce. En revanche, les rivières pourraient ne pas constituer des barrières géographiques majeures pour *S. longicaudatus*. En effet, la forte ressemblance entre les haplotypes de Kisangani provenant de la rive gauche et de la rive droite du fleuve Congo laisse supposer que des individus ont traversé le fleuve récemment. Cette traversée pourrait avoir été effectuée à la faveur des jacinthes d'eau qui couvrent saisonnièrement le fleuve sur toute sa largeur. Cette hypothèse est d'autant plus vraisemblable que *S. longicaudatus* a été décrite comme affectionnant les zones marécageuses (Rosevear, 1969 ; Kingdon, 1997), mais il n'est pas exclu que cette traversée soit d'origine anthropique.

Les haplotypes de la région d'intergradation sont polyphylétiques pour les deux espèces, ce qui est en accord avec l'hypothèse d'une colonisation multiple de cette région.

Facteurs pouvant expliquer la différence entre les deux modèles

Dans l'article précédent, nous avons avancé l'hypothèse de l'importance de la taille des représentants d'une espèce pour sa dispersion et sa survie dans des refuges d'étendue limitée ; les espèces de plus grande taille étant probablement plus mobiles, mais moins aptes à survivre en milieu fragmenté. Pour les deux exemples traités ici, la situation est inversée puisque l'espèce pour laquelle il existe une nette ségrégation géographique des haplotypes (*S. longicaudatus*) est quatre fois plus lourde et presque deux fois plus longue que l'autre. Le rapport élevé entre poids et taille chez

S. longicaudatus pourrait éventuellement constituer une entrave à la dispersion. Cependant, cette espèce semble être capable d'une colonisation rapide, puisque les haplotypes trouvés à Ngotto et à Kisangani diffèrent par une seule substitution, en dépit des 900 kilomètres qui les séparent. Il pourrait exister entre les deux espèces des différences biologiques et écologiques méconnues affectant leurs capacités de survie en milieu fragmenté ou entraînant une préférence pour un micro-habitat. Dans le cas de *S. longicaudatus*, certains auteurs ayant relevé une préférence pour le bord de l'eau et les zones marécageuses (Rosevear, 1969 ; Kingdon, 1997), nous pouvons supposer que sa survie en période de fragmentation de l'habitat n'a pu se faire que dans les refuges les plus humides, le long des cours d'eau. Nous pouvons émettre l'hypothèse qu'il existe une influence des bassins drainants sur la structure génétique de cette espèce. Les haplotypes devraient alors se grouper selon le bassin drainant dont ils proviennent, soit celui de l'Ogooué, soit celui du Congo. Pour tester cette hypothèse, nous avons adopté la même démarche que pour tester la correspondance entre schémas phylogéographiques et régions fauniques. Il apparaît que les haplotypes ne se regroupent pas par bassin drainant ($p = 0,006$). D'ailleurs, la présence de *S. longicaudatus* dans des zones de régénération et de cultures près des villages (Rosevear, 1969) indique que sa dispersion ne doit pas se faire uniquement le long des cours d'eau.

L'étude phylogéographique de *H. stella* fait apparaître une expansion récente en Afrique centrale qui peut soit correspondre à une première colonisation de la région, soit avoir effacé les traces d'événements plus anciens. Les cycles glaciaires postérieurs à cette expansion pourraient avoir eu une influence sur la structuration génétique de l'espèce, qui est compatible avec les régions fauniques. L'existence d'haplotypes identiques dans différentes localités laisse supposer qu'il pourrait exister un flux génique actuel. Le schéma phylogéographique obtenu pour cette espèce rappelle celui qui a été obtenu pour *Sylvisorex ollula*, qui présente également les traces d'une expansion.

L'étude phylogéographique de *S. longicaudatus* révèle une ségrégation géographique des haplotypes qui résulterait d'une fragmentation passée. Cette ségrégation pourrait avoir été modelée par les phases de contraction et d'expansion forestière de la fin du Pliocène et de la première moitié du Pléistocène, puis avoir perduré en dépit des cycles glaciaires ultérieurs et d'une capacité de colonisation non négligeable. Le schéma phylogéographique obtenu pour *S. longicaudatus* rappelle celui qui a été obtenu pour *Sylvisorex johnstoni*, qui semble lui aussi résulter d'événements de vicariance Plio-Pléistocène.

III - ETUDE DE LA SUPER-ESPECE *CERCOPITHECUS CEPHUS*

Parmi les primates forestiers africains, nous avons choisi d'étudier la super-espèce *Cercopithecus cephus*, qui appartient à la Famille des Cercopithecidae, et dont la distribution couvre l'ensemble des forêts de plaine d'Afrique tropicale (**doc. 2.5**). Cette super-espèce comprendrait 5 ou 6 espèces ou semi-espèces (*C. ascanius*, *C. cephus*, *C. erythrogaster*, *C. erythrotis*, *C. petaurista*, et *C. sclateri* ?) et 14 sous-espèces aux distributions allopatriques (Lernould, 1988 ; Grubb *et al.*, 1999 ; Colyn, 1999). La reconnaissance de *C. sclateri* en tant qu'espèce, proposée par Kingdon (1980), ne fait pas l'unanimité. *C. sclateri* est le plus souvent considérée comme une sous-espèce de *C. erythrotis* (cf. Napier, 1981 ; Oates, 1985 ; Lernould, 1988). Sa distribution limitée et son faible nombre de représentants font qu'elle est généralement omise dans les études comparatives, ce qui est le cas des études présentées ci-après.

La super-espèce *C. cephus* présente un large éventail de colorations de la face et du pelage (Kingdon, 1980, 1988) en dépit d'une variabilité anatomique pratiquement nulle (Verheyen, 1962 ; Gebo et Sargis, 1994). Elle se distingue des autres Cercopithecini par sa formule chromosomique qui est la même pour les 4 espèces dont le caryotype a été déterminé ($2N = 66$ et $NF = 110$; Dutrillaux *et al.*, 1980). Une comparaison d'études écologiques portant sur *C. ascanius*, *C. cephus* et *C. petaurista* montre que ces espèces ont des caractéristiques écologiques assez semblables, et que pour certains paramètres, tels que la taille des groupes, l'occupation verticale de l'espace et la taille du domaine vital, la variabilité intra-spécifique est parfois supérieure à la variabilité inter-spécifique (Gautier-Hion, 1988). Une étude portant sur les vocalisations soutient la monophylie de la super-espèce par trois synapomorphies et sépare les deux espèces occidentales, *C. erythrogaster* et *C. petaurista*, des trois espèces d'Afrique centrale, *C. ascanius*, *C. cephus* et *C. erythrotis*, par deux synapomorphies (Gautier, 1988). Parallèlement, une étude portant sur la mobilité électrophorétique de cinq protéines sépare ces deux mêmes groupes d'espèces, par une seule synapomorphie (Walker *et al.*, 1988). Malgré la concordance entre ces deux études, les relations phylétiques au sein du groupe restent imprécises. Un scénario évolutif hypothétique a été proposé sur la base des colorations faciales et corporelles (Kingdon, 1988, 1990). *C. erythrogaster* serait l'espèce ressemblant le plus à la forme ancestrale, dont auraient dérivé *C. petaurista* à l'Ouest, et *C. ascanius*, *C. cephus* et *C. erythrotis* à l'Est. Le groupe aurait ainsi colonisé les forêts d'Afrique centrale à partir de l'Ouest, et plus précisément de la Haute Guinée, en profitant des zones dégradées et des chablis.

Dans le but d'éclaircir les relations entre espèces et de proposer un scénario évolutif pour la super-espèce, nous avons séquencé la seconde moitié de l'ARN 16s de 18 spécimens couvrant les

cinq espèces habituellement reconnues et 11 des 14 sous-espèces. Nous avons supposé que, en dépit d'un échantillonnage restreint, la confrontation des relations phylétiques unissant les différentes espèces et sous-espèces avec leur aire de distribution nous permettrait de proposer une interprétation phylogéographique.

1) Matériel et méthodes

Les 18 échantillons analysés englobent 4 des 5 sous-espèces de *C. ascanius*, les 3 sous-espèces de *C. cephus*, 1 des 2 sous-espèces de *C. erythrogaster*, 1 des 2 sous-espèces de *C. erythrotis* et les 2 sous-espèces de *C. petaurista* (tab. 4.8). Leur provenance n'est pas toujours connue avec précision puisque certains spécimens, hébergés dans des parcs zoologiques, proviennent de saisies de douanes, et que d'autres ont été obtenus sur des marchés de viande de chasse, en Afrique.

Deux groupes externes ont été choisis parmi les plus proches voisins de la super-espèce, au vu des résultats d'une analyse moléculaire incluant un plus grand nombre d'espèces (cf. chapitre 5). Ce sont des spécimens de *C. lhoesti* (Pri 2) et de *C. preussi* (spécimen provenant de Mulhouse).

Tableau 4.8 : Liste des échantillons séquencés

Espèce	Sous-espèce	Provenance	Références	Donateur
<i>C. ascanius</i>	<i>schmidti</i> <i>ascanius</i> <i>whitesidei</i> <i>katangae</i>	RCA RDC RDC RDC	Pat 1 SMP 1 SMP 2 Pri 11	N. et P. Thivillon N. et P. Thivillon
<i>C. cephus</i>	<i>cephodes</i> <i>cephus</i> <i>ngottoensis</i>	Gabon Lefini, Congo Gabon Makokou, Gabon Ngotto, RCA	GA313, GA314 R16515, R16530, R16533 LBV3 Pri 7 R12115, R12117, R12470	
<i>C. erythrogaster</i>	<i>pokoeki</i>	(Lagos?), Nigeria	Mulhouse 1	J.-M. Lemould
<i>C. erythrotis</i>	<i>camerounensis</i>	Cameroun	Pri 16	
<i>C. petaurista</i>	<i>buettikoferi</i> <i>petaurista</i>	Mont Nimba, Guinée Taï, RCI	1908.1 Tai	P. Formenty

Toutes les séquences étant de même longueur (539 bases), l'alignement a été réalisé manuellement. Les 23 premières positions ont été éliminées.

Nous avons construit des arbres des haplotypes selon deux méthodes : sous forme d'arbres d'étendue minimale établis à partir des distances corrigées par la formule de Kimura à deux paramètres,

et sous forme de réseaux parcimonieux d'haplotypes, à partir des séquences elles-mêmes.

Dans le but de tester la monophylie des espèces et de procéder à des datations, nous avons également construit des arbres phylogénétiques. Nous avons analysé les données par la méthode MP, en procédant à une recherche heuristique sans pondération avec dix additions aléatoires des séquences. Cette analyse a été complétée par un test de bootstrap avec 1000 répliques et une seule addition aléatoire de séquences. Nous avons également employé la méthode ML, en utilisant le modèle d'évolution moléculaire le mieux supporté par les données, identifié au moyen d'un test du taux de vraisemblance. La monophylie des espèces a été testée, en comparant la longueur des arbres MP et ML non contraints à celle d'arbres contraints, par un test de Kishino-Hasegawa. Pour dater les principaux événements de divergence au sein du groupe, nous avons construit un second arbre ML en imposant le respect de l'horloge moléculaire, et comparé les arbres par un test de Kishino-Hasegawa. Le principe de l'horloge n'étant pas rejeté, nous avons alors procédé à une datation en nous basant sur les distances ML obtenues.

La calibration de l'horloge a été effectuée en deux temps. Comme la divergence entre Cercopithecini et Papionini, datée à partir des fossiles, est supposée ancienne par rapport à la super-espèce *C. cephus*, et que nous avons séquencé l'ARN 16s pour peu d'espèces extérieures à ce groupe, nous avons recherché un point d'ancrage intermédiaire. Nous avons utilisé le temps de divergence moyen entre les groupes externes et les représentants de la super-espèce *C. cephus* calculé dans le chapitre 5.I à partir de séquences d'ARN 12s (soit 3,6 MA) pour calibrer une horloge propre à l'ARN 16s. Les temps de divergence moyens et leur écart-type ont été calculés à partir des distances entre clades repérés sur l'arbre ML respectant l'horloge moléculaire.

2) Résultats

Relations entre haplotypes

L'alignement des séquences ne nécessite aucune insertion ni délétion. Parmi les 516 caractères analysés, 467 sont constants, 49 sont variables et 25 sont informatifs pour la parcimonie. Si on exclut les groupes externes, il reste 39 caractères variables dont 21 informatifs pour la parcimonie. Les trois spécimens de *C. c. cephus* de la Léfini (R16515, R16530 et R16533) partagent le même haplotype. Nous nous y référerons sous le nom de *C. c. cephus* 1, tandis que les haplotypes des spécimens Pri7 et LBV3 porteront respectivement les numéros 2 et 3. Les trois spécimens de *C. c. ngottoensis* provenant de Ngotto (R12115, R12117, R12470) partagent eux aussi un même haplotype.

La longueur des 26 arbres MP non contraints est de 68 pas (**fig. 4.10 A**). Ces arbres sont caractérisés par un indice de cohérence relativement élevé, mais un indice de cohérence normé moyen, ce qui reflète sans doute le petit nombre de spécimens inclus dans l'analyse (CI = 0,721, RC = 0,503 et RI = 0,698). Ils ne sont pas significativement plus courts que l'arbre contraint pour la monophylie des espèces ($l = 73$ pas, $p = 0,095$). De même, les 44 arbres ML non contraints ne sont pas significativement plus courts que l'arbre ML contraint ($-\ln L = 1132,6$ contre $1148,8$ - $p = 0,104$).

L'arbre des haplotypes obtenu par la méthode phénétique (**fig. 4.11 A**) est sensiblement plus long que l'arbre obtenu par la méthode de parcimonie cladistique (**fig. 4.11 B**, $l = 64$ pas contre 49). Les deux arbres diffèrent essentiellement dans la position de *C. c. cephus 1* et de *C. p. petaurista*. Ils diffèrent aussi par leur enracinement, qui se fait au niveau de *C. p. petaurista* pour l'arbre d'étendue minimale (**fig. 4.11 A**) et de *C. a. ascanius* pour le réseau parcimonieux d'haplotypes (**fig. 4.11 B**). Dans ces deux arbres comme dans l'arbre MP, les espèces *C. c. cephus* et *C. petaurista*, ainsi que la sous-espèce *C. c. cephus*, sont polyphylétiques.

Datation

Les 5 arbres ML obtenus en imposant le respect de l'horloge moléculaire ne sont pas significativement plus long que les 44 arbres non contraints ($-\ln L = 1141,9$ contre $1141,36$, $p = 0,616$). Le principe de l'horloge est donc accepté.

En considérant que la divergence entre les groupes externes et la super-espèce *C. cephus* a eu lieu il y a 3,6 MA, le taux de substitution serait de 1,439 substitutions par million d'années pour l'ARN 16s. Les temps de divergence estimés et leur écart-type sont reportés sur l'un des arbres ML respectant l'horloge moléculaire (**fig. 4.10 B**).

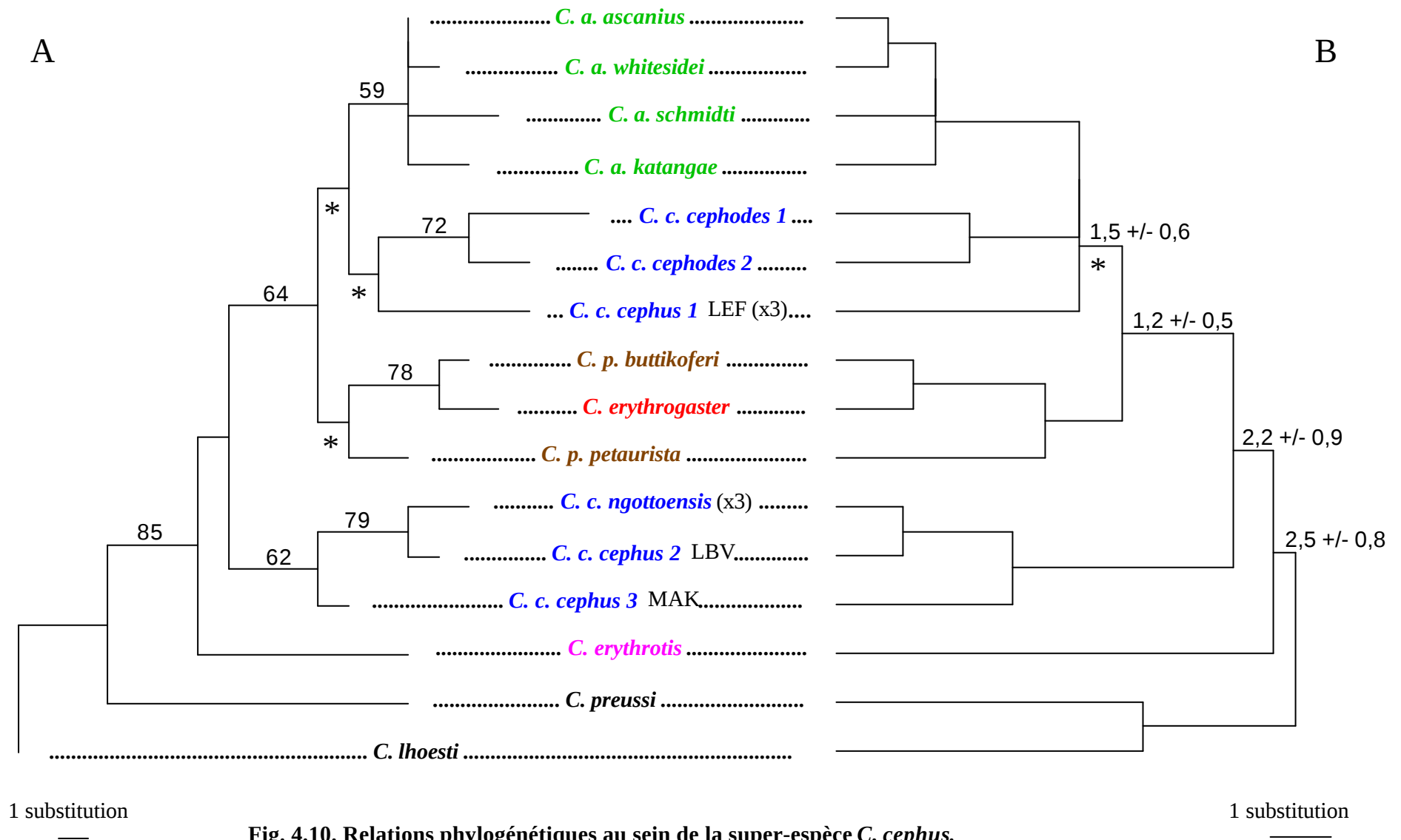


Fig. 4.10. Relations phylogénétiques au sein de la super-espèce *C. cephus*.

A. un des 26 arbres MP, avec support bootstrap (1000 réplifications).

B. un des 5 arbres ML obtenus en imposant le respect de l'horloge moléculaire, avec datation.

Les arbres présentés ont été choisis parce que leur topologie était voisine de celle du consensus à la majorité.

* : branche présente dans moins de 100% des arbres.

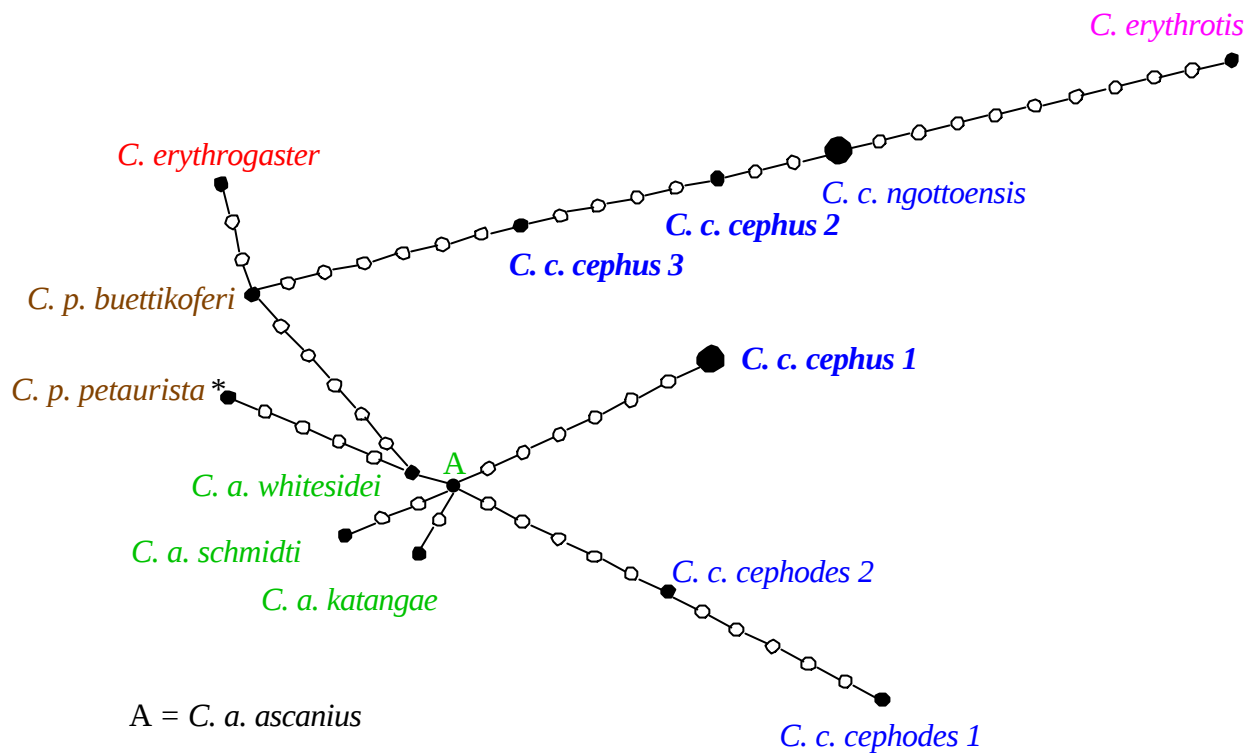
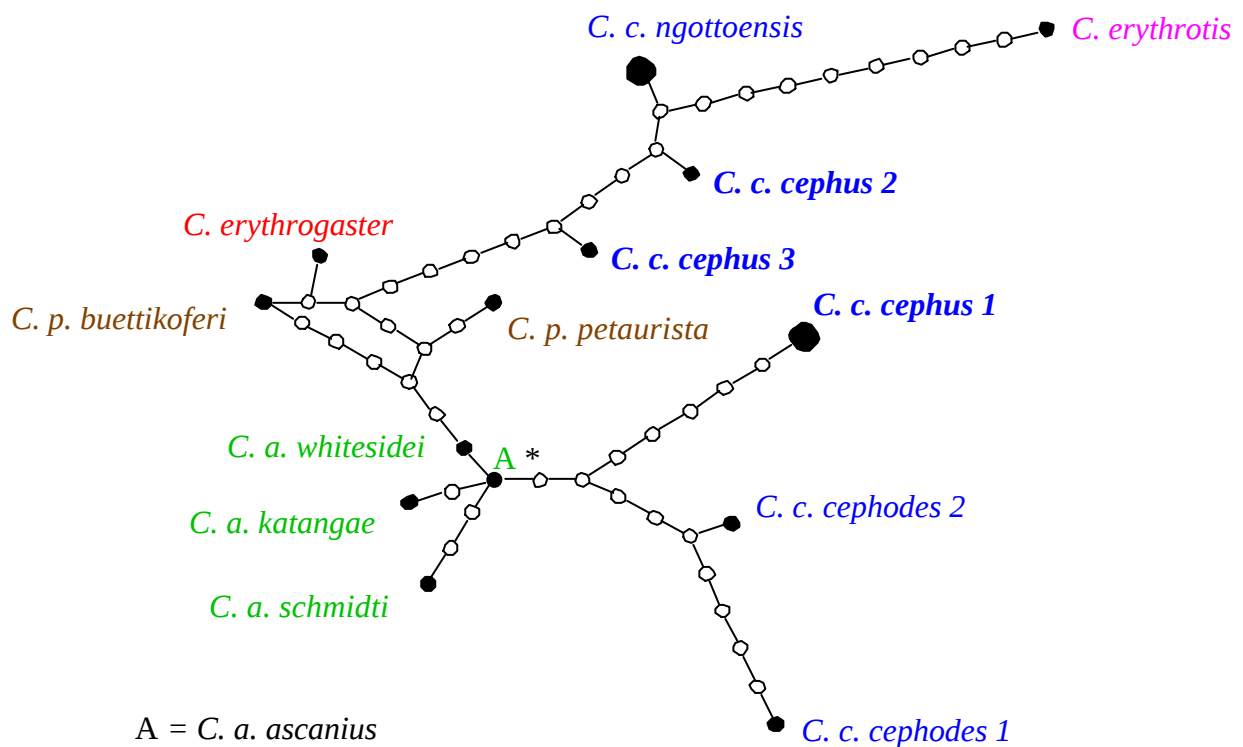


Figure 4.11. Arbres des haplotypes de la super-espèce *C. cephus*.

A. arbre d'étendue minimale ; B. réseau parcimonieux d'haplotypes.

Chaque espèce est représentée par une couleur différente. Les astérisques repèrent les haplotypes supposés "ancestraux". La taille des points noirs est proportionnelle au nombre d'haplotypes. Les points blancs représentent des haplotypes intermédiaires potentiels, non échantillonnés.



Phylogénie et phylogéographie

Dans les deux cas où des haplotypes identiques ont été trouvés, les spécimens proviennent de la même localité. Ce sont les deux seuls cas où plusieurs spécimens d'une même localité ont été analysés. Le peu de spécimens analysés ne nous permet pas de généraliser ce résultat à partir de nos seuls travaux. En revanche, van der Kuyl et collaborateurs (2000b) ont montré qu'il existe une certaine ségrégation géographique des haplotypes chez les Cercopithecini et les Papionini, en conséquence de quoi l'ADN mitochondrial peut être utilisé pour retrouver l'origine géographique de spécimens captifs. Cette ségrégation de l'ADN mitochondrial s'explique par le fait que, chez les cercopithèques, les mâles migrent alors que les femelles ont tendance à rester dans leur groupe natal (e.g. Pusey et Packer, 1987; Gautier-Hion *et al.*, 1999).

La grande diversité des haplotypes trouvés pour une même espèce ou sous-espèce fait qu'il n'est pas possible de considérer qu'un haplotype est représentatif de toute l'aire de distribution de l'espèce ou de la sous-espèce à laquelle il appartient. Ce résultat était prévisible pour la raison évoquée ci-dessus. Il a déjà été rapporté pour d'autres espèces de Primates, notamment chez des Macaques, chez qui la dispersion est fortement biaisée en faveur des mâles (Melnick et Hoelzer, 1993). Il en résulte que, étant donné le faible échantillonnage géographique dont nous disposons, les relations phylogénétiques entre haplotypes ne peuvent pas être interprétées en terme de relations phylogéographiques.

Datation

Nous avons tenté de limiter le biais dû à une grande distance temporelle entre les taxons utilisés pour calibrer l'horloge et les taxons étudiés, en utilisant une datation intermédiaire pour calibrer l'horloge et en appliquant le modèle de correction des distances qui semblait le plus adapté aux données. D'après nos estimations, le polymorphisme génétique au sein de la super-espèce *C. cephus* daterait d'il y a 1 à 3 millions d'années. Toutefois, nous sommes conscients de l'imprécision des datations proposées, qui est d'autant plus grande que le nombre d'échantillons est limité, le nombre de substitutions faible et le polymorphisme spécifique élevé. Ce dernier facteur fait que les temps de divergence entre espèces proches risquent d'être fortement surestimés (Nichols, 2001).

Polyphyly et hybridation

Bien que la monophylie des espèces ne soit pas rejetée par le test de Kishino-Hasegawa, les espèces *C. c. cephus* et *C. petaurista*, ainsi que la sous-espèce *C. c. cephus*, apparaissent polyphylétiques dans tous les arbres (fig. 4.10 et 4.11). L'haplotype de *C. p. petaurista* est plus proche de celui de *C. erythrogaster* que de celui de *C. p. buettikoferi*. Ceci ne peut s'expliquer par une proximité géographique, car l'aire de distribution de la sous-espèce *C. p. buettikoferi* est située entre celles des deux autres sous-espèces. Dans le cas de *C. cephus*, les haplotypes sont extrêmement éloignés et se regroupent d'une part avec *C. p. petaurista*, sous-espèce ayant la distribution la plus occidentale, et d'autre part avec *C. ascanius*, espèce géographiquement voisine de *C. cephus* mais dont la distribution s'étend au Sud du fleuve Congo. Une telle situation pourrait résulter d'une convergence phénotypique entre individus provenant de lignées et de régions éloignées se rejoignant dans une même région géographique, peut-être en raison d'un phénomène de sélection sexuelle, mais cette explication nous paraît très improbable.

Une autre interprétation est envisageable. Dans les deux arbres obtenus par des méthodes de parcimonie, l'haplotype *C. c. cephus 1* se groupe avec les deux haplotypes de *C. c. cephodes* (fig. 4.10 A et 4.11 B). Les deux sous-espèces se distinguent par la couleur de leur queue qui est rouge pour *C. c. cephus* et grise pour *C. c. cephodes*. Les trois spécimens correspondant à l'haplotype *C. c. cephus 1* avaient bien la queue rouge comme leurs conspécifiques. Si les groupements mis en évidence sur les arbres les plus parcimonieux sont justes, cette opposition entre caractères morphologiques et caractères moléculaires suggère qu'il y a peut-être eu hybridation entre *C. c. cephus* et *C. c. cephodes*. Cette hypothèse est d'autant plus vraisemblable que les trois individus de *C. c. cephus* concernés proviennent de la réserve de la Lefini, une galerie forestière située entre les aires de répartition de *C. c. cephus* et *C. c. cephodes*. De la même façon, l'existence de deux lignées mitochondriales chez *C. cephus*, pourrait avoir pour cause une hybridation entre *C. cephus* et *C. ascanius*. Pour tester cette hypothèse, il faudrait séquencer de l'ADN nucléaire, et de préférence des gènes portés par le chromosome Y, de façon à comparer l'information génétique transmise par la mère à celle transmise par le père ou par les deux parents.

Chez les Cercopithèques, plusieurs cas d'hybridation entre sous-espèces et espèces d'une même super-espèce ont été repérés dans la nature (e.g. Lemoult, 1988). Les hybrides entre espèces n'appartenant pas à la même super-espèce ne sont pas rares en captivité (Lemoult, 1988), et des hybrides F1 et F2 entre *C. ascanius schmidtii* et *C. mitis stuhlmanni* ont été observés dans la nature, à Kibalé au Kenya (Struhsaker *et al.*, 1988). Une étude éco-éthologique de ces hybrides indique qu'alors que les individus hybrides de première génération présentent des traits intermédiaires entre les deux

espèces parentes, les individus issus de mères hybrides ressemblent fortement à l'espèce à laquelle appartient le père. L'hybridation n'est donc presque plus décelable dès la deuxième génération. Comme les femelles restent généralement dans leur groupe social, si les femelles hybrides se reproduisent avec des individus appartenant à l'espèce à laquelle elles appartiennent socialement, leurs descendants présenteront le pattern de coloration de l'espèce au sein de laquelle ils vivent. Leur ADN mitochondrial, transmis par la mère, sera aussi celui de l'espèce d'accueil, en conséquence de quoi l'hybridation ne pourra pas être décelée par comparaison de séquences mitochondriales. Nos résultats ne peuvent donc être interprétés comme résultant d'une hybridation que si cette situation n'est pas générale et que, dans certains cas, les femelles, hybrides ou non, migrent pour rejoindre un groupe social qui n'appartient pas à leur espèce maternelle. Il est à noter que, même si cette situation se produit rarement, l'haplotype peut se répandre rapidement dans la population par sélection des lignées (cf. *Avise et al.*, 1987). Il serait intéressant de mener des études éthologiques et génétiques dans des régions charnières entre deux espèces ou sous-espèces de façon à savoir si ce cas de figure se produit.

La théorie de la coalescence nous fournit une autre hypothèse permettant d'expliquer la polyphylie des espèces. La séparation entre espèces pourrait être trop récente, le polymorphisme ancestral trop élevé, et le processus de coalescence trop lent pour que les espèces aient atteint le stade de monophylie génétique (e.g. *Excoffier*, 1997 ; *Maddison*, 1997). Le schéma observé résulterait alors d'une répartition plus ou moins aléatoire du stock d'haplotypes ancestraux. Pour tester cette hypothèse, il faudrait analyser beaucoup plus d'individus, en provenance de localités connues couvrant l'ensemble de l'aire de répartition de la super-espèce, et utiliser l'ADN nucléaire. L'utilisation de l'ADN mitochondrial en phylogéographie a été critiquée en raison de la rapidité et de la stochasticité du processus de sélection des lignées, du fait qu'un allèle avantageux peut se répandre rapidement, et des risques d'introgression inter-spécifique (e.g. *Moritz et Faith*, 1998).

En raison de la polyphylie et du taux de polymorphisme élevé des haplotypes d'une même espèce ou sous-espèce, nous ne pouvons considérer un haplotype obtenu pour un individu d'une sous-espèce comme représentatif de toute l'aire de répartition de ladite sous-espèce. Il en résulte que, compte-tenu du peu de spécimens analysés, nous ne pouvons tirer les interprétations phylogéographiques escomptées de l'étude de ce modèle.

Cet exemple probant de non-adéquation entre phénotype et génotype est cependant très instructif. Il nous rappelle qu'il faut rester prudent lors de l'utilisation de données moléculaires pour retracer une phylogénie.

IV - COMPARAISON DES MODELES DANS UNE PERSPECTIVE BIOGEOGRAPHIQUE

Nous avons tenté de synthétiser nos données phylogéographiques de façon à en tirer des informations sur les relations entre aires géographiques. Dans un premier temps, nous avons comparé les schémas phylogéographiques obtenus avec les assemblages floristiques, de façon à déceler une éventuelle influence de l'environnement sur la différenciation génétique. Dans un second temps, nous avons tenté une synthèse biogéographique. Pour cela, nous avons adopté deux démarches, l'une traitant les données par localité dans le but d'en inférer des événements de vicariance entre aires géographiques identifiées par l'analyse, l'autre analysant les relations entre aires géographiques prédéfinies en fonction des résultats de la première analyse. Nous n'avons pas pu incorporer l'étude de la super-espèce *C. cephus* dans cette synthèse, du fait de la polyphylie des haplotypes correspondant à une même sous-espèce et du faible échantillonnage géographique. En conséquence, nous n'avons traité que les schémas phylogéographiques obtenus pour les deux modèles musaraignes et les deux modèles rongeurs présentés dans les chapitres 4.I et 4.II. Ce nombre limité de modèles ne nous a pas permis d'avoir recours à des approches telles que la parcimonie de Brooks ou la "Component Analysis".

1) Matériel et méthodes

Pour rechercher une éventuelle corrélation entre groupements d'haplotypes et phytorégions, nous avons superposé les schémas phylogéographiques mis en évidence sur les arbres d'étendue minimale avec la carte des phytorégions de Letouzey (1968).

Dans le but de déceler un éventuel message biogéographique commun à plusieurs modèles, nous avons d'abord analysé nos données sans prédéfinir les aires géographiques, selon la méthode de Hovenkamp (1997). Pour cela, nous avons identifié les événements cladogénétiques opposant deux ensembles de localités sur les arbres phénétiques utilisés pour les datations, et nous les avons traduits en séquences d'événements. Pour plus de simplicité, nous avons supprimé de ces séquences les localités représentées dans un seul modèle. Les événements cladogénétiques étant peu nombreux, nous avons essayé d'apporter une amélioration à la méthode en utilisant les temps de divergence estimés pour ordonner les événements propres à chaque modèle, puis pour les intercaler d'un modèle à l'autre.

Nous avons ensuite tenté d'évaluer le support apporté par nos modèles aux régions biogéographiques prédéfinies d'après des travaux biogéographiques antérieurs (cf. fig. 1.3) et

remodelées en fonction des résultats de l'analyse précédente. Les régions Ouest Cameroun et Est Congo n'ont pas été considérées car elles n'ont été traitées que pour un ou deux modèles. Pour chacune des espèces, nous avons construit des arbres MP contraints en imposant une stricte monophylie des régions retenues. Nous avons compté le nombre de synapomorphies soutenant chaque région, à l'aide du programme MacClade 3.06 (Maddison et Maddison, 1996). Nous avons aussi cherché s'il existait des associations préférentielles entre ces régions, en construisant un diagramme de Venn. Pour cela, nous avons mesuré les distances génétiques, corrigées par la formule de Kimura à 2 paramètres, associées à chacune des branches identifiées sur des arbres NJ. Nous avons synthétisé les résultats sous forme de diagramme de Venn et construit l'arbre des relations entre aires associé.

Pour finir, nous avons estimé la diversité génétique au sein de chacune des régions prédéfinies. Nous avons calculé la distance génétique moyenne entre tous les haplotypes de chaque région, à partir des distances corrigées par la formule de Kimura à deux paramètres. Afin de supprimer le biais dû à la taille de la zone géographique couverte et à la validité ou non des régions selon les modèles, nous avons également comparé la diversité génétique moyenne entre les localités les mieux échantillonnées de chaque région.

2) Résultats

Recherche de corrélation entre schémas phylogéographiques et phytorégions

Certains groupements d'haplotypes s'inscrivent entièrement dans une zone d'homogénéité floristique ou de forêt mixte. Chez *S. johnstoni*, en particulier, les haplotypes des localités situées au Sud de l'Ogoué sont génétiquement proches des haplotypes trouvés en Centrafrique, et toutes les localités concernées se trouvent dans une forêt semi-caducifoliée ou mixte (**fig. 4.12 A**). Cependant, pour tous les modèles, on observe des groupements qui se superposent avec deux ou trois types d'associations végétales (**fig. 4.12 A à D**). Pour aucun des quatre modèles il n'existe de correspondance directe entre les principaux groupements d'haplotypes et l'environnement floristique.

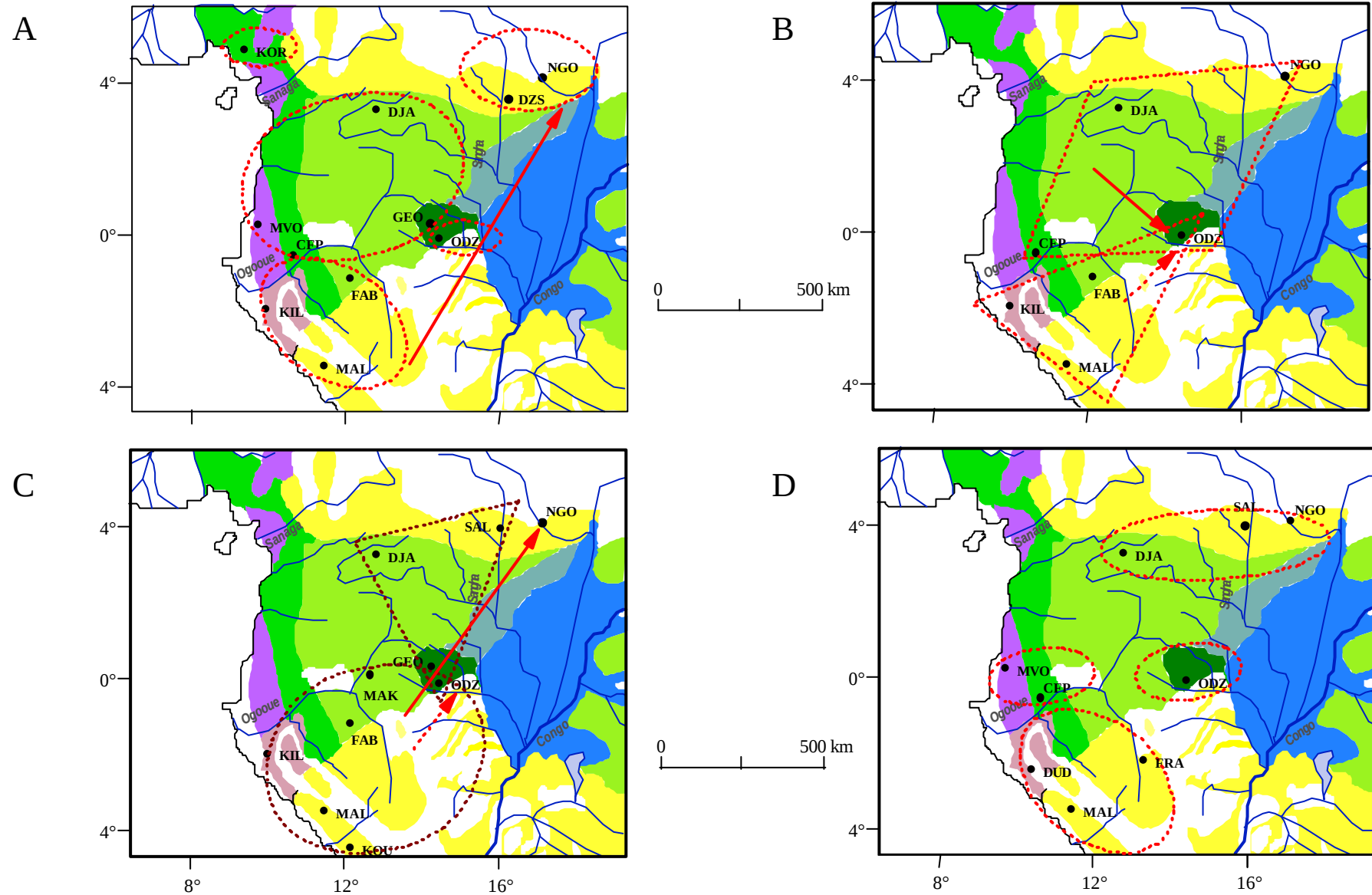


Figure 4.12. Comparaison des groupements phylogénétiques observés chez les quatre modèles avec les phytorégions définies par Letouzey (1968).
A. *S. johnstoni* ; B. *S. ollula* ; C. *H. stella* ; D. *S. longicaudatus*. Les pointillés rassemblent les localités dont les haplotypes se groupent dans les analyses. Les flèches rouges indiquent des événements probables de colonisation et de dispersion sur longue distance - cf. carte 1.5 pour la légende de la végétation.

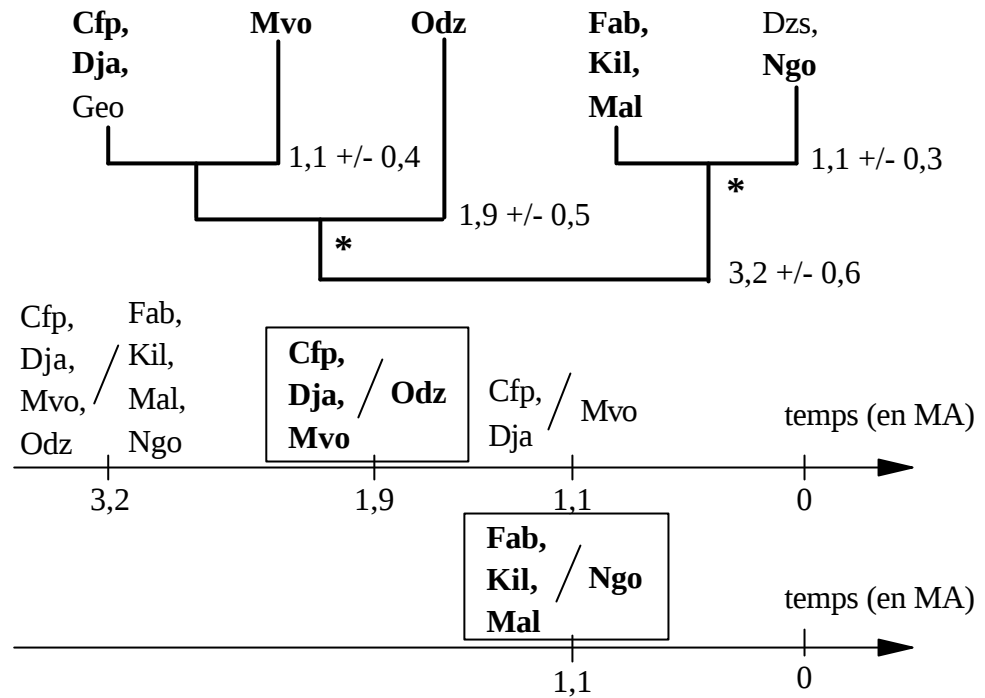
Synthèse biogéographique selon la méthode de Hovenkamp

Le modèle *H. stella* n'a pas été retenu pour l'analyse parce qu'il n'est pas possible de définir des relations phylogénétiques précises entre les haplotypes d'Afrique centrale. Pour les autres modèles, nous avons reconstitué la séquence des événements cladogénétiques selon leur ordre chronologique (**fig. 4.13**). Dans le cas de *S. johnstoni*, nous avons dû construire deux séquences indépendantes en raison de l'existence de deux événements simultanés. Nous n'avons pu identifier aucun événement de divergence exactement commun à deux modèles. Cependant, la divergence entre Cfp-Dja-Mvo et Odzala observée pour *S. johnstoni* est incluse dans la divergence entre Cfp-Dja-Mvo-Ngo et Fra-Mal-Odz observée pour *S. longicaudatus*. Elle peut donc être retenue. Les datations proposées pour les deux événements sont d'ailleurs assez proches (1,9 et 2,0 MA). De même, la divergence entre Fab-Kil-Mal et Ngotto observée pour *S. johnstoni* est incluse dans la divergence entre Fab-Kil-Mal et Cfp-Dja-Ngo observée pour *S. ollula*. Les datations proposées pour les deux événements sont elles aussi voisines (1,1 et 1,3 MA). Toutefois, ce dernier événement fait intervenir les localités Cfp et Dja, qui sont impliquées dans des événements de divergence incompatibles survenus vers la même période pour *S. johnstoni* (1,1 MA) et *S. longicaudatus* (0,8 MA). Le premier des deux événements confirmés oppose trois localités situées dans la région d'intergradation à Odzala, et le second oppose trois localités de la région Sud Ogooué à Ngotto (**fig. 4.14 A**).

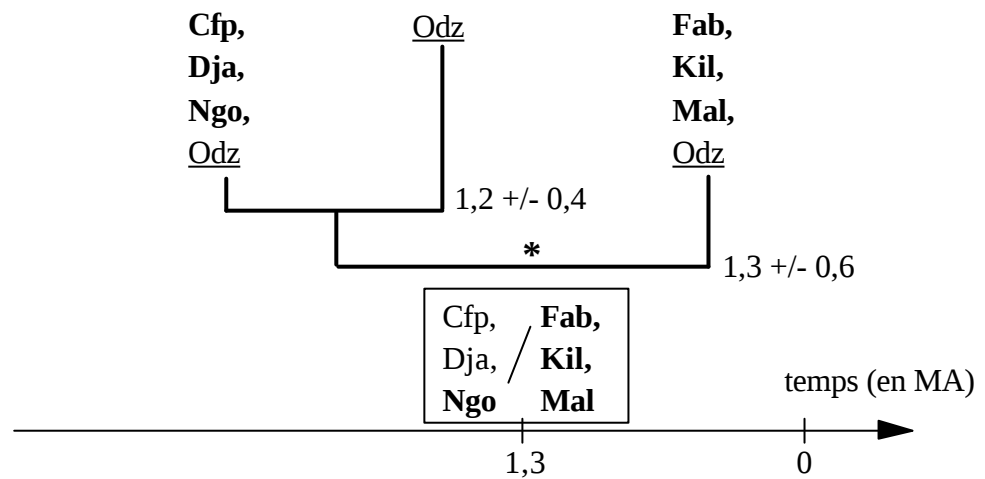
De ce fait, nous avons retenu les trois régions mises en évidence par des travaux biogéographiques antérieurs (cf. **fig. 1.3**), à savoir, les régions Ouest Congo, Sud Ogooué et d'intergradation. Nous avons scindé la région Ouest Congo en deux régions appelées Ouest Congo 1 et Ouest Congo 2, la première correspondant à Odzala et l'autre à Ngotto, Dzangha et Salo² (**fig. 4.14 B**).

Afin de mieux visualiser en quoi chacun des schémas phylogéographiques obtenus diffère des régions fauniques telles qu'elles ont été définies par Colyn et collaborateurs, nous avons recherché la meilleure adéquation entre ces régions et les groupements d'haplotypes (**fig. 4.15**). Ces derniers se superposent assez bien avec les régions ou sous-régions fauniques pour *S. johnstoni* (**fig. 4.15 A**), et dans une moindre mesure pour *S. longicaudatus* (**fig. 4.15 D**). Pour ce second modèle, les haplotypes du Dja se groupent avec ceux de la région Ouest Congo 2 plutôt qu'avec ceux de la région d'intergradation. Pour *S. ollula* et *H. stella*, en revanche, les ensembles d'haplotypes couvrent plusieurs régions (**fig. 4.15 B et C**). Pour ces deux modèles, Odzala est présent dans les deux principaux groupements d'haplotypes mis en évidence et constituerait donc une zone de convergence.

S. johnstoni



S. ollula



S. longicaudatus

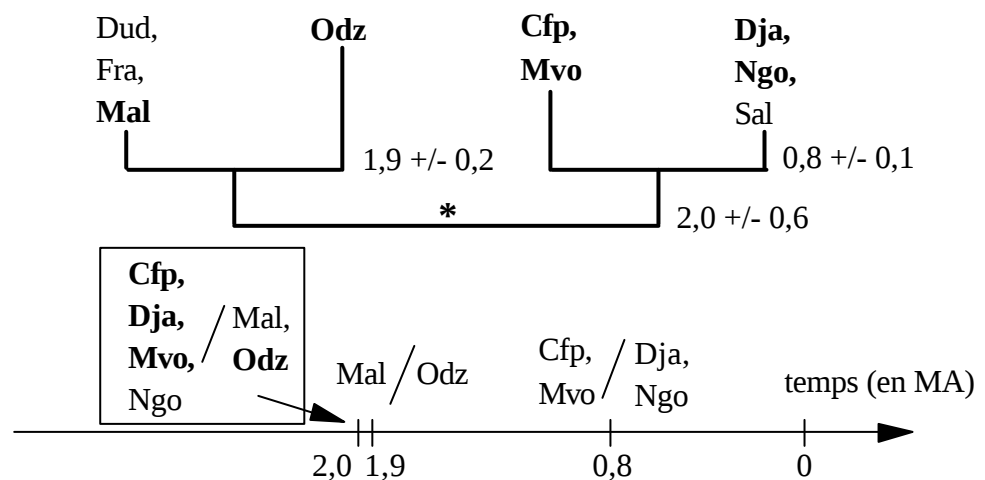


Figure 4.13. Traduction des événements de vicariance ayant affecté trois des modèles sous forme de séquence linéaire, selon la méthode de Hovenkamp.

Pour *S. ollula*, la localité Odzala est supprimée de la séquence car elle est présente dans tous les clades. Les astérisques et les encadrés repèrent les événements supportés par 2 modèles.

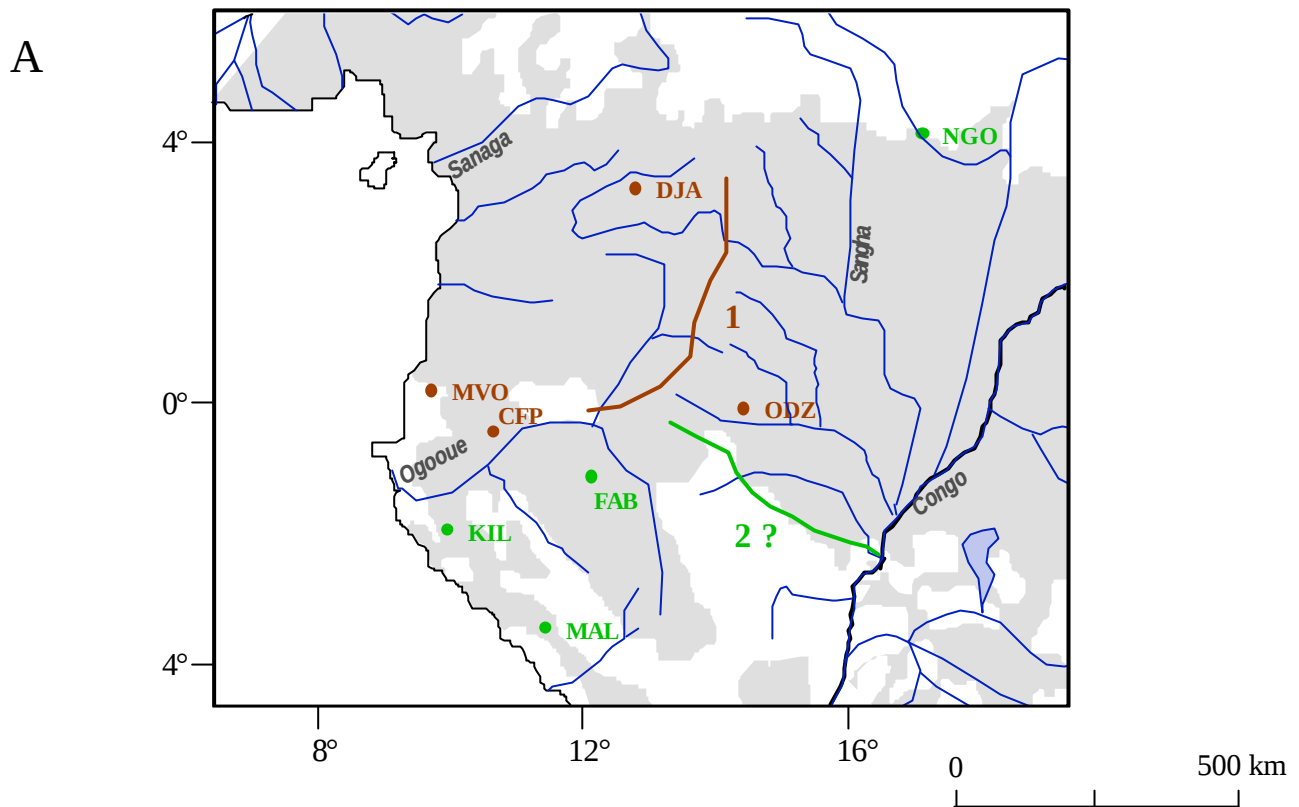
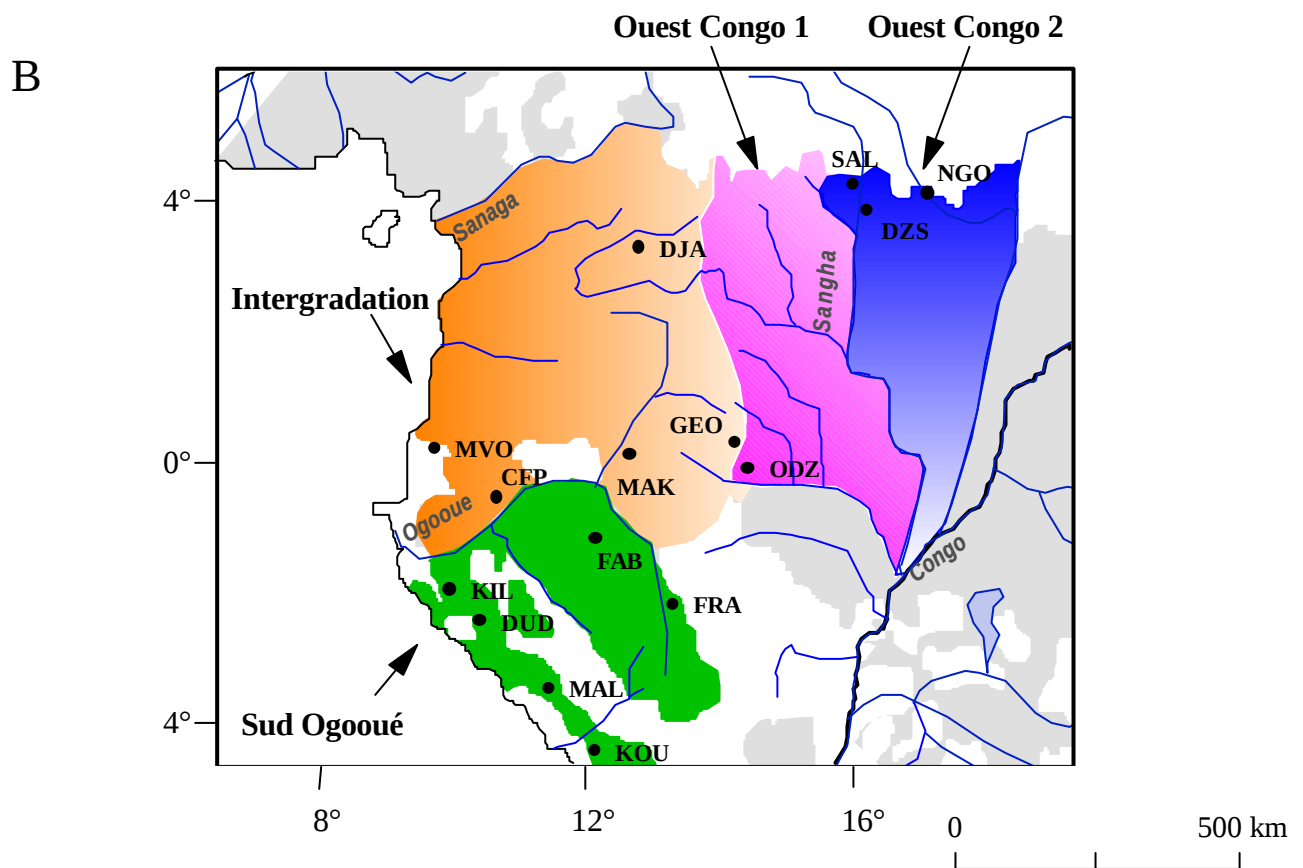


Figure 4.14. A. Événements cladogénétiques partagés par au moins deux modèles mis en évidence en appliquant la méthode de Hovenkamp. Chaque événement est repéré par une couleur et par une ligne brisée dont la localisation est arbitraire.

B. Régions fauniques redéfinies à l'issue des analyses biogéographiques. Les limites des régions sont inspirées des travaux de Colyn et collaborateurs (cf. carte 1.3).



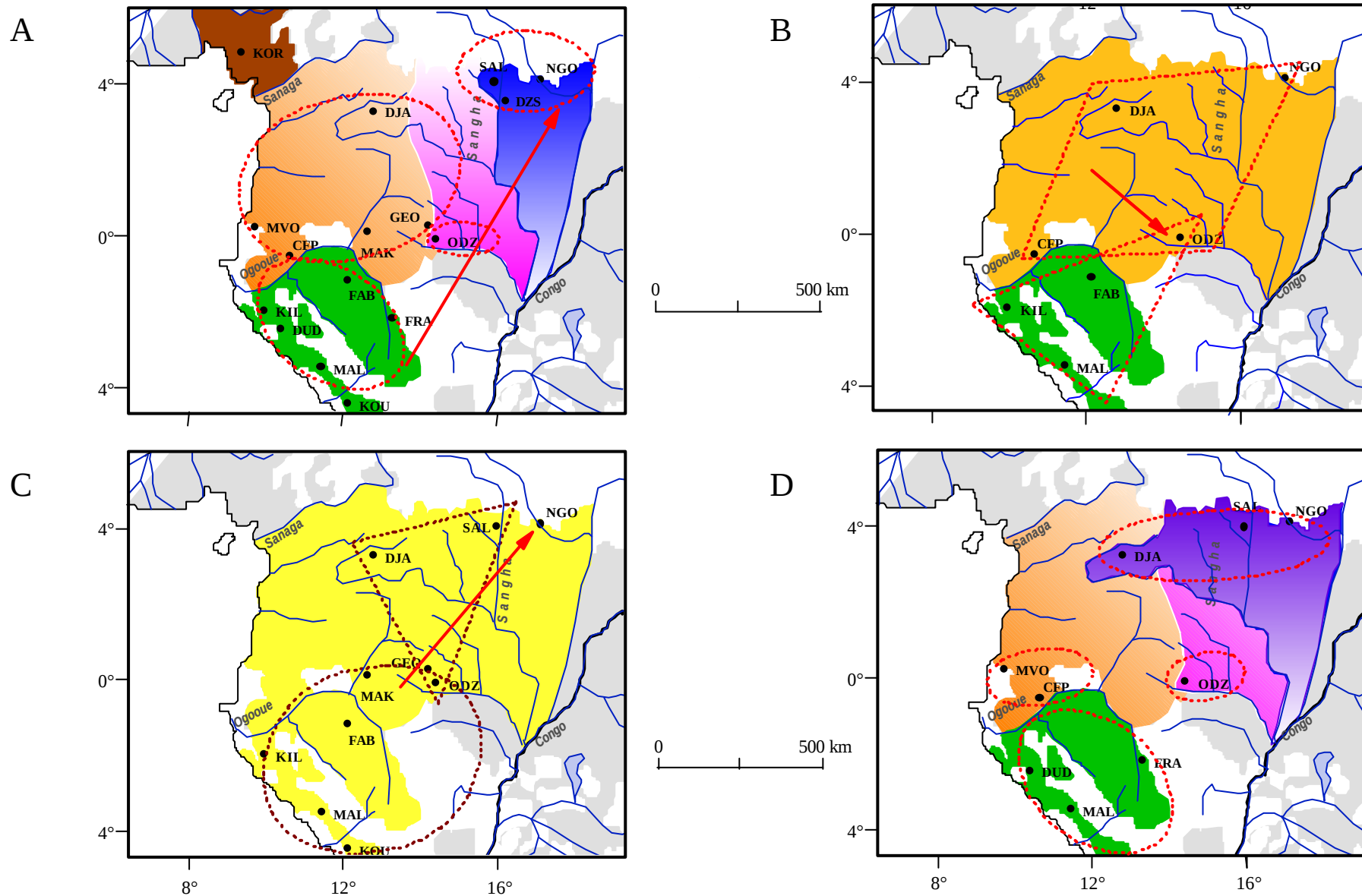


Figure 4.15. Comparaison des groupements phylogénétiques observés chez les quatre modèles avec les régions fauniques définies par M. Colyn. A. *S. johnstoni* ; B. *S. ollula* ; C. *H. stella* ; D. *S. longicaudatus*. Les pointillés rassemblent les localités dont les haplotypes se groupent dans les analyses. Les flèches rouges indiquent des événements probables de colonisation et de dispersion sur longue distance. Les surfaces colorées représentent des régions d'homogénéité haplotypique, dont les limites hypothétiques sont adaptées des travaux de M. Colyn.

Evaluation des régions fauniques et des relations entre régions

La mesure du nombre de synapomorphies soutenant chaque région prédéfinie révèle que seule la monophylie des régions Ouest Congo 1 et Sud Ogooué est bien supportée (14 synapomorphies contre 2 à 4 pour les deux autres régions ; tab. 4.9). La monophylie de la région Ouest Congo 1 est soutenue par deux modèles seulement, alors que cette région n'inclut qu'une seule localité. Celle de la région Sud Ogooué est soutenue par trois modèles, mais le support apporté par l'un de ces modèles est très faible.

Tableau 4.9 : Nombre de synapomorphies soutenant chacune des régions prédéfinies.

Régions	<i>S. johnstoni</i>	<i>S. ollula</i>	<i>H. stella</i>	<i>S. longic.</i>	Total
Ouest Congo 1	7	0	0	7	14
Ouest Congo 2	3	0	1	0	4
Sud Ogooué	1	6	0	7	14
Intergradation	2	0	0	0	2

S. longic. = *S. longicaudatus*

Les distances génétiques mesurées sur les arbres phénétiques construits pour chacun des modèles (**fig. 4.16**) sont résumées dans le tableau 4.10. Le cumul des distances génétiques régionales est globalement plus élevé pour *S. johnstoni* que pour les autres modèles. Le couple de régions le mieux supporté (16 synapomorphies) est le couple WCongo 2 / S Ogooué. Toutefois, il n'est sérieusement soutenu que par *S. johnstoni*. Trois autres couples sont moyennement supportés (8-9 synapomorphies), mais par deux modèles à la fois. Ce sont W Congo 1 / Intergradation, W Congo 1 / S Ogooué, et W Congo 2 / Intergradation. Parmi les associations de trois régions, seul le trio W Congo 1 / W Congo 2 / S Ogooué est un peu supporté, par deux modèles.

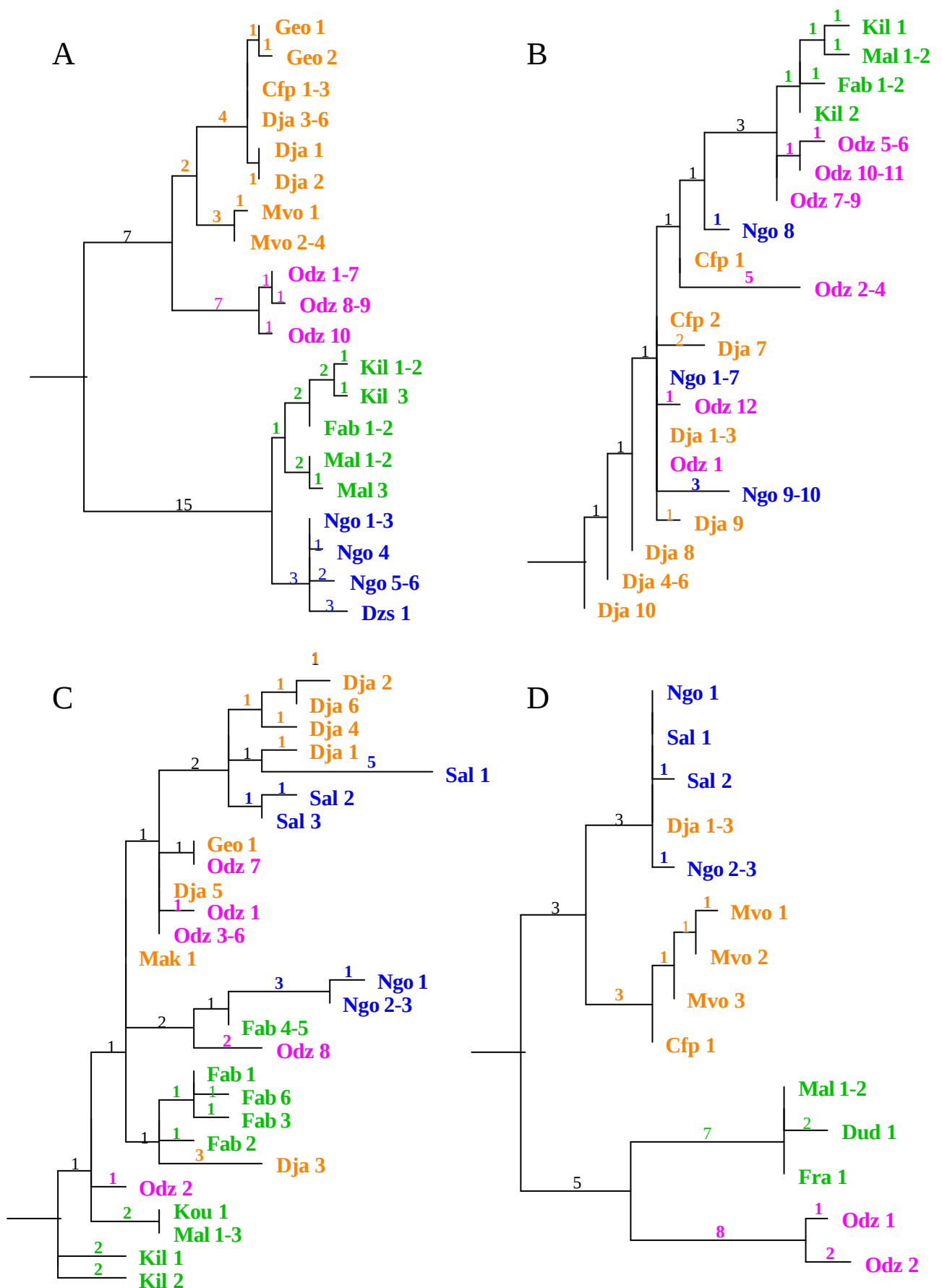


Figure 4.16. Arbres NJ avec longueur des branches exprimées en nombre de substitutions.

A. *S. johnstoni* ; B. *S. ollula* ; C. *H. stella* ; D. *S. longicaudatus*. Chaque couleur fait référence à une région différente (violet : Ouest Congo 1, bleu : Ouest Congo 2, vert : Sud Ogooué, et orange : intergradation).

Tableau 4.10 : Distances génétiques intra- et inter- régions, mesurées sur des arbres NJ et exprimées en nombre de substitutions.

Régions	<i>S. johnstoni</i>	<i>S. ollula</i>	<i>H. stella</i>	<i>S. longic.</i>	Total
Ouest Congo 1	10	8	4	11	33
Ouest Congo 2	9	4	11	2	26
Sud Ogooué	10	5	10	9	34
Intergradation	13	3	8	6	30
W Congo 1, W Congo 2	0	0	0	0	0
W Congo 1, S Ogooué	0	3	0	5	8
W Congo 1, Intergradation	7	0	1	0	8
W Congo 2, S Ogooué	15	0	1	0	16
W Congo 2, Intergradation	0	0	3	6	9
S Ogooué, Intergradation	0	0	1	0	1
W Congo 1, W Congo 2, S Ogooué	0	1	2	0	3
W Congo 1, W Congo 2, Intergradation	0	0	1	0	1
W Congo 1, S Ogooué, Intergradation	0	0	0	0	0
W Congo 2, S Ogooué, Intergradation	0	0	0	0	0
Toutes	0	4	2	0	6

S. longic. = *S. longicaudatus*

Le diagramme de Venn (**fig. 4.17 A**) révèle que, malgré les variations du nombre de localités et d'haplotypes par région (cf. tab. 4.11), toutes les régions, prises isolément, ont une diversité comparable. Sa traduction sous forme de cladogramme des aires (**fig. 4.17 B**) met en évidence une association préférentielle entre les régions Ouest Congo 2 et Sud Ogooué.

Diversité génétique régionale et locale

La comparaison de la diversité génétique dans chacune des quatre régions retenues révèle une diversité élevée dans la région Ouest Congo 1 pour *S. ollula* et dans la région Ouest Congo 2 pour *H. stella* (1,1 % de divergence moyenne ; tab. 4.11). Pour ce dernier modèle, la diversité génétique est importante (Δ 0,7 %) dans trois régions sur quatre. La diversité moyenne par région est relativement plus faible dans la région Ouest Congo 2 que dans les trois autres régions, pour lesquelles elle est identique. Bien que la région Ouest Congo 1 ne comprenne qu'une localité, la diversité moyenne y est aussi importante que dans les régions Sud Ogooué et d'intergradation.

A

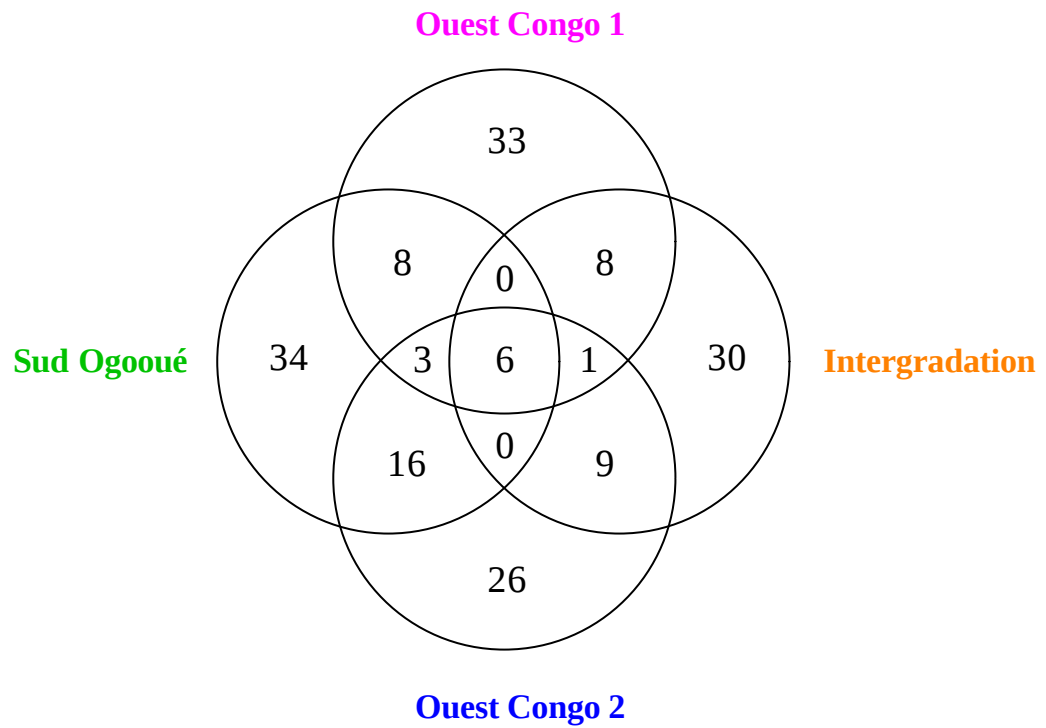


Figure 4.17. Diagramme de Venn montrant la fraction de la diversité totale partagée par plusieurs régions ou unique à chaque région (A), et sa traduction sous forme de cladogramme des aires (B).

B

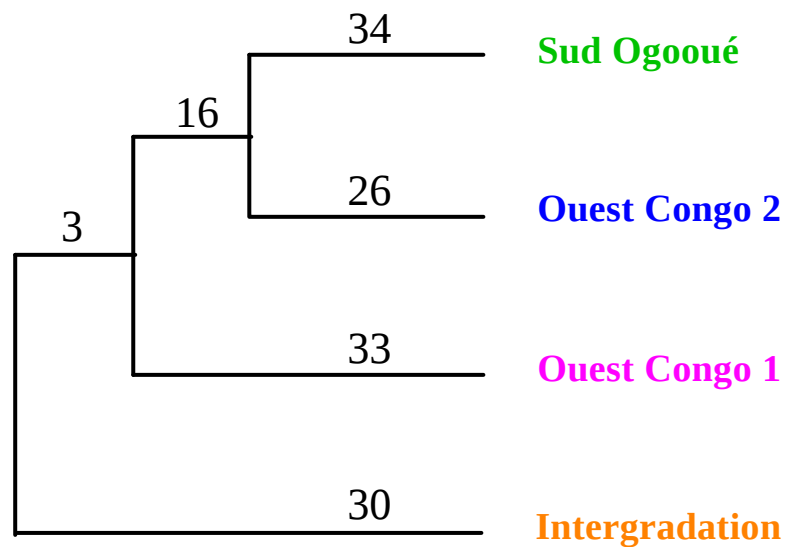


Tableau 4.11 : Nombre de localités échantillonnées (N. loc.), nombre de spécimens séquencés (N. seq.) et diversité génétique moyenne (D. moy., exprimée en pourcentage, avec écart type), par modèle et par région.

Régions		<i>S. johnstoni</i>	<i>S. ollula</i>	<i>H. stella</i>	<i>S. longic.</i>	Moyenne
Ouest Congo 1	N. loc.	1	1	1	1	1
	N. seq.	10	12	8	2	n = 32
	D. moy.	0,2 +/- 0,2	1,1 +/- 0,6	0,4 +/- 0,4	0,6	0,6
Ouest Congo 2	N. loc.	2	1	2	2	2
	N. seq.	7	10	6	5	n = 28
	D. moy.	0,4 +/- 0,3	0,2 +/- 0,2	1,1 +/- 0,7	0,2 +/- 0,1	0,4
Sud Ogooué	N. loc.	3	3	4	3	3
	N. seq.	8	6	12	4	n = 30
	D. moy.	0,7 +/- 0,5	0,3 +/- 0,2	0,8 +/- 0,3	0,2 +/- 0,2	0,6
Intergradation	N. loc.	4	2	3	3	3
	N. seq.	15	12	8	7	n = 42
	D. moy.	0,7 +/- 0,7	0,3 +/- 0,2	0,7 +/- 0,4	0,8 +/- 0,6	0,6

S. longic. = *S. longicaudatus*

La comparaison de la diversité génétique pour les localités les mieux échantillonnées de chaque région révèle une opposition entre la forte diversité moyenne à Odzala (0,6 % - n= 32) et la faible diversité moyenne à Malounga (0,03 % - n= 10 ; tab. 4.12). Notons que cette dernière localité est la moins bien échantillonnée des quatre.

Tableau 4.12 : Nombre de spécimens séquencés (N. seq.) et diversité génétique moyenne (D. moy., exprimée en pourcentage, avec écart type), pour les quatre modèles et les quatre localités les mieux échantillonnées.

Régions		<i>S. johnstoni</i>	<i>S. ollula</i>	<i>H. stella</i>	<i>S. longic.</i>	Moyenne
Odzala	N. seq.	10	12	8	2	n = 32
Ouest Congo 1	D. moy.	0,2 +/- 0,2	1,1 +/- 0,6	0,4 +/- 0,4	0,6	0,6
Ngotto	N. seq.	6	10	3	3	n = 22
Ouest Congo 2	D. moy.	0,3 +/- 0,2	0,2 +/- 0,2	0,1 +/- 0,1	0,1 +/- 0,1	0,2
Malounga	N. seq.	3	2	3	2	n = 10
Sud Ogooué	D. moy.	0,1 +/- 0,2	0	0	0	0,03
Dja	N. seq.	6	10	6	3	n = 25
Intergradation	D. moy.	0,1 +/- 0,1	0,3 +/- 0,2	0,8 +/- 0,4	0	0,3

S. longic. = *S. longicaudatus*

3) Discussion

Limites de la méthode de Hovenkamp

Les trois modèles retenus ont permis de construire quatre séquences d'événements. Seuls deux événements de divergence, présents dans l'arbre obtenu pour *S. johnstoni*, seraient confirmés par les

autres modèles. Le plus ancien de ces événements oppose les localités Four-Place, Dja et Mvoum, toutes trois situées dans la région supposée d'intergradation, à Odzala. Le plus récent oppose les localités Forêt des Abeilles, Kili et Malounga, situées dans la région Sud Ogooué, à Ngotto. Ces deux événements de vicariance nous ont permis de confirmer le découpage en trois régions déjà adopté dans les chapitres 4.I et 4.II, et de scinder la région Ouest Congo en deux. Cependant, le second de ces événements a été identifié comme résultant d'une colonisation sur grande distance et non d'une vicariance, lors de l'analyse des groupements emboîtés effectuée pour *S. johnstoni* (chapitre 4.I.2). L'absence de distinction entre événements de colonisation et de vicariance nous apparaît comme une limitation de la méthode de Hovenkamp. Nous proposons donc de commencer par rechercher les causes de chaque événement, par exemple à l'aide d'une analyse des groupements emboîtés, et de ne retenir que les événements de divergence identifiés comme résultant d'une fragmentation passée. Si nous appliquons ce principe à l'analyse réalisée ci-dessus, un seul événement de divergence confirmé subsiste.

La méthode de Hovenkamp, tout comme les autres méthodes de biogéographie cladistique, ne semble pas adaptée pour traiter à la fois peu de taxons et beaucoup de localités. Le problème du faible nombre de taxons était prévisible puisque la méthode ne retient que les événements apparaissant dans au moins deux séquences. L'utilisation des temps de divergence pour déterminer la séquence des événements nous semble un moyen utile pour conserver le maximum d'information, puisqu'elle permet d'insérer dans une même séquence des événements affectant des branches distinctes d'un cladogramme. La datation permet aussi d'envisager de retenir des taxons pour lesquels un seul événement est mis en évidence et donc aucune "séquence" d'événements ne peut être identifiée, comme c'est le cas pour *S. ollula*. Cependant, étant donné l'imprécision des datations et les variations de l'horloge moléculaire d'un taxon à l'autre, il nous semble préférable de ne pas se servir des temps de divergence pour insérer les événements d'un modèle à l'autre. Les événements isolés ne peuvent donc être utilisés que pour conforter un événement identifié dans un autre cladogramme, et survenu à une période comparable.

Evaluation des régions fauniques et des relations entre régions

Sur les quatre régions dont nous avons évalué la monophylie pour chacun de nos jeux de données, seule la monophylie des régions Ouest Congo 1 et Sud Ogooué est bien supportée, par respectivement deux et trois modèles (tab. 4.9). Celle de la région Ouest Congo 2 est également supportée par deux modèles, mais par quelques synapomorphies seulement. L'homogénéité de la région d'intergradation n'est soutenue que par un modèle et deux synapomorphies, ce qui confirme qu'elle pourrait être de nature composite. Cependant, le nombre total de synapomorphies soutenant une région n'est pas

forcément une bonne mesure de son homogénéité. Il convient également d'analyser en détail la correspondance entre régions et schémas phylogéographiques pour chacun des modèles (fig. 4.15). Parmi les quatre régions qui ressortent des analyses, seules les régions Sud Ogooué et Ouest Congo 2 apparaissent relativement homogènes et différenciées, pour respectivement trois et deux modèles (*S. johnstoni* et *S. longicaudatus* +/- *S. ollula*). Odzala apparaît comme une région différenciée pour deux modèles (*S. johnstoni* et *S. longicaudatus*) et comme une zone de convergence des faunes pour les deux autres modèles. Quand à la région dite "d'intergradation", elle présente des haplotypes monophylétiques et uniques à cette région pour toutes les localités échantillonnées chez *S. johnstoni*, et pour deux des trois localités échantillonnées chez *S. longicaudatus*. Pour les deux autres modèles, non seulement le nombre de localités est très faible, mais en plus il y aurait eu expansion de populations sur une vaste zone géographique incluant ces localités.

Nous avons également évalué le support apporté à chaque groupement de régions possibles, par la construction d'un diagramme de Venn (fig. 4.17). L'association la plus supportée est l'association entre les régions Ouest Congo 2 et Sud Ogooué qui apparaît sur le cladogramme des aires associé au diagramme de Venn (fig. 4.17 B). Ces régions sont géographiquement éloignées, mais peuvent être jointes sans traverser de barrière fluviale majeure. Cependant, cette association n'est fortement supportée que pour *S. johnstoni* et n'a donc aucun caractère général. Ce résultat pourrait découler d'un biais lié au fait que le nombre de substitutions est plus élevé pour cette espèce que pour les autres (cf. chapitre 4.I) et que les plans d'échantillonnages ne sont pas équilibrés. Cet exemple confirme la nécessité de se référer aux phylogénies de départ pour interpréter le diagramme global (cf. Moritz et Faith, 1998). Par ailleurs, deux associations ne sont pas supportées, ce sont d'une part, l'association entre les régions Ouest Congo 1 et 2, ce qui justifie la scission que nous avons effectuée au sein de cette région, et d'autre part, l'association entre les régions Sud Ogooué 1 et Intergradation, qui sont adjacentes mais séparées par le fleuve Ogooué. Le groupement des régions Sud Ogooué et Ouest Congo 1 et 2, à l'exclusion de la région d'intergradation, pourrait résulter du caractère composite de cette dernière.

Diversité génétique régionale et locale

La diversité génétique régionale a été mesurée selon deux méthodes. Dans le diagramme de Venn, seule la diversité apparue au sein de chaque région est comptabilisée (tab. 4.10), alors que la mesure de la diversité génétique régionale globale (tab. 4.11) prend en compte les distances entre tous les haplotypes présents dans la région, même s'ils sont le fruit d'un événement de colonisation. Si l'on compare les résultats obtenus modèle par modèle, les deux méthodes donnent des résultats

différents. Par contre, la variabilité régionale cumulée sur l'ensemble des modèles (tab. 4.10) et la diversité régionale moyenne (tab. 4.11) sont comparables. La diversité au sein de la région Ouest Congo 2 est légèrement inférieure aux autres. La diversité présente dans la région d'intergradation n'est pas supérieure à celle des autres régions (tab. 4.11), bien que les haplotypes qu'elle abrite semblent être d'origine polyphylétique pour tous les modèles (tab. 4.9). Celle de la région Ouest Congo 1, qui ne comprend que la localité Odzala, est aussi forte que celle des autres régions qui incluent pourtant plusieurs localités. De plus, la diversité présente à Odzala est nettement plus élevée que la diversité trouvée dans les autres localités relativement bien échantillonnées (tab. 4.12 ; mais ceci pourrait être une conséquence du biais d'échantillonnage en faveur d'Odzala). Deux interprétations sont envisageables : Odzala pourrait avoir été une zone de refuge ou bien constituer actuellement une zone de convergence des faunes. La situation d'Odzala en périphérie du bloc forestier et l'absence d'espèces endémiques dans cette localité (Colyn et Perpète, 1995) rendent peu crédible la première hypothèse. En revanche, la forte polyphylye des haplotypes trouvée chez deux des quatre taxons étudiés soutient la seconde hypothèse.

Nos résultats soutiennent la reconnaissance de la région Sud Ogooué. Ils nous incitent à scinder la région Ouest Congo en deux, dans la mesure où les haplotypes d'Odzala sont très différents des haplotypes des autres localités de la région Ouest Congo. De plus, la diversité génétique présente à Odzala est supérieure à celle des autres localités, tout en étant comparable à celle des autres régions. Pour ce qui est de la région d'intergradation, les résultats ne permettent pas de dire s'il s'agit effectivement d'une zone de convergence des faunes.

Quelle que soit la méthode employée, et que les aires soient prédéfinies ou non, les relations entre aires qui ressortent des analyses sont essentiellement les relations trouvées chez *S. johnstoni*. Il est probable que ce phénomène résulte du peu de modèles traités.

CHAPITRE 5: EVOLUTION DE TRAITS D'HISTOIRE DE VIE ET SPECIATION

Dans ce chapitre, nous avons tenté d'identifier les processus impliqués dans l'évolution des Primates africains de la tribu des Cercopithecini. Ce modèle a été choisi parce qu'il est relativement bien documenté, ce qui n'est pas le cas des petits Mammifères en général. Nous avons recherché le mode géographique de spéciation prédominant et essayé d'identifier des traits impliqués dans la spéciation. Nous nous sommes notamment intéressés aux facteurs écologiques impliqués dans la diversification des espèces et à la façon dont ils ont évolué en relation avec certaines adaptations morphologiques.

I -MODE GEOGRAPHIQUE DE SPECIATION CHEZ LES CERCOPITHEQUES (ARTICLE 3)

Article : Test du mode géographique de spéciation chez les Primates africains de la tribu des Cercopithecini.

Quérrouil, S., Verheyen, E. et Colyn, M. (soumis à *Proc. R. Soc. London, ser. B*) Testing the geographic mode of speciation in African primates of the tribe Cercopithecini.

Résumé : Suite à la reprise du débat sur le mode géographique de spéciation, relancé par des études mettant en évidence des cas de spéciation sympatrique, un test du mode géographique de spéciation a été proposé. Il repose sur le principe que des espèces ayant divergé récemment en allopatrie devraient présenter un recouvrement nul de leurs aires de distribution, et que le recouvrement devrait augmenter au cours du temps. Au contraire, pour des espèces ayant divergé en sympatrie, donc en situation de recouvrement total des aires de distribution, le recouvrement devrait diminuer au cours du temps. Il devrait alors exister une corrélation, positive ou négative, entre taux de recouvrement et temps de divergence entre taxons, selon que la spéciation s'est faite essentiellement en allopatrie ou en sympatrie. En conséquence, connaissant la phylogénie et les aires de distribution de l'ensemble des espèces qui constituent un taxon, il est possible d'en inférer le mode géographique de spéciation prédominant. Nous avons réalisé ce test pour les primates africains de la tribu des Cercopithecini. Une phylogénie moléculaire, incluant la plupart des espèces, a été obtenue à partir de séquences d'ARN 12s. Après avoir vérifié que l'hypothèse de l'horloge moléculaire pouvait être acceptée, nous avons daté les événements de divergence entre clades et évalué le recouvrement entre les aires de distribution cumulées de l'ensemble des espèces constituant chaque clade.

Les résultats indiquent que le mode de spéciation prédominant est l'allopatrie et que le taux de dispersion est élevé. La confrontation de l'arbre phylogénétique sur lequel est basé ce test avec d'autres reconstructions phylogénétiques met en évidence des différences importantes qui suggèrent que l'arbre obtenu pourrait ne pas correspondre à l'arbre des espèces. Cependant, même si certaines des relations phylogénétiques observées s'avéraient erronées, il est probable que le résultat du test reste inchangé. En effet, les espèces d'une même super-espèce, qui apparaissent monophylétiques à une exception près, sont allopatriques, tandis que les différentes super-espèces ont des distributions fortement recouvrantes. Il existe donc bien une opposition entre espèces allopatriques et phylogénétiquement proches et espèces sympatriques et phylogénétiquement éloignées.

Dans le but de mieux comprendre les processus de spéciation, il est également important de connaître les traits impliqués dans la différenciation des espèces. Cette question peut être abordée en comparant des espèces phylogénétiquement proches, voire en cours de spéciation, ce qui serait le cas des espèces regroupées au sein d'une même super-espèce. Nous avons donc recherché dans la littérature les différences écologiques (habitat), phénotypiques (coloration), caryologiques (réarrangements chromosomiques) et comportementales (communication vocale) affectant les espèces des super-espèces les mieux étudiées. Les résultats montrent que chaque espèce occupe une gamme d'habitats relativement large, et qu'il y a peu de différences entre les espèces d'une même super-espèce. Ceci confirme que l'hypothèse d'un mode de spéciation sympatrique selon des gradients environnementaux ne semble pas adaptée pour les Cercopithèques. Trois des cinq super-espèces étudiées présentent des différences de coloration faciale et corporelle très marquées. Ce trait pourrait être impliqué dans un processus de sélection sexuelle et constituer un facteur d'isolement pré-zygotique. L'existence d'un certain nombre de remaniements chromosomiques chez deux des super-espèces permet d'envisager également la possibilité d'un isolement post-zygotique. Enfin, au sein de chaque super-espèce, la différenciation des traits étudiés apparaît plus marquée lorsqu'il existe une zone de sympatrie entre espèces, et plus limitée lorsque les espèces sont séparées par des centaines de kilomètres. La différenciation pourrait donc augmenter lorsque le degré d'isolement géographique diminue.

Le test du mode géographique de spéciation chez les Cercopithèques indique une prédominance de l'allopatrie. Le fait que les espèces d'une même super-espèce tendent à occuper les mêmes habitats et présentent des traits morphologiques et génétiques d'autant plus différenciés que leur degré d'isolement géographique est faible va également dans le sens d'une spéciation allopatrique.

Testing the geographic mode of speciation in African primates of the tribe Cercopithecini

Sophie QUÉROUIL¹, Erik VERHEYEN² and Marc COLYN¹

¹ Université de Rennes I, CNRS - U.M.R. 6552, Laboratoire Ethologie - Evolution - Ecologie, Station Biologique, F-35380 Paimpont, FRANCE.

² Royal Institute for Natural Sciences, Section of Biochemical Systematics and Taxonomy, Vautier street 29, B-1000 Brussels, BELGIUM.

Corresponding author

Sophie Quéroutil, UMR 6552 CNRS, Station Biologique, F-35380 Paimpont, France. Tel / Fax: +00 33 2 99 61 81 63 / 88; e-mail: sophie.querouil@univ-rennes1.fr

Abstract

Our attempts to understand speciation processes requires that we identify the geographic modes of speciation and the traits involved in species differentiation. The first can be addressed providing sufficient information on species phylogeny and distribution, while the second requires intraspecific molecular, phenotypic and reproductive studies in combination with interspecific phylogenies. Among the few taxa which have been studied so extensively, the African primates of the tribe Cercopithecini (guenons) are likely good candidates for investigating speciation in tropical mammals. First, we have tested their geographic mode of speciation. Second, we have searched the literature in attempt to identify the traits which were differentiating the closely related members of four guenon super-species. Results indicate that speciation in guenons has been predominantly allopatric. The main characters affected would have been colour patterns, although other pre-zygotic and post-zygotic factors may be involved.

Key-words: speciation, African primates, Cercopithecini, guenons, allopatry

Short title: Allopatric speciation in guenons

Introduction

Speciation has long been considered to proceed primarily in allopatry, but recent studies on the geographic modes of speciation suggest that sympatric speciation may be more common than previously thought (e.g. Bush 1994; Schluter 1998). One way to test which geographic mode of speciation prevails in a taxon is to analyse the geography of sister clades across all nodes in a species level phylogeny. The prediction is that, for allopatric speciation, genetic distances between clades should be positively correlated with range overlap; while it should be the opposite for sympatric speciation (Berlocher 1998). Unfortunately, data on species ranges are usually too scarce for this analysis to be performed (Berlocher 1998). A few examples of birds, fishes and insects have been analysed (Barraclough & Vogler 2000). To our knowledge, no mammalian taxa have been tested so far, whereas large mammals are amongst the better documented taxa in terms of species identification and range definition.

Recent phylogenetic analyses of African guenons (tribe Cercopithecini, Primates) based on 12s rRNA sequences (van der Kuyl *et al.* 1995, 2000) together with the existence of reliable distribution maps for guenons (e.g. Lernould 1988) provide the opportunity to test the geographic modes of speciation in a diversified group of mammals. Actually, the tribe Cercopithecini constitute a very speciose lineage with moderate morphological differentiation between and within species, but marked differences in colour patterns, behaviour and ecology (Napier 1981). Species are not separated by absolute reproductive barriers and a number of wild and captive hybrids have been recorded (Lernould 1988). Hence, species are defined owing to the Cohesion Species Concept (Templeton 1989) rather than the Biological Species Concept (Mayr 1963). The occurrence of gene flow between species and the diversity of habitats they use, ranging from savanna to swamp forest, suggest that they could have evolved by sympatric speciation. The reticulate evolution of chromosomes also suggests sympatric speciation by hybridisation (Dutrillaux *et al.* 1988). However, the existence of groups of morphologically similar syntopic species or semi-species with allopatric distributions indicates that speciation could have been primarily allopatric or parapatric (cf. Hamilton 1988). These groups of species are regarded as "super-species" and named after one of their representatives. We will follow Lernould (1988) in distinguishing seven super-species: *Cercopithecus cephus*, *C. diana*, *C. lhoesti*, *C. nictitans*, *C. mona*, *Chlorocebus* (*Ch.*) *aethiops* (which was classified in the genus *Cercopithecus* by Lernould, but see Groves 1993), and *Miopithecus talapoin*. As the incipient species constituting these super-species are thought to be in the process of speciation (Lernould 1988), they can help us with identifying the forces that drive the evolution of isolating traits (Howard 1998). This question is addressed by a comparison of ecological, morphological, behavioural and cytogenetic traits documented in the literature.

Material and methods

Molecular data

We used published guenon 12s sequences (van der Kuyl *et al.* 1995, 2000) and completed the data set with 4 new sequences from 4 additive species, so as to have 21 out of the 28 extant cercopithecine species or semi-species (table 1). A fifth sequence from a species that was already sampled (*C. cephus*) was analysed to ascertain that the sequences were the same. The additional sequences were deposited in GenBank, under accession numbers AF394268 to AF394272. As specimens from a same species usually differed by no more than one substitution (van der Kuyl *et al.* 1995), one specimen per species was retained. Two Papionini, sister tribe to the Cercopithecini, were chosen as outgroups: *Cercocebus (Ce.) galeritus* (L35208) and *Cercocebus torquatus* (L35204; van der Kuyl *et al.* 1995).

Methodological outlines were as described by van der Kuyl and collaborators (1995). However, the primers used for sequencing the additional specimens were slightly different from the original ones. We used the 12Sai 5'-AAACTAGGATTAGATACCCTATTAT-3' and 12Sbi 5'-AAGAGCGACGGGCGATGTGT-3' primers (Kessing *et al.* 1989). In accordance with Disotell (1996)'s warnings, and following van der Kuyl and collaborators (2000), we took care that the sequences analysed represented the mitochondrial gene rather than a nuclear copy. All sequences were clearly readable, suggesting that nuclear and mitochondrial DNA were not amplified jointly. In addition, our *C. cephus* sequence proved to be identical to the published one.

Phylogenetic analyses

We first used Modeltest 3.04 (Posada & Crandall 1998) to select the substitution model which best fitted the data according to a hierarchical ratio test. We performed two maximum likelihood (ML) analyses with ten random additions of taxa under the selected model, using PAUP 4.0d64 (Swofford 1997). The first analysis was conducted without assuming a molecular clock, while for the second analysis the molecular clock was enforced. Likelihood-ratio tests were computed to evaluate branch length significance. The monophyly of super-species was tested, whenever relevant, by evaluating the likelihood length difference between constrained and unconstrained trees, using a Kishino-Hasegawa test (1989).

Time estimates

First, the ML trees constructed assuming and not assuming a molecular clock were compared by means of a Kishino-Hasegawa test (1989), to estimate whether the data were consistent with the

molecular clock hypothesis. Second, given that the null hypothesis of no length difference between the two trees was not rejected, we estimated divergence times between clades using ML distances. We calibrated the molecular clock according to the fossil-based divergence for Papionini and Cercopithecini, which is estimated to have occurred 9 Million years (Myr) ago (Delson 1992).

Table 1. List of extant guenon species (after Lernould, 1988 except for *M. talapoin ogoouensis* named by Kingdon in 1997, and for the recognition of the genus *Chlorocebus*, following Groves, 1993), with number of specimens analysed (N. seq.) and references (Ref.) of specimens selected for phylogenetic analyses.

(Reference studies are coded as follows: 1 = van der Kuyl *et al.* 1995; 2 = van der Kuyl *et al.* 2000; 3 = present study)

Genus / super-species	Species		N. seq.	Ref. of selected specimens	Study
<i>Allenopithecus</i>	<i>nigroviridis</i>	(Pocock, 1907)	0		
<i>Cercopithecus</i> / <i>cephus</i>	<i>ascanius</i>	(Audebert, 1799)	3	L35202	1
	<i>cephus</i>	(Linnaeus, 1758)	3	L35191, R12115	1 3
	<i>erythrogaster</i>	Gray, 1866	0		
	<i>erythrotis</i>	Waterhouse, 1838	1	Pri16	3
	<i>petaurista</i>	(Schreber, 1774)	1	AF164879	2
	<i>sclateri</i> ?	Pocock, 1904	0		
/ <i>diana</i>	<i>diana</i>	(Linnaeus, 1758)	2	L35193	1
	<i>dryas</i>	Schwarz, 1932	0		
/ <i>lhoesti</i>	<i>lhoesti</i>	Sclater, 1899	1	BbPalma	3
	<i>preussi</i>	Matschie, 1898	1	Mulhouse	3
	<i>solatus</i>	Harrison, 1988	1	Pri3	3
/ <i>nictitans</i>	<i>albogularis</i>	(Sykes, 1831)	0		
	<i>mitis</i>	Wolf, 1822	3	L35197	1
	<i>nictitans</i>	(Linnaeus, 1766)	2	L35199	1
/ <i>mona</i>	<i>campbelli</i>	Waterhouse, 1838	0		
	<i>mona</i>	(Schreber, 1774)	1	AF164880	2
	<i>pogonias</i>	Bennett, 1833	2	L35198	1
	<i>wolfi</i>	Meyer, 1891	0		
(none)	<i>hamlyni</i>	Pocock, 1907	1	AF164881	2
(none)	<i>neglectus</i>	Schlegel, 1876	2	L35182	1
<i>Chlorocebus</i>					
/ <i>aethiops</i>	<i>aethiops</i>	(Linnaeus, 1758)	1	L35185	1
	<i>pygerythrus</i>	(Cuvier, 1821)	4	L35207	1
	<i>sabaeus</i>	(Linnaeus, 1766)	8	L35187	1
	<i>tantalus</i>	(Ogilby, 1841)	2	L35189	1
<i>Erythrocebus</i>	<i>patas</i>	(Schreber, 1775)	6	L35186	1
<i>Miopithecus</i>					
/ <i>talapoin</i>	<i>ogoouensis</i>	Kingdon, 1997	7	L35205	1
	<i>talapoin</i>	(Schreber, 1774)	1	AF164872	2

Test of the geographic mode of speciation

We performed a test of allopatric versus sympatric speciation following the method of Barraclough and Vogler (2000). We plotted range overlap as a function of estimated time of divergence between clusters identified on the ML tree assuming a molecular clock. Range overlap was estimated after Lermould (1988)'s distribution maps corrected according to Oates (1988) and Colyn (1999) for the *C. mona* and the *C. cephus* super-species. Interestingly, as most guenon species are syntopic, geographic overlap truly reflects sympatry.

We did not perform the peripatric test of speciation proposed by Barraclough and Vogler (2000), because we presumed that range shifts had obscured potential geographic asymmetries (see results).

Potential isolating traits

We documented ecological, morphological, behavioural and cytogenetic variations within super-species, based on the literature record. Among ecological characters only habitat preferences were detailed, as this trait can be directly involved in speciation (Schluter 1998). We recorded the occurrence of species in eight main habitat types (after Kingdon 1971, 1997; Wolfheim 1983; Cords 1987; Lawes 1992; Oates 1996; Gautier-Hion *et al.* 1999 - Appendix 1). Then, for each super-species, we used the mean number of habitats occupied (with standard error) as an index of intra-specific variability. We also compiled an index of habitat overlap between species, expressed as the number of cells where a species was present (marked with "x") divided by the total number of cells included in the range of the super-species (marked with "x", "?" or "-"). Uncertainties were attributed half a point for the calculation of the first index, and caused the second index to fluctuate between a lower limit ($x / (x + ? + -)$) and an upper limit $((x + ?) / (x + ? + -))$.

We focused only on the most variable morphological traits, i.e. face and coat colour patterns (Kingdon 1980, 1988). We roughly evaluated their range of variation by comparing tones, contrast and complexity on good quality colour drawings and on black-and-white photocopies. With respect to behavioural traits, we only considered vocal communication, by means of a comparison of the species vocal repertoires (Gautier 1988). We counted the total number of differences within each super-species, among 22 characters describing the structure of cohesion-contact, loud and warning calls. We evaluated cytogenetic variations by counting the number of chromosomal rearrangements between karyotypes (Dutrillaux *et al.* 1988). In that study, the total number of rearrangements affecting guenons is 51.

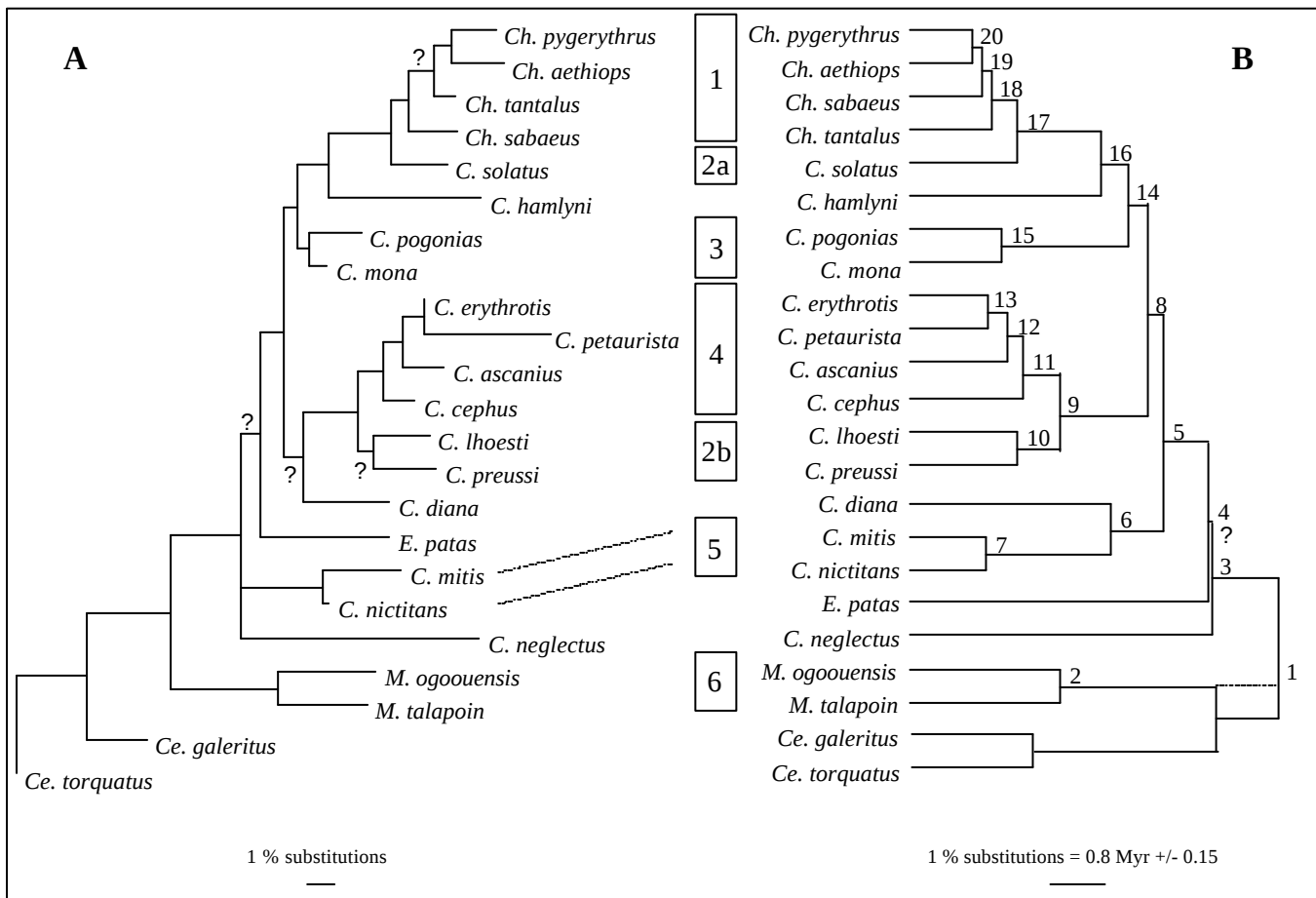
Results and discussion

Phylogenetic analyses

Among the 395 molecular characters analysed, 294 are constant, 101 are variable and 63 are parsimony informative. The hierarchical ratio test reveals that the model which best fits the data is the Tamura-Nei model (1993) with a gamma parameter of 0.21 and a null proportion of invariable sites. Estimated substitution rates are 7 for A-G transversions, 26 for C-T transversions, and 1 for all classes of transitions. Base frequencies are equal to 0.34, 0.28, 0.18 and 0.2 respectively for A, C, G and T.

The ML tree constructed without assuming a molecular clock has a $-\ln L$ value of 1636.6 (figure 1A), while the ML tree constructed under the molecular clock hypothesis has a $-\ln L$ of 1657.2 (figure 1B). The two trees differ only by the position of *C. diana* and *E. patas*, but the branches which unite these species to other guenons are not significantly different from zero. Five out of the six super-species examined actually appear monophyletic. The *C. lhoesti* super-species appears to be diphyletic (figure 1), and its monophyly is rejected by the Kishino-Hasegawa test (length of the constrained tree = 1664.9, $p = 0.01$).

Figure 1. Maximum Likelihood phylogenetic trees: **A**. without assuming a molecular clock; **B**. assuming a molecular clock. Branches not significantly longer than zero are indicated by question marks. Genera are abbreviated as follows: *C.* = *Cercopithecus*, *Ce.* = *Cercocebus*, *Ch.* = *Chlorocebus*, *E.* = *Erythrocebus* and *M.* = *Miopithecus*. Vertical bars delineate the super-species: 1 = *Ch. aethiops*, 2a and 2b = *C. lhoesti*, 3 = *C. mona*, 4 = *C. cephus*, 5 = *C. nictitans* and 6 = *M. talapoin*.



A comparison of these trees with published cladograms based on chromosomes (Dutrillaux *et al.* 1988; Disotell 1996), proteins (Ruvolo 1988) and vocalisations (Gautier 1988) reveals that no two trees are alike. Our phylogeny is based on a higher number of informative characters than previous cladograms, but it is also the most divergent. It has been suggested that, in the case of primates, mitochondrial DNA will not always reflect the true phylogeny due to ancient polymorphism and to geographical segregation of haplotypes in species where dispersal is strongly biased in favour of males (Melnick & Hoelzer 1992, 1993), as it is the case in guenons (Pusey & Packer 1987). However, this phenomenon, which should mainly affect lower level phylogenies, is presumably insufficient to explain the polyphyly of the *C. lhoesti* super-species. In addition, when 12s rRNA sequences were obtained for more than one specimen, the sequences were very conservative (van der Kuyl *et al.* 1995). Another drawback with guenons is the possibility of hybridisation (Dutrillaux *et al.* 1988; Lemould 1988). Hybridisation may falsify nuclear-based phylogenies dealing with few specimens, such as that of karyotypes. It should be less of an impediment to mitochondrial DNA phylogenies because guenon females, who inherit mitochondrial DNA from their mothers, usually remain in their natal groups (Pusey & Packer 1987; Gautier-Hion *et al.* 1999).

Time estimates

The Kishino-Hasegawa test does not reject the null hypothesis that the two ML trees are of equivalent length (the length of the ML tree following a molecular clock is 1638.8 when relaxing this constraint, $p = 0.707$). Hence, the molecular clock hypothesis is not rejected, in spite of the general statement that it does not apply to primate mitochondrial DNA (Melnick & Hoelzer 1993).

Divergence time estimates range from 1.5 (± 0.2) Myr within the *C. mona* super-species (node 15; figure 2) to 8.1 (± 1.4) Myr between *Miopithecus* and the rest of the Cercopithecini (node 1), and thus encompass the Pliocene and late Miocene. Within super-species, species would have diverged between 1.5 (± 0.2 ; node 15) Myr and 3.8 (± 0.6 ; node 2) Myr, with a mean of 2.3 (± 0.6). Although time estimates based on a short fragment of mitochondrial DNA cannot be precise, our results suggest that guenon diversification had started long before 2.9 Myr ago, which is the age of the first fossil attributed to the Cercopithecini (Leakey 1988; Pickford & Senut 1988).

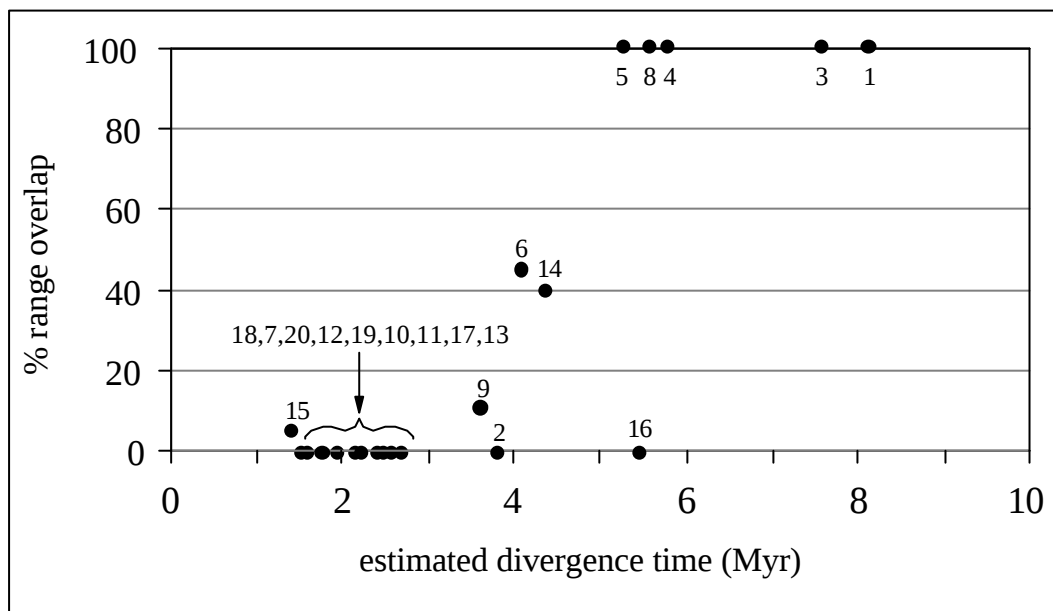
Test of the geographic mode of speciation

The plot of geographic range overlap as a function of estimated time since divergence shows that recently split clades have non-overlapping ranges (except for *C. mona* versus *C. pogonias*), while anciently split clades exhibit up to 100% range overlap (figure 2). The Arcsin transformed range overlap is

correlated to the estimated divergence time ($\rho = 0.723$, $p = 0.002$) and the regression intercept is negative. These results meet the expectations for primarily allopatric speciation with significant dispersal and range shift (cf. Barraclough & Vogler 2000). The outcome of the test would be unchanged even if *C. solatus* would be the sister species of *C. lhoesti* and *C. preussi* because all three species are allopatric.

The predominance of the allopatric mode of speciation in Cercopithecini is not surprising. It had already been hypothesised (e.g. Hamilton 1988), and is to be expected in mammals, who are neither food nor habitat specialists (Patton & Da Silva 1998). This geographic mode of speciation was also found to predominate in the Vertebrate taxa tested by Barraclough and Vogler (2000).

Figure 2. Rate of geographic range overlap as a function of estimated time since divergence (in Myr). Dots are labelled in reference to ML tree node numbers (cf. figure 1B).



Potential isolating traits

Among the seven cercopithecine super-species, the ones including *Cercopithecus diana* and *Miopithecus talapoin* are not documented enough for inter-specific comparisons. Within the five remaining super-species, we restricted the *C. lhoesti* super-species to *C. lhoesti* and *C. preussi* because it appeared to be diphyletic in our analyses (figure 1). Our literature-based synthesis of ecological, morphological, karyotypic and vocal characters reveals a number of differences within super-species (table 2). The index of habitat polymorphism indicates that the studied species are not habitat restricted. The index of habitat overlap is high and rather homogenous among the five super-species. It is indicative of a lack of habit divergence within each super-species. A direct comparison of habitat occurrences

shows that there is no example of sister species inhabiting completely distinct habitats within a super-species (Appendix 1). These arguments point to the conclusion that the Cercopithecini do not fit a habitat-based model of sympatric speciation. Whatsoever, in the small area where *C. mona* and *C. pogonias* occur in sympatry, there seems to be partial niche partition between the two closely related species (Gartlan & Struhsaker 1972). In addition, while two super-species are specialised on forest habitat, two are tolerant to wooded savanna habitats, and the last one is mostly specialised on savanna habitats, which suggests that habitat preferences have shifted during the course of evolution.

Although our list of potentially isolating traits is far from exhaustive, it appears that evolution in the investigated super-species could be dependent on face and body colour patterns (at least in *C. cephus*, *C. mona* and *C. nictitans*). These traits, in conjunction with visual communication gestures (cf. Kingdon 1980), are thought to be directly involved in mate recognition and selection, and constitute prezygotic factors of speciation. In the case of the *C. mona* super-species, the existence of chromosomal rearrangements suggests that postzygotic factors could also have promoted differentiation, as a result of genetic incompatibility or reduced fitness of heterozygotes.

Table 2. Variability of potentially differentiating traits within five guenon super-species.

(n = number of species within each super-species; # indicates that *Ch. aethiops* and *C. tantalus* are not dealt with respectively in the vocal and karyotype studies; * indicates that the super-species includes only *C. lhoesti* and *C. preussi*; ° indicates that *C. albogularis* is not dealt with in the vocal and karyotype studies; "poly" means that the super-species is, or might be, polyphyletic relative to the trait considered)

trait		habitat		body color (Kingdon 1980)	face pattern (Kingdon 1980, 88)	vocal (Gautier 1988)	karyotype (Dutrillaux <i>et al.</i> 1988)
super-species	n	index of variability	index of overlap				
<i>Ch. aethiops</i> [#]	4	4.2 (+/- 0.5)	0.85	+	+	0	0
<i>C. cephus</i>	6	4.2 (+/- 0.4)	0.83 - 0.87	++	+++	2	0
<i>C. lhoesti</i> *	2	4.7 (+/- 0.7)	0.9	+	+	2 or 3, poly	0
<i>C. mona</i>	4	4.0 (+/- 0.0)	1	+++	++	2	7, poly?
<i>C. nictitans</i> °	3	5.2 (+/- 0.8)	0.86 - 0.96	+++	++	0	2
maximum variation	-	8	1	+++	+++	22	51

Synthesis

The main cause of speciation appears to be the isolation of populations in allopatry. Divergence events are estimated to have occurred during the Miocene, Pliocene and early Pleistocene (between

about 8.1 and 1.5 Myr ago). Distribution ranges could have been fragmented as a result of climatic shifts generating habitat fragmentation, orogenetic events and / or competition with other taxa. It is reasonable to assume that the break-off of gene flow caused a genetic drift (Mayr 1963) and the modification of heritable traits, such as colour patterns. After dispersal events, as the differentiating populations met in secondary contact zones, specific adaptations and incompatibilities were possibly reinforced (Dobzhanski 1940; Howard 1998). Actually, within super-species, species whose distribution ranges are separated by hundreds of kilometres (*C. lhoesti* super-species) are among the less differentiated with respect to the investigated traits, while species separated only by rivers (*C. cephus* and *C. pogonias* super-species) are more differentiated, and those exhibiting limited sympatry (*C. mona* super-species) are the most differentiated (table 2). The *Ch. aethiops* super-species does not conform to this rule, but the extensive intergradation of characters between species suggest that they could actually be sub-species within a single species (Napier 1981; Groves 1993).

The test does not rule out the possibility that hybridisation has played a role in cercopithecine speciation. The reticulate evolution of chromosome rearrangements is suggestive of hybridisation (Dutrillaux *et al.* 1988). The occurrence of small populations exhibiting intermediate characters between adjacent species or subspecies reinforce this assumption (e.g. Dandelot 1959; Kingdon 1971). Finally, we join Gautier-Hion and collaborators (1988) in advocating that field and laboratory studies of genetic and behavioural barriers are required, for a better understanding of the maintenance of species distinctiveness and of the role of hybridisation and introgression in cercopithecine speciation.

Acknowledgements

Authors are very grateful to J.M. Lermould and P. Masson from Mulhouse Zoological Park (France) for sending them hairs from a captive *C. preussi* specimen. They acknowledge M. Dillen for sequencing the extra specimens and the Belgian Federal Services for Scientific, Cultural and Technical Affairs for funding the molecular analyses (DWTC project 31.64). Finally, they wish to thank P. Deleporte and A. Gautier-Hion for useful comments on the manuscript.

References

- Barracough, T.G. & Vogler, A.P. 2000 Detecting the geographical pattern of speciation from species-level phylogenies. *Am. Nat.* **4**, 419-434.
- Berlacher, S.H. 1998 Can sympatric speciation via host or habitat shift be proven from phylogenetic and biogeographic evidence? In *Endless forms, species and speciation* (ed. D.J. Howard & S.H. Berlacher), pp. 99-113. New York: Oxford University Press, Inc.

- Bush, M.B. 1994 Amazonian speciation: a necessarily complex model. *J. Biogeogr.* **21**, 5-17.
- Colyn, M. 1988 Distribution of guenons in the Zaïre-Lualaba-Lomami river system. In *A Primate radiation: evolutionary biology of the African guenons* (ed. A. Gautier-Hion, F. Bourlière, J.P. Gautier & J. Kingdon), pp. 104-124. Cambridge: Cambridge University Press.
- Colyn, M. 1999 Population study of the superspecies *Cercopithecus cephus* living in the Sangha-Ubangi forest enclave (Central African Republic) - Description of *C. cephus ngottoensis* subsp. nov. *Mammalia* **63**, 137-147.
- Cords, M. 1987 Forest guenons and patas monkeys: male-male competition in one-male groups. In *Primate Societies* (ed. B.B. Smuts, D.L. Cheney & R.M. Seyfarth), pp. 98-111. Chicago: The University of Chicago Press.
- Dandelot, P. 1959 Note sur la classification des cercopithèques du groupe *aethiops*. *Mammalia* **23**, 357-368.
- Delson, E. 1992 Evolution of old world monkeys. In *Cambridge encyclopedia of human evolution* (ed. J.S. Jones, R.D. Martin, D. Pilbeam & S. Bunney), pp. 217-222. Cambridge: Cambridge University Press.
- Disotell, T.R. 1996 The phylogeny of Old World monkeys. *Evol. Anthropol.* **5**, 18-24.
- Dobzhanski, T. 1940 Speciation as a stage in evolutionary divergence. *Am. Nat.* **74**, 312-321.
- Dutrillaux, B., Muleris, M. & Couturier, J. 1988 Chromosomal evolution of Cercopithecinae (chapter II.9). In *A Primate radiation: evolutionary biology of the African guenons* (ed. A. Gautier-Hion, F. Bourlière, J.P. Gautier & J. Kingdon), pp. 150-159. Cambridge: Cambridge University Press.
- Gartlan, J.S. & Struhsaker, T.T. 1972 Polyspecific associations and niche separation of rain-forest anthropoids in Cameroon, West Africa. *J. Zool. Lond.* **168**, 221-266.
- Gautier, J.P. 1988 Interspecific affinities among guenons as deduced from vocalizations (chapter II.12). In *A Primate radiation: evolutionary biology of the African guenons* (ed. A. Gautier-Hion, F. Bourlière, J.P. Gautier & J. Kingdon), pp. 194-226. Cambridge: Cambridge University Press.
- Gautier-Hion, A., Bourlière, F., Gautier, J.P. & Kingdon, J. 1988 *A primate radiation: evolutionary biology of the African guenons*. Cambridge: Cambridge University Press.
- Gautier-Hion, A., Colyn, M. & Gautier, J.P. 1999 *Histoire naturelle des Primates d'Afrique centrale*. Libreville: ECOFAC.
- Groves, C.P. 1993 Order Primates. In *Mammal species of the world: a taxonomic and geographic reference* (ed. D.E. Wilson & D.A.M. Reeder), pp. 243-278. Washington, D.C.: Smithsonian Institution.
- Hamilton, A.C. 1988 Guenon evolution and forest history. In *A Primate radiation: evolutionary biology of the African guenons* (ed. A. Gautier-Hion, F. Bourlière, J.P. Gautier & J. Kingdon), pp. 13-34. Cambridge: Cambridge University Press.
- Howard, D.J. 1998 Unanswered questions and future directions in the study of speciation. In *Endless*

- forms, species and speciation* (ed. D.J. Howard & S.H. Berlocher), pp. 439-448. New York: Oxford University Press, Inc.
- Kessing, B. Croom, H., Martin, A., McIntosh, C., McMillan, W.O. & Palumbi, S. 1989 *The Simple Fool's guide to PCR*, version 1.0. Department of Zoology, University of Hawaii.
- Kingdon, J. 1971 *East African Mammals: an atlas of evolution in Africa*, Vol. I London: Academic Press.
- Kingdon, J.S. 1980 The role of visual signals and face patterns in African forest monkeys (guenons) of the genus *Cercopithecus*. *Trans. Zool. Soc. Lond.* **35**, 425-475.
- Kingdon, J. 1988 What are face patterns and do they contribute to reproductive isolation in guenons? In *A Primate radiation: evolutionary biology of the African guenons* (ed. A. Gautier-Hion, F. Bourlière, J.P. Gautier & J. Kingdon), pp. 227-245. Cambridge: Cambridge University Press.
- Kingdon, J. 1997 *The Kingdon field guide to African mammals*. San Diego: Academic Press.
- Kishino, H. & Hasegawa, M. 1989 Evaluation of the maximum likelihood estimates of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. *J. Mol. Evol.* **29**, 170-179.
- Lawes, M.J. 1992 Estimates of population density and correlates of the status of the samango monkey *Cercopithecus mitis* in Natal, South Africa. *Biol. Conserv.* **60**, 197-210.
- Leakey, M. 1988 Fossil evidence for the evolution of the guenons. In *A primate radiation: evolutionary biology of the African guenons* (ed. A. Gautier-Hion, F. Bourlière, J.P. Gautier & J. Kingdon), pp. 7-12. Cambridge: Cambridge University Press.
- Lernould, J.-M. 1988 Classification and geographical distribution of guenons: a review. In *A primate radiation: evolutionary biology of the African guenons* (ed. A. Gautier-Hion, F. Bourlière, J.P. Gautier & J. Kingdon), pp. 54-78. Cambridge: Cambridge University Press.
- Mayr, E. 1963 *Animal Species and Evolution*, 1st edn. Cambridge, Massachusetts: The Belknap Press of Harvard University Press.
- Melnick, D.J. & Hoelzer, G.A. 1992 Differences in male and female macaque dispersal lead to contrasting distributions of nuclear and mitochondrial DNA variation. *Int. J. Primatol.* **13**, 379-393.
- Melnick, D.J. & Hoelzer, G.A. 1993 What is mtDNA good for in the study of primate evolution? *Evol. Anthropol.* **2**, 191-199.
- Napier, P.H. 1981 *Catalogue of primates in the British Museum (Natural History) and elsewhere in the British Isles. Part 2. Family Cercopithecidae, subfamily Cercopithecinae*. London: British Museum Natural History.
- Oates, J.F. 1988 The distribution of *Cercopithecus* monkeys in West African forests. In *A primate radiation: evolutionary biology of the African guenons* (ed. A. Gautier-Hion, F. Bourlière, J.P. Gautier & J. Kingdon), pp. 79-103. Cambridge: Cambridge University Press.

- Oates, J.F. 1996 *African Primates. Status survey and conservation action plan*. Revised edition. Gland, Switzerland: IUCN.
- Patton, J.L. & da Silva, M.N.F. 1998 Rivers, refuges, and ridges: the geography of speciation of Amazonian mammals. In *Endless forms, species and speciation* (ed. D.J. Howard & S.H. Berlocher), pp. 202-213. New York: Oxford University Press, Inc.
- Pickford, M. & Senut, B. 1988 Habitat and locomotion in Miocene cercopithecoids. In *A primate radiation: evolutionary biology of the African guenons* (ed. A. Gautier-Hion, F. Bourlière, J.P. Gautier & J. Kingdon), pp. 35-53. Cambridge: Cambridge University Press.
- Posada, D. & Crandall, K.A. 1998 MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**, 817-818.
- Pusey, A.E. & Packer, C. 1987 Dispersal and philopatry. In *Primate societies* (ed. D.L. Cheney, R.M. Seyfarth, R.W. Wrangham & T.T. Struhsaker), pp. 250-266. Chicago: University of Chicago Press.
- Ruvolo, M. 1988 Genetic evolution in the African guenons (chapter II.7). In *A Primate radiation: evolutionary biology of the African guenons* (ed. A. Gautier-Hion, F. Bourlière, J.P. Gautier & J. Kingdon), pp. 127-139. Cambridge: Cambridge University Press.
- Schluter, D. 1998 Ecological causes of speciation. In *Endless forms, species and speciation*. (ed. D.J. Howard & S.H. Berlocher), pp. 114-129. New York: Oxford University Press, Inc.
- Swofford, D. 1997 *PAUP: Phylogenetic analysis using parsimony. Version 4.0d60*, Smithsonian Institution, Washington, DC.
- Tamura, K. & Nei, M. 1993 Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* **10**, 512-526.
- Templeton, A.R. 1989 The meaning of species and speciation: a genetic perspective. In *Speciation and its consequences* (ed. D. Otte & J.A. Endler), pp. 3-27. Sunderland, Massachusetts: Sinauer.
- van der Kuyl, A.C., Kuiken, C.L., Dekker, J.T. & Goudsmit, J. 1995 Phylogeny of African monkeys based upon mitochondrial 12S rRNA sequences. *J. Mol. Evol.* **40**, 173-180.
- van der Kuyl, A.C., Dekker, J.T. & Goudsmit, J. 2000 Primate genus *Miopithecus*: evidence for the existence of species and subspecies of dwarf guenons based on cellular and endogenous viral sequences. *Mol. Phylogenet. Evol.* **14**, 403-413.
- Wolfheim, J.H. 1983 *Primates of the world. Distribution, abundance, and conservation*. Seattle: University of Washington Press.

Appendix 1. Distribution of *Cercopithecus* species across vegetation zones, for the super-species *Ch. aethiops*, *C. cephus*, *C. lhoesti* (to the exclusion of *C. solatus*), *C. mona* and *C. nictitans*.

("x" indicates that the species is present, "?" that it might be present, and "-" that it is absent from a habitat type where other members of the super-species are present; preferred habitats are ticked in bold face)

type	FOREST					SAVANN A		
altitude	highland	lowland				lowland		
habitat	montane > 1500 m	riverine	primary	secondary	fragmented	riverine woodland	woodland	shrubland
<i>Ch. aethiops</i>								
<i>Ch. aethiops</i>	x				x	x	x	x
<i>C. pygerythrus</i>	-				x	x	x	x
<i>C. sabaeus</i>	-				x	x	x	x
<i>C. tantalus</i>	-				x	x	x	x
<i>C. cephus</i>								
<i>C. ascanius</i>	x	x	x	x	x			
<i>C. cephus</i>	-	x	x	x	x			
<i>C. erythrogaster</i>	-	x	x	x	x			
<i>C. erythrotis</i>	x	x	x	x	?			
<i>C. petaurista</i>	-	x	x	x	x			
<i>C. sclateri</i> ?	-	x	x	x	x			
<i>C. lhoesti</i>								
<i>C. lhoesti</i>	x		x	x	x		x	
<i>C. preussi</i>	x		x	x	x		-	
<i>C. mona</i>								
<i>C. campbelli</i>		x	x	x	x			
<i>C. mona</i>		x	x	x	x			
<i>C. pogonias</i>		x	x	x	x			
<i>C. wolfi</i>		x	x	x	x			
<i>C. nictitans</i>								
<i>C. albogularis</i>	x	x	?	?	x		?	
<i>C. mitis</i>	x	x	x	x	x		x	
<i>C. nictitans</i>	x	x	x	x	x		-	

Marginal and poorly explored habitats, such as mangroves and coastal scrubs, were voluntarily excluded. Exceptional sighting of animals in habitat presumably not suitable to sustain a population were also discarded (e.g. occurrence of *C. ascanius* and *C. campbelli* in woodlands, cited by Wolfheim, 1983 and Cords, 1987, as personal communications).

II -PHYLOGENIE ET EVOLUTION DE L'ARBORICOLIE CHEZ LES CERCOPITHEQUES (ARTICLE 4)

Article : Evolution de l'arboricolie chez les Cercopithèques : analyse combinée de données moléculaires, morpho-anatomiques et comportementales.

Quérrouil, S., Deleporte, P., Gautier-Hion, A., Verheyen, E. et Colyn, M. Cercopithecini up and down the trees ? combined phylogenetic analysis of mitochondrial-DNA, morpho-anatomical and behavioural data.

Résumé : La tribu des Cercopithecini regroupe un grand nombre d'espèces aux caractéristiques écologiques variées. Tandis que la plupart des espèces sont forestières et arboricoles ou semi-terrestres, d'autres sont savanicoles et semi-terrestres. Nous nous sommes intéressés à l'évolution de ces caractéristiques écologiques et des adaptations morphologiques potentiellement liées au mode de locomotion. Dans le but de retracer l'évolution de ces traits, nous avons tenté d'obtenir une phylogénie fiable des Cercopithecini, à partir de séquences d'ADN mitochondrial, de caryotypes, de l'anatomie crânienne et post-cranienne, de la morphologie externe et de traits comportementaux. En nous basant sur le principe d'évidence totale, nous avons analysé les données d'abord séparément, puis conjointement. Les données moléculaires, caryologiques et morphologiques se sont révélées incohérentes. Comme il n'est pas possible de déterminer *a priori* si un jeu de données est plus fiable qu'un autre, nous avons privilégié l'analyse simultanée qui met en jeu un plus grand nombre de caractères. L'arbre résultant de cette analyse a été utilisé pour reconstruire les scénarios évolutifs les plus parcimonieux pour différents caractères écologiques et morphologiques.

Nos résultats suggèrent que les ancêtres des Cercopithèques étaient forestiers et semi-terrestres, ce qui est en accord avec les données fossiles et avec les hypothèses formulées par plusieurs morpho-anatomistes. Quelques espèces seraient retournées à un habitat plus ouvert, qui était vraisemblablement celui des premiers Cercopithecoïdea et Cercopithecidae, leurs lointains ancêtres. Il se serait produit deux événements d'adaptation à l'arboricolie, sans changements morphologiques majeurs, et au moins un événement d'adaptation à une terrestrialité plus marquée associé à des changements morphologiques importants. Bien que pouvant se déplacer au sol et dans les arbres, l'ancêtre des Cercopithèques aurait été morphologiquement plus proche des espèces arboricoles actuelles. De plus, des données de la littérature indiquent que des espèces ayant des morphotypes semblables peuvent adopter des modes de

locomotion différents. Il semble donc que ce ne soit pas tant l'adaptation morphologique que la plasticité écologique et comportementale qui représente le mieux l'évolution des Cercopithèques.

L'analyse simultanée des données moléculaires, caryologiques et morpho-anatomiques actuellement disponibles nous a permis de proposer une phylogénie des Cercopithecini différente de celle qui a été obtenue avec l'ARN 12s seul, mais potentiellement plus fiable. La reconstitution de scénarios évolutifs sur cet arbre révèle un nombre limité de changements de niches écologiques (habitat, strate). La plupart des espèces phylogénétiquement proches appartenant à une même super-espèce occupent la même niche écologique et sont allopatriques. Parmi les espèces phylogénétiquement plus éloignées, toutes n'occupent pas le même habitat, mais celles qui sont sympatriques et syntopiques présentent un léger décalage dans l'occupation des strates.

Cercopithecini up and down the trees ? combined phylogenetic analysis of mitochondrial-DNA, morpho-anatomical and behavioural data

Sophie QUÉROUIL¹, Pierre DELEPORTE¹, Annie GAUTIER-HION¹, Erik VERHEYEN² and Marc COLYN¹

¹ Université de Rennes I, CNRS - U.M.R. 6552, Laboratoire Ethologie - Evolution - Ecologie, Station Biologique, F-35380 Paimpont, FRANCE. Tel / Fax: 00 33 2 99 61 81 62 / 88.

² Royal Institute for Natural Sciences, Section of Biochemical Systematics and Taxonomy, Vautier street 29, B-1000 Brussels - BELGIUM. Tel / Fax: 00 32 2 627 42 86 / 41 41.

Running title: Evolution of locomotion in guenons

Corresponding author:

Sophie Quéroutil, UMR 6552 CNRS, Station Biologique, F-35380 Paimpont, France. Tel / Fax: +00 33 2 99 61 81 59 / 88; e-mail: sophie.querouil@univ-rennes1.fr

Abstract

The African primates of the tribe Cercopithecini (or guenons) exhibit a wide range of ecological characteristics. While most extant species inhabit the forest and are either arboreal or semi-terrestrial, some species are semi-terrestrial savanna-dwellers. Competing hypotheses have been formulated for the ecological preferences of the ancestral guenons and for the number of times arborealism and terrestriality have appeared during the course of their evolution. To test these hypotheses, we evaluated the phylogenetic information embedded in different character sets: mitochondrial DNA, karyotypes, cranial and post-cranial anatomy, external morphology, and behavioural traits. Following the principle of total evidence, we analysed the data first separately, then by a simultaneous analysis. As the phylogenetic informativeness of the different data sets could not be determined *a priori*, we used the cladogram resulting from the combined analysis of all data sets to reconstruct the most parsimonious scenarios for the evolution of habitat preferences, locomotor behaviour, and morphological adaptations. Our results suggest that the guenon ancestors were semi-terrestrial forest-dwellers. There would have been at least three main events of transition between arboreal and semi-terrestrial habits, two towards increased arborealism, without much morphological adaptations, and one towards increased terrestriality, involving a series of morphological adaptations. However, the high plasticity of locomotor behaviour relative to morphological adaptations found in the Cercopithecini suggests that their evolution is characterised by behavioural plasticity rather than by morphological specialisation.

Key-words: African primates, guenons, arborealism, terrestriality, evolutionary scenarios

Introduction

The Cercopithecidae, or Old World monkeys, are the most successful group of extant non-human primates. They occur on two continents, Africa and Asia, and were present in Europe during Pliocene and Pleistocene times (Delson, 1992). They account for approximately 25% of the extant genera and 40% of the living primate species (Groves, 1993; Disotell, 1996). Two sub-families, the Colobinae and the Cercopithecinae, are recognised (e.g. Szalay & Delson, 1979; Zhang & Ryder, 1998), but their taxonomy and systematics are debated (Disotell, 1996). Early studies aimed at establishing a classification of Old World monkeys through comparative analysis of morpho-anatomical features (e.g. Jolly, 1970; Delson & Andrews, 1975). More recently, a number of karyological (Dutrillaux et al., 1986), molecular (Disotell et al., 1992; Page et al., 1999) and morpho-anatomical (Strasser, 1988; Martin, 1990) studies have investigated their phylogenetic relationships. They confirmed the distinction between Asian and African Colobinae (e.g. Delson & Andrews, 1975, Stewart & Disotell, 1998; Page & Goodman, 2001) and the existence of two tribes within the Cercopithecinae: Papionini and Cercopithecini (e.g. Napier, 1981; Adkins & Honeycutt, 1994). Whatsoever, the tribal allocation of some genera remains difficult. Based on morphological characters, the genera *Allenopithecus* and *Miopithecus*, usually assigned to the Cercopithecini (e.g. Verheyen, 1962; Napier, 1981; Strasser & Delson, 1987; Gautier-Hion et al., 1988), have been suggested to belong to the Papionini for the former and to be the sister taxon to the two cercopithecine tribes for the latter (Groves, 1989).

While the phylogenetic relationships of the Papionini have been extensively studied (e.g. Disotell, 1994; Harris & Disotell, 1998; Harris, 2000), the phylogeny of the Cercopithecini (or guenons) requires further analysis (Disotell, 1996). As pointed out by early studies, guenon classification is tricky (Schwartz, 1928; Raven & Hill, 1942). In spite of the great diversity of facial and coat coloration patterns (Kingdon, 1980, 1988a), morpho-anatomical and especially cranial diversity is very low (Verheyen, 1962; Napier, 1981). The generic allocation of some species has long been controversial, but five genera are presently recognised within the Cercopithecini: *Allenopithecus*, *Cercopithecus*, *Chlorocebus*, *Erythrocebus* and *Miopithecus* (Napier, 1981, updated according to Groves, 1993, for the recognition of the genus *Chlorocebus*). Some allopatric species are morpho-anatomically so similar that they have been considered semi-species and grouped into super-species (e.g. Hill, 1966; Larnould, 1988).

Most extant species of guenons are forest-dwellers, a few species inhabit gallery forests and wooded savanna, and one (*Erythrocebus patas*) is a true savanna-dweller. Some species are arboreal while others are semi-terrestrial. Competing hypotheses have been formulated for the evolution of ecological preferences and locomotor habits in guenons. The fossil record suggests that the early

Cercopithecoidea and Cercopithecidae probably occurred in wooded savanna during lower and middle Miocene, 15-20 million years (Myr) ago (Delson & Andrews, 1975; Andrews & Aiello, 1984; Leakey, 1988; Pickford & Senut, 1988; Ciochon, 1993). There is virtually no fossil for the period comprised between 14 to 4 Myr ago that preceded the evolution of the Cercopithecini, which is supposed to have started at least 3 Myr ago (Leakey, 1988; Ciochon, 1993). As bone preservation is poor in forest habitat (Andrews & Aiello, 1984) and few excavations were conducted in present-day tropical forests, this gap in the fossil record could partially result from a lack of fossil evidence of forest-dwelling species. The first fossils attributed to the Cercopithecini, which are designated as *Cercopithecus* because no generic distinction can be made for fossils within the tribe Cercopithecini, were found in forest habitat and are 2.9 Myr old (Leakey, 1988). There is no information on when the transition from savanna to forest habitat occurred.

The fossil record suggests that the early Cercopithecoidea and Cercopithecidae were semi-terrestrial (Delson & Andrews, 1975; Andrews & Aiello, 1984; Leakey, 1988; Pickford & Senut, 1988; Ciochon, 1993). The most ancient fossils of Cercopithecini also presented adaptations to both terrestriality and arborealism (Pickford & Senut, 1988). However, a recent comparison of morphological traits in extant species (Gebo & Sargis, 1994) has re-emphasised the assumption that the ancestors of guenons were certainly arboreal as most of the extant species are (Simons, 1970; Delson & Andrews, 1975; Manaster, 1979). Actually, even the most terrestrial species of Cercopithecini climb trees for resting and feeding (e.g. Chism & Rowell, 1988, Fedigan & Fedigan, 1988) and there are only minor post-cranial skeleton differences between arboreal and semi-terrestrial species (Rollinson & Martin, 1981; Gebo & Sargis, 1994). Whatever the ancestral condition of locomotor behaviour in Cercopithecini, another debated question is whether there have been multiple switches between arboreal and semi-terrestrial habits or not. Comparisons of the post-cranial morphology of extant species have led to different interpretations. On the one hand, the ancestral guenons could have occupied the intermediate strata and occasionally descended on the ground, while fine branch arborealism and extensive terrestriality would be derived conditions (Manaster, 1979; Kingdon, 1988b). On the other hand, there might have been multiple transitions between arboreal and terrestrial habits (Rollinson & Martin, 1981; Gebo & Sargis, 1994; Terrassin-Tholance, 1995).

Hypotheses concerning the ancestral state of ecological traits (forest or savanna habitat) and locomotion (semi-terrestriality or arborealism), and the complexity of their evolution, constitute evolutionary scenarios in need to be tested on a robust phylogeny (e.g. Carpenter, 1989; Coddington, 1990). The available phyletic reconstructions based on morpho-anatomical (Kingdon, 1988b; Martin &

MacLarnon, 1988), karyological (Dutrillaux et al., 1988), allozymic (Ruvolo, 1988) and behavioural characters (Gautier, 1988) are partially incongruent. Thus, the interrelationships between some guenon species and super-species are not decisively resolved. One way to acquire numerous informative characters for phylogenetic analyses is to sequence mitochondrial DNA gene fragments. However, the analysis of DNA sequences will give a gene tree, which does not necessarily reflect the species tree (Pamilo & Nei, 1988). In the case of Primates, the use of mitochondrial DNA for phylogeny reconstruction has been debated because mitochondrial and nuclear markers often lead to conflicting phylogenies (Melnick & Hoelzer, 1992, 1993). Caution should be taken, especially when dispersal is biased in favour of males, which is the case in guenons (cf. Pusey & Packer, 1987). Whatsoever, in higher primates, such as anthropoids and papionines, it seems that molecular phylogenies could be better representative of the true species phylogeny than morphological ones (Collar & Wood, 2000). Most dental and some cranial characters appear to be highly homoplastic and thus potentially misleading, while post-cranial and soft-tissue characters could be more reliable (Collar & Wood, 2000; Gibbs et al., 2000). Behavioural and ecological characters constitute another source of evidence, which is rarely used for phylogenetic inferences, but seems to be no more homoplastic than morphological data (Miller & Wenzel, 1995). Finally, it cannot be determined *a priori* what sorts of characters are more homoplastic nor phylogenetically reliable (Miller & Wenzel, 1995). Any inheritable character may carry a phylogenetic signal, and even homoplastic data may be phylogenetically informative (Philippe et al., 1996). Consequently, we tried to gather information on as many inheritable characters as possible, following the principle of total evidence (Kluge, 1989).

We analysed the phylogenetic information contained in mitochondrial DNA, morphological, karyological, and behavioural data. First, we completed an existing molecular phylogeny of guenons based on a 12s rRNA gene fragment (van der Kuyl et al., 1995, 2000). As one super-species was not represented in that study, we sequenced the same gene fragment for additional species. Second, we sequenced a portion of the 16s rRNA gene that had already been used to infer New World monkey phylogeny (Horovitz & Meyer, 1995). Third, we completed Verheyen's (1962) craniological revision of the group. Fourth, we compiled the information available relative to chromosome rearrangements (Dutrillaux et al., 1982, 1988), post-cranial anatomy (Gebo & Sargis, 1994) and other morphological, and behavioural features (e.g. Kingdon, 1997; Gautier-Hion et al., 1999). We then performed independent and simultaneous analyses of the different data sets. The simultaneous analysis tree was considered to be more reliable than any separate analysis tree, given that it extracts the congruent phylogenetic signal shared by all available and putatively informative characters (Nixon & Carpenter,

1996 ; Grandcolas et al., 2001). Thus, it was used to infer optimal evolutionary scenarios for locomotor behaviour, habitat use and morphological adaptations. The morpho-anatomical characters traced on the phylogeny had been used to construct it, but there is no circularity in doing so because the phylogeny has to be independent of the evolutionary hypotheses to be tested, not of the characters themselves (Deleporte, 1993).

Material and methods

Species analysed

We followed Lernould's (1988) classification, but considered the *aethiops* super-species to pertain to the genus *Chlorocebus* (Ch.) rather than *Cercopithecus* (Groves, 1993), *Miopithecus nsp* to be *M. ogoouensis* (Kingdon, 1997), and *C. salongo* not to exist (Colyn et al., 1991). We recognised six out of the seven super-species defined by Lernould: *C. cephus*, *C. lhoesti*, *C. nictitans*, *C. mona*, *Ch. aethiops* and *M. talapoin*. The *C. diana* super-species was not considered as such. The poorly known species *C. dryas* was omitted due to insufficient data. Four pairs of species were treated jointly, because they were not differentiated in most studies: *Ch. djadjamensis* and *Ch. aethiops*, *C. albogularis* and *C. mitis*, *C. sclateri* and *C. erythrogaster*, *M. ogoouensis* and *M. talapoin*. In the case of *Miopithecus*, most field work and documented material concern *M. ogoouensis*, but we will refer to the two species under the name *M. talapoin*, which is that of the super-species. Two Papionini, sister tribe to the Cercopithecini, were chosen as outgroups: *Lophocebus albigena* and *Cercocebus* (Ce.) *galeritus*. The present name of the latter species is actually *C. agilis*, while *C. galeritus* is confined to coastal Kenya (Groves, 1993), but we will employ the name *C. galeritus* which is still used by most authors.

Not all species were represented in each data set (Table 1).

Table 1. Representation of the species in the data sets, with number of specimens considered (replaced by a "x" when it is not known), and number of variable and informative characters per data set.

Species, listed per super-species	12s	16s	cranial m / f	postcranial m / f	karyotype	"additional"
Outgroups						
<i>Cercocebus galeritus</i>	2		6 / 6		x	x
<i>Lophocebus albigena</i>		1	16 / 12		x	x
<i>Allenopithecus nigroviridis</i>			3 / 0	1 / 1	x	x
<i>Cercopithecus cephus</i> super-species						
<i>C. ascanius</i>	3	4	38 / 15	7 / 3	x	x
<i>C. cephus</i>	3	10	11 / 15	3 / 1	x	x
<i>C. erythrogaster</i>		1				x
<i>C. erythrotis</i>	1	1			x	x
<i>C. petaurista</i>	1	2	1 / 1	1 / 0	x	x
<i>C. diana</i>	2		0 / 2	3 / 3	x	x
<i>C. hamlyni</i>	1		10 / 4	1 / 0	x	x
<i>Cercopithecus mona</i> super-species						
<i>C. campbelli</i>		1	1 / 0	1 / 1	x	x
<i>C. mona</i>	1			1 / 0	x	x
<i>C. pogonias</i>	2	1	3 / 3	0 / 1	x	x
<i>C. wolfi</i>			9 / 3		x	x
<i>C. neglectus</i>	2		13 / 10	3 / 3	x	x
<i>Cercopithecus nictitans</i> super-species						
<i>C. mitis</i>	3		29 / 7	7 / 7	x	x
<i>C. nictitans</i>	2	1	22 / 14	1 / 0	x	x
<i>Cercopithecus lhoesti</i> super-species						
<i>C. lhoesti</i>	1	1	23 / 12	0 / 1	x	x
<i>C. preussi</i>	1	2		1 / 1	x	x
<i>C. solatus</i>	1	2	4 / 2		x	x
<i>Chlorocebus aethiops</i> super-species						
<i>Ch. aethiops</i>	1				x	x
<i>Ch. pygerythrus</i>	4		0 / 1	6 / 5	x	x
<i>Ch. sabaesus</i>	8				x	x
<i>Ch. tantalus</i>	2		2 / 1			x
<i>Erythrocebus patas</i>	6		2 / 0	4 / 2	x	x
<i>Miopithecus talapoin</i> super-species						
<i>Miopithecus sp.</i>	8		7 / 1	1 / 4	x	x
N. species	21	12	20	17	24	26
N. variable characters	96	109	28	26	57	11
N. informative characters	59	46	18	22	37	11

Acquisition of molecular sequences

Most of the samples were taken from captive specimens hosted at the Station Biologique de Paimpont (hair or tissues collected on dead animals) or dead specimens on African bush-meat markets (tissues). Other samples were obtained as hair, tissues or DNA extract, from the following persons: N.

and P. Thivillon (Espace zoologique de Saint Martin la Plaine, France) for *C. a. ascanius* and *C. a. whitesidei*, J.-M. Lernoux and P. Moisson (Zoo Mulhouse, France) for *C. preussi* and *C. erythrogaster*, B. Beer (Laboratory of Molecular Microbiology, Rockville, USA) for another *C. preussi*, and P. Formenty (OMS, Abidjan, Côte d'Ivoire) for *C. p. petaurista*. Whenever possible, we analysed several subspecies per species in order to evaluate the intra-specific variability (Table 2).

Table 2. Number and references of specimens sequenced per species and subspecies for each molecular marker (sequences obtained by van der Kuyl and collaborators, 1995, 2000, are in italic while new sequences are in plain face and sequences selected for this study are in bold)

Species	Subspecies	12s	Reference numbers	Subspecies	16s	Reference numbers
<i>Cercocebus galeritus</i>		2	<i>L35208</i>			
<i>Lophocebus albigena</i>					1	R14031
<i>Ceropithecus ascanius</i>		3	<i>L35202</i>	<i>schmidt</i>	1	Pat1
				<i>ascan</i>	1	SMP1
				<i>whiteside</i>	1	SMP2
				<i>katangae</i>	1	Pri11
<i>C. cephus</i>		2	<i>L35191</i>	<i>cephodes</i>	2	GA313-314
				<i>cephus</i>	5	R16515-530-533, LBV3, Pri7
	<i>ngottoensis</i>	1	R12115	<i>ngottoensis</i>	3	R12115-117-470
<i>C. erythrogaster</i>					1	Mulhouse1
<i>C. erythrotis</i>		1	Pri16		1	Pri16
<i>C. petaurista</i>		1	<i>AF164879</i>	<i>buettikoferi</i>	1	1908.1
				<i>petaurista</i>	1	Tai
<i>C. diana</i>	<i>roloway</i>	2	<i>L35193</i>			
<i>C. hamlyni</i>		1	<i>AF164881</i>			
<i>C. campbelli</i>					1	1908.2
<i>C. mona</i>		1	<i>AF164880</i>			
<i>C. pogonias</i>		2	<i>L35198</i>	<i>grayi</i>	1	Pri13
<i>C. neglectus</i>		2	<i>L35182</i>			
<i>C. mitis</i>		3	<i>L35197</i>			
<i>C. nictitans</i>		2	<i>L35199</i>	<i>nictitans</i>	1	NC162
<i>C. lhoesti</i>		1	BbPalma		1	Pri2
<i>C. preussi</i>		1	Mulhouse2		2	Mulhouse2 , BBeer
<i>C. solatus</i>		1	Pri3		2	Pri3 , G10002
<i>Chlorocebus aethiops</i>		1	<i>L35185</i>			
<i>Ch. pygerythrus</i>		4	<i>L35207</i>			
<i>Ch. sabaeus</i>		8	<i>L35187</i>			
<i>Ch. tantalus</i>		2	<i>L35189</i>			
<i>Erythrocebus patas</i>		6	<i>L35186</i>			
<i>Miopithecus ogoouensis</i>		7	L35205			
<i>M. talapoin</i>		1	<i>AF164872</i>			

DNA was isolated from ethanol-preserved muscle, liver tissues or freshly plucked hair by standard proteinase-K dissolution, followed by phenol-chloroform extraction and ethanol precipitation (Sambrook *et al.*, 1989), or by the Chelex method (Walsh *et al.*, 1991).

The selected 12s rRNA gene fragment was the same as that sequenced by van der Kuyl and collaborators (1995). It was amplified using the 12Sai 5'-AAACTAGGATTAGATACCCTATTAT -3' and 12Sbi 5'-AAGAGCGACGGGCGATGTGT-3' primers (Kessing *et al.*, 1989). The 16s rRNA gene fragment comprised much of the second half of the 16s rRNA gene. It was amplified using the 16Sar L-primer, 5'-CGCCTGTTTAACAAAAACAT-3' (Palumbi *et al.*, 1991) and the 16S-Hm H-primer, 5'-AGATCACGTAGGACTTTAAT-3' (Qu  rouil *et al.*, 2001). After alignment, the two fragments were respectively 395 and 533 base pair (bp) long. Sequences were deposited in GenBank, under accession numbers AF394268 to AF394272 and AF435501 to AF435527.

The PCR reaction consisted of 30 cycles: 60 s at 94  C, 60 s at 49  C and 90 s at 72  C. The double-stranded PCR product was purified using an EasyPrep kit from Pharmacia. The samples were sequenced with both primers using an automated Alf-sequencer from Pharmacia and an Amersham sequencing kit, following the manufacturer's protocol. The cycle-sequencing reaction consisted of 30 cycles: 36 s at 94  C, 36 s at 52  C for the L-primer or 55  C for the H-primer, and 80 s at 72  C.

As insertion and duplication of mitochondrial DNA in the nuclear genome have been reported in Primates (e.g. Collura & Stewart, 1995), we checked for absence of nuclear copies. We preferably extracted DNA from mitochondria-rich tissues (muscle). We used a 16s-Hm primer that was designed specifically for mammals given that specific primers are expected not to amplify ancient nuclear copies. We verified that there was no noise in the sequences, and that base composition, transition / transversion ratio and number of indels were homogeneous between sequences (cf. Arctander, 1995; Zhang & Hewitt, 1996).

Craniological data

Three hundred and nine skulls from the University of Rennes 1 collection were analysed (Table 1). After a preliminary analysis of intra- and inter-populational polymorphism in the better represented species, twenty eight characters described by Verheyen (1962) were retained (Table 3, Appendix 1). When two states of a binary character were found within the same species, they were coded with "X" rather than "?", but both symbols were treated as missing data for the analysis.

Table 3. List of cranial characters with description, and character state definition.

N°	Character	States
1	Upper incisors	0 = long, 1 = short
2	Number of tubercles on lower third molar	0 = 5, 1 = 4
3	Size of skull	0 = small, 1 = medium, 2 = large, 3 = very large
4	Temporal lines (Medio-sagittal crest; males only)	0 = parallel, 1 = convergent and joined posteriorly, 2 = convergent but not joined
5	Fossa pterygoidea	0 = very shallow, 1 = medium, 2 = deep
6	Prognathous face	0 = moderate, 1 = pronounced
7	External edge of auditory meatus	0 = short, 1 = long
8	Foramen of occipito-mastoidea sutura	0 = small, 1 = large
9	Hamulus pterygoideus	0 = narrow, 1 = wide
10	Vomer posterior edge	0 = curved and slightly concave, 1 = indented
11	Palatal vault	0 = round, 1 = U-shaped
12	Medio-sagittal crest on palatal vault	0 = absent, 1 = present
13	Lamina lateralis / medialis ratio	0 = low, 1 = high
14	Premaxillary-maxillary suture	0 = faint, 1 = marked
15	Lower edge of orbital bone	0 = below supra-orbital bone, 1 = backward, 2 = very backward
16	Posterior palatine foramen	0 = medium size, 1 = deep and wide
17	Nucal crest	0 = absent, 1 = present
18	External edge of auditory meatus	0 = round, 1 = V-shaped, 2 = with inferior crest
19	Glabella	0 = common, 1 = higher than supra-orbital edge
20	Superciliary fossa	0 = faint, 1 = marked, 2 = pronounced
21	Infra orbital maxillary	0 = deep fossa, 1 = shallow fossa, 2 = flat, 3 = convex
22	Elongated facial skull (males only)	0 = low, 1 = medium, 2 = high, 3 = very high
23	Suborbital arch	0 = low, 1 = prominent
24	Upper canine heap (males only)	0 = convergent, 1 = parallel and very bulging
25	Molar shape and structure	0 = as in most Cercopithecidae, 1 = as in <i>Allenopithecus</i>
26	Mandibular angle	0 = perpendicular, 1 = oblique
27	Height of the horizontal branch of the mandible	0 = high, 1 = low
28	Height of the vertical branch of the mandible	0 = high, 1 = low

Post-cranial data

Post-cranial anatomical characters were coded after Gebo and Sargis (1994). We retained the 26 variable characters listed in their table number 6 and coded them by directly translating words into numbers (Table 4, Appendix 1), thus following the possible discretization of continuous characters chosen by the authors. For characters number 7 and 8, which exhibited four character states, we pooled character states two by two in order not to include excessive polymorphism in the parsimony analysis. For some species, the authors analysed only males or females (Table 1), but no sex-related differentiation in

limb structure (Gebo & Sargis, 1994) nor in positional behaviour (Gebo & Chapman, 1995) have been evidenced.

Table 4. List of post-cranial characters with character state definition (after Gebo & Sargis, 1994). Character states assumed to constitute adaptations to terrestriality are highlighted in bold.

N°	Character	States
1	Scapular vertebral width	0 = very narrow , 1 = narrow , 2 = moderate
2	Supraspinatus fossa height	0 = short , 1 = moderate
3	Medial curvature of the humerus	0 = slight, 1 = strong
4	Humeral head width	0 = narrow, 1 = wide
5	Angle of humeral head	0 = perpendicular, 1 = oblique
6	Lateral trochlear ridge of humerus	0 = slight, 1 = smooth, 2 = prominent
7	Radial facet size of ulna	0 = tiny or small , 1 = wide or very wide
8	Humeral facet shape on the ulna	0 = narrow or very narrow , 1 = wide or very wide
9	Distal twist of ulna	0 = absent, 1 = present
10	Distal facets of radius	0 = flat , 1 = curved
11	Brachial index of radius	0 = lower, 1 = moderate, 2 = higher
12	Ischium / innominate	0 = lower, 1 = moderate
13	Ilium / innominate	0 = lower, 1 = moderate
14	Angle of femoral head	0 = perpendicular, 1 = oblique
15	Articular surface of femoral head	0 = reduced , 1 = moderate
16	Suprapatellar sesamoid depression of femur	0 = shallow, 1 = deep
17	Patellar facet width of femur	0 = very narrow , 1 = narrow , 2 = wide
18	Knee height	0 = moderate, 1 = tall
19	Medial condyle width of femur	0 = narrow , 1 = moderate, 2 = wide
20	Crural index of tibia	0 = short, 1 = moderate, 2 = long
21	Talar head shape	0 = oval, 1 = narrow and tall
22	Trochlear width of talus	0 = narrow, 1 = moderate
23	Talar neck length	0 = short, 1 = moderate, 2 = long
24	Anterior calcaneal facet size	0 = reduced , 1 = small , 2 = moderate
25	Peroneal tubercle of calcaneus	0 = small, 1 = moderate, 2 = large
26	Calcaneocuboid pivot length	0 = shallow , 1 = moderate

Karyological data

Karyological characters were inferred from the work of Dutrillaux and collaborators (Dutrillaux et al., 1982, 1988). The 58 chromosomal rearrangements affecting the ingroup and outgroup species were coded present or absent following the authors' description of chromosome evolution (Appendix 1). The number of individuals analysed per species was not mentioned (Table 1).

Additional data

Additional characters were chosen among morphological, and behavioural characteristics potentially carrying a phylogenetic signal (Table 5, Appendix 1). Ecological characters, such as habitat preferences and strata use, were not included because they were broadly defined and probably comprised a series of heritable components (cf. Grandcolas et al., 2001). Furthermore, their coding depended on the interpretation of a series of continuous data and was somehow subjective. Information on the 11 selected characters was obtained from the literature (mainly Kingdon, 1997; Gautier-Hion et al., 1999). Before defining discrete categories for continuous characters, i.e. sexual weight dimorphism and relative length of the tail, we verified that their distributions were multi-modal. A buffer interval was defined between the modes of the curve, and measurements within that interval were coded "X", as polymorphic characters. These heterogeneous characters were not treated as a data set, but were grouped under the term "additional" for convenience.

Table 5. List of morphological, behavioural and ecological characters with character state description. Most information is derived from Kingdon (1997) and Gautier-Hion et al. (1999), additional sources of information are listed in the character column. Unless specified, 0 equals "no" and 1 equals "yes". Character states usually associated with terrestriality are highlighted in bold.

N°	Character	States
1	Sexual weight dimorphism (Ξ / X ratio) (Napier, 1981; Colyn, 1994)	0 = weak (<1.5), 1 = high (>1.6)
2	Catamenial skin swelling (Napier, 1981)	0 / 1
3	Ischial callosities	0 = fused, 1 = wide and separate, 2 = small and separate
4	Blue scrotum	0 / 1
5	Relative length of tail (tail / head & body ratio) (Napier, 1981)	0 = short (<1.3) , 1 = long (>1.4)
6	Ear shape (Hill, 1966)	0 = round, 1 = sharp, 2 = as in <i>Miopithecus</i>
7	Scent marking (Loireau & Gautier-Hion, 1988)	0 / 1
8	Display of the tail	0 = not specific, 1 = as a question mark, 2 = backward
9	Social structure (Quris, 1975; Nakagawa, 1989)	0 = family group or harem < 10 individuals, 1 = one-male group, 2 = multi-male group
10	Locomotor habits	0 = arboreal / 1 = semi-terrestrial
11	Pledged to water	0 / 1

Separate phylogenetic analysis of each data set

The two mitochondrial sequences were analysed in the same way. Alignments were performed with ClustalW (Thompson et al., 1994) and appeared not to be ambiguous. Gaps were few and consisted mainly of point insertions or deletions (12s: seven positions with gaps for the ingroups, plus three related

to outgroups; 16s: one position with gap for the ingroups, no additional gap related to outgroups). After ensuring that specimens from a same species clustered, a single sequence per species was retained (cf. van der Kuyl et al., 1995 for the 12s RNA sequences). However, in the case of *C. c. cephus*, the sequences appeared not to be monophyletic and one specimen was chosen arbitrarily. The sequences were analysed by maximum likelihood (ML) because this method generally outperforms parsimony for DNA sequences (e.g. Yang, 1996). We used Modeltest 3.04 (Posada & Crandall, 1998) to select the substitution model which best fitted the data according to a hierarchical ratio test. We performed ML analyses with ten random additions of taxa under the selected model, using PAUP 4.0d64 (Swofford, 1997), and evaluated branch length significance by means of a likelihood-ratio test. We tested for congruence between the two phylogenies by computing the incongruence length difference index (ILD - Farris et al., 1994) using Winclada 0.9.9 beta version (Nixon, 1999). Settings were as follows: 1000 replications with two random additions of taxa, and one tree held at each step. As the two data sets were congruent (see results), we combined them in a single matrix and performed a new ML analysis on the whole set of characters.

The other data sets were analysed using Maximum Parsimony (MP). In order to avoid assigning extra weight to some types of change, characters were considered unordered. Heuristic searches were performed with ten random additions of taxa. Branch significance was evaluated by bootstrapping, with 1000 replications and a single random addition of taxa. Congruence between all pairs of data sets was estimated by computing ILD as described above. Mitochondrial sequences were treated as a single data set for that purpose.

Simultaneous analysis

The data sets can be combined even when they are incongruent (Nixon & Carpenter, 1996). Thus, simultaneous analysis of all available data was performed by a MP heuristic search with ten random additions of taxa. Gaps, which have been shown to hold valuable phylogenetic signal (Simmons et al., 2001), were considered as fifth state. A weighting scheme was applied to compensate for heterogeneous rates of substitution in DNA sequences. After estimation of the frequency of each class of substitution (see results), transversions were weighted four, A-G transitions three, and C-T transitions one. In order not to penalise the morpho-anatomical, karyological and behavioural data relative to DNA sequences, the former were given a weight of two. Branch significance was evaluated by a MP bootstrap, with 1000 replications and a single random addition of taxa. The number of characters from each data set supporting each node in the final cladogram was counted with MacClade 3.06 (Maddison & Maddison, 1996).

Reconstruction of most parsimonious scenarios

We optimised the evolution and ancestral states of ecological, behavioural and morphological traits on the simultaneous analysis cladogram, using MacClade 3.06 (Maddison & Maddison, 1996). Optimisations were performed with and without the outgroups, in order to verify that the outcome was not influenced by outgroup choice. First, we traced the evolution of savanna and forest habits (Table 5; Appendix 1). Second, we differentiated the species which used the ground (semi-terrestrial species) from those which did not or only for a negligible amount of time (arboreal species; Table 5; Appendix 1).

Table 6. Habitat (F = Forest or S = Savanna), locomotor category (A = Arboreal, ST = Semi-Terrestrial), and vertical repartition of Cercopithecinae (* = not quantified).

Species	Habitat	Category	Index of strata use	References
<i>Ce. galeritus</i>	F	ST	1	17
<i>L. albigena</i>	F	A	4	13, 21
<i>A. nigroviridis</i>	F	ST	0	6, A.G.-H., pers. obs.
<i>C. ascanius</i>	F	A	3	1, 20, 21
<i>C. cephus</i>	F	A	3	9
<i>C. erythrogaster</i>	F	A	3	16
<i>C. erythrotis</i>	F	A	3	5, 11
<i>C. petaurista</i>	F	A	3	4, 14
<i>C. diana</i>	F	A	4	4, 14
<i>C. hamlyni</i>	F	ST	1*	8, 21
<i>C. campbelli</i>	F	A	3	4, 14
<i>C. mona</i>	F	A	3	5, 11
<i>C. pogonias</i>	F	A	4	5, 9, 21
<i>C. wolffi</i>	F	A	4	A.G.-H., pers. obs.
<i>C. neglectus</i>	F	ST	1	10, 18
<i>C. mitis</i>	F > S	A	3	1, 12, 19, 20, 21
<i>C. nictitans</i>	F	A	4	5, 9, 11
<i>C. lhoesti</i>	F	ST	0	12
<i>C. preussi</i>	F	ST	0	5, 8
<i>C. solatus</i>	F	ST	0	8
<i>Chlorocebus sp.</i> (<i>aethiops</i> and <i>sabaeus</i>)	S > F	ST	0	2, 3
<i>E. patas</i>	S	ST	0	15
<i>Miopithecus sp.</i> (<i>M. ogoouensis</i>)	F	A	2	7

References:

1. Cords, 1987; 2. Dunbar & Dunbar, 1979; 3. Galat & Galat-Luong, 1976; 4. Galat & Galat-Luong, 1985; 5. Gartan & Struhsaker, 1972; 6. Gautier, 1985; 7. Gautier-Hion, 1971; 8. Gautier-Hion, 1988; 9. Gautier-Hion & Gautier, 1974; 10. Gautier-Hion & Gautier, 1978; 11. Howard, 1977; 12. Kaplin & Moermond, 2000; 13. Ham, 1994; 14. McGraw, 2000; 15. Nagakawa, 1989; 16. Oates, 1985; 17. Quris, 1975; 18. Quris, 1976; 19. Schlitch, 1978; 20. Struhsaker, 1978; 21. Thomas, 1991.

Third, we classified the species according to their use of vertical strata, on a scale from 0 to 4 relative to increasing arborealism (Table 6). Semi-terrestrial species using the ground at least 35% of time were coded 0, and those using it 10 to 25% of time were coded 1. Arboreal species showing marked preference for the lower strata (at least 70% of time) were coded 2, and those using essentially lower and middle strata or middle and higher strata were coded 3 and 4, respectively. Fourth, we considered the evolution of morphological characters assumed to be involved in terrestrial or arboreal adaptations, i.e. 14 post-cranial characters (Gebo & Sargis, 1994; Table 4) and two "additional" morphological characters: sexual dimorphism and relative length of the tail (Gautier-Hion et al., 1999; Table 5; Appendix 1).

Results

Mitochondrial sequences

For the 12s sequences, among the 395 characters analysed, 299 are constant, 96 are variable and 59 are parsimony informative. The hierarchical ratio test reveals that the model which best fits the data is the Tamura-Nei model (1993) with a gamma parameter of 0.21 and a null proportion of invariable sites. Estimated substitution rates are 7 for A-G transitions, 26 for C-T transitions, and 1 for all classes of transversions. Base frequencies are equal to 0.34, 0.28, 0.18 and 0.2 respectively for A, C, G and T. The ML tree has a $-\ln L$ value of 1636.6, a Rescaled Consistency index (RC) of 0.308 and a Retention Index (RI) of 0.541 (Fig. 1A).

For the 16s sequences, among the 533 characters analysed, 422 are constant, 109 are variable and 46 are parsimony informative. The hierarchical ratio test reveals that the model which best fits the data is the Tamura-Nei model (1993) with a gamma parameter of 0.22 and a null proportion of invariable sites. Estimated substitution rates are 7 for A-G transitions, 23 for C-T transitions, and 1 for all classes of transversions. Base frequencies are equal to 0.33, 0.26, 0.19 and 0.22 respectively for A, C, G and T. The ML tree has a $-\ln L$ value of 1547.9, a RC of 0.369 and a RI of 0.511 (Fig. 1B).

According to the ILD test, the two data sets are fairly congruent ($p = 0.936$), hence they can be combined. As settings were quite similar for the two separate analyses and because the 12s data set contained more taxa, we used the same settings as for the 12s for the combined analysis. The combined ML tree has a $-\ln L$ value of 3097.2 (Fig. 1C), a RC of 0.332 and a RI of 0.524 and closely resembles the 12s tree. In that tree, as in the 12s tree, all super-species except the *C. lhoesti* super-species are monophyletic. The *C. cephus* super-species clusters with *C. lhoesti* and *C. preussi*, and more distantly with the *C. mona* super-species. The papionine species *L. albigena* clusters with ingroups rather than with *Cercocebus galeritus*, but all the internal branches leading to *L. albigena* are not significantly different from zero.

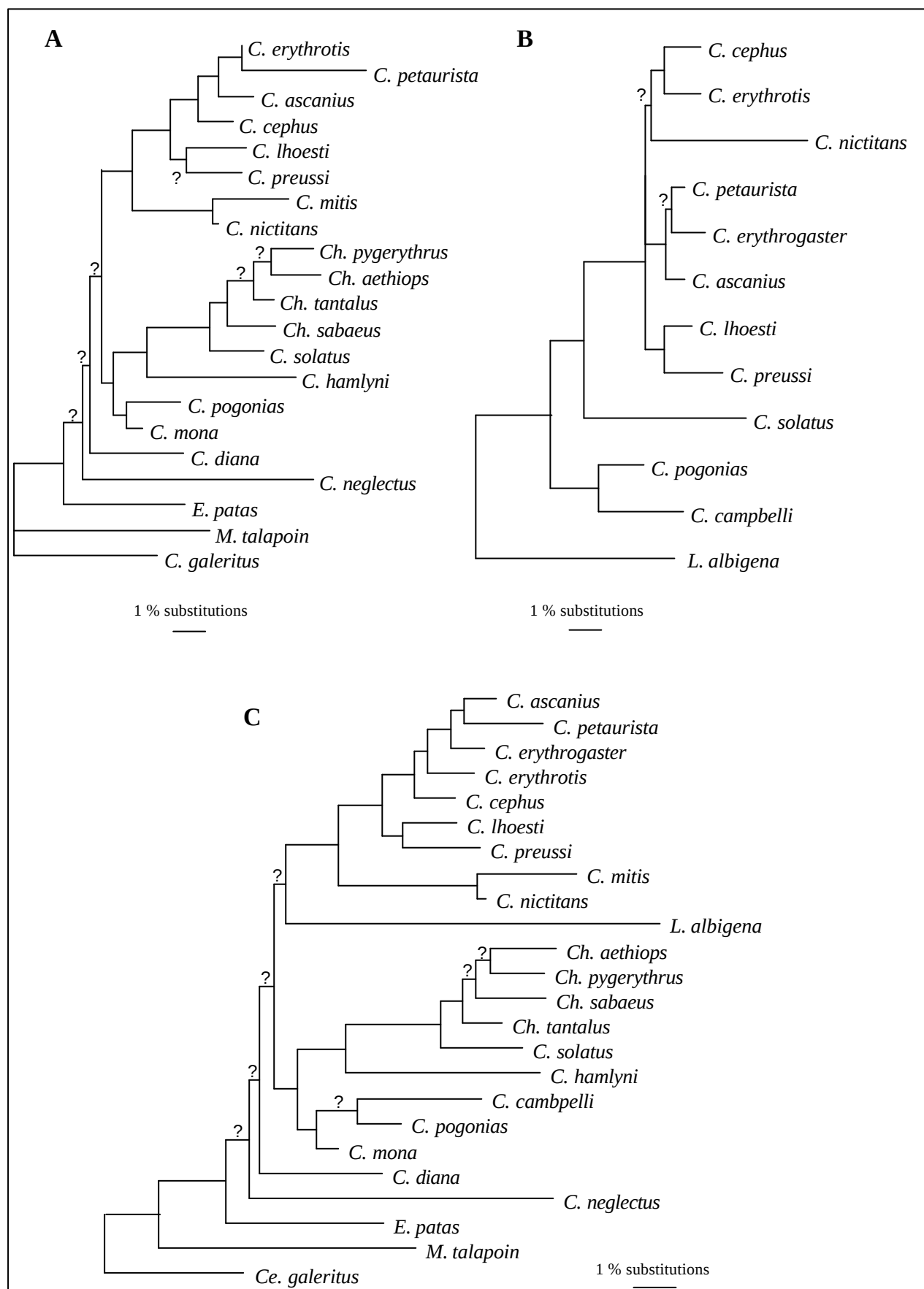


Fig. 1 ML trees obtained for the 12S rRNA sequences (A), 16S rRNA sequences (B) and combined sequences (C). ? = branch not significantly different from zero.

Other sources of data

The 18 parsimony informative cranial characters lead to 6 MP trees of 48 steps, with a RC of 0.675 and a RI of 0.830. Within each super-species except the *C. lhoesti* one, species are not differentiated (Fig. 2A). The data do not differentiate the *C. cephus* from the *C. mona* super-species. Two poorly supported associations between species and super-species appear. The first one links *A. nigroviridis* with the *Chlorocebus* species, and the second one links *C. hamlyni*, *E. patas* and the *C. lhoesti* super-species.

The 22 parsimony informative post-cranial characters lead to 12 MP trees of 52 steps, with a RC of 0.485 and a RI of 0.721. They separate the two species of the *C. nictitans* super-species, but show little differentiation between the *C. cephus*, *C. mona* and *C. nictitans* super-species (Fig. 2B). They also support an association between *C. hamlyni*, *E. patas* and the *C. lhoesti* super-species.

The 37 parsimony informative chromosomal rearrangements lead to 6 MP trees of 67 steps, with a RC of 0.826 and a RI of 0.955. They indicate close relationships between members of all super-species, with the exception of *C. mona* which is potentially polyphyletic (Fig. 2C). They group *C. hamlyni* and the *C. cephus*, *C. mona* and *C. nictitans* super-species with low support, and *E. patas* with the *Ch. aethiops* and *C. lhoesti* super-species with high support.

Comparison between data sets

The ILD tests performed for each pairwise comparison of data sets indicate that all comparisons involving either DNA sequences or karyotypes are significantly incongruent ($p = 0.001$, except for the comparison between post-cranial data and mitochondrial DNA, where $p = 0.049$). The cranial and post-cranial data sets are not significantly incongruent ($p = 0.112$).

Simultaneous analysis

The combined analysis of all available data leads to a single most parsimonious tree of 969 steps, with a RC of 0.414 and a RI of 0.673. It supports the monophyly of each super-species with a very high bootstrap support (higher than 90 %), except for the *C. lhoesti* super-species which has a bootstrap support of 59 % (Fig. 3). All nodes but two have a bootstrap support higher than 50 %, and half of the nodes have a high bootstrap support (higher than 70 %). Nine out of 22 nodes are supported by only one data set, and six of these are within super-species. DNA sequences and karyotypes support respectively four and three clades which are not supported by other data. The data set which lends most support to the nodes is the mitochondrial DNA data set, which is also the larger one. However, for each data set, the total number of synapomorphies supporting the clades is not proportional to the number of informative characters.

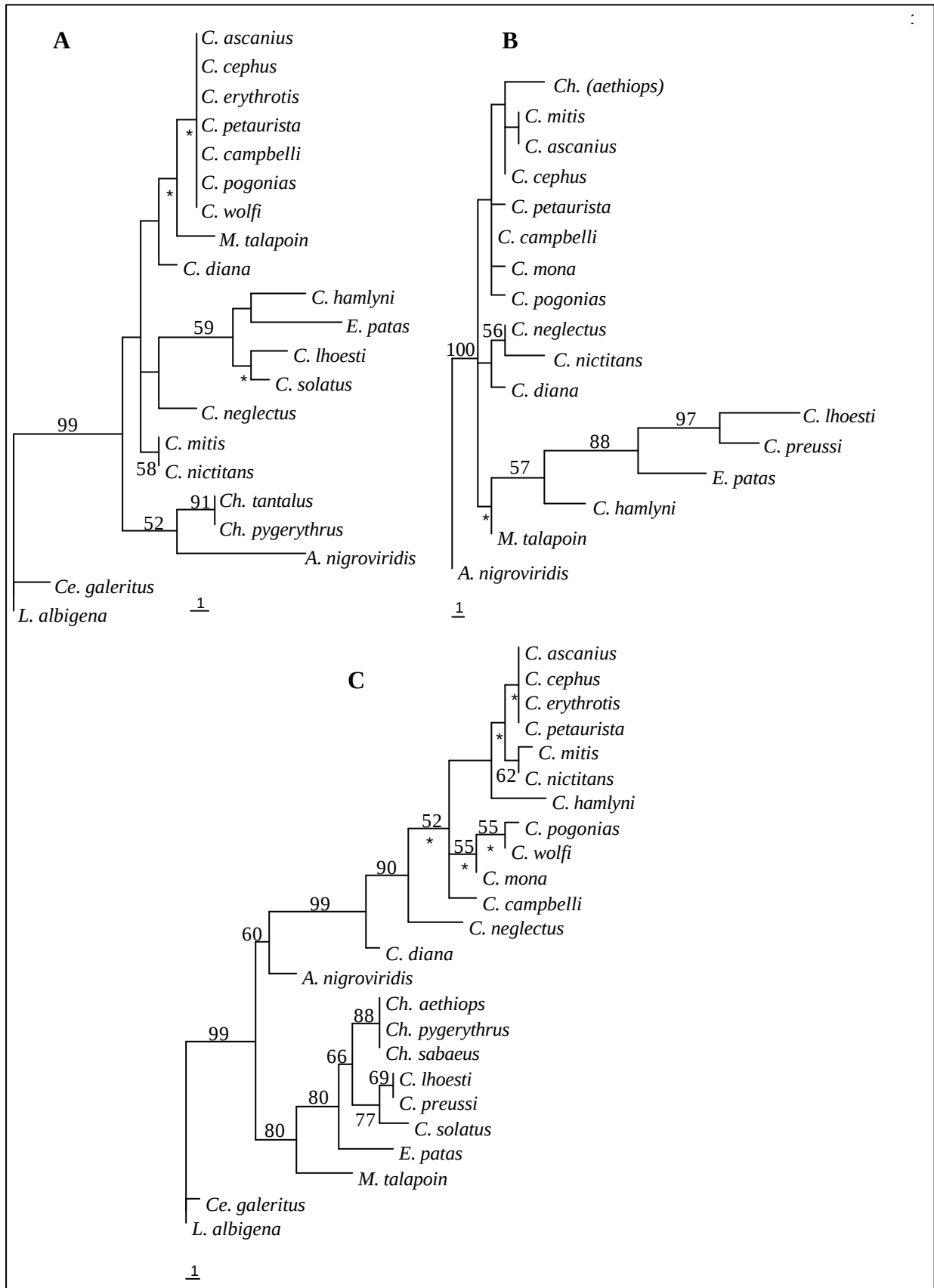


Fig. 2 MP trees obtained for the cranial (A), post-cranial (B), chromosomal rearrangements data sets (C) and for "additional" characters (D), with bootstrap support whenever greater than 50%. MP trees were selected for their resemblance with the majority-rule consensus tree. * = nodes found in only 50 to 95% of the MP trees.

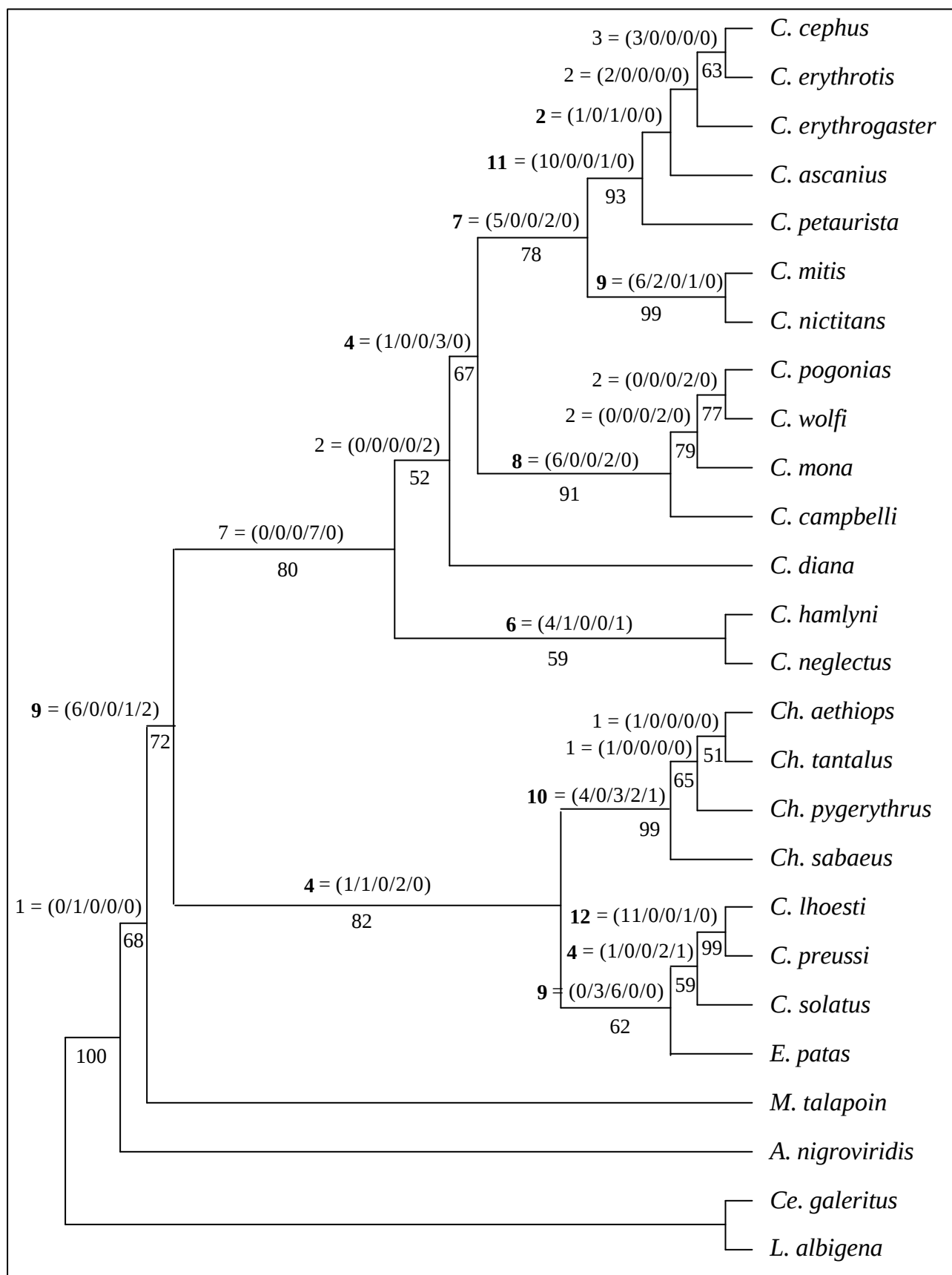


Fig. 3 Strict consensus of the tree simultaneous analysis cladograms based on the combined analysis of all data sets, with bootstrap support (below branches). The numbers above the branches, i.e. $N = (n/n/n/n/n)$, refer to the total number of synapomorphies supporting each clade (N) and the number of synapomorphies from the mitochondrial DNA, cranial, post-cranial, karyotype and "additional" data sets, in that order. N is in bold when only one data set supports the node.

The combined analysis groups the *C. cephus* and *C. nictitans* super-species with a high bootstrap support (Fig. 3). It tends to group these two super-species first with the *C. mona* one, further with *C. diana*, and then with *C. hamlyni* and *C. neglectus*, but the first two relationships are poorly supported. *E. patas* and the *Ch. aethiops* and *C. lhoesti* super-species are grouped with a high support, while *M. talapoin* and *A. nigroviridis* hold basal positions.

Most parsimonious scenarios for the evolution of arborealism and terrestriality

Optimal reconstruction of evolutionary scenarios on the simultaneous analysis cladogram indicates that the most parsimonious ancestral state for guenon habitat is forest (Fig. 4). Three switches to savanna habitats (or two switches to savanna and one reversal to forest habitat in the *C. lhoesti* super-species) could have occurred. The most parsimonious ancestral state for locomotion is semi-terrestriality. Two events of specialisation for arboreal habits would have occurred. As regards the use of vertical strata, the most parsimonious ancestral state is undefined when the outgroups are included, but it becomes "0" when they are excluded. *A. nigroviridis* and the lineage leading to *E. patas* and the *Ch. aethiops* and *C. lhoesti* super-species would have maintained extensive use of the ground (index 0). The *C. hamlyni* and *C. neglectus* lineage would have continued to use the ground, but less often than their ancestors (index 1). A switch to lower strata would have occurred in *Miopithecus* (index 2). The remaining species would have adopted fine branch arborealism, with a mix of lower and middle strata and middle and higher strata preferences (index 3 and 4).

The reconstruction of the evolution of morphological characters thought to be related with locomotor habits reveals that, among the investigated post-cranial characters, only secondary adaptations to terrestriality would have occurred (Fig. 4). Actually, the inferred ancestral states are always the ones better adapted to arborealism. For characters 2 (supraspinatus fossa height) and 21 (talar head shape), it cannot be excluded that reversals did occur in *C. preussi* and *E. patas* respectively, but these changes would have been counter-adaptive. We did not represent the evolution of character number eight (humeral facet shape on the ulna), because it appeared very homoplastic and not related with terrestriality. The evolution of sexual dimorphism must have been complex and cannot be decisively resolved. Actually, figure 4 shows a scenario based on the assumption that the ancestral state was highly dimorphic, but the opposite scenario is just as plausible. The relative length of the tail would have had an ancestral state associated with terrestriality (short tail), and would have changed once, in arboreal *Cercopithecus*.

Discussion

Phylogenetic analyses

The 12s RNA and 16s RNA sequences, which are linked on the same mitochondrial DNA molecule, are congruent. Their combined ML analysis generates a tree close to that obtained for 12s rRNA, but where several branches are not significantly longer than zero (Fig. 1). This tree is characterised by relatively low consistency and retention indexes. The MP analysis of karyotypes generates a highly resolved tree, where most branches have a bootstrap support greater than 50%. This cladogram has the highest consistency and retention indexes, which suggests that chromosomal rearrangements are not as homoplastic and saturated as previously thought (cf. Dutrillaux et al., 1982, 1988). The cranial and post-cranial data sets fail to differentiate a number of species (Fig. 2). This could reflect either the limited number of characters investigated or an actual lack of differentiation. Unexpectedly, the cladogram we obtained for cranial characters (Fig. 2A) is very different from that of Groves (2000). The two cladograms differ essentially by the position of *E. patas*, which clusters respectively with *C. hamlyni* or *Ch. aethiops*, and that of *Miopithecus*, which is either closely related to the *C. cephus* and *C. mona* super-species or holds a rather basal position. One potential explanation is that Groves (2000) included more dental characters and analysed less specimens per species than we did.

The ILD tests reveal that the cranial and post-cranial data sets, which both contain morphological characters, are not significantly incongruent, while the mitochondrial sequences and karyotypes are incongruent with all other data sets. Differences between molecular and morphological phylogenies are not unexpected, because these types of characters are submitted to different constraints (e.g. Patterson et al., 1993; Collar & Wood, 2000; Gibbs et al., 2000).

The simultaneous analysis of all available data, based on a total of 190 informative characters, conducted to a single tree, where all super-species are monophyletic and all branches but two have a bootstrap value higher than 50% (Fig. 3). Forty percent of the nodes are supported by only one data set, which are in most instances mitochondrial DNA or karyotypes, i.e. the most informative ones. The obtained cladogram has a low Rescaled Consistency index, due to the combination of incongruent data sets and to the inclusion of relatively homoplastic DNA sequences. It roughly resembles the cladogram derived from 19 informative vocal characters (Gautier, 1988, fig. 12.12, p. 216), from which it differs mainly by the position of *M. talapoin*, *C. hamlyni*, *C. neglectus*, *C. solatus* and *C. preussi*. It is quite different from the most parsimonious tree based on 10 informative protein sites (Ruvolo, 1988, fig. 7.4, p. 135), although *C. neglectus* and the *C. cephus*, *C. mona*, and *C. nictitans* super-species cluster in both cladograms.

Finally, there is a taxonomic problem concerning the generic status of the *C. lhoesti* super-species. Given that *E. patas* and the *Ch. aethiops* super-species are given full generic status, if the *C. lhoesti* super-species is actually the sister taxa to these species, then it would be relevant to classify them in a new genus.

Evolution of locomotion in guenons

Providing that we accept the simultaneous analysis cladogram despite the relatively low support of some branches, we can tentatively trace the evolution of habitat preferences, locomotion and use of vertical strata (Fig. 4).

The ancestral guenons would have been semi-terrestrial forest dwellers. The inferred forest habit for ancestral Cercopithecini is not in disagreement with assumptions based on the fossil record. The switch from savanna to forest must have taken place in the poorly documented time-frame comprised between the divergence of the Colobinae and Cercopithecinae (about 14 Myr ago, Stewart & Disotell, 1998) and the first appearance of forest guenons in the fossil record (about 3 Myr ago, Leakey, 1988; Pickford & Senut, 1988). Reversion from forest to savanna would have occurred three times, or twice followed by one switch back to forest, indicating that this trait is variable.

The ancestral state for locomotion, which we inferred to be semi-terrestriality, is also in agreement with the fossil record. It is probable that for most of their history the Cercopithecidae were semi-terrestrial animals (Leakey, 1988). The Plio-Pleistocene cercopithecine fossils which have been excavated were also semi-terrestrial (Pickford & Senut, 1988). Arborealism would have evolved twice, once in *Cercopithecus* and once in *Miopithecus*. Among the arboreal *Cercopithecus* species, switches from lower and middle strata (index 3) to middle and higher strata (index 4) are frequent. The strata occupied by sympatric species are highly overlapping, but stratification could limit inter-specific competition (Gautier-Hion, 1988). *Miopithecus* guenons differ from other arboreal Cercopithecini in that they favour the lower strata. Their locomotor habits, involving digitigrady, are closer to those of semi-terrestrial monkeys (Rollinson & Martin, 1981). Among semi-terrestrial species, their would have been one increase in arborealism, in *C. hamlyni* and *C. neglectus*. The fact that these species cluster with the arboreal *Cercopithecus* suggests that this increase might have affected their common ancestor.

Among the investigated traits, all morphological changes but one or two (i.e. relative length of the tail, and possibly sexual dimorphism) would have constituted adaptations to terrestriality. Hence, although potentially adapted to both an arboreal and a terrestrial life-style, the ancestral guenon morphology was that of an arboreal species. This ancestral phenotype could lead to either arboreal (*Miopithecus* and

arboreal *Cercopithecus* species) or semi-terrestrial species (*A. nigroviridis*, *C. hamlyni*, *C. neglectus* and *Chlorocebus* species) without extensive limb adaptations, although a number of adaptations to terrestrially occurred in *E. patas* and the *C. lhoesti* super-species. Other studies have shown that phenotypically similar species can occupy different substrates (Rollinson & Martin, 1981; Gebo & Sargis, 1994; Terrassin-Tholance, 1995). Even phylogenetically close and morphologically comparable species, such as *C. petaurista* and *C. ascanius*, exhibit a high behavioural diversity in locomotor patterns (McGraw, 2000). In South-eastern Africa, where there is no dense forest, the otherwise arboreal *C. mitis* is often seen on the ground (Moreno-Black & Maples, 1977). Thus, locomotor behaviour appears not to be strictly constrained by morphology. Conversely, there is a relative plasticity of phenotypic characters in relation with habitat use, as illustrated by the increase of sexual dimorphism in *C. mitis* populations living in open habitats (Terrassin-Tholance, 1995; M.C., pers. obs.).

Conclusion

Our results suggest that the Cercopithecini have evolved in a forest habitat. Their common ancestor would have been semi-terrestrial, but morphologically more similar to arboreal species. Their history has certainly been complex, with three events of colonisation of wooded or open savanna, and at least three events of locomotor evolution. We identified two increases in arborealism without much morphological adaptations, in *Miopithecus* and in arboreal *Cercopithecus* species, and a series of morphological adaptations to terrestriality, in *E. patas* and the *C. lhoesti* super-species. Other species would have retained the ancestral semi-terrestrial life-style without extensive morphological adaptations. Our results re-emphasise the potential behavioural plasticity of species sharing similar morphotypes and suggest that guenon evolution would be characterised by plasticity rather than by specialisation. Our interpretations must be regarded as provisional, because they depend on the reliability of the phylogenetic reconstruction they are inferred from. Further morphological and molecular data, including nuclear DNA sequences, are needed to confirm or improve the obtained phylogeny.

Acknowledgements

Authors are very grateful to B. Beer (Laboratory of Molecular Microbiology, Rockville, USA), P. Formenty (OMS, Abidjan, Ivory Coast), J.M. Lernoux and P. Moisson (Mulhouse Zoo, France) and N. and P. Thivillon. (Espace zoologique de Saint Martin la Plaine, France) for providing samples, and to

M. Dillen for carrying out part of the molecular analyses. They wish to thank P. Barrière, S. Dufour and O. Perpète for collecting specimens. They acknowledge the Belgian Federal Services for Scientific, Cultural and Technical Affairs for funding the molecular analyses (DWTC project 31.64 entitled "Evaluating the Effect of Pleistocene Climate Changes on Speciation Patterns in African Vertebrates"), the EU-DGVIII ECOFAC program "Conservation et Utilisation Rationnelle des Ecosystèmes Forestiers en Afrique Centrale" (managed by AGRECO, GEIE, BDPA, SCETAGRI, SECA, CIRAD-FORET), and the Ebola project "Forêt de Taï, OMS Abidjan", for funding some of the field collection trips.

References

- Adkins RM, and Honeycutt RL. 1994. Evolution of the primate cytochrome *c* oxydase subunit II gene. *J Mol Evol* 38: 215-231.
- Andrews P, and Aiello L. 1984. An evolutionary model for feeding and positional behaviour. In: Chivers DJ, Wood BA, Bilsborough A, editors. New York & London: Plenum Press. p 429-466.
- Arctander P. 1995. Comparison of a mitochondrial gene and a corresponding nuclear pseudogene. *Proc R Soc Lond B* 262: 13-19.
- Carpenter JM. 1989. Testing scenarios: wasp social behavior. *Cladistics* 5: 131-144.
- Chism J, and Rowell TE. 1988. The natural history of patas monkeys. In: Gautier-Hion A, Bourlière F, Gautier J-P, Kingdon J, editors. A primate radiation: evolutionary biology of the African guenons. p 412-438.
- Ciochon RL. 1993. Evolution of the cercopithecoid forelimb: Phylogenetic and functional implications from morphometric analysis. *Geol Sci* 138: 1221-1251.
- Coddington JA. 1990. Bridges between evolutionary pattern and process. *Cladistics* 6: 379-386.
- Collard M, and Wood B. 2000. How reliable are human phylogenetic hypotheses? *Proc Natl Acad Sci USA* 97: 5003-5006.
- Collura RV, and Stewart CB. 1995. Insertions and duplications of mtDNA in the nuclear genomes of Old World monkeys and hominoids. *Nature* 378: 485-489.
- Colyn M, Gautier-Hion A, and Thys Van den Audenaerde D. 1991. *Cercopithecus dryas* Schwarz 1932 and *C. salongo* Thys van den Audenaerde 1977 are the same species with an age-related coat pattern. *Folia Primatol* 56: 167-170.

- Colyn M. 1994. Données pondérales sur les primates Cercopithecidae d'Afrique Centrale (bassin du Zaïre/Congo). *Mammalia* 58: 483-487.
- Cords M. 1987. Mixed species association of *Cercopithecus* monkeys in the Kakamega forest. University of California: Publications in zoology 117: 1-109.
- Deleporte P. 1993. Characters, attributes and test of evolutionary scenarios. *Cladistics* 9: 427-432.
- Delson E, and Andrews P. 1975. Evolution and interrelationships of the catarrhine primates. In: Lockett WP, Szalay FS, editors. *Phylogeny of the primates*. New York: Plenum Press. p 405-446.
- Delson E. 1992. Evolution of Old World monkeys. In: *Cambridge encyclopedia of human evolution*. Jones JS, Martin RD, Pilbeam D, Bunney S, editors. Cambridge: Cambridge University Press. p 217-222.
- Disotell TR, Honeycutt RL, and Ruvolo M. 1992. Mitochondrial DNA phylogeny of the Old World monkey tribe Papionini. *Mol Biol Evol* 9: 1-13.
- Disotell TR. 1994. Generic level relationships of the Papionini (Cercopithecidae). *Am J Phys Anthropol* 94: 47-57.
- Disotell TR. 1996. The phylogeny of Old World monkeys. *Evol Anthropol* 5: 18-24.
- Dunbar RIM, and Dunbar EP. 1977. Ecological relations and niche separation between sympatric terrestrial primates in Ethiopia. In: Sussman RW, editor. *Primate ecology: problem-oriented field studies*. New York: John Wiley & Sons. p 187-210.
- Dutrillaux B, Couturier J, Muleris M, Lombard M, and Chauvier G. 1982. Chromosomal phylogeny of forty-two species or subspecies of Cercopithecoids (Primates, Catarrhini). *Ann Genet* 25: 96-109.
- Dutrillaux B, Muleris M, and Couturier J. 1988. Chromosomal evolution of Cercopithecinae. In: Gautier-Hion A, Bourlière F, Gautier J-P, Kingdon J, editors. *A Primate radiation: evolutionary biology of the African guenons*. Cambridge: Cambridge University Press. p 150-159.
- Farris JS, Källersjö M, Kluge AG, and Bult C. 1994. Testing significance of incongruence. *Cladistics* 10: 315-319.
- Fedigan L, and Fedigan LM. 1988. *Cercopithecus aethiops*: a review of field studies. In: Gautier-Hion A, Bourlière F, Gautier JP, Kingdon J, editors. *A Primate radiation: evolutionary biology of African guenons*. Cambridge: Cambridge University Press. p 389-411.
- Galat G, and Galat-Luong A. 1976. La colonisation de la mangrove par *Cercopithecus aethiops sabaes* au Sénégal. *Rev Ecol (Terre Vie)* 30: 3-30.
- Galat G, and Galat-Luong A. 1985. La communauté de primates diurnes de la forêt de Tai, Côte d'Ivoire. *Rev Ecol (Terre Vie)* 40: 3-32.

- Gartlan JS, and Struhsaker TT. 1972. Polyspecific associations and niche separation of rain-forest anthropoids in Cameroon, West Africa. *J Zool Lond* 168: 221-266.
- Gautier J-P. 1985. Quelques caractéristiques écologiques du singe des marais: *Allenopithecus nigroviridis* Lang 1923. *Rev Ecol (Terre Vie)* 40: 331-342.
- Gautier JP. 1988. Interspecific affinities among guenons as deduced from vocalizations. In: Gautier-Hion A, Bourlière F, Gautier J-P, Kingdon J, editors. *A Primate radiation: evolutionary biology of the African guenons*. Cambridge: Cambridge University Press. p 194-226.
- Gautier-Hion A, and Gautier J-P. 1974. Les associations polyspécifiques de cercopithèques du plateau de M'passa (Gabon). *Folia Primatol* 22: 134-177.
- Gautier-Hion A, and Gautier J-P. 1978. Le singe de Brazza: une stratégie originale. *Z Tierpsychol* 46: 84-104.
- Gautier-Hion A, Bourlière F, Gautier JP, and Kingdon J. 1988. *A primate radiation: evolutionary biology of the African guenons*. Cambridge: Cambridge University Press.
- Gautier-Hion A, Colyn M, and Gautier JP. 1999. *Histoire naturelle des Primates d'Afrique centrale*. Libreville: ECOFAC.
- Gautier-Hion A. 1971. L'écologie du talapoin du Gabon. *Rev Ecol (Terre Vie)* 25: 427-490.
- Gautier-Hion A. 1988. Polyspecific associations among forest guenons: ecological, behavioural and evolutionary aspects. In: Gautier-Hion A, Bourlière F, Gautier JP, Kingdon J, editors. *A Primate radiation: evolutionary biology of the African guenons*. Cambridge University Press. p 452-475.
- Gebo DL, and Chapman CA. 1995. Positional behavior in five sympatric Old World monkeys. *Am J Phys Anthropol* 97: 49-76.
- Gebo DL, and Sargis EJ. 1994. Terrestrial adaptations in the postcranial skeletons of guenons. *Am J Phys Anthropol* 93: 341-371.
- Gibbs S, Collard M, and Wood B. 2000. Soft-tissue characters in higher primate phylogenetics. *Proc Natl Acad Sci USA* 97: 11130-11132.
- Grandcolas P, Deleporte P, Desutter-Grandcolas L, and Daugeron C. 2001. Phylogenetics and ecology: As many characters as possible should be included in the cladistic analysis. *Cladistics* 17: 104-110.
- Groves CP. 1989. *A theory of human and primate evolution*. Oxford: Clarendon Press.
- Groves CP. 2000. The phylogeny of the Cercopithecoidea. In: Whitehead PF, Jolly CJ, editors. *Old World monkeys*. Cambridge: Cambridge University Press. p 77-98.
- Groves CP. 1993. Order Primates. In: *Mammal species of the world: a taxonomic and geographic reference*. Wilson DE, Reeder DAM, editors. Washington DC: Smithsonian Institution. p 243-278.

- Ham RM. 1994. Behaviour and ecology of grey-cheeked mangabeys (*Cercocebus albigena*) in the Lope Reserve, Gabon. PhD Thesis. University of Stirling.
- Harris EE, and Disotell TR. 1998. Nuclear gene trees and the phylogenetic relationships of the mangabeys (Primates: Papionini). *Mol Biol Evol* 15: 892-900.
- Harris EE. 2000. Molecular systematics of the Old World monkey tribe Papionini: analysis of the total available genetic sequences. *J Human Evol* 38: 235-256.
- Hill WCO. 1966. Primates. Comparative anatomy and taxonomy. Vol VI: Catarrhini, Cercopithecoidea, Cercopithecinae. Edimburg: University Press.
- Horovitz I, and Meyer A. 1995. Systematics of New World monkeys (Platyrrhini, Primates) based on 16S mitochondrial DNA sequences: A comparative analysis of different weighting methods in cladistic analysis. *Mol Phylogenet Evol* 4: 448-456.
- Howard BA. 1977. Niche separation among three sympatric species of *Cercopithecus* monkeys. PhD dissertation. The University of Texas at Austin.
- Jolly CJ. 1970. The large African monkeys as an adaptive array. In: Napier JR, Napier PH, editors. *Old World monkeys*. New York & London: Academic Press. p 139-174.
- Kaplin BA, and Moermond TC. 2000. Foraging ecology of the mountain monkey (*Cercopithecus lhoesti*): implications for its evolutionary history and use of disturbed forest. *Am J Primatol* 50: 227-246.
- Kessing B, Croom H, Martin A, McIntosh C, McMillan WO, and Palumbi S. 1989. The Simple Fool's guide to PCR, version 1.0. Department of Zoology, University of Hawaii.
- Kingdon JS. 1980. The role of visual signals and face patterns in African forest monkeys (guenons) of the genus *Cercopithecus*. *Trans Zool Soc Lond* 35: 425-475.
- Kingdon J. 1988. Comparative morphology of hands and feet in the genus *Cercopithecus*. In: Gautier-Hion A, Bourlière F, Gautier J-P, Kingdon J, editors. *A Primate radiation: evolutionary biology of the African guenons*. Cambridge: Cambridge University Press. p 184-193.
- Kingdon J. 1988a. What are face patterns and do they contribute to reproductive isolation in guenons? In: Gautier-Hion A, Bourlière F, Gautier J-P, Kingdon J, editors. *A Primate radiation: evolutionary biology of the African guenons*. Cambridge: Cambridge University Press. p 227-245.
- Kingdon J. 1997. *The Kingdon field guide to African mammals*. San Diego: Academic Press.
- Kluge AG. 1989. A concern for evidence and a phylogenetic hypothesis of relationships among *Epicrates* (Boidae, serpentes). *System Zool* 38: 7-25.

- Leakey M. 1988. Fossil evidence for the evolution of the guenons. In: Gautier-Hion A, Bourlière F, Gautier J-P, Kingdon J, editors. A Primate radiation: evolutionary biology of the African guenons. Cambridge: Cambridge University Press. p 7-12.
- Lernould J-M. 1988. Classification and geographical distribution of guenons: a review. In: Gautier-Hion A, Bourlière F, Gautier J-P, Kingdon J, editors. A Primate radiation: evolutionary biology of the African guenons. Cambridge: Cambridge University Press. p 54-78.
- Loireau JN, and Gautier-Hion A. 1988. Olfactory marking behaviour in guenons and its implications. In: Gautier-Hion A, Bourlière F, Gautier JP, Kingdon J, editors. A Primate radiation: evolutionary biology of the African guenons. Cambridge: Cambridge University Press. p 246-253.
- Maddison WP, and Maddison DR. 1996. MacClade: Analysis of phylogeny and character evolution, Version 3.06. Sinauer, Sunderland, MA.
- Manaster BJ. 1979. Locomotor adaptations with the *Cercopithecus* genus: a multivariate approach. Am J Phys Anthropol 50: 169-182.
- Martin RD, and MacLarnon AM. 1988. Quantitative comparisons of the skull and teeth in guenons. In: Gautier-Hion A, Bourlière F, Gautier J-P, Kingdon J, editors. A Primate radiation: evolutionary biology of the African guenons. Cambridge: Cambridge University Press. p 160-183.
- Martin RD. 1990. Primate origins and evolution: a phylogenetic reconstruction. Princeton: Princeton University Press.
- McGraw WS. 2000. Positional behavior of *Cercopithecus petaurista*. Int J Primatol 21: 157-182.
- Melnick DJ, and Hoelzer GA. 1992. Differences in male and female macaque dispersal lead to contrasting distributions of nuclear and mitochondrial DNA variation. Int J Primatol 13: 379-393.
- Melnick DJ, and Hoelzer GA. 1993. What is mtDNA good for in the study of primate evolution? Evol Anthropol 2: 191-199.
- Miller JS, and Wenzel JW. 1995. Ecological characters and phylogeny. Annu Rev Entomol 40: 389-415.
- Moreno-Black G, and Maples WR. 1977. Differential habitat utilization of four Cercopithecidae in a Kenyan forest. Folia Primatol 27: 85-107.
- Nakagawa N. 1989. Feeding strategies of Japanese monkeys against deterioration of habitat quality. Primates 30: 1-16.
- Napier PH. 1981. Catalogue of primates in the British Museum (Natural History) and elsewhere in the British Isles. Part 2. Family Cercopithecidae, subfamily Cercopithecinae. London: British Museum Natural History.
- Nixon KC, and Carpenter JM. 1996. On simultaneous analysis. Cladistics 12: 221-242.

- Nixon KC. 1999. Winclada version 0.9.9+ (beta). Program published by the author, Ithaca, NY, available at http://www.cladistics.com/about_winc.htm.
- Oates JF. 1985. The Nigerian guenon, *Cercopithecus erythrogaster*: Ecological, behavioral, systematic and historical observations. *Folia Primatol* 45: 25-43.
- Page SL, and Goodman M. 2001. Catarrhine phylogeny: Noncoding DNA evidence for a diphyletic origin of the mangabeys and for a human-chimpanzee clade. *Mol Phylogenet Evol* 18: 14-25.
- Page SL, Chiu CH, and Goodman M. 1999. Molecular phylogeny of Old World monkeys (Cercopithecidae) as inferred from gamma-globin DNA sequences. *Mol Phylogenet Evol* 13: 348-359.
- Palumbi SR, Martin A, Romano S, McMillan WO, Stice L, and Grabowski G. 1991. The simple fool's guide to PCR. University of Hawaii Press, Honolulu.
- Pamilo P, and Nei M. 1988. Relationships between gene trees and species trees. *Mol Biol Evol* 5: 568-583.
- Patterson C, Williams DM, and Humphries CJ. 1993. Congruence between molecular and morphological phylogenies. *Annu Rev Ecol Syst* 24: 153-188.
- Philippe H, Lecointre G, Vàn Le HL, Le Guyader H. 1996. A critical study of homoplasy in molecular data with the use of a morphologically based cladogram, and its consequences for character weighting. *Mol Biol Evol* 13: 1174-1186.
- Pickford M, and Senut B. 1988. Habitat and locomotion in Miocene cercopithecoids. In: Gautier-Hion A, Bourlière F, Gautier J-P, Kingdon J, editors. *A Primate radiation: evolutionary biology of the African guenons*. Cambridge: Cambridge University Press. p 35-53.
- Posada D, and Crandall KA. 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14: 817-818.
- Pusey AE, and Packer C. 1987. Dispersal and philopatry. In: *Primate societies*. Cheney DL, Seyfarth RM, Wrangham RW, Struhsaker TT, editors. Chicago: University of Chicago Press. p 250-266.
- Quérouil S, Hutterer R, Barrière P, Colyn M, Peterhans JCK, and Verheyen E. 2001. Phylogeny and evolution of African shrews (Mammalia : Soricidae) inferred from 16s rRNA sequences. *Mol Phylogenet Evol* 20: 185-195.
- Quris R. 1975. Ecologie et organisation sociale de *Cercocebus galeritus agilis* dans le Nord-Est du Gabon. *Rev Ecol (Terre Vie)* 29: 337-398.

- Quris R. 1976. Données comparatives sur la socio-écologie de huit espèces de Cercopithecidae vivant dans une même zone de forêt primitive périodiquement inondée (Nord-Est du Gabon). *Rev Ecol (Terre Vie)* 30: 193-209.
- Raven HC, and Hill JE. 1942. Notes on *Cercopithecus hamlyni* Pocock. *Am Mus Nov* 1177: 1-6.
- Rollinson JM, and Martin RD. 1981. Comparative aspects of primate locomotion with special reference to arboreal Cercopithecines. In: Day MH, editor. *Vertebrate Locomotion*. London: Academic Press. p 377-427.
- Ruvolo M. 1988. Genetic evolution in the African guenons. In: Gautier-Hion A, Bourlière F, Gautier J-P, Kingdon J, editors. *A Primate radiation: evolutionary biology of the African guenons*. Cambridge: Cambridge University Press. p 127-139.
- Sambrook J, Fritsch EF, and Maniatis T. 1989. *Molecular cloning: A laboratory manual*, Second Edition. Chapter 14: In vitro amplification of DNA by the polymerase chain reaction. Cold Spring Harbour Laboratory Press. p 14.1-14.35.
- Schlichte HJ. 1978. The ecology of two groups of blue monkeys, *Cercopithecus mitis stuhlmanni*, in an isolated habitat of poor vegetation. In: Montgomery GG, editor. *Ecology of arboreal folivores*. Washington DC: Smithsonian Institution Press. p 505-517.
- Schwarz E. 1928. Notes on the classification of the African monkeys in the genus *Cercopithecus*, *Erleben*. *Ann Mag Nat Hist* 10: 649-663.
- Simmons MP, Ochoterena H, and Carr TG. 2001. Incorporation, relative homoplasy, and effect of gap characters in sequence-based phylogenetic analyses. *Syst Biol* 50: 454-462.
- Simons EL. 1970. The deployment and history of Old World monkeys (Cercopithecidae, Primates). In: Napier JR, Napier PH, editors. *Old World monkeys*. New York & London: Academic Press. p 97-137.
- Stewart CB, and Disotell TR. 1998. Primate evolution - In and out of Africa. *Curr Biol* 8: 582-588.
- Strasser E. 1988. Pedal evidence for the origin and diversification of cercopithecoid clades. *J Hum Evol* 17: 225-245.
- Struhsaker TT. 1978. Food habits of five monkey species in the Kibale forest, Uganda. In: Chivers DJ, Herbert J, editors. *Recent advances in Primatology*. Vol 1. Behaviour. New York: Academic Press. p 225-248.
- Swofford D. 1997. *PAUP: Phylogenetic analysis using parsimony*. Version 4.0d60. Smithsonian Institution, Washington, DC.

- Szalay FS, and Delson E. 1979. Catarrhiniens. In: Evolutionary history of the primates. New York: Academic Press.
- Tamura K, and Nei M. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol* 10: 512-526.
- Terrassin-Tholance C. 1995. De la terrestrialité à l'arboricolisme: anatomie fonctionnelle comparée de la ceinture scapulaire et du membre antérieur des Cercopithèques. PhD Thesis. University of Rennes.
- Thomas SC. 1991. Population densities and patterns of habitat use among anthropoid primates of the Ituri Forest, Zaire. *Biotropica* 23: 68-83.
- Thompson JD, Higgins DG, and Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22: 4673-4680.
- van der Kuyl AC, Kuiken CL, Dekker JT, and Goudsmit J. 1995. Phylogeny of African monkeys based upon mitochondrial 12S rRNA sequences. *J Mol Evol* 40: 173-180.
- van Der Kuyl AC, Dekker JT, and Goudsmit J. 2000. Primate genus *Miopithecus*: evidence for the existence of species and subspecies of dwarf guenons based on cellular and endogenous viral sequences. *Mol Phylogenet Evol* 14: 403-413.
- Verheyen WN. 1962. Contribution à la craniologie comparée des primates. Les genres *Colobus* Illiger 1811 et *Cercopithecus* Linné 1758. *Ann Mus Afr Centr Tervuren Belg Ser 8 Sci Zoo* 105: 1-256.
- Walsh PS, Metzger DA, and Higuchi R. 1991. Chelex 100 as a medium for simple extraction of DNA for PCR based typing from forensic material. *Biotechniques* 10: 506-513.
- Yang Z. 1996. Phylogenetic analysis using parsimony and likelihood methods. *J Mol Evol* 42: 294-307.
- Zhang D-X, and Hewitt GM. 1996. Nuclear DNA: challenges for mitochondrial DNA markers. *TREE* 11: 247-251.
- Zhang YP, and Ryder OA. 1998. Mitochondrial cytochrome *b* gene sequences of Old-World monkeys: with special reference on evolution of Asian Colobines. *Primates* 39: 39-49.

Appendix 1. Matrix of morpho-anatomical, chromosomal, behavioural and ecological character states (? = undefined, X = 0 or 1).

[illegible]

I - LIMITES METHODOLOGIQUES DES ANALYSES PHYLOGENETIQUES

1) Limites des données moléculaires, et autres sources potentielles d'information

Dans les chapitres 3.II et 5.II, nous avons comparé l'information phylogénétique contenue dans deux fragments d'ADN mitochondrial distincts (ARN 16s et cytochrome *b*, pour le genre *Hylomyscus* ; ARN 12s et ARN 16s pour la tribu des Cercopithecini). Les différentes portions d'ADN séquencées tendent à donner le même signal phylogénétique et leur réunion permet d'obtenir un signal phylogénétique plus fort. Cependant, les gènes choisis étant liés au sein d'une même molécule d'ADN mitochondrial, ils apportent des informations non indépendantes. Même si plusieurs séquences mitochondriales tendent à supporter un même arbre, celui-ci peut ne pas correspondre à l'arbre des espèces (e.g. Nei, 1987 ; Moore, 1995).

L'étude de la super-espèce *Cercopithecus cephus* nous met en garde contre l'utilisation de données moléculaires pour retracer une phylogénie lorsque la monophylie des unités taxinomiques élémentaires n'est pas atteinte ou qu'il y a eu hybridation. Il est donc nécessaire de disposer d'un échantillonnage suffisant pour vérifier au préalable la monophylie de chacune des espèces. Si des disparités entre ADN mitochondrial et phénotypes peuvent être mises en évidence chez des espèces pour lesquelles la détermination spécifique est relativement aisée, il est probable que de telles disparités existent aussi chez les petits mammifères dont la détermination est parfois très délicate, mais que celles-ci ne soient pas toujours détectées. Les analyses moléculaires ne peuvent constituer un critère de distinction des espèces en soi. Pour distinguer les espèces selon le concept d'espèce biologique, il serait nécessaire d'étudier la dynamique et la génétique des populations afin de déterminer s'il existe un flux génique entre populations.

Il en résulte qu'il est important de confronter plusieurs sources d'information, par exemple des données moléculaires et morpho-anatomiques, comme nous l'avons fait de façon succincte pour le genre *Hylomyscus* et de façon plus approfondie pour les Cercopithecini. Pour ces derniers, nous avons mis en évidence une discordance entre les relations phylogénétiques supportées par les données génétiques et morphologiques. Ce phénomène n'a rien d'exceptionnel (e.g. Rüber *et al.*, 1999 ; Collard et Wood, 2000). D'ailleurs, selon Patterson et collaborateurs (1993), "la phylogénie moléculaire contribue là où elle n'est pas contestée, mais est rarement incontestée lorsqu'elle est confrontée à la morphologie". Trois causes principales peuvent expliquer ces discordances : les spécimens choisis pour représenter une même unité taxinomique peuvent en réalité ne pas appartenir au même taxon, le choix des méthodes et des modèles

évolutifs pour les reconstructions phylogénétiques peut ne pas être optimal, et le taux d'homoplasie peut être élevé dans l'un ou l'autre des jeux de données. Les données moléculaires présentent un risque de saturation (cf. chapitre 2.IV.3) et les données morphologiques un risque de convergence entre espèces ayant des préférences écologiques voisines (e.g. Collard et Wood, 2000). Dans le but de retrouver la "vraie" phylogénie, il apparaît nécessaire d'analyser autant de caractères potentiellement porteurs de signal phylogénétique que possible (Grandcolas *et al.*, 2001). Les données écologiques et comportementales peuvent s'avérer intéressantes à cet égard (Wenzel, 1992 ; Miller et Wenzel, 1995). Dans le cas des Cercopithèques, l'analyse combinée de données moléculaires, caryologiques, morpho-anatomiques et comportementales selon le principe d'évidence totale nous a permis de proposer une phylogénie relativement bien résolue. Ce type d'analyse n'a pas pu être réalisé pour les *Hylomyscus* et les Crocidurinae, du fait d'un manque de données morpho-anatomiques, écologiques et comportementales.

2) Traitement des données moléculaires

a) Comment obtenir une phylogénie plus fiable ?

L'étude du genre *Hylomyscus* nous montre que les différentes méthodes de reconstruction phylogénétique produisent souvent des cladogrammes différents. Par contre, pour ce genre comme pour la révision phylogénétique des Crocidurinae, les groupements fortement supportés sont généralement les mêmes quelles que soient les méthodes. Les relations peu supportées résultent vraisemblablement d'une cladogénèse rapide ou de l'insuffisance des données (cf. Slowinski, 2001). Nous avons choisi de séquencer un ou deux fragment d'ADN pour un grand nombre de taxons et de spécimens plutôt qu'un grand nombre de fragments pour un nombre d'échantillons plus limité, car des simulations ont montré que plus de taxons valent mieux que plus de nucléotides (e.g. Lecointre *et al.*, 1993 ; Hillis, 1996). Cependant, nous sommes conscients que le nombre limité de positions nucléotidiques analysées est une des limitations de cette étude.

A l'instar de la majorité des biologistes moléculaires, nous avons accordé une plus grande confiance à la méthode du maximum de vraisemblance (e.g. Yang, 1996), et nous avons parfois présenté uniquement les résultats obtenus par cette méthode. Toutefois, dans l'étude du genre *Hylomyscus*, nous avons montré qu'elle est effectivement sujette à la répulsion des longues branches (cf. Farris, 1999) et qu'elle gagnerait à traiter également les indels. Ces derniers constituent une source d'information non négligeable, ils sont un peu moins homoplasiques que les caractères nucléotidiques et leur prise en compte augmente souvent le support des branches (Simmons *et al.*, 2001). Malheureusement, deux problèmes majeurs restent à

résoudre avant de pouvoir les incorporer dans une analyse de vraisemblance (Graham *et al.*, 2000) : comment définir la fonction de coût associée aux indels, et, lorsqu'il existe des indels chevauchants, comment construire la séquence la plus vraisemblable conduisant à cette distribution ?

b) Comment affiner les datations ?

Avant d'envisager de dater des événements de divergence, nous avons toujours vérifié que l'horloge moléculaire était respectée par un "relative rate test" ou un test de Kishino-Hasegawa. Les tests effectués n'ont révélé de différence significative que lors de la comparaison entre les deux espèces de *Sylvioorex* étudiées dans le chapitre phylogéographie. Le "relative rate test" a été critiqué pour son faible pouvoir résolutif (Voelker, 1999) et le test de Kishino-Hasegawa pour ses fluctuations statistiques (Goldman, 1993). Il est donc possible que l'hypothèse de l'horloge moléculaire ait parfois été acceptée à tort.

Nous avons pris la précaution de toujours calibrer l'horloge moléculaire pour les taxons et les séquences analysés, à partir de temps de divergence inférés des fossiles. La calibration de l'horloge à partir des fossiles est limitée par la richesse des données fossiles, par la qualité des interprétations qui en ont été faites, et par la justesse des datations. Les temps de divergence risquent d'être sous-estimés parce qu'il y a peu de chance de retrouver des fossiles correspondant exactement à une période de divergence entre taxons. D'autre part, lorsque nous ne disposons que d'un seul point de calibration, et / ou que celui-ci était ancien par rapport aux événements à dater, comme dans le cas de la super-espèce *Cercopithecus cephus*, nous n'avons pu obtenir une estimation précise des temps de divergence.

Nous avons modulé les corrections des distances en fonction de la saturation probable des différents types de substitutions, depuis l'utilisation d'une formule de correction des distances ne compensant que les différences de taux de transitions et transversions pour les phylogénies intra-spécifiques, jusqu'à l'élimination totale des transitions pour les Crocidurinae. Le choix de la méthode de correction des distances nous semble déterminant pour la qualité des estimations.

Des datations plus précises auraient pu être obtenues si nous avions analysé plusieurs loci indépendants. La comparaison de phylogénies nous aurait permis de déceler d'éventuels artefacts dus à un flux génique temporaire postérieur à la divergence, et nous aurait évité une surestimation des temps de divergence liée au fait que la divergence des gènes précède celle des espèces (Nichols, 2001).

Les analyses réalisées nous confirment certaines mises en garde mentionnées dans le chapitre "matériel et méthodes". Il est important de disposer d'un nombre suffisant de spécimens et de vérifier la monophylie des espèces sur critères génétiques et morpho-anatomiques. Il est nécessaire de séquencer

plusieurs gènes indépendants et de comparer l'information génétique transmise par l'un et l'autre des parents, donc d'analyser de l'ADN mitochondrial et de l'ADN nucléaire. Les données doivent être analysées par des méthodes fiables, même si celles-ci requièrent beaucoup plus de temps, et le support des branches doit toujours être évalué.

Comme nous l'avions supposé en commençant cette étude, les analyses phylogénétiques sur base génétique et morpho-anatomique constituent une étape préliminaire indispensable aux études phylogéographiques. Elles permettent de choisir les taxons se prêtant à de telles études et d'évaluer le cadre temporel de leur évolution. En outre, elles peuvent servir de support à la reconstruction de scénarios évolutifs pour des traits morphologiques, écologiques ou comportementaux, et par conséquent aident à mieux comprendre les processus évolutifs.

II - DE LA PHYLOGEOGRAPHIE A LA BIOGEOGRAPHIE

1) Diversité des schémas phylogéographiques

Les schémas phylogéographiques obtenus pour deux des espèces étudiées, *Sylvisorex johnstoni* et *Stochomys longicaudatus*, présentent une ségrégation géographique marquée des haplotypes qui serait d'origine ancienne. Les schémas obtenus pour les deux autres espèces, *Sylvisorex ollula* et *Hylomyscus stella*, font apparaître une expansion récente dans la zone d'étude. Le contraste entre ces deux types de schémas phylogéographiques pourrait dépendre du cadre temporel et spatial de l'évolution de ces espèces. Le paramètre temporel ne semble pas prédominant puisque au moins trois des espèces seraient apparues soit vers la fin du Pliocène (*S. johnstoni* et *S. longicaudatus*), soit au début du Pléistocène (*S. ollula*). La quatrième espèce, *H. stella 2* (c'est-à-dire la forme de *H. stella* présente en Afrique Ouest centrale), pourrait être apparue plus tard dans la zone d'étude ; nous supposons seulement que son origine se situe entre 1,9 MA et 0,6 MA. En revanche, comme les distributions géographiques des espèces au cours du Pléistocène ne sont pas connues, il est possible que les différences phylogéographiques observées résultent, au moins en partie, de différences d'extension et de localisation des aires de distribution passées.

Une autre possibilité est que toutes les espèces aient bien été sympatriques et soumises à des événements similaires durant la majeure partie du Pléistocène, mais n'aient pas répondu de la même façon aux fluctuations paléo-environnementales. Plusieurs hypothèses ont été proposées pour tenter d'expliquer la variabilité des réponses aux changements paléo-environnementaux selon les taxons. Elles incluent : la survie

différentielle des espèces dans différents refuges, la sensibilité aux barrières et aux pressions environnementales, les capacités de colonisation, le pouvoir compétiteur, ainsi que des paramètres directement impliqués dans la dynamique des populations, tels que le flux génique, le taux d'évolution moléculaire, la taille efficace des populations, et le temps de génération (Zink, 1996 ; Taberlet *et al.*, 1998). Toutes ces hypothèses font intervenir les caractéristiques écologiques et biologiques des espèces, or celles-ci sont peu connues dans le cas des modèles étudiés. Pour les deux musaraignes, *S. johnstoni* et *S. ollula*, nous avons évoqué l'influence possible de la taille des individus sur leurs capacités de dispersion et de survie en période de fragmentation forestière. Pour les deux muridés, *H. stella* et *S. longicaudatus*, le rapport de taille est inversé par rapport à la situation rencontrée pour les deux musaraignes. Ces deux espèces ne sont pas congénériques, ni proches parentes (Musser et Carleton, 1993 ; Chevret, 1994), et il est possible qu'elles diffèrent pour quelques caractères écologiques et biologiques méconnus. En effet, peu d'études ont été consacrées à *S. longicaudatus* (cf. Rosevear, 1969), et les caractéristiques de *H. stella* sont d'autant moins bien connues que cette espèce ne peut être diagnostiquée *in vivo*.

Un contraste entre espèces présentant un degré de structuration génétique élevé et espèces génétiquement homogènes a aussi été mis en évidence chez des poissons d'eau douce (Bernatchez et Wilson, 1998), des rongeurs forestiers sud Américains (Patton *et al.*, 1996) et chez des oiseaux des forêts tropicales australiennes et des forêts tempérées américaines (Joseph *et al.*, 1995 ; Zink, 1996 ; Moritz et Faith, 1998). Pour les espèces forestières, les différences phylogéographiques observées dans ces travaux semblent essentiellement imputables à la mobilité des espèces et à leur dépendance vis-à-vis de l'habitat forestier, et par conséquent à leur sensibilité à la fragmentation de cet habitat. La comparaison de communautés de Primates habitant des fragments forestiers de tailles différentes a permis de déceler des différences de sensibilité à la fragmentation de l'habitat. Au Gabon, certaines espèces forestières, comme *C. cephus* et *C. nictitans*, conservent des chances de survie élevées en milieu fragmenté, tandis que d'autres, comme *C. pogonias* et les mangabeys, subiraient d'importantes pressions de sélection liées à l'augmentation de la densité et à la compétition (Tutin et White, 1999). D'autre part, les espèces capables de survivre dans un environnement savanicole, comme les mandrills et les chimpanzés, auraient plus de chances de survivre en milieu fragmenté que les espèces ne pouvant vivre qu'en milieu forestier, comme les gorilles par exemple. Cependant, à Kibale, en Ouganda, aucune caractéristique biologique ou écologique des espèces de Primates ne permet de prédire leur capacité à vivre dans des isolats forestiers de taille et de structure différentes (Onderdonk et Chapman, 2000).

2) Tentative de synthèse biogéographique

Nous avons cherché s'il existait une concordance entre les quatre schémas phylogéographiques obtenus. La plupart des méthodes classiques de biogéographie cladistique ne sont pas applicables pour un nombre aussi limité de modèles. La méthode de Hovenkamp, améliorée en tenant compte des temps de divergence estimés, ne nous a permis de conforter qu'un seul événement de vicariance, observé dans deux cladogrammes (fig. 4.14 A). La construction d'un diagramme de Venn (fig. 4.17) pour des aires biogéographiques prédéfinies selon des études biogéographiques antérieures et adaptées en fonction de nos résultats (cf. fig. 4.14 B) a révélé un seul événement de divergence majeur, qui n'est conforté que par un des modèles (*S. johnstoni*). Cette tentative de combinaison des résultats phylogéographiques nous met en garde sur le risque de biaiser l'analyse par un modèle porteur d'un signal cohérent plus fort que les autres (cf. Moritz et Faith, 1998).

Ces résultats permettent de remettre en cause la prédiction que, s'il existe une histoire commune des aires, elle devrait être reflétée par la majeure partie des peuplements (Nelson et Platnick, 1981). Il peut exister un signal dominant sans que celui-ci apparaisse chez la majorité des taxons. De ce fait, une étude n'impliquant que quelques modèles est d'un intérêt limité lorsqu'il s'agit de proposer un scénario biogéographique. En revanche, la comparaison de schémas phylogéographiques discordants permet de s'interroger sur la généralité des processus impliqués, et notamment sur la réalité de la prédominance de la vicariance sur la dispersion (cf. Voelker, 1999). Nous suggérons que la part relative de ces deux mécanismes soit évaluée avant de combiner les données selon une méthode de biogéographie cladistique qui privilégie la vicariance sur la dispersion.

La datation des événements de vicariance a permis d'aborder l'évolution dans sa dimension temporelle. Malgré l'ampleur des intervalles de confiance associés aux datations, seuls quelques-uns des événements de divergence observés semblent avoir été synchrones. Ce résultat pourrait tout aussi bien être un artefact lié à l'imprécision des datations. En fait, la théorie de la coalescence prédit que, pour un même événement historique, les estimations du temps de divergence intra-spécifique ne devraient pas coïncider entre espèces (Taberlet *et al.*, 1998). En pratique, les temps de divergence sont rarement concordants entre taxons (e.g. Joseph *et al.*, 1995).

3) Comparaison avec les hypothèses biogéographiques actuelles et avec d'autres modèles

Nous avons essayé de savoir si la diversité génétique intra-spécifique a été affectée par l'histoire supposée des aires biogéographiques, telle qu'elle a été inférée d'études menées par notre équipe (cf. chapitre 1.II). Seul le schéma phylogéographique obtenu pour *H. stella* pourrait coïncider avec l'histoire des aires envisagée. La ségrégation géographique des haplotypes observée chez *S. johnstoni* et *S. longicaudatus* présente une certaine adéquation avec les régions fauniques, mais les groupements ne suivent pas les relations présumées entre aires. Pour aucun des modèles il n'apparaît de dichotomie entre région côtière et continentale, comme cela a pu être observé chez d'autres taxons (grenouilles : Ohler, 1996 ; oiseaux : Beresford et Cracraft, 1999 ; rongeurs : Renaud, 1999).

Les résultats obtenus par la méthode de Hovenkamp font apparaître des groupements de localités correspondant aux régions fauniques et ceci bien que cette méthode ne nécessite pas la prédéfinition des aires géographiques. Le nombre limité de localités incluses dans la présente étude ne nous permet pas de définir des régions biogéographiques (dont les limites ne peuvent de toutes façons pas être définies avec précision), mais seulement de préciser l'appartenance de chaque localité à une région ou une autre. Ainsi, les résultats obtenus pour deux des modèles (*S. johnstoni* et *S. longicaudatus*) nous incitent à scinder la région Ouest Congo en deux, selon un axe différent de celui qui délimite actuellement les deux sous-régions reconnues pour la région Ouest Congo. Une de ces subdivisions, que nous avons appelée Ouest Congo 1, ne contient pour l'instant qu'une seule localité (Odzala). Un échantillonnage plus conséquent de la région permettrait de mieux définir la limite entre sous-régions, et le rôle de la rivière Sangha dans cette séparation.

Parmi les quatre régions qui ressortent des analyses, seules les régions Sud Ogooue et Ouest Congo 2 apparaissent relativement homogènes et différenciées. Odzala apparaît comme une région différenciée pour deux modèles (*S. johnstoni* et *S. longicaudatus*) et comme une zone de convergence des faunes pour les deux autres modèles. Quand à la région dite "d'intergradation", elle présente des haplotypes monophylétiques et uniques à cette région pour au moins un des modèles (*S. johnstoni*). Des petits Mammifères ont donc pu survivre dans cette zone. Nos résultats ne permettent pas de valider l'hypothèse d'un corridor de savane ayant perduré pendant la majeure partie du Pléistocène (Bigalke, 1978 ; Deleporte et Colyn, 1999), mais il se peut que des isolats forestiers aient servi de refuge à des petits Mammifères sans permettre la survie de Mammifères de plus grande taille. Il serait nécessaire d'analyser d'autres modèles et de suivre un plan d'échantillonnage plus fin pour préciser si la région "d'intergradation" constitue effectivement une zone de convergence des faunes.

Contrairement à ce qui apparaît dans d'autres études comparatives portant sur les rongeurs du bassin amazonien (Patton *et al.*, 2000) ou sur des régions géographiques limitées d'Amérique du Nord et d'Australie (Taberlet, 1998), les schémas phylogéographiques obtenus ne sont pas concordants, et ceci même pour des espèces phylogénétiquement proches. En fait, il semble que la concordance phylogéographique entre taxons serait l'exception plutôt que la règle (Zink, 1996 ; Taberlet *et al.*, 1998), car elle dépend non seulement de l'âge des taxons et de leur distribution ancestrale, mais aussi de leur écologie, de leur plasticité, et de leurs capacités de survie et de colonisation.

Il est difficile d'extraire une information biogéographique cohérente de modèles aussi différents et aussi peu nombreux. Cependant, les résultats obtenus montrent une certaine concordance avec les régions fauniques et les hypothèses biogéographiques qui les sous-tendent (refuges, barrières fluviales), et ceci bien que la diversité génétique semble avoir une origine beaucoup plus ancienne que celle qui a été envisagée pour les régions fauniques. D'autre part, la comparaison de schémas phylogéographiques diversifiés nous instruit sur la complexité des processus macro-évolutifs impliqués.

III - MODALITES DE LA SPECIATION

1) Spéciation sympatrique ou allopatrique ?

Le test du mode géographique de spéciation prédominant chez les Cercopithèques (chapitre 5.I) indique que l'allopatrie constituerait la règle générale. Le fait que la différenciation de certains traits phénotypiques et caryotypiques augmente avec le degré de sympatrie suggère aussi une divergence en allopatrie suivie d'un renforcement des différences dans les zones de contact secondaires. Ce modèle évolutif est également soutenu par l'étude de l'évolution des préférences pour un habitat forestier ou savanicole, de l'occupation des différentes strates de végétation, et des adaptations morphologiques à l'arboricole et à la terrestrialité (chapitre 5.II). En effet, il apparaît que les espèces proches parentes tendent à partager les mêmes caractéristiques tandis que des espèces plus éloignées diffèrent pour certains traits. Le choix de l'habitat et les modalités d'exploitation de cet habitat ne constitueraient donc pas des facteurs de spéciation (cf. Barraclough et Nei, 2001), comme le suppose l'hypothèse de spéciation sympatrique selon des gradients environnementaux (cf. Moritz *et al.*, 2000).

Le test du mode géographique de spéciation n'a pas pu être réalisé pour les musaraignes et les muridés en raison du manque de données sur la distribution et l'écologie de la plupart des espèces. Il ne

nous est donc pas possible de dire si l'allopatrie constitue aussi le mode géographique de spéciation prédominant chez les Muridae et les Soricidae. L'analyse des groupements emboîtés réalisée pour les quatre modèles étudiés dans le chapitre 4 (chapitres 4.I.2 et 4.II) révèle toutefois l'importance de la fragmentation passée et donc de la vicariance dans la différenciation génétique intra-spécifique. Le mode de spéciation allopatrique est attendu chez les espèces qui ne sont pas spécialisées pour un type d'habitat ou pour certains items alimentaires (Patton et da Silva, 1998). Certaines espèces de petits mammifères sont spécialisées, comme par exemple le *Prionomys batesi* qui ne consomme que quelques espèces de fourmis arboricoles (Genest-Villard, 1980), mais la plupart des espèces, et notamment celles qui ont été étudiées ici, semblent être relativement tolérantes vis-à-vis de l'habitat et avoir un régime alimentaire diversifié (Rosevear, 1969 ; Kingdon, 1971, 1997). En effet, les Rongeurs ne consomment pas uniquement des graines et des tubercules, de même que les Insectivores ne se nourrissent pas exclusivement d'insectes ni même d'arthropodes.

Actuellement, des cas de spéciation sympatrique ont été mis en évidence essentiellement chez les Insectes, notamment des Insectes phytophages inféodés à leurs plantes hôtes (e.g. Berlocher, 1998 ; Feder, 1998 ; Barraclough et Vogler, 2000), et chez les Poissons d'eau douce (cf. revue dans McCune et Lovejoy, 1998). Pour ce qui est des Mammifères, ce mode de spéciation n'a pas été prouvé, mais une différenciation sympatrique pourrait expliquer en partie la distribution géographique du polymorphisme génétique et phénotypique observé chez deux espèces de rongeurs amazoniens (Smith *et al.*, 2001). En fait, il n'est pas aisé de prouver qu'un événement de spéciation s'est produit en sympatrie, notamment parce que les mécanismes impliqués peuvent être les mêmes que pour un événement qui aurait été initié en allopatrie (Via, 2001). Il pourrait être plus pertinent de distinguer les modes de spéciation selon les mécanismes impliqués, tels que sélection naturelle ou dérive génétique, que selon leur mode géographique (Via, 2001).

2) Ancienneté des événements de spéciation et de divergence intra-spécifique

La diversité génétique intra-spécifique semble remonter au Pliocène pour au moins deux des modèles. Les phénomènes à l'origine de la différenciation intra-spécifique seraient donc beaucoup plus anciens que les derniers cycles glaciaires (contra Grubb, 1990 et Colyn, 1991 pour les Primates ; Wieczorek *et al.*, 2000 pour les grenouilles), lesquels peuvent par contre expliquer la distribution géographique actuelle des lignées, en conjonction avec les conditions écologiques, géo-morphologiques et climatiques (e.g. Heyer et Maxson, 1982 ; Endler, 1982). Nos résultats sont en accord avec les datations

obtenues pour d'autres taxons. Des divergences intra-spécifiques datant de la fin du Pliocène et de la première moitié du Pléistocène ont été mises en évidence chez des rongeurs amazoniens (Patton *et al.*, 2000), des grenouilles d'Amérique du Sud (Heyer et Maxson, 1982), et plusieurs espèces de Vertébrés et de plantes européennes (Taberlet *et al.*, 1998). Une synthèse d'études portant sur les Mammifères du monde entier indique que près des trois quarts des événements majeurs de divergence intra-spécifique se seraient produits au Pléistocène, selon une loi de répartition assez uniforme (Avice, 2000). Un quart de ces événements seraient antérieurs au Pléistocène.

Dans la mesure où les changements climatiques et la fragmentation de l'habitat ont vraisemblablement joué un rôle sur la différenciation intra-spécifique, ces temps de divergence moléculaire anciens peuvent paraître surprenants. Cependant, un épisode de vicariance ne laissera de signature génétique que s'il se produit une dérive génétique, phénomène dépendant de la durée de l'isolement et de la taille des populations. En outre, la mise en évidence d'événements anciens ne remet pas en cause la possibilité d'un renforcement par des événements ultérieurs.

Deux modèles, *S. ollula* et *H. stella*, nous permettent d'appréhender les conséquences de l'augmentation de l'amplitude des cycles glaciaires au cours des derniers 800.000 et surtout 250.000 ans (cf. Maley, 1996). Les signes d'expansion observés chez ces espèces permettent de supposer qu'elles sont passées par un goulot d'étranglement il y a près de 250.000 et 600.000 ans, respectivement. Chez *S. ollula*, on n'observe pas de ségrégation géographique des haplotypes subséquente à l'expansion (Fig. 4.2). Deux interprétations sont possibles : soit les phases de fragmentation forestière ultérieures, bien que très sévères, pourraient avoir été trop courtes pour qu'il y ait différenciation au niveau des marqueurs étudiés, soit il existe un flux génique actuel. A l'inverse, chez *H. stella*, espèce pour laquelle l'expansion semble plus ancienne, on observe une certaine ségrégation géographique des haplotypes qui se serait produite après l'expansion (Fig. 4.8). Certains des cycles climatiques compris entre 600.000 ans et l'actuel auraient donc laissé une marque.

Les événements de divergence entre espèces mis en évidence aussi bien chez les Primates que chez les petits Mammifères remonteraient pour la plupart au Pliocène. Ils seraient donc antérieurs aux grandes fluctuations climatiques du Pléistocène. Pour les Soricidés, à l'instar de ce qui a été observé pour le genre *Sorex* (Fumagalli *et al.*, 1999), il semble que l'origine de certains genres, tels que *Paracrocidura*, *Suncus* et *Sylvisorex*, soit bien antérieure à l'âge des fossiles les plus anciens (Butler, 1998). Ceci n'est pas surprenant puisque les données fossiles sont incomplètes et ne peuvent fournir qu'un âge minimum. Pour le genre *Hylomyscus*, les divergences entre espèces seraient plus anciennes que ce qui a été proposé pour les genres voisins *Praomys* et *Mastomys* sur la base d'hybridation d'ADN (Chevret *et al.*, 1994). Pour les

Cercopithèques, la diversification des espèces forestières serait plus ancienne que ce qui est suggéré par les fossiles (Leakey, 1988 ; Pickford et Senut, 1988). Au sein des super-espèces, la divergence entre espèces sœurs se serait produite autour de 2,3 +/- 0,6 MA. De même, au sein du genre *Hylomyscus*, les deux couples d'espèces sœurs observés auraient divergé vers 2,6 et 3,0 MA. Ces résultats sont compatibles avec ceux qui ont été obtenus pour d'autres Mammifères. Une synthèse portant sur un grand nombre de couples d'espèces sœurs indique que seul un quart des événements de divergence se serait produit au Pléistocène, et que la valeur médiane est de 3,2 MA (Avice, 2000).

En Amérique du Sud, la plupart des événements de spéciation semblent remonter au Pliocène pour de nombreux genres d'oiseaux (Zink et Slowinski, 1995 ; Voelker, 1999 ; Loughheed *et al.*, 2000), de rongeurs (Riddle, 1995), et de grenouilles (Heyer et Maxson, 1982). Ainsi, contrairement à ce qui a pu être envisagé (e.g. Lynch, 1988 ; Moritz *et al.*, 2000), les événements de spéciation conduisant aux espèces actuelles ne seraient pas plus anciens en Amérique tropicale qu'en Afrique tropicale. L'orogénèse andine et ses répercussions géologiques dans le bassin Amazonien sont souvent invoquées pour expliquer ces spéciations relativement anciennes en Amérique du Sud (e.g. Clough et Summers, 2000 ; Loughheed *et al.*, 2000 ; Patton *et al.*, 2000). En Afrique centrale, les phénomènes orogéniques auraient été moins importants, mais la formation du rift africain pourrait avoir eu des conséquences indirectes tant géomorphologiques que climatiques (cf. Partridge *et al.*, 1995). En outre, il semble que les périodes arides du Tertiaire aient été beaucoup plus marquées en Afrique qu'en Amérique du Sud (Maley, 1996).

3) Retour sur les différentes théories de l'évolution

De façon à évaluer si les caractéristiques évolutives des modèles étudiés soutiennent une théorie de l'évolution en particulier, nous avons cherché si certaines des prédictions associées à chaque théorie (Tab. 6.1, cf. chapitre 1.II) étaient vérifiées pour les musaraignes et les muridés analysés, d'une part, et pour les Cercopithèques, d'autre part (Tab. 6.2). Il est à noter que plus ces théories sont complexes, moins elles sont faciles à valider ou à falsifier. Le problème se pose notamment pour la théorie des perturbations qui fait intervenir à la fois la notion d'écotones dynamiques, dont la localisation change au cours du temps, et la notion de vicariance. La théorie des barrières fluviales est également difficile à tester puisque même lorsqu'il existe une différenciation entre populations situées de part et d'autre d'une rivière, cela ne signifie pas que la rivière en soit la cause.

Tableau 6.1 : Prédictions associées aux différentes théories de l'évolution

Prédictions	Gradients	Paléogéographie	Barrières fluviales	Refuges	Perturbations
concordance géographique entre taxons / et par rapport aux ...	non	oui / événements géologiques	oui / rivières	oui / régions fauniques	non
concordance temporelle entre taxons / et par rapport aux ...	non	oui / événements géologiques	non	oui / périodes arides	oui / périodes froides
temps de divergence entre espèces	Holocène, Pléistocène >	Miocène, Pliocène	Holocène, Pléistocène >	surtout Pléistocène	Pléistocène
contraction / expansion des populations		oui		oui	oui
taxons frères syntopiques	non			oui	non
différentiation morphologique liée à l'habitat	oui				oui
mode géographique de spéciation	sympatrie	allopatrie	allopatrie	allopatrie	allopatrie

Tableau 6.2 : Prédictions vérifiées par les modèles étudiés

Résultats	Soricidés et Muridés	Cercopithèques
concordance géographique entre taxons / et par rapport aux ...	(schémas phylogéographiques) limitée / régions fauniques, rivières	(distributions) oui / régions fauniques, rivières
concordance temporelle entre taxons / et par rapport aux ...	partielle / périodes froides et arides ?	?
temps de divergence entre espèces	surtout Pliocène	surtout Pliocène
contraction / expansion des populations	oui	?
taxons frères syntopiques	?	oui, au sein des super-espèces
différentiation morphologique liée à l'habitat	non décelée	limitée
mode géographique de spéciation	?	allopatrie
théories supportées	perturbations, refuges, barrières fluviales	refuges, barrières fluviales

Il ne semble pas que la théorie des gradients rende compte de l'évolution des modèles étudiés, notamment parce qu'elle ne s'applique pas pour des taxons dont le mode géographique prédominant de spéciation est l'allopatrie, comme c'est le cas au moins pour les Cercopithèques. La théorie paléogéographique pourrait être éliminée en raison de la rareté des événements tectoniques et orogéniques

ayant affecté la zone d'étude, mais une influence de la formation du rift africain ne peut être exclue (cf. Partridge *et al.*, 1995 pour l'Afrique de l'Est et du Sud). L'importance des refuges et des barrières fluviales pour les Cercopithèques était déjà connue (e.g. Grubb, 1990 ; Colyn, 1991 ; Deleporte et Colyn, 1999). La localisation potentielle des refuges a d'ailleurs été inférée en partie de travaux portant sur les distributions des Primates (cf. Maley, 1996).

Pour les petits mammifères étudiés, dans la mesure où les schémas phylogéographiques ne coïncident pas avec les phytorégions (fig. 4.12) et où aucune différenciation morphologique entre zones géographiques n'a été observée, nous sommes tentés de rejeter les deux hypothèses faisant intervenir une spécialisation pour un type d'habitat, à savoir la théorie des gradients environnementaux et celle des perturbations. Une révision plus minutieuse des caractères crâniens et corporels serait toutefois nécessaire avant de se prononcer. Il est difficile de trancher entre les différentes théories de l'évolution impliquant un mode de spéciation allopatrique, notamment parce que trop peu de modèles ont été analysés pour pouvoir conclure sur la concordance des schémas phylogéographiques et que les distributions et les préférences écologiques des espèces sont mal connues. Il est bien sûr possible que plusieurs théories expliquent la spéciation des musaraignes et des muridés, d'autant plus que ces deux familles couvrent un très large éventail d'espèces.

Nos résultats indiquent une prédominance de l'allopatrie et des conditions historiques du Plio-Pléistocène et du Miocène dans la différenciation intra-spécifique et la spéciation des Mammifères forestiers d'Afrique tropicale. Ils confortent surtout les prédictions associées à la théorie des refuges, sans exclure la théorie paléogéographique, ni celle des barrières fluviales, qui auraient nécessité un échantillonnage géographique plus spécifique pour être évaluées. D'une façon générale, ils limitent le rôle des conditions écologiques actuelles et récentes à leur impact sur le maintien des espèces au sein des communautés.

IV - APPLICATIONS ET PERSPECTIVES

1) Intérêt pour la conservation

a) Diversité morphologique ou diversité génétique ?

Le choix des aires prioritaires pour la conservation se fait généralement sur base de la diversité spécifique, de façon à choisir les aires où le nombre d'espèces endémiques (ou menacées) et la richesse

spécifique sont maximales (Hacker *et al.*, 1998). Les zones les plus riches en espèces ne sont pas forcément les zones d'endémisme ; elles peuvent au contraire correspondre à des zones de convergence des faunes (Deleporte et Colyn, 1999 ; Gautier-Hion *et al.*, 1999). Par exemple, dans le cas de l'Afrique centrale, la richesse spécifique des Primates simiens est importante dans la réserve du Dja, située dans la région d'intergradation, alors que le taux d'endémisme est plus élevé à la Forêt des Abeilles et à la Lopé (Gautier-Hion *et al.*, 1999).

Un autre critère d'évaluation de la biodiversité consiste à prendre en compte la diversité génétique (e.g. Moritz et Faith, 1998 ; Smith *et al.*, 2000). La diversité génétique et la diversité morpho-anatomique ne vont pas forcément de paire et ne suivent pas toujours les mêmes schémas de distribution géographique. Cela a été montré chez des oiseaux australiens et africains (Schneider et Moritz, 1999 ; Schneider *et al.*, 1999 ; Moritz et Faith, 1998 ; Smith *et al.*, 2000), et semble être aussi le cas ici pour *S. johnstoni*. Ces deux types de diversité intègrent des dimensions spatiales et temporelles différentes. La diversité spécifique dépend de la distribution des espèces, donc à la fois des conditions écologiques actuelles et des variations de l'environnement qui ont modelé les assemblages d'espèces au cours des derniers millions d'années, tandis que la diversité génétique résulte des événements de vicariance du Plio-Pléistocène (Moritz et Faith, 1998). La diversité génétique présente l'avantage d'intégrer les espèces cryptiques et les désaccords entre classification et phylogénèse (Moritz et Faith, 1998). De plus, le fait de conserver un pool de diversité génétique permet de limiter la consanguinité et de préserver le potentiel adaptatif des espèces. Parmi les localités relativement bien échantillonnées dans le cadre de cette étude, c'est à Odzala que la richesse génétique est la plus élevée (cf. chapitre 4.IV). Cette localité présente également une richesse spécifique de Mammifères (Colyn et Perpète, 1995) et d'Oiseaux (Dowsett-Lemaire, 1995) parmi les plus élevées d'Afrique centrale, mais aucune espèce de Mammifères endémique n'y a été observée.

b) Comment conserver la biodiversité d'Afrique centrale à long terme ?

Pour conserver durablement, il est nécessaire d'anticiper à la fois les fluctuations environnementales et les fluctuations démographiques. Les zones choisies pour la conservation doivent être suffisamment grandes pour que se perpétuent les processus dynamiques naturels qui régissent les communautés (Moritz *et al.*, 2000). La taille des populations doit être suffisamment élevée pour que le brassage génétique soit maintenu (Hacker *et al.*, 1998). Il semble donc préférable de conserver une grande zone plutôt que plusieurs sites éparpillés. Le degré de fragmentation de l'habitat est aussi un paramètre important. Dans les régions où l'habitat est fragmenté, les îlots d'habitat dont l'aire est inférieure à une certaine valeur seuil peuvent disparaître rapidement (da Silveira Lobo Sternberg, 2001). Enfin, dans une perspective de

conservation à long terme, il est nécessaire de privilégier les sources de diversité (Taberlet, 1998, Deleporte et Colyn, 1999). Celles-ci, reconnaissables à leur taux d'endémisme élevé, sont supposées être génératrices de diversité et devraient correspondre aux zones où l'habitat se maintient en dépit des fluctuations climatiques.

Dans notre zone d'étude, Odzala ne répond pas à ces critères, car les densités de petits Mammifères y sont faibles et l'habitat y est fragmenté (Gautier-Hion et Colyn, 2000). De plus, nos résultats tendent à indiquer qu'il s'agit d'une zone de convergence des faunes. En revanche, la région Sud Ogooué et la portion de forêt appartenant à la région Ouest Congo comprise entre la Sangha et l'Oubangui pourraient abriter des sources de diversité (cf. chapitre 4.IV). Ces régions abritent déjà les réserves contiguës de Dzanga-Sangha et de Nouabalé-Ndoki, et la réserve de la Lopé. Une solution efficace consisterait à étendre ces aires protégées, ou à les relier par des corridors. C'est ce que propose le "projet tri-national" qui inclut non seulement ces trois réserves mais aussi celle du Dja et le Parc National D'Odzala.

2) Implications biogéographiques

a) Choix des modèles

Notre étude met en évidence l'importance, non seulement de la quantité, mais aussi de la qualité des modèles utilisés en biogéographie. Il apparaît que seule la confrontation d'un grand nombre de schémas phylogéographiques pourrait permettre de révéler un schéma dominant dans l'histoire du peuplement. De plus, les modèles ne doivent pas être choisis au hasard, mais en fonction des hypothèses à tester. Des travaux ayant pour but d'évaluer l'impact des changements paléo-environnementaux devraient se limiter à des organismes ayant des capacités de dispersion modérées et les mêmes préférences écologiques, donc ayant potentiellement réagi de la même façon à ces changements paléo-environnementaux. Par contre, l'histoire des aires résultant d'événements géologiques et tectoniques pourra être révélée par des modèles couvrant la plus large gamme de préférences écologiques possible (cf. Zink, 1996 ; Taberlet *et al.*, 1998). De la même façon, il serait possible de tester l'impact des capacités de dispersion sur les schémas biogéographiques en comparant des modèles ayant des modes de dispersion différents (Ditchfield, 2000).

Outre la nécessité de sélectionner des espèces dont les caractéristiques écologiques sont connues et adaptées aux questions posées, le choix des modèles met en œuvre un certain nombre de critères qui sont autant de contraintes. La taxinomie doit être fiable, les aires de distribution doivent être connues et suffisamment étendues, et le nombre de spécimens disponibles doit être élevé et bien réparti entre les

localités de collecte. L'existence de fossiles constitue un atout supplémentaire. Dans le cadre de cette étude, nous avons mis de côté beaucoup de modèles potentiels faute de connaissances taxinomiques fiables ou d'échantillonnage suffisant, et nous avons consacré une partie de nos efforts à des questions taxinomiques et phylogénétiques. Les mises au point méthodologiques et les analyses préliminaires ajoutent un coût supplémentaire, aussi bien temporel que financier.

Nous avons choisi nos modèles parmi trois familles de Mammifères de façon à couvrir une gamme assez large de caractéristiques écologiques et biologiques. Dans la perspective d'une étude biogéographique à large échelle, il conviendrait d'étudier d'autres familles. Un groupe d'Ongulés, les Céphalophes, est déjà en cours d'analyse (M. Colyn et E. Verheyen, comm. pers.). Il pourrait également être intéressant de comparer des modèles ayant des modes de dispersion très différents, comme des Mammifères terrestres et des Mammifères volants. En Amazonie, les schémas phylogéographiques obtenus pour quelques espèces de chauves-souris diffèrent fortement de ceux qui ont été révélés pour des rongeurs et marsupiaux (Ditchfield, 2000). Pour l'Afrique tropicale, ces petits Mammifères volants n'ont fait pour l'instant l'objet que de travaux de révision de leur systématique et des aires de distribution (e.g. Bergmans, 1994, 1997).

Il serait également intéressant de traiter des modèles appartenant à d'autres classes. De nombreux travaux biogéographiques portant sur la distribution des espèces d'Oiseaux ont été réalisées (e.g. Diamond et Hamilton, 1980 ; Prigogine, 1988 ; Louette, 1992), mais il s'avère que les distributions des espèces sont parfois insuffisamment documentées (Beresford et Cracraft, 1999). La classe des Amphibiens pourrait aussi fournir de bons modèles, comme l'ont montré des études réalisées dans d'autres régions du monde (e.g. Heyer et Maxson, 1982 ; Clough et Summers, 2000). Les Amphibiens étant souvent inféodés aux rivières, ils pourraient permettre de révéler l'histoire des bassins drainants (Avice, 2000). Les Insectes se prêtent plus difficilement à ce genre d'étude car leur extrême diversité est un obstacle à leur identification morphologique, aux inventaires de faunes et aux révisions phylogénétiques indispensables. Cependant, des travaux préliminaires ont été réalisés sur les glossines, ou mouches tsé-tsé, en Afrique centrale (e.g. Gouteux *et al.*, 1991).

L'éventail des taxons pouvant être analysés en biogéographie comprend également la flore. Les bégoniacées africaines ont été étudiées dans une perspective biogéographique (Sosef, 1994, 1996). Ces plantes pourraient constituer un modèle privilégié en raison de leurs capacités de dispersion limitées, mais une dispersion des graines par les cours d'eau ne peut être exclue. De plus, leur distribution pourrait s'expliquer par une niche écologique étroite plutôt que par une faible dispersion hors des zones de refuge.

b) Echantillonnage géographique

Le plan d'échantillonnage est limité par des contraintes temporelles, financières et politiques (certaines régions n'ont pu être échantillonnées en raison des conditions politiques qui y règnent). L'effort réalisé par notre équipe en Afrique centrale a nécessité plus d'une cinquantaine de missions réparties sur près de dix ans, et cependant, il s'est révélé insuffisant pour aborder un certain nombre de questions. Nous n'avons pu inclure la région Ouest Cameroun dans nos analyses que pour *S. johnstoni*, et ceci grâce à un envoi de matériel extérieur. L'absence de collectes effectuées dans cette région par notre équipe demande à être comblée, afin de pouvoir considérer l'ensemble des processus ayant affecté l'Afrique "Ouest Centrale". Cette région doit d'autant moins être négligée que la richesse spécifique et le nombre d'espèces menacées y sont très élevés, du moins chez les Primates (Hacker *et al.*, 1998). D'autre part, il serait intéressant de tester l'effet de barrière des rivières, comme cela a été fait en Amazonie (Patton *et al.*, 1994 ; Peres *et al.*, 1996). Ceci nécessite un échantillonnage minutieux sur les deux rives d'un cours d'eau important. Les zones de contact secondaire mériteraient aussi d'être étudiées, en adoptant un plan d'échantillonnage suffisamment fin pour aborder la génétique des populations, dans le but d'identifier les mécanismes impliqués dans la spéciation. Pour les petits mammifères, la localisation de ces zones reste à déterminer. En conséquence, il semble plus important de multiplier les localités de collecte dans une zone géographiquement limitée que de chercher à étendre cette zone en Afrique de l'Est ou de l'Ouest.

c) Approche multidisciplinaire

Les travaux présentés sont basés sur des séquences d'ADN mitochondrial. Les avantages de l'ADN mitochondrial (cf. chapitre 2.III) font qu'il est souvent employé en phylogénie moléculaire (e.g. Horovitz et Meyer, 1995 ; Hassanin *et al.*, 1998b ; Matthee et Robinson, 1999 ; Delpero *et al.*, 2001) et en phylogéographie (e.g. Avise *et al.*, 1987 ; Joseph *et al.*, 1995 ; Schneider *et al.*, 1998 ; Patton *et al.*, 2000). Nos résultats font ressortir la nécessité d'analyser également de l'ADN nucléaire de façon à disposer de plusieurs sources d'information indépendantes pour retracer la phylogénie des organismes (e.g. Moore, 1995), à déceler s'il existe un biais dans la dispersion des mâles et des femelles (e.g. Melnich et Hoelzer, 1992) et à détecter et localiser d'éventuels phénomènes d'hybridation ou d'introggression (e.g. Ruedi *et al.*, 1997). L'utilisation de marqueurs moléculaires évoluant très rapidement, tels que les micro-satellites, permettrait notamment d'étudier les processus mis en jeu dans des zones de contact secondaires ou d'écotone.

Dans le but de mieux comprendre les processus évolutifs, plusieurs études complémentaires devraient être menées en parallèle avec les études phylogéographiques basées sur la variabilité génétique.

Premièrement, lorsque les modèles choisis sont des taxons dont les caractéristiques biologiques et écologiques sont méconnues, ils devraient faire l'objet d'études de terrain à long terme afin de combler cette lacune. Deuxièmement, il apparaît indispensable d'étudier les variations géographiques des caractères morphologiques et de les comparer à celles de la diversité génétique. Troisièmement, dans les régions où des espèces sœurs se retrouvent en sympatrie, des études écologiques permettraient d'observer ce qui se produit en situation de compétition entre taxons phylétiquement proches. Par exemple, au Cameroun, de part et d'autre de la Sanaga, il existe une zone de sympatrie entre *C. pogonias* et *C. mona*, où les deux espèces se partageraient l'espace grâce à un léger décalage de niches écologiques (Gartlan et Stuhlsaker, 1972). Quatrièmement, une étude des différences comportementales entre espèces sœurs, en laboratoire et sur le terrain, permettrait de mieux connaître les mécanismes d'isolement interspécifiques et les conditions dans lesquelles se produit l'hybridation. Dans le cas des Cercopithèques, des observations de terrain portant sur trois individus hybrides et leur descendance ont montré que l'hybridation est possible en dépit de différences de coloration, de formule chromosomique et de taille corporelle (Struhsaker *et al.*, 1988). La barrière spécifique ne serait cependant rompue que dans des situations particulières, par exemple lors d'un fort déséquilibre densitaire. D'autres études sont nécessaires pour préciser les conditions dans lesquelles peut se produire l'hybridation.

Ces travaux peuvent être considérés comme une étude exploratoire permettant d'évaluer les intérêts et les limites des analyses phylogéographiques dans l'étude de l'évolution des faunes forestières d'Afrique tropicale, et de mettre en évidence des pistes qui mériteraient d'être explorées à l'avenir. Dans l'optique de tester des hypothèses biogéographiques d'ordre général, telles que l'influence de la latitude, l'impact de barrières biogéographiques prédéfinies ou l'importance du mode de dispersion, il apparaît nécessaire d'étudier un grand nombre de modèles diversifiés. Cependant, une analyse "en profondeur" privilégiant quelques modèles semble nécessaire pour évaluer les processus mis en jeu. Nous suggérons de multiplier les points de collecte en Afrique centrale et d'adopter une approche multidisciplinaire, de façon à déterminer avec précision le statut taxinomique et l'aire de distribution des espèces étudiées, et à pouvoir trier les espèces selon un filtre écologique. Cette perspective ne peut être envisagée que comme un travail d'équipe et sur le long terme.

REFERENCES

- Adams, J.M. et Faure, H. (1997) Review and Atlas of Palaeovegetation: Preliminary land ecosystem maps of the world since the Last Glacial Maximum. QEN project, Oak Ridge National Laboratory, TN, USA, disponible à l'adresse : <http://www.esd.ornl.gov/projects/gen/adams1.html>.
- Arctander, P. (1995) Comparison of a mitochondrial gene and a corresponding nuclear pseudogene. *Proc. R. Soc. Lond. B*, **262**, 13-19.
- Aubreville, A. (1949) *Climats, forêts et désertification de l'Afrique tropicale*. Société des Editions Géographiques et Coloniales, Paris.
- Avise, J.C. (1989) Gene trees and organismal histories: a phylogenetic approach to population biology. *Evolution*, **43**, 1192-1208.
- Avise, J.C. (1992) Molecular population structure and the biogeographic history of a regional fauna: a case study with lessons for conservation biology. *Oikos*, **638**, 62-76.
- Avise, J.C. (2000) *Phylogeography*. Harvard University Press, Cambridge, MA.
- Avise, J.C., Arnold, J., Ball, R.M., Bermingham, E., Lamb, T., Neigel, J.E., Reeb, C.A. et Saunders, N.C. (1987) Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. *Annu. Rev. Ecol. Syst.*, **18**, 489-522.
- Ayres, J.M. et Clutton Brock, T.H. (1992) River boundaries and species range size in Amazonian primates. *Am. Nat.*, **140**, 531-537.
- Bahuchet, S. (1986) Linéaments d'une histoire humaine de la forêt du bassin congolais. In : Mémoires du Museum National d'Histoire Naturelle, Série A, Zoologie , Tome 132 - Edition du Museum, Paris, pp. 312-313.
- Barbault, R. (1995) *Ecologie des peuplements: structure et dynamique de la biodiversité* (2e tirage). Masson, Paris.
- Barker, F.K. et Lanyon, S.M. (2000) The impact of parsimony weighting schemes on inferred relationships among toucans and neotropical barbets (Aves: piciformes). *Mol. Phylogenet. Evol.*, **15**, 215-234.
- Barracough, T.G. et Vogler, A.P. (2000) Detecting the geographical pattern of speciation from species-level phylogenies. *Am. Nat.*, **4**, 419-434.
- Barracough, T.G. et Nee, S. (2001) Phylogenetics and speciation. *T.R.E.E.*, **16**, 391-399.
- Baum, B.R. (1992) Combining trees as a way of combining data sets for phylogenetic inference, and the desirability of combining gene trees. *Taxon*, **41**, 3-10.
- Bena, G., Prosper, J.-M., Lejeune, B. et Olivieri, I. (1998) Evolution of annual species of the genus *Medicago*: A molecular phylogenetic approach. *Mol. Phylogenet. Evol.*, **9**, 552-559.
- Bennett, K.D. (1990) Milankovitch cycles and their effects on species in ecological and evolutionary times. *Paleobiology*, **16** (1), 11-21.
- Beresford, P. et Cracraft, J. (1999) Speciation in African forest robins (*Stiphornis*): species limits, phylogenetic relationships, and molecular biogeography. *Am. Mus. Novitates*, **3270**, 1-22.
- Bergmans, W. (1994) Taxonomy and biogeography of african fruit bats (Mammalia, megachiroptera). 4. The genus *Rousettus* Gray, 1821. *Beaufortia*, **44**, 79-126.
- Bergmans, W. (1997) Taxonomy and biogeography of African fruit bats (Mammalia, Megachiroptera) 5. *Beaufortia*, **47** (2), 11-11.

- Berlocher, S.H. (1998) Can sympatric speciation via host or habitat shift be proven from phylogenetic and biogeographic evidence? *In* : Endless forms, species and speciation. Howard, D.J. et Berlocher, S.H. (eds), Oxford University Press, Inc. New York, pp. 99-113.
- Bernatchez, L. et Wilson, C.C. (1998) Comparative phylogeography of Nearctic and Palearctic fishes. *Mol. Ecol.*, **7**, 431-452.
- Bibb, M. (1981) Sequence and gene organization of mouse mitochondrial DNA. *Cell*, **26**, 167-180.
- Bigalke, R.C. (1978) Present day mammals of Africa (Chapter 1). *In* : Evolution of African Mammals. Maglio, J. et Cooke, H.B.S. (eds), Harvard University Press, Cambridge, pp. 1-16.
- Blondel, J. (1995) *Biogéographie. Approche écologique et évolutive*. Masson, Paris.
- Bonnefille, R. (1987) Evolution des milieux tropicaux africains depuis le début du Cénozoïque. *Mémoire et travaux de l'Institut de Montpellier*, **17**, 101-110.
- Bonnefille, R., Roeland, J.C. et Guiot, J. (1990) Temperature and rainfall estimates for the past 40.000 years in equatorial Africa. *Nature*, **346**, 347-349.
- Booth, A.H. (1954) The Dahomey Gap and the mammalian fauna of the West African forests. *Rev. Zool. Bot. Afr.*, **50**, 305-314.
- Booth, A.H. (1958a) The zoogeography of West African primates. A review. *Bull. Inst. Fond. Afr. Noire.*, **20**, 587-622.
- Booth, A.H. (1958b) The Niger, the Volta and the Dahomey Gap as geographical barriers. *Evolution*, **12**, 48-62.
- Bremer, K. (1992) Ancestral areas: a cladistic reinterpretation of the center of origin concept. *Syst. Biol.*, **41**, 436-445.
- Bremer, K. (1994) Branch support and tree stability. *Cladistics*, **10**, 295-304.
- Bricquet, J.P. (1990) Régimes et bilans hydrologiques de l'Afrique Centrale. Les apports à l'océan, du Golfe du Biafra à la Pointa do Dande. *In* : Paysages quaternaires de l'Afrique centrale atlantique. Lanfranchi, R. et Schwartz, D. (eds), ORSTOM, Paris, pp. 42-51.
- Bromham, L., Rambaut, A. et Harvey, P.H. (1996) Determinant of rate variation in mammalian DNA sequence evolution. *J. Mol. Evol.*, **43**, 610-621.
- Brooks, D.R. (1981) Hennig's parasitological method: a proposed solution. *Syst. Zool.*, **30**, 229-249.
- Brooks, D.R. (1990) Parsimony analysis in historical biogeography and coevolution: methodological and theoretical update. *Syst. Zool.*, **39**, 14-30.
- Brooks, D.R., van Veller, M.G.P. et McLennan, D.A. (2001) How to do BPA, really. *J. Biogeogr.*, **28**, 345-358.
- Brosset, A., Dubost, G. et Heim de Balsac, H. (1965) Mammifères inédits récoltés au Gabon. *Biol. Gabon.*, **1** (2), 147-147.
- Broughton, R.E., Stanley, S.E. et Durrett, R.T. (2000) Quantification of homoplasy for nucleotide transitions and transversions and a reexamination of assumptions in weighted phylogenetic analysis. *Syst. Biol.*, **49** (4), 617-627.
- Brown, W.M., George, M. Jr. et Wilson, A.C. (1979) Rapid evolution of animal mitochondrial DNA. *Proc. Natl. Acad. Sci. U.S.A.*, **76**, 1967-1971.
- Brown, W.M., Prager, E.M., Wang, A. et Wilson, A.C. (1982) Mitochondrial DNA sequences of Primates: tempo and mode of evolution. *J. Mol. Evol.*, **18**, 225-239.
- Bruno, W.J., Socci, N.D. et Halpern, A.L. (2000) Weighted neighbor joining: A likelihood-based approach to distance-based phylogeny reconstruction. *Mol. Biol. Evol.*, **17**, 189-197.

- Buffon, G.L.L., comte de (1761) *Histoire naturelle, générale et particulière*. Imprimerie Royale, Paris.
- Bush, G.L. (1975) Modes of animal speciation. *Annu. Rev. Ecol. Syst.*, **6**, 339-364.
- Bush, M.B. (1994) Amazonian speciation: a necessarily complex model. *J. Biogeogr.*, **21**, 5-17.
- Butler, P.M. (1998) Fossil history of shrews in Africa. In : Evolution of Shrews. Wojcik, J.M. et Wolsan, M. (eds), Mammal Research Institute, Polish Academy of Science, Bialowieza, pp. 121-132.
- Caballé, G. (1978) Essai sur la géographie forestière du Gabon. *Adansonia*, **17**, 425-440.
- Camin, J.H. et Sokal, R.R. (1965) A method for deducing branching sequences in phylogeny. *Evolution*, **19**, 311-326.
- Capparella, A.P. (1991) Neotropical avian diversity and riverine barriers. *Ibid.*, **20**, 307-316.
- Carcasson, R.H. (1964) A preliminary survey of the zoogeography of African butterflies. *East Afr. Wildl. J.*, **2**, 122-157.
- Castelloe, J. et Templeton A.R. (1994) Root probabilities for intra-specific gene trees under neutral coalescent theory. *Mol. Phylogenet. Evol.*, **3**, 102-113.
- Cerling, T.E., Wang, Y. et Quade, J. (1993) Expansion of C4 ecosystems as an indicator of global ecological change in the late Miocene. *Nature*, **361**, 344-345.
- Chevret, P. (1994) Etude évolutive des Murinae (Rongeurs: Mammifères) africains par hybridation ADN/ADN. Comparaison avec les approches morphologiques et paléontologiques. Thèse Univ. Montpellier II.
- Chevret, P., Granjon, L., Duplantier, J.-M., Denys, C; et Catzeflis, F. (1994) Molecular phylogeny of the *Praomys* complex (Rodentia: Murinae): a study based on DNA/DNA hybridization experiments. *Zool. J. Linn. Soc.*, **112**, 425-442.
- Chippendale, P.T. et Wins, J.J. (1994) Weighting, partitioning and combining characters in phylogenetic analysis. *Syst. Biol.*, **43**, 278-287.
- Clement, M., Posada, D. et Crandall, K.A. (2000) TCS: a computer program to estimate gene genealogies. *Mol. Ecol.*, **9** (10), 1657-1660.
- Clough, M. et Summers, K. (2000) Phylogenetic systematics and biogeography of the poison frogs: evidence from mitochondrial DNA sequences. *Biol. J. Linn. Soc.*, **70** (3), 515-540.
- Coetzee, J.A. (1964) Evidence for a considerable depression of the vegetation belts during the Upper Pleistocene on the East African mountains. *Nature*, **204**, 564-566.
- Colinvaux, P. (1993) Pleistocene biogeography and diversity in tropical forests of South America. In : Biological Relationships Between Africa and South America, Goldblatt, P. (ed.), New Haven, CT : Yale Univ. Press, pp. 437-499.
- Collard, M. et Wood, B. (2000) How reliable are human phylogenetic hypotheses? *Proc. Natl. Acad. Sci. USA*, **97** (9), 5003-5006.
- Collins, T.M., Kraus, F. et Estabrook, G. (1994) Compositional effect and weighting of nucleotide sequences for phylogenetic analyses. *Syst. Biol.*, **43** (3), 449-459.
- Colyn, M.M. (1987) Les Primates des forêts ombrophiles de la Cuvette du Zaïre : interprétations zoogéographiques des modèles de distribution. *Rev. Zool. Afr.*, **101**, 183-196.
- Colyn, M. (1991) L'importance géographique du bassin du fleuve Zaïre pour la spéciation : le cas des Primates simiens. *Ann. Sci. Zool. Mus. R. Afr. Cent. Tervuren, Belgique*, **264**, 1-250.
- Colyn, M. (1993) Coat colour polymorphism of red colobus monkeys (*Colobus badius*, Primates, Colobinae) in eastern Zaire: taxonomic and biogeographic implications. *J. Afr. Zool.*, **107**, 301-320.

- Colyn, M. (1999) Population study on the superspecies *Cercopithecus cephus* living in the Sangha-Ubangi forest enclave (Central African Republic) - Description of *C. cephus ngottoensis* subsp.nov. *Mammalia*, **63**, 137-147.
- Colyn, M., Gautier Hion, A. et Verheyen, W. (1991) A re-appraisal of palaeoenvironmental history in Central Africa: evidence for a major fluvial refuge in the Zaïre Basin. *J. Biogeogr.*, **18**, 403-407.
- Colyn, M. et Perpète, O. (1995) Missions zoologiques, Parc National d'Odzala. Projet ECOFAC – Composante Congo, groupement AGRECO/CTFT.
- Colyn, M. et Van Rompaey, H. (1994a) Morphometric evidence of the monotypic status of the african long-nosed mongoose *Xenogale naso* (Carnivora, Herpestidae). *Belg. J. Bot.*, **124**, 175-192.
- Colyn, M. et Van Rompaey, H. (1994b) A biogeographic study of cusimanses (*Crossarchus*) (Carnivora, Herpestidae) in the Zaire basin. *J. Biogeogr.*, **21**, 479-489.
- Cornelius, C., Cofre, H. et Marquet, P.A. (2000) Effects of habitat fragmentation on bird species in a relict temperate forest in semiarid Chile. *Conserv. Biol.*, **14**, 534-543.
- Cornish-Bowden, A. (1983) Phenetic methods of classification use information that is disregarded by minimum-length methods. *J. Theor. Biol.*, **101**, 317-319.
- Cowling, S.A., Maslin, M.A. et Sykes, M.T. (2001) Paleovegetation simulations of lowland Amazonia and implications for Neotropical allopatry and speciation. *Quaternary research*, **55**, 140-149.
- Cox, C.B. et Moore, P.D. (1993) *Biogeography: an ecological and evolutionary approach*. Blackwell Sciences, Oxford.
- Cracraft, J. (1988) Deep-history biogeography: retrieving the historical pattern of evolving continental biotas. *Syst. Zool.*, **37**, 221-236.
- Cracraft, J. et Prum, R.O. (1988) Patterns and processes of diersification: speciation and historical congruence in some Neotropical birds. *Evolution*, **42**, 603-620.
- Crandall, K.A. (1996) Multiple interspecies transmissions of human and simian T-cell leukemia/lymphoma virus type I sequences. *Mol. Biol. Evol.*, **13** (1), 115-131.
- Craw, R.C., Grehan, J.R. et Heads, M.J. (1999) *Panbiogeography : tracking the history of life*. Oxford University Press, New York.
- Croizat, L. (1976) Biogeografia analitica y sintética ("Panbiogeografia") de las Américas. *Biblioteca Acad. Cienc. Fis. Mat. Nat.*, Caracas.
- Croizat, L., Nelson, G. et Rosen, D.E. (1974) Centres of origin and related concepts. *Syst. Zool.*, **23**, 265-287.
- Cunningham, C.W. et Collins, T.M. (1994) Developing model systems for molecular biogeography: vicariance and interchange in marine invertebrates. In : *Molecular ecology and evolution: approahes and implications*. Schierwater, B., Streit, B., Wagner, G.P. et DeSalle, R. (eds), Birkhäuser Verlag, Basel, Switzerland, pp. 405-433.
- Cunningham, C.W., Omland, K.E. et Oakley, T.H. (1998) Reconstructing ancestral character states: a critical reappraisal. *T.R.E.E.*, **13** (9), 361-366.
- Da Silva, M.N.F. et Patton, J.L. (1998) Molecular phylogeography and the evolution and conservation of Amazonian mammals. *Mol. Ecol.*, **7**, 475-486.
- Da Silveira Lobo Sternberg, L. (2001) Savanna-forest hysteresis in the tropics. *Global Ecol. Biogeogr.*, **10** (4), 369-378.
- Darwin, C.R. (1859) *On the origin of species*. John Murray, London.
- de Candolle, A.-P. (1820) Géographie botanique. In : *Dictionnaire des sciences naturelles*, **18**, 359-422, Strasbourg.

- de Foresta, H. (1990) Origine et évolution des savanes intramayombiennes (R.P. du Congo). II. Apports de la botanique forestière. *In* : Paysages quaternaires de l'Afrique Centrale Atlantique. Lanfranchi, R. et Schwartz, D. (eds), ORSTOM, Paris, pp. 326-334.
- de Maret, P. (1990) Le "néolithique" et l'âge du fer ancien dans le sud-ouest de l'Afrique Centrale. *In* : Paysages quaternaires de l'Afrique centrale atlantique. Lanfranchi, R. et Schwartz, D. (eds), ORSTOM, Paris, pp. 447-457.
- de Menocal, P.B. (1995) Plio-pleistocene african climate. *Science*, **270**, 53-59.
- De Rijk, P. (1995) Optimisation of a database for ribosomal RNA structure and application in structural and evolutionary research, Ph.D. dissertation, University of Antwerp, Antwerp.
- Dechamps, R., Lanfranchi, R., Le Cocq, A. et Schwartz, D. (1990) Les macrofossiles végétaux du pays Bateke et de la bordure de la cuvette congolaise (R.P. du Congo). *In* : Paysages quaternaires de l'Afrique centrale atlantique. Lanfranchi, R. et Schwartz, D. (eds), ORSTOM, Paris, pp. 224-227.
- Delany, M.J. et Happold, D.C.D. (1979) *Ecology of African mammals*. Longman, London.
- Deleporte, P. (1993) Characters, attributes and tests of evolutionary scenarios, *Cladistics*, **9**, 427-432.
- Deleporte, P. et Colyn, M. (1999) Biogéographie et dynamique de la biodiversité: application de la "PAE" aux forêts planitiaies d'Afrique centrale. *Biosystema*, **17**, 37-43.
- Delpero, M., Masters, J.C., Cervella, P., Crovella, S., Ardito, G. et Rumpler, Y. (2001) Phylogenetic relationships among the Malagasy lemuriforms (Primates : Strepsirrhini) as indicated by mitochondrial sequence data from the 12S rRNA gene. *Zool. J. Linn. Soc.*, **133**, 83-103.
- Denys, C., Chorowicz, J. et Tiercelin, J.J. (1986) Tectonic and environmental control on the rodent diversity in the Plio-Pleistocene sediments of the African Rift System. *In* : Sedimentation in the African Rifts, Frostick, L.E. *et al.* (eds), *Geol. Soc. Spec. Publ.*, **25**, 363-372.
- Dercourt, J., Zonenshain, L.P., Ricou, L.E., *et al.* (1986) Geological evolution of the Thetys belt from the Atlantic to the Pamirs since the Lias. *Tectonophysics*, **123**, 241-341.
- Diamond, A.W. et Hamilton, A.C. (1980) The distribution of forest passerine birds and quaternary climatic change in tropical Africa. *J. Zool. Lond.*, **191**, 379-402.
- Dieterlen, F. (1990) Rodents of the African rain forest - ecological and evolutionary aspects. *In* : Vertebrates in the Tropics. Peters, G. et Hutterer, R. (eds), Museum Alexander Koenig, Bonn, pp. 215-219.
- Dieterlen, F. et Heim de Balsac, H. (1979) Zur Ökologie und Taxonomie der Spitzmäuse (Soricidae) des Kivu-Gebietes. *Säugetierkundliche Mitteilungen*, **27 (4)**, 241-287.
- Ditchfield, A.D. (2000) The comparative phylogeography of Neotropical mammals: patterns of intraspecific mitochondrial DNA variation among bats contrasted to nonvolant small mammals. *Mol. Ecol.*, **9**, 1307-1318.
- Dobzhansky, T. (1940) Speciation as a stage in evolutionary divergence. *Am. Nat.*, **74**, 312-321.
- Dobzhansky, T. (1951) *Genetics and the origin of species*. Columbia University Press, New York.
- Dowsett-Lemaire, F. (1995) Inventaire ornithologique du Parc National d'Odzala. Projet ECOFAC – Composante Congo, groupement AGRECO/CTFT.
- Doyle, J.J. et Sicks, E. (1987) Preservation of plant samples for DNA restriction endonuclease analysis. *Taxon*, **36**, 715-722.
- Ducroz, J.F. (1998) Contribution des approches cytogénétiques et moléculaires à l'étude systématique et évolutive des genres de rongeurs murinae de la "division" *Arvicanthis*. Thèse de doctorat, Museum National d'Histoire Naturelle, Paris.

- Duplantier, J.M. (1982) Les rongeurs myomorphes forestiers du nord-est du Gabon : peuplement, utilisation de l'espace et des ressources alimentaires, rôle dans la dispersion et la germination des graines. Thèse de doctorat, UST, Montpellier.
- Dupont, L.M. et Agwu, C.O.C. (1992) Latitudinal shifts of forest and savanna in NW Africa during the Brunhes chron: further marine palynological results for site M 16415 (9°N-19°W). *Veg. Hist. Archaeobot.*, **1**, 163-175.
- Dupont, L.M., Ning, S., Jahns, S. et Marret, F. (1996) *Podocarpus* in West Africa during the Late Pleistocene. In : Palaeoecology of Africa and the surrounding Islands. van Zinderen Bakker E.M. Sr et Heine K. (eds), **24**, pp. 85-101.
- Dutrillaux, B., Couturier, J. et Chauvier, G. (1980) Chromosomal evolution of 19 species or subspecies of Cercopithecinae. *Ann. Genet.*, **3**, 133-143.
- EDICEF (1983) *Géographie et cartographie du Gabon*. EDICEF, Paris, France.
- Elton, C.S. (1927) *Animal ecology*. Sidgwick et Jackson, London.
- Emsley, M.G. (1965) Speciation in *Heliconius* (Lep., Nymphalidae): morphology and geographic distribution. *Zoologica (NY)*, **50**, 191-254.
- Endler, J.A. (1977) *Geographic variation, speciation and clines*. Princeton Univ. Press, Princeton, N.J.
- Endler, J.A. (1982) Pleistocene forest refuges: fact or fancy. In : Biological diversification in the tropics. Prance, G.T. (ed.), Columbia University Press, New York, pp. 641-657.
- Evrard, C. (1968) *Recherches écologiques sur le peuplement forestier des sols hydromorphes de la Cuvette centrale Congolaise*. Série Scientifique n° 110. Wellens-Pay, Bruxelles.
- Excoffier, L. (1993) *MinSpNet*. Genetics and Biometry Laboratory, University of Geneva, Switzerland.
- Excoffier, L. (1997) Ce que nous dit la généalogie des gènes. La théorie de la coalescence porte un regard nouveau sur l'histoire de notre espèce. *La Recherche*, **302**, 82-90.
- Faith, D.P. (1992) Conservation evaluation and phylogenetic diversity. *Biol. Conserv.*, **61**, 1-10.
- Farris, J.S. (1969) A successive approximations approach to character weighting. *Syst. Zool.*, **18**, 374-385.
- Farris, J.S. (1970) Methods for computing Wagner trees. *Syst. Zool.*, **19**, 83-92.
- Farris, J.S. (1977) Phylogenetic analysis under Dollo's Law. *Syst. Zool.*, **26**, 77-88.
- Farris, J.S. (1989) The retention index and the rescaled consistency index. *Cladistics*, **5**, 417-419.
- Farris, J.S. (1999) Likelihood and inconsistency. *Cladistics*, **15**, 199-204.
- Farris, J.S., Källersjö, M., Kluge, A.G. et Bult, C. (1994) Testing significance of incongruence. *Cladistics*, **10** (3), 315-319.
- Feder, J.L. (1998) The apple maggot fly, *Rhagoletis pomonella*: flies in the face of conventional wisdom about speciation. In : Endless forms, species and speciation. Howard, D.J. et Berlocher, S.H. (eds), Oxford University Press, Inc. New York, pp. 130-144.
- Fedorov, A.N., Fedorova, L.V., Grechko, V.V., Ryabinin, D.M., Sheremet'eva, V.A., Bannikova, A.A., Lomov, A.A., Ryskov, A.P. et Darevsky, I.S. (1999) Variable and invariable DNA repeat characters revealed by taxonprint approach are useful for molecular systematics. *J. Mol. Evol.*, **48**, 69-76.
- Felsenstein, J. (1978) Cases in which parsimony or compatibility methods will be positively misleading. *Syst. Zool.*, **27**, 401-410.
- Felsenstein, J. (1981a) Evolutionary trees from DNA sequences: a maximum likelihood approach. *J. Mol. Evol.*, **17**, 368-376.

- Felsenstein, J. (1981b) A likelihood approach to character weighting and what it tells us about parsimony and compatibility. *Biol. J. Linn. Soc.*, **16**, 183-196.
- Felsenstein, J. (1984) The statistical inference approach to inferring evolutionary trees and what it tells us about parsimony and compatibility. *In* : Cladistics: perspectives on the reconstruction of evolutionary history. Duncan, T. et Stuessy, T.F. (eds), Columbia University Press, New York, pp. 169-191.
- Felsenstein, J. (1985) Confidence limits on phylogenies with a molecular clock. *Syst. Zool.*, **34**, 152-161.
- Felsenstein, J. (1988) Phylogenies from molecular sequences. *Annu. Rev. Genet.*, **22**, 521-565.
- Felsenstein, J. (1995) *PHYLIP (Phylogeny Inference Package) version 3.57c*. Programme et manuel disponibles à l'adresse : <http://evolution.genetics.washington.edu/phylip.html>.
- Ferris, S.D., Sage, R.D., Prager, E.M., Ritte, U. et Wilson, A.C. (1983) Mitochondrial DNA evolution in mice. *Genetics*, **105**, 681-721.
- Fielhouse, D., Yadzani, F. et Golding, G.B. (1997) Substitution rate variation in closely related rodent species. *Heredity*, **78**, 21-31.
- Fitch, W.M. (1971) Rate of change of concomitantly variable codons. *J. Mol. Evol.*, **1**, 84-96.
- Fitch, W.M. (1984) Cladistic and other methods: problems, pitfalls, and potentials. *In* : Cladistics: perspectives on the reconstruction of evolutionary history Duncan, T. et Stuessy, T.F. (eds), Columbia University Press, New York, 221-252.
- Fjeldsa, J. et Lovett, J. (1997) Geographical patterns of old and young species in African forest biota: the significance of specific mountain areas as evolutionary centers. *Biodiv. Conserv.*, **6**, 325-346.
- Frye, M.S. et Hedges, S.B. (1995) Monophyly of the order rodentia inferred from mitochondrial DNA sequences of the genes for 12S rRNA, 16S rRNA, and tRNA-valine. *Mol. Biol. Evol.*, **12**, 168-176.
- Fumagalli, L., Taberlet, P., Stewart, D.T., Gielly, L., Hausser, J. et Vogel, P. (1999) Molecular phylogeny and evolution of *Sorex* shrews (Soricidae: Insectivora) inferred from mitochondrial DNA sequence data. *Mol. Phylogenet. Evol.*, **11** (2), 222-235.
- Gartlan, J.S. et Struhsaker, T.T. (1972) Polyspecific associations and niche separation of rain-forest anthropoids in Cameroon, West Africa. *J. Zool. Lond.*, **168**, 221-266.
- Gascuel, O. (1997) BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. *Mol. Biol. Evol.*, **14** (7), 685-695.
- Gautier, J.P. (1988) Interspecific affinities among guenons as deduced from vocalizations (chapter II.12). *In* : A Primate radiation: evolutionary biology of the African guenons. Gautier-Hion, A., Bourlière, F., Gautier, J.-P. et Kingdon, J. (eds), Cambridge University Press, Cambridge, pp. 194-226.
- Gautier-Hion, A. (1988) Polyspecific associations among forest guenons: ecological, behavioural and evolutionary aspects. *In* : A primate radiation: evolutionary biology of the African guenons. Gautier-Hion, A., Bourlière, F., Gautier, J.P. et Kingdon, J. (eds), Cambridge University Press, pp. 452-475.
- Gautier-Hion, A., Colyn, M. et Gautier, J.P. (1999) *Histoire naturelle des Primates d'Afrique centrale*. Libreville: ECOFAC.
- Gautier-Hion, A. et Colyn, M. (2000) Rapport final de synthèse – ECOFAC II. Termes de référence Biodiversité animale et Ecologie. Projet ECOFAC – Composante Congo, Union européenne, groupement AGRECO-G.E.I.E.
- Gautier-Hion, A. et Gautier, J.-P. (1978) Le singe de Brazza: une stratégie originale. *Z. Tierpsychol.*, **46**, 84-104.
- Gebo, D.L. et Sargis, E.J. (1994) Terrestrial adaptations in the postcranial skeletons of guenons. *Am. J. Phys. Anthropol.*, **93**, 341-371.

- Genest-Villard, H. (1980) Régime alimentaire des rongeurs myomorphes de forêt équatoriale (région de M'Baiki, République Centrafricaine). *Mammalia*, **44**, 423-426.
- Givnish, T.J. et Sytsma, K.J. (1997) Consistency, characters and the likelihood of correct phylogenetic inference. *Mol. Phylogenet. Evol.*, **7**(3), 320-330.
- Goldman, N. (1993) Statistical models of DNA substitution. *J. Mol. Evol.*, **36**, 182-198.
- Goldman, N. et Whelan, S. (2000) Statistical tests of gamma-distributed rate heterogeneity in models of sequence evolution in phylogenetics. *Mol. Biol. Evol.*, **17**, 975-978.
- Golobov, P.A. (1993) Estimating character weights during tree search. *Cladistics*, **9** (1), 83-92.
- Gonzales, I.L., Sylvester, J.E., Smith, T.F., Stambolian, D. et Schmickel, R.D. (1990) Ribosomal RNA gene sequences and hominoid phylogeny. *Mol. Biol. Evol.*, **7**, 203-219.
- Goodman, S.M. et Rakotondravony, D. (2000) The effects of forest fragmentation and isolation on insectivorous small mammals (Lipotyphla) on the Central High Plateau of Madagascar. *J. Zool.*, **250**, 193-200.
- Gouteux, J.P., Sinda, D. et de Foresta, H. (1991) Répartition et écodistribution de *Glossina frezili* (Diptera : Glossinidae) au Congo. *Ann. Soc. Ent. Fr.*, **27** (4), 483-494.
- Graham, S.W., Reeves, P.A., Burns, A.C.E. et Olmstead, R. (2000) Microstructural changes in non-coding chloroplast DNA: interpretation, evolution, and utility of indels and inversions in basal angiosperm phylogenetic inference. *Int. J. Plant Sci.*, **161** (6, Suppl.), S83-S96.
- Grandcolas, P., Deleporte, P., Desutter-Grandcolas, L. et Daugeron, C. (2001) Phylogenetics and ecology: As many characters as possible should be included in the cladistic analysis. *Cladistics*, **17**, 104-110.
- Granjon, L. (1991) Les rongeurs myomorphes du bassin du Kouilou (Congo). In : Flore et faune du bassin du Kouilou, Congo et leur exploration. Tauraco Research Report n°4. Dowsett, R.J. et Dowsett-Lemaire, F. (eds), Tauraco Press, Belgique, pp. 265-278.
- Grinnel, J. (1917) The niche relationship of the California Thrasher. *The Auk*, **34**, 427-433.
- Groves, C.P. (1993) Order Primates. In : Mammal species of the world: a taxonomic and geographic reference. Wilson, D.E. et Reeder, D.M. (eds), Smithsonian Institution Press, Washington D.C. pp. 262-271.
- Grubb, P. (1978) Patterns of speciation in African mammals. *Bull. Carnegie Mus. Nat. Hist.*, **6**, 152-167.
- Grubb, P. (1982) Refuges and dispersal in the speciation of African forest mammals. In : Biological diversification in the tropics. Prance, G.T. (ed.), Columbia University Press, New York, pp. 537-553.
- Grubb, P. (1990) Primate geography in the Afro-Tropical forest biome. In : Vertebrates in the Tropics. Peters, G. et Hutterer, R. (eds), Museum Alexander Koenig, Bonn, pp. 187-214.
- Grubb, P., Larnould, J.M. et Oates, J.F. (1999) Validation of *Cercopithecus erythrogaster pococki* as the name for the Nigerian white-throated guenon. *Mammalia*, **63**, 389-392.
- Guiraud, R. et Maurin, J.C. (1991) Le rifting en Afrique au Crétacé inférieur : synthèse structurale, mise en évidence de deux étapes dans la genèse des bassins, relations avec les ouvertures océaniques péri-africaines. *Bull. Soc. Géol. France*, **162**, 811-823.
- Hacker, J.E., Cowlishaw, G. et Williams, P.H. (1998) Patterns of African primate diversity and their evaluation for the selection of conservation areas. *Biol. Conserv.*, **84**, 251-262.
- Haffer, J. (1969) Speciation in Amazonian forest birds. *Science*, **165**, 131-136.
- Haffer, J. (1990) Geoscientific aspects of allopatric speciation. In : Vertebrates in the Tropics. Peters, G. et Hutterer, R. (eds), Museum Alexander Koenig, Bonn, pp. 45-60.

- Haffer, J. (1993) Time's cycle and time's arrow in the history of Amazonia. *Biogeographica*, **69**, 15-45 (C.R. de séances de la Société de Biogéographie, Paris).
- Haffer, J. (1997) Alternative model of vertebrate speciation in Amazonia: an overview. *Biodiv. Conserv.*, **6**, 451-476.
- Hallam, A. (1994) *An outline of phanerozoic biogeography*. Oxford Biogeography Series n°10, Oxford University Press, Oxford.
- Hamilton, A.C. (1988) Guenon evolution and forest history (chapter I.2). *In* : A Primate radiation: evolutionary biology of the African guenons. Gautier-Hion, A., Bourlière, F., Gautier, J.-P. et Kingdon, J. (eds), Cambridge University Press, Cambridge, pp. 13-34.
- Hamilton, A.C. (1989) African forests. *In* : Tropical rain forest ecosystems: biogeographical and ecological studies (Ecosystems of the world 14B). Leith, H. et Werger, M.J.A. (eds), Elsevier, Amsterdam, pp. 155-182.
- Happold, D.C.D. (1996) Mammals of the Guinea-Congo rain forest. *Proc. R. Soc. Edinburgh*, **104B**, 243-284.
- Harris, E.E. et Disotell, T.R. (1998) Nuclear gene trees and the phylogenetic relationships of the mangabeys (primates: Papionini). *Mol. Biol. Evol.*, **15** (7), 892-900.
- Harrison, M.J.S. (1988) A new species of guenon (genus *Cercopithecus*) from Gabon. *J. Zool. Lond.*, **215**, 561-575.
- Hasegawa, M., Kishino, H. et Yano, T. (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.*, **22**, 160-174.
- Hassanin, A., Lecomte, G. et Tillier, S. (1998a) The 'evolutionary signal' of homoplasy in protein-coding gene sequences and its consequences for a priori weighting in phylogeny. *Evolution*, **321**, 611-620.
- Hassanin, A., Pasquet, E. et Vigne, J.D. (1998b) Molecular systematics of the subfamily caprinae (Artiodactyla, bovidae) as determined from cytochrome *b* sequences. *J. Mammal. Evol.*, **5**, 217-236.
- Hecketsweiler, P., Doumenge, C. et Mokoko, J. (1991) *Le Parc National d'Odzala, Congo*. U.I.C.N, Cambridge.
- Hein, J. (1989) A new method that simultaneously aligns and reconstructs ancestral sequences for any number of homologous sequences, when phylogeny is given. *Mol. Biol. Evol.*, **6**, 649-668.
- Hennig, W. (1966) *Phylogenetic systematics*. University of Illinois Press, Urbana.
- Hey, J. (1994) Bridging phylogenetics and population genetics with gene tree models. *In* : Molecular ecology and evolution: approaches and applications. Schierwater, B., Streit, B., Wagner, G.P. et DeSalle, R. (eds), Birkhäuser Verlag, Basel, pp. 435-449.
- Heyer, W.R. et Maxson, L.R. (1982) Distributions, relationships, and zoogeography of lowland frogs : the *Leptodactylus* complex in South America, with special reference to Amazonia. *In* : Biological diversification in the Tropics. Prance, G.T. (ed.), Columbia University Press, New York, pp. 375-388.
- Hickman, G.C., Nevo, E. et Heth, G. (1983) Geographic variation in the swimming ability of *Spalax ehrenbergi* (Rodentia: Spalacidae) in Israel. *J. Biogeogr.*, **10**, 29-36.
- Hillis, D.M. (1991) Discriminating between phylogenetic signal and random noise in DNA sequences. *In* : Phylogenetic analysis of DNA sequences Miyamoto, M.M. et Cracraft, J. (eds). Oxford University Press, New York, 278-294.
- Hillis, D.M. (1996) Inferring complex phylogenies. *Nature*, **383**, 130-131.
- Hillis, D.M. et Dixon, M.T. (1991) Ribosomal DNA: molecular evolution and phylogenetic inference. *Quarter. Rev. Biol.*, **66** (4), 411-453.

- Hillis, D.M. et Huelsenbeck, J.P. (1992) Signal, noise and reliability in molecular phylogenetic analysis. *J. Hered.*, **83**, 189-195.
- Hooghiemstra, H. (1999) Tropical rain forest refugia during the last glacial in Africa and South America. XVIe symposium de l'AFLP (association des palynologues de langue française), Institut de Géographie de l'Université de Liège, 13-17/09/99, pp. 3-5.
- Hooghiemstra, H. et Agwu, C.O. (1988) Changes in the vegetation and trade winds in equatorial northwest Africa 140,000-70,000 year B.P. as deduced from two marine pollen records. *Palaeogeogr. Palaeoclimatol. Palaeoecol.*, **66**, 173-213.
- Horowitz, I. et Meyer, A. (1995) Systematics of new world monkeys (Platyrrhini, Primates) based on 16S mitochondrial DNA sequences: A comparative analysis of different weighting methods in cladistic analysis. *Mol. Phylogenet. Evol.*, **4** (4), 448-456.
- Hovenkamp, P. (1997) Vicariance events, not areas, should be used in biogeographical analysis. *Cladistics*, **13**, 67-79.
- Howard, D.J. (1998) Unanswered questions and future directions in the study of speciation. In : Endless forms, species and speciation. Howard, D.J. et Berlocher, S.H. (eds), Oxford University Press, Inc. New York, pp. 439-448.
- Hsü, K., Montadert, J.L., Bernoulli, D., Cita, M.B., Erikson, A., Garrison, R.E., Kidd, R.B., Mélières, F., Müller, C. et Wright, R. (1977) History of the Mediterranean salinity crisis. *Nature.*, **267**, 399-403.
- Huelsenbeck, J.P. et Hillis, D.M. (1993) Success of phylogenetic methods in the four-taxon case. *Syst. Biol.*, **42**, 247-264.
- Huelsenbeck, J.P., Hillis, D.M. et Nielsen, R. (1996) A likelihood-ratio test of monophyly. *Syst. Biol.*, **45**, 546-558.
- Hutchinson, G.E. (1957) Concluding remarks. *Cold Spring Harbor Symposium Quant. Biol.*, **22**, 415-427.
- Hutterer, R. (1986). Diagnosen neuer spitzmäuse aus Tansania (Mammalia: Soricidae). *Bonn. Zool. Beiträge*, **37**, 23-33.
- Huxley, J.S. (1942) *Evolution: The Modern Synthesis*. Harper, New York.
- Irwin, D.M., Kocher, T.D. et Wilson, A.C. (1991) Evolution of the cytochrome *b* gene of Mammals. *J. Mol. Evol.*, **32**, 128-144.
- Iskandar, D. Duplantier, J.M. Bonhomme, F. Petter, F. et Thaler, L. (1988) Mise en évidence de deux espèces jumelles sympatriques du genre *Hylomyscus* dans le nord-est du Gabon. *Mammalia*, **52** (1), 126-130.
- Jacobs, L.L. et Downs, W.R. (1994) The evolution of murine rodents in Asia. In : Rodents and Lagomorph families of Asian origin and diversification. Tomida, Y., Li, C. et Setogushi, T. (eds), National Science Museum Monographs, Tokyo, pp. 149-156.
- Jaeger, J.J., Tong, H. et Buffetaut, E. (1986) The age of *Mus-Rattus* divergence: Paleontological data compared with the molecular clock. *C. R. Acad. Sci. Paris Ser. 2*, **302**, 917-922.
- Jahns, S. (1996) Vegetation history and climate changes in West Equatorial Africa during the Late Pleistocene and Holocene, based on a marine pollen diagram from the Congo fan. *Veget. Hist. Archaeobot.*, **5**, 207-213.
- Jahns, S., Hüls, M. et Sarntheim, M. (1998) Vegetation and climate history of West Equatorial Africa based on a marine pollen record off Liberia (site GIK 16776) covering the last 400,000 years. *Rev. Palaeobot. Palynol.*, **102**, 277-288.
- Jansa, S.A. Goodman, S.M. et Tucker, P.K. (1999) Molecular phylogeny and biogeography of the native rodents of Madagascar (Muridae: nesomyinae). *Cladistics*, **15**, 253-270.

- Jansen, J.H.F. (1990) Glacial-interglacial Oceanography of the Southeastern Atlantic Ocean and the Paleoclimate of West Central Africa. *In* : Paysages quaternaires de l'Afrique centrale atlantique. Lanfranchi, R. et Schwartz, D. (eds), ORSTOM, Paris, pp. 110-123.
- Jarne, P. et Lagoda, J.L. (1996) Microsatellites, from molecules to populations and back. *T.R.E.E.*, **11**, 424-428.
- Jin, L. et Nei, M. (1990) Limitations of the evolutionary parsimony method of phylogenetic analysis. *Mol. Biol. Evol.*, **7**, 82-102.
- Johnson, A. et Gullberg, U. (1998) Theory and models of sympatric speciation. *In* : Endless forms, species and speciation. Howard, D.J. et Berlocher, S.H. (eds), Oxford University Press, Inc. New York, pp. 79-89.
- Joseph, L., Moritz, C. et Hugall, A. (1995) Molecular support for vicariance as a source of diversity in rainforest. *Proc. R. Soc. Lond. Ser. B*, **260**, 177-182.
- Jukes, T.H. et Cantor, C.R. (1969) Evolution of protein molecules. *In* : Mammalian Protein Metabolism. Munro, H.M. (ed.), Academic press, New York, pp. 21-132.
- Juo, A.S.J. et Wilding, L.P. (1996) Soils of the lowland forests of West and Central Africa. *Proc. R. Soc. Edinburgh*, **104B**, 15-29.
- Juste, J., Alvarez, Y., Tabares, E., Garrido Pertierra, A. et Ibanez, C. (1999) Phylogeography of African fruitbats (Megachiroptera). *Mol. Phylogenet. Evol.*, **13**, 596-604.
- Kessing, B., Croom, H., Martin, A., McIntosh, C., McMillan, W.O., et Palumbi, S. (1989) *The Simple Fool's guide to PCR, version 1.0*. Department of Zoology, University of Hawaii.
- Kim, J. (1993) Improving the accuracy of phylogenetic estimation by combining different methods. *Syst. Biol.*, **42** (3), 331-340.
- Kimura, M. (1980) A simple method for estimating evolutionary rate of base substitution through comparative studies of nucleotide sequences. *J. Mol. Evol.*, **16**, 111-120.
- Kingdon, J. (1971) *East African mammals. Vol. 1*. Academic Press, London.
- Kingdon, J. (1980) The role of visual signals and face patterns in African forest monkeys (guenons) of the genus *Cercopithecus*. *Trans. Zool. Soc. Lond.*, **35**, 425-475.
- Kingdon, J. (1988) What are face patterns and do they contribute to reproductive isolation in guenons? (Chapter II.13). *In* : A Primate radiation: evolutionary biology of the African guenons. Gautier-Hion, A., Bourlière, F., Gautier, J.-P. et Kingdon, J. (eds), Cambridge University Press, Cambridge, pp. 227-245.
- Kingdon, J. (1990) An evolutionary whirlpool: the Zaïre Basin (ch. 10). *In* : Island Africa. Kingdon, J.S. (ed.), Collins, London, UK, pp.191-213.
- Kingdon, J. (1997) *The Kingdon field guide to African mammals*. Academic Press, San Diego, California.
- Kingman, J.F.C. (1982) The coalescent. *Stoch. Process Appl.*, **13**, 235-248.
- Kishino, H. et Hasegawa, M. (1989) Evaluation of the maximum likelihood estimates of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. *J. Mol. Evol.*, **29**, 170-179.
- Kishino, H. et Hasegawa, M. (1990) Converting distance to time: Application to human evolution. *Methods in Enzymology*, **183**, 550-570.
- Kluge, A.G. et Farris, J.S. (1969) Quantitative phyletics and the evolution of anurans. *Syst. Zool.*, **18**, 1-32.
- Kluge, A.G. (1989) A concern for evidence and a phylogenetic hypothesis of relationships among *Epicrates* (Boidae, serpentes). *Syst. Zool.*, **38**, 7-25.

- Kocher, T.D., Thomas, W.K., Meyer, A., Edwards, S.V., Pääbo, S., Villablanca, F.X. et Wilson, A.C. (1989) Dynamics of mitochondrial DNA evolution in animals: Amplification and sequencing with conserved primers. *Proc. Nat. Acad. Sci. USA.*, **86**, 6196-6200.
- Kumar, S. (1996) *PHYLTEST version 2.0 : A program for testing phylogenetic hypothesis, version 2.0*, distributed by the author, Institute of Molecular Evolutionary Genetics and department of Biology, The Pennsylvania State University from ftp.bio.indiana.edu in directory molbio/ibmpc.
- Kumar, S., Tamura, K. et Nei, M. 1993. *MEGA : Molecular Evolutionary Genetics Analysis*. Pennsylvania State University, University Park, PA.
- Lanfranchi, R. et Schwartz, D. (1990) *Paysages quaternaires de l'Afrique centrale atlantique*. ORSTOM, Paris.
- Laurent, R.F. (1973) A parallel survey of equatorial amphibians and reptiles in Africa and south America. *In : Tropical ecosystem in Africa and South America: a comparative review*. Meggers, B.J., Ayensu, E.S. et Duckworth, W.D. (eds), Smithsonian Inst. Press, Washington DC, pp. 259-266.
- Lawson, G.W. (1996) The Guinea-Congo lowland rain forest: an overview. *Proc. R. Soc. Edinburgh*, **104B**, 5-13.
- Le Quesne, W.J. (1969) A method of selection of characters in numerical taxonomy. *Syst. Zool.*, **18**, 201-205.
- Le Turdu, C., Tiercelin, J.-J., Gibert, E. *et al.* (1999) The Ziway-Shala lake basin system, Main Ethiopian Rift: Influence of volcanism, tectonics, and climatic forcing on basin formation and sedimentation. *Palaeogeogr. Palaeoclimatol. Palaeoecol.*, **150 (3-4)**, 135-177.
- Leakey, M. (1988) Fossil evidence for the evolution of guenons (chapter I.1). *In : A Primate radiation: evolutionary biology of the African guenons*. Gautier-Hion, A., Bourlière, F., Gautier, J.-P. et Kingdon, J. (eds), Cambridge University Press, Cambridge, pp. 1-12.
- Lecointre, G., Philippe, H., Van Le, H.L. et Le Guyader, H. (1993) Species sampling has a major impact on phylogenetic inference. *Mol. Phylogenet. Evol.*, **2 (3)**, 205-224.
- Lepersonne, J. (1974) *Notice explicative de la carte géologique du Zaïre au 1/2.000.000*. Direction de la Géologie, Bruxelles.
- Lernould, J-M (1988) Classification and geographical distribution of guenons: a review. *In : A primate radiation: evolutionary biology of the African guenons*. Gautier-Hion, A., Bourlière, F., Gautier, J.P. et Kingdon, J. (eds), Cambridge University Press, Cambridge, pp. 54-78.
- Letouzey, R. (1968) *Etude phytogéographique du Cameroun*. Lechevalier, Paris.
- Li, W.H., Gouy, M., Sharp, P.M., O'huigin, C. et Yang, Y.W. (1990) Molecular phylogeny of Rodentia, Lagomorpha, Primates, Artiodactyla and Carnivora and molecular clocks. *Proc. Nat. Acad. Sci. USA*, **87**, 6703-6707.
- Linder, H.P. (2001) Plant diversity and endemism in sub-Saharan tropical Africa. *J. Biogeog.*, **28**, 169-182.
- Linnaeus, C. (1781) *Selected Dissertations from the Amoenitates Academicæ, a supplement to Mr. Stillingfleet's Tracts relating to Natural History* 2. On the increase of the Habitable Earth, London.
- Lloyd, D.G. et Calder, V.L. (1991) Multi-residue gaps, a class of molecular characters with exceptional reliability for phylogenetic analyses. *Syst. Biol.*, **42**, 576-581.
- Louette, M. (1992) Barriers, contact zones and subspeciation in Central Equatorial Africa. *Bull. B.O.C. Centenary Suppl.*, **112A**, 209-216.
- Lougheed, S.C., Freeland, J.R., Handford, P. et Boag, P.T. (2000) A molecular phylogeny of warbling-finches (Poospiza): Paraphyly in a neotropical emberizid genus. *Mol. Phylogenet. Evol.*, **17**, 367-378.

- Lynch, J.D. (1988) Refugia (chapter 10). *In* : Analytical biogeography: an integrated approach to the study of animal and plant distributions. Myers, A.A. et Giller, P.S. (eds), Chapman & Hall, London, pp. 311-342.
- Mac Arthur, R.H. (1942) *Geographical Ecology*. Harper & Row, New York.
- Mac Arthur, R.H. et Wilson, E.O. (1967) *The theory of island biogeography*. Princeton University Press, Princeton, NJ.
- Maddalena, T., Van der Straeten, E., Ntahuga, L. et Sparti, A. (1989) Nouvelles données et caryotypes des rongeurs du Burundi. *Rev. Suisse Zool.*, **96** (3), 939-948.
- Maddison, W.P. (1997) Gene trees in species trees. *Syst. Biol.*, **46**, 523-536.
- Maddison, W.P. et Maddison, D.R. (1996) *MacClade: Analysis of phylogeny and character evolution, Version 3.06*. Sinauer, Sunderland, MA. .
- Maglio, J. (1978) Patterns of faunal evolution (Chapter 30). *In* : Evolution of African Mammals. Maglio, J. et Cook, H.B.S. (eds), Harvard University Press, Cambridge, pp. 603-619.
- Maley, J. (1990) Conclusions de la quatrième partie. Synthèse sur le domaine forestier africain au quaternaire récent. *In* : Paysages quaternaires de l'Afrique centrale atlantique. Lanfranchi, R. et Schwartz, D. (eds), ORSTOM, Paris, pp. 383-389.
- Maley, J. (1996) The African rain forest - main characteristics of changes in vegetation and climate from the Upper Cretaceous to the Quaternary. *Proc. R. Soc. Edinburgh*, **104B**, 31-73.
- Maley, J. et Brenac, P. (1998) Vegetation dynamics, palaeoenvironments and climatic changes in the forests of western cameroon during the last 28,000 years BP. *Rev. Palaeobot. Palynol.*, **99** (2), 157-187.
- Marchant, S. (1954) The relation of the southern Nigerian avifauna to those of Upper and Lower Guinea. *Ibis*, **96**, 371-379.
- Martin, A.P., Naylor, G.J.P. et Palumbi, S.R. (1992) Rates of mitochondrial DNA evolution in sharks are slow compared with mammals. *Nature*, **357**, 153-155.
- Martin, A.P. et Palumbi, S.R. (1993) Body Size, metabolic rate, generation time, and the molecular clock. *Proc. Natl. Acad. Sci. USA*, **90** (9), 4087-4091.
- Matthee, C.A. et Robinson, T.J. (1999) Cytochrome b phylogeny of the family bovidae: Resolution within the Alcelaphini, Antilopini, Neotragini, and Tragelaphini. *Mol. Phylogenet. Evol.*, **12**, 31-46.
- Matthew, W.D. (1915) Climate and evolution. *Ann. N.Y. Acad. Sci.*, **24**, 171-318.
- Matthey, R. (1963) La formule chromosomique chez sept espèces et sous-espèces de Murinae africains. *Mammalia*, **27** (2), 157-176.
- Mayr, E. (1942) *Systematics and the origin of species*. Columbia University Press, New York.
- Mayr, E. (1963) *Animal species and evolution*. Harvard University Press, Cambridge.
- Mayr, E. et O'Hara, R. (1986) The biogeographic evidence supporting the pleistocene forest refuge hypothesis. *Evolution*, **40** (1), 55-67.
- McCune, A.R. et Lovejoy, N.R. (1998) The relative rate of sympatric and allopatric speciation in fishes: tests using DNA sequence divergence between sister species and among clades. *In* : Endless forms, species and speciation. Howard, D.J. et Berlocher, S.H. (eds), Oxford University Press, Inc. New York, pp. 172-185.
- Meester, J. et Setzer, H.W. (1971) *The Mammals of Africa. An identification manual*. Smithsonian Institution Press, Washington D.C.
- Melnick, D.J. et Hoelzer, G.A. (1992) Differences in male and female macaque dispersal lead to contrasting distributions of nuclear and mitochondrial DNA variation. *Int. J. Primatology*, **13** (4), 379-393.

- Melnick, D.J. et Hoelzer, G.A. (1993) What is mtDNA good for in the study of primate evolution?. *Evolutionary Anthropology*, **2**, 191-199.
- Michaux, J. et Catzeflis, F. (2000) The bushlike radiation of muroid rodents is exemplified by the molecular phylogeny of the LCAT nuclear gene. *Mol. Phylogenet. Evol.*, **17**, 280-293.
- Miller, J.S. et Wenzel, J.W. (1995) Ecological characters and phylogeny. *Annu. Rev. Entomol.*, **40**, 389-415.
- Mindell, D.P. et Honeycutt, R.L. (1990) Ribosomal RNA in vertebrates: evolution and phylogenetic applications. *Annu. Rev. Ecol. Syst.*, **21**, 541-566.
- Misonne, X. (1969) African and Indo-Australian Muridae: Evolutionary trends. *Ann. Mus. R. Afr. Centr., Tervuren, Belgique, Ser. IN-8, Sci. Zool.*, **172**, 1-219.
- Moore, W.S. (1995) Inferring phylogenies from mt-DNA variation: mitochondrial-gene versus nuclear-gene tree. *Evolution*, **49** (4), 718-726.
- Moreau, R.E. (1966) The Ethiopian lowland forest bird fauna. In : The bird faunas of Africa and its islands. Academic Press, London & New York, pp. 159-188.
- Moreau, R.E. (1969) Climatic changes and the distribution of forest vertebrates in west Africa. *J. Zool. Lond.*, **158**, 39-61.
- Moritz, C. et Faith, D.P. (1998) Comparative phylogeography and the identification of genetically divergent areas for conservation. *Mol. Ecol.*, **7**, 419-429.
- Moritz, C., Patton, J.L., Schneider, C.J. et Smith, T.B. (2000) Diversification of rainforest faunas: An integrated molecular approach. *Annu. Rev. Ecol. Syst.*, **31**, 533-563.
- Morrone, J.J. et Crisci, J.V. (1995) Historical biogeography: introduction to methods. *Annu. Rev. Ecol. Syst.*, **26**, 373-401.
- Mpounza, M. et Samba-Kimbata, M.J. (1990) Aperçu sur le climat de l'Afrique Centrale Occidentale. In : Paysages quaternaires de l'Afrique centrale atlantique. Lanfranchi, R. et Schwartz, D. (eds), ORSTOM, Paris, pp. 31-41.
- Müller, J., Irion, G., Nunes de Mello, J. et Junk, W. (1995) Hydrological changes of the amazon during the last glacial-interglacial cycle in central Amazonia. *Naturwissenschaften*, **82**, 232-235.
- Musser, G.G. et Carleton, M.D. (1993) Order Rodentia. In : Mammal species of the world: a taxonomic and geographic reference. Wilson, D.E. et Reeder, D.M. (eds), Smithsonian Institution Press, Washington D.C, pp. 598-599.
- Myers, A.A. et Giller, P.S. (1988) Analytical biogeography: an integrated approach to the study of animal and plant distributions, Chapman & Hall, London.
- Napier, P.H. (1981) *Catalogue of primates in the British Museum (Natural History) and elsewhere in the British Isles. Part 2. Family Cercopithecidae, subfamily Cercopithecinae*. British Museum Natural History, London.
- Needleman, S.B. et Wunsch, C.D. (1970) A general method applicable to the search for similarities in the amino acid sequences of two proteins. *J. Mol. Biol.*, **48**, 443-453.
- Nei, M. (1987) Molecular evolutionary genetics. Columbia university press, New York.
- Nelson, G. et Platnick, N. (1981) *Systematics and biogeography. Cladistics and vicariance*. Columbia University Press, New York.
- Nichols, R. (2001) Gene trees and species trees are not the same. *T.R.E.E.*, **16**, 358-364.
- Nixon, K.C. et Carpenter, J.M. (1996) On simultaneous analysis. *Cladistics*, **12**, 221-241.
- Oates, J.F. (1985) The Nigerian guenon, *Cercopithecus erythrogaster* : Ecological, behavioral, systematic and historical observations. *Folia Primatol.*, **45**, 25-43.

- Oates, J.F. (1988) The distribution of *Cercopithecus* monkeys in West African forests. *In* : A Primate radiation: evolutionary biology of the African guenons. Gautier-Hion, A., Bourlière, F., Gautier, J.-P. et Kingdon, J. (eds), Cambridge University Press, Cambridge, pp. 79-103.
- Ohler, A. (1996) Systematics, morphometrics and biogeography of the genus *Aubria* (Ranidae, Pyxicephalinae). *Alytes*, **13**(4), 141-166.
- Omi, G. et Kato, Y. (1982) Paleolithic implements in Cameroun. *In* : Geomorphology and environmental changes in Cameroon forest and savanna. Kadomera, H. (ed.), Hokkaido Univ., Sapporo, pp. 105-133.
- Omland, K.E. (1999) The assumptions and challenges of ancestral state reconstructions, *Syst. Biol.*, **48** (3), 6604-611.
- Onderdonk, D.A. et Chapman, C.A. (2000) Coping with forest fragmentation: The primates of Kibale National Park, Uganda. *Int. J. Primatol.*, **21**, 587-611.
- Oosterbroek, P. (1994) Biodiversity of the Mediterranean region. *In*: Systematics and conservation evaluation, Forey, P.L., Humphries, C.J. et Vane-Wright, R.I. (eds), Clarendon Press, Oxford, pp. 289-307.
- Osliisly, R. (1995) The middle Ogooue Valley, Gabon, from 3.500 years BP: cultural changes and palaeoclimatic implications. *In* : International Conference on "The growth of farming communities in Africa from the Equator southwards", Br. Inst. in Eastern Africa (ed.), Newnham College, Cambridge.
- Page, R.D.M. (1990) Component analysis: a valliant failure? *Cladistics*, **6** (2), 119-136.
- Page, R.D.M. (1993) *Component version 2.0*, programme distribué par l'auteur, disponible à l'adresse : <http://taxonomy.zoology.gla.ac.uk/rod/cpw.html>
- Page, R.D.M. (1994a) Maps between trees and cladistic analysis of historical associations among genes, organism and areas. *Syst. Biol.*, **43**, 58-57.
- Page, R.D.M. (1994b) Parallel phylogenies: reconstructing the history of host parasite assemblages *Cladistics*, **10**, 155-173.
- Page, R.D.M. (1995) *Treemap version 1.0*, programme distribué par l'auteur, disponible à l'adresse : <http://taxonomy.zoology.gla.ac.uk/rod/treemap.html>
- Pagel, M. (1999a) *Discrete*, programme distribué par l'auteur. University of Reading, UK.
- Pagel, M. (1999b) The maximum likelihood approach to reconstructing ancestral character states of discrete characters on phylogenies, *Syst. Biol.*, **48** (3), 612-622.
- Pagney, P. (1976) *Les climats de la terre*. Masson, Paris.
- Palumbi, S.R., Martin, A., Romano, S., McMillan, W.O., Stice, L. et Grabowski, G. (1991). *The simple fool's guide to PCR*. University of Hawaii Press, Honolulu.
- Pamilo, P. et Nei, M. (1988) Relationships between gene trees and species trees. *Mol. Biol. Evol.*, **5**, 568-583.
- Partridge, T.C., Wood, B.A. et Demenocal, P.B. (1995) The influence of global climatic change and regional uplift on large-mammalian evolution in East and Southern Africa. *In* Paleoclimate and Evolution with Emphasis on Human Origins. Vrba, E. S., Denton, G. H., Partridge, T. C. et Burckle, L. H. (eds). Yale University Press, London, pp. 331-355.
- Patterson, C., Williams, D.M. et Humphries, C.J. (1993) Congruence between molecular and morphological phylogenies. *Annu. Rev. Ecol. Syst.*, **24**, 153-188.
- Patton, J.L., Da Silva, M.N.F. et Malcolm, J.R. (1994) Gene genealogy and differentiation among arboreal spiny rats (Rodentia: Echimyidae) of the Amazon basin: a test of the riverine barrier hypothesis. *Evolution*, **48** (4), 1314-1323.
- Patton, J.L., Da Silva, M.N.F. et Malcolm, J.R. (1996) Hierarchical genetic structure and gene flow in three sympatric species of Amazonian rodents. *Molec. Ecol.*, **5**, 229-238.

- Patton, J.L. et da Silva, M.N.F. (1998) Rivers, refuges, and ridges: the geography of speciation of Amazonian mammals. *In* : Endless forms, species and speciation. Howard, D.J. et Berlocher, S.H. (eds), Oxford University Press, Inc. New York, pp. 202-213.
- Patton, J.L., Da Silva, M.N.F. et Malcolm, J.R. (2000) Mammals of the Rio Jurua and the evolutionary and ecological diversification of Amazonia. *Bull. Am. Mus. Nat. Hist.*, **244**, 13-306.
- Penny, D. (1982) Towards a basis for classification: The incompleteness of distance measures, incompatibility analysis and phenetic classification. *J. Theor. Biol.*, **96**, 129-142.
- Peres, C.A., Patton, J.L., et Da Silva, M.N.F. (1996) Riverine barriers and gene flow in Amazonian saddle-back tamarins. *Folia Primatol.*, **67**, 113-124.
- Petit, M. (1990) Les grands traits morphologiques de l'Afrique Centrale Atlantique. *In* : Paysages quaternaires de l'Afrique centrale atlantique. Lanfranchi, R. et Schwartz, D. (eds), ORSTOM, Paris, pp. 20-30.
- Philippe, H., Lecointre, G., Vàn Le, H.L. et Le Guyader, H. (1996) A critical study of homoplasy in molecular data with the use of a morphologically based cladogram, and its consequences for character weighting. *Mol. Biol. Evol.*, **13**, 1174-1186.
- Pickford, M et Senut, B. (1988) Habitat and locomotion in Miocene cercopithecoids. *In* : A Primate radiation: evolutionary biology of the African guenons. Gautier-Hion, A., Bourlière, F., Gautier, J.-P. et Kingdon, J. (eds), Cambridge University Press, Cambridge, pp. 35-53.
- Platnick, N.I. et Nelson, G. (1978) A method of analysis for historical biogeography. *Syst. Zool.*, **27**, 1-16.
- Pollock, D.D. et Goldstein, D.B. (1995) A comparison of two methods for constructing evolutionary distances from a weighted contribution of transition and transversion differences. *Mol. Biol. Evol.*, **12**, 713-717.
- Posada, D. et Crandall, K.A. (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics*, **14** (9), 817-818.
- Posada, D., Crandall, K.A. et Templeton, A.R. (2000) GeoDis: A program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. *Mol. Ecol.*, **9** (4), 487-488.
- Prigogine, A. (1988) Speciation pattern of birds in the Central African forest refugia and their relationships with other refugia. *Acta XIX Congressus Int. Ornithol.*, **2**, 144-157.
- Prim, R.C. (1957) Shortest connections networks and some generalizations. *Bell Syst. Tech. J.*, **36**, 1389-1401.
- Pusey, A.E. et Packer, C. (1987) Dispersal and philopatry. *In* : Primate societies. Cheney, D.L. Seyfarth, R.M. Wrangham, R.W. et Struhsaker, T.T. (eds), University of Chicago Press, Chicago, pp. 250-266.
- Quinn, J.F. et Harrison, S.P. (1988) Effects of habitat fragmentation and isolation on species richness: evidence from biogeographical patterns. *Oecologia*, **75**, 132-140.
- Ragan, M.A. (1992) Phylogenetic inference based on matrix representation of trees. *Mol. Phylogenet. Evol.*, **1** (1), 53-58.
- Reitsma, J.M. (1988) *Végétation forestière du Gabon*. Tropenbos Foundation, London.
- Renaud, S. (1999) Size and shape variability in relation to species differences and climatic gradients in the African rodent *Oenomys*. *J. Biogeogr.*, **26**, 857-865.
- Reumer, J.W.F. (1989) Speciation and evolution in the Soricidae (Mammalia: Insectivora) in relation with the paleoclimate. *Rev. Suisse Zool.*, **96** (1), 81-90.
- Reyes, A., Gissi, C., Pesole, G. et Saccone, C. (1998) Asymmetrical directional mutation pressure in the mitochondrial genome of mammals. *Mol. Biol. Evol.*, **15** (8), 957-966.

- Riddle, B.R. (1995) Molecular biogeography in the pocket mice (*Perognathus* and *Chaetopidus*) and grasshopper mice (*Onychomys*): the late cenozoic development of a North American aridlands rodent guild. *J. Mammal.*, **76** (2), 238-301.
- Ridley, M. (1997) *Evolution biologique*. DeBoeck Université, Paris.
- Robbins, C.B. (1978) The Dahomey Gap - A reevaluation of its significance as a faunal barrier to West African high forest Mammals. *Bull. Carnegie Mus. Nat. Hist.*, **6**, 168-174.
- Robbins, L.W., Choate, J.R. et Robbins, R.L. (1980) Nongeographic and interspecific variation in four species of *Hylomyscus* (Rodentia: Muridae) in Southern Cameroon. *Ann. Carn. Museum*, **49** (2), 31-48.
- Robert, C. (1980) Climats et courants cénozoïques dans l'Atlantique Sud d'après l'étude des minéraux argileux (legs 3, 39 et 40). *Oceanologica Acta*, **3**, 369-376.
- Robert, C. et Chamley, H. (1987) Cenozoic evolution of continental humidity and paleoenvironment, deduced from kaolinite content of oceanic sediments. *Palaeogeogr. Palaeoclimatol. Palaeoecol.*, **60**, 171-187.
- Robinson-Rechavi, M. et Huchon, D. (2000) RRTree: Relative-Rate Tests between groups of sequences on a phylogenetic tree. *Bioinformatics*, **16** (3), 296-297.
- Ronquist, F. (1996) DIVA version 1.1. Computer program and manual available by anonymous ftp from Uppsala university (ftp.uu.se or ftp.systbot.uu.se). .
- Rosen, D.E. (1978) Vicariant patterns and historical explanation in biogeography. *Syst. Zool.*, **27**, 159-188.
- Rosen, D.E. (1985) Geological hierarchies and biogeographic congruence in the Caribbean. *Ann. Missouri Bot. Gdn.*, **72**, 636-659.
- Rosevear, D.R. (1969) The rodents of West Africa. British Museum, London.
- Rüber, L., Verheyen, E. et Meyer, A. (1999) Replicated evolution of trophic specializations in an endemic cichlid fish lineage from Lake Tanganyika. *Proc. Natl. Acad. Sci. USA*, **96**, 10230-10235.
- Ruddiman, W., Sarnthein, M. et al. (1986) Paleoclimate linkage between high and low latitudes. *Nature*, **322**, 211-212.
- Ruddiman, W.F. et Kutzbach, J.E. (1989) Forcing of late Cenozoic northern hemisphere climate by plateau uplift in southeast Asia and the American southwest. *J. Geophys. Research*, **94**, 18409-18427.
- Ruedi, M., Smith, M.F. et Patton, J.L. (1997) Phylogenetic evidence of mitochondrial DNA introgression among pocket gophers in New Mexico (family Geomyidae). *Mol. Ecol.*, **6**, 453-462.
- Rzhetsky, A. et Nei, M. (1993) Theoretical foundation of the minimum-evolution method of phylogenetic inference. *Mol. Biol. Evol.*, **10**, 1073-1095.
- Saitou, N. et Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.*, **4**, 406-425.
- Sambrook, J., Fritsch, E.F. et Maniatis, T. (1989) Molecular cloning: A laboratory manual. Second Edition, Chapter 14 : In vitro amplification of DNA by the polymerase chain reaction, pp. 14.1-14.35. Cold Spring Harbour Laboratory Press.
- Sayer, J.A., Harcourt, C.S. et Collins, N.M. (1992) *The conservation atlas of tropical forest Africa*. Macmillan, London.
- Sbisa, E., Tanzariello, F., Reyes, A., Pesole, G. et Saccone, C. (1997) Mammalian mitochondrial D-loop region structural analysis: identification of new conserved sequences and their functional and evolutionary implications. *Gene*, **205** (1-2), 125-140.
- Schlitter, D.A., Hutterer, R. et Maddalena, T. (1999) New karyotypes of shrews (Mammalia : Soricidae) from Cameroon and Somalia. *Ann. Carn. Museum*, **68**, 1-13.

- Schluter, D. (1998) Ecological causes of speciation. *In* : Endless forms, species and speciation. Howard, D.J. et Berlocher, S.H. (eds), Oxford University Press, New York, pp. 114-129.
- Schneider, C.J., Cunningham, M. et Moritz, C. (1998) Comparative phylogeography and the history of vertebrates endemic to the Wet Tropics rainforest of Australia. *Mol. Ecol.*, **7**, 487-498.
- Schneider, C.J. et Moritz, C. (1999) Rainforest refugia and evolution in Australia's wet tropics. *Proc. R. Soc. Lond. Ser. B*, **226**, 191-196.
- Schneider, C.J., Smith, T.B., Larison, B. et Moritz, C. (1999) A test of alternative models of diversification in tropical rainforests: Ecological gradients versus rainforest refugia. *Proc. Natl. Acad. Sci. USA*, **96**, 13869-13873.
- Schwartz, D. (1990) La couverture pédologique de l'Afrique Centrale Atlantique. *In* : Paysages quaternaires de l'Afrique centrale atlantique. Lanfranchi, R. et Schwartz, D. (eds), ORSTOM, Paris, pp. 52-59.
- Shackleton, N.J., Backman, J., Zimmerman, H. *et al.* (1984) Oxygen isotope calibration of the onset of ice-rafting and history of glaciation in the North Atlantic region. *Nature*, **307**, 620-623.
- Shannon, C.E. (1948) The mathematical theory of communication. *Bell System Technology J.*, **27**, 379-423.
- Siddall, M.E. (1998) Success of parsimony in the four-taxon case: long-branch repulsion by likelihood in the Farris zone. *Cladistics*, **14**, 209-220.
- Simmons, M.P. et Ochoterena, H. (2000) Gaps as characters in sequence-based phylogenetic analyses. *Syst. Biol.*, **49**, 369-381.
- Simmons, M.P., Ochoterena, H. et Carr, T.G. (2001) Incorporation, relative homoplasy, and effect of gap characters in sequence-based phylogenetic analyses. *Syst. Biol.*, **50** (3), 454-462.
- Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H. et Flook, P. (1994) Evolution, weighting and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Ann. Entomol. Soc. Am.*, **87** (6), 651-701.
- Slowinski, J.B. (2001) Molecular polytomies. *Mol. Phylogenet. Evol.*, **19** (1), 114-120.
- Smith, A.B. (1994) Rooting molecular trees: problems and strategies. *Biol. J. Linn. Soc.*, **51**, 279-292.
- Smith, A.B. (1989) RNA sequence data in phylogenetic reconstruction: testing the limits of its resolution. *Cladistics*, **5**, 321-344.
- Smith, M.F., Kelt, D.A. et Patton, J.L. (2001) Testing models of diversification in mice in the *Abrothrix olivaceus/xanthorhinus* complex in Chile and Argentina. *Mol. Ecol.*, **10**, 397-405.
- Smith, T.B., Wayne, R.K., Girman, D.J. et Bruford, M.W. (1997) A role for ecotones in generating rainforest biodiversity. *Science*, **276**, 1855-1857.
- Smith, T.B., Holder, K., Girman, D., O'Keefe, K., Larison, B. et Chan, Y. (2000) Comparative avian phylogeography of Cameroon and Equatorial Guinea mountains: implications for conservation. *Mol. Ecol.*, **9**, 1505-1516.
- Sosef, M.S.M. (1994) Refuge begonias. Taxonomy, phylogeny and historical biogeography of *Begonia* section *Loasibegonia* and section *Scutobegonia* in relation to glacial rain forest refuges in Africa. PhD thesis, Wageningen Agric. Univ. Papers, **94-1**, 1-306.
- Sosef, M.S.M. (1996) Begonias and African rain forest refuges: general aspects and recent progress. *In* : The biodiversity of African plants. Van der Maesen, L.J.G., Van der Burgt, X.M., Van Medenbach de Roy, J.M. (eds), Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 602-611.
- Sourdis, J. et Nei, M. (1988) Relative efficiencies of the maximum parsimony and distance matrix methods in obtaining the correct phylogenetic tree. *Mol. Biol. Evol.*, **5**, 298-311.
- Stewart, C.-B. (1993) The powers and pitfalls of parsimony. *Nature*, **361**, 603-607.

- Steward, D.T. et Baker, A.J. (1994) Patterns of sequence variation in the mitochondrial D-loop region of Shrews. *Mol. Biol. Evol.*, **11** (1), 9-21.
- Stoneking, M. et Soodyall, H. (1996) Human evolution and the mitochondrial genome. *Curr. Opin. Genet. Develop.*, **6** (6), 731-736.
- Strimmer, K. et von Haeseler, A. (1996) Quartet puzzling: a quartet maximum likelihood method for reconstructing tree topologies. *Mol. Biol. Evol.*, **13**, 964-969.
- Strimmer, K. et von Haeseler, A. (1997) Likelihood-mapping: a simple method to visualize phylogenetic content of a sequence alignment. *PNAS (USA)*, **94**, 6815-6819.
- Struhsaker, T.T., Butynski, T.M. et Lwanga, J.S. (1988) Hybridization between redbellied (Cercopithecus ascanius schmidtii) and blue (C. mitis stuhlmanni) monkeys in the Kibale forest, Uganda. In : A primate radiation. Evolutionary biology of the african guenons. Gautier-Hion, A., Bourlière, F., Gautier, J.P. et Kingdon, J. (eds), Cambridge University Press, Cambridge, pp. 477-497.
- Sullivan, J., Arellano, E. et Rogers, D.S. (2000) Comparative phylogeography of Mesoamerican highland rodents: concerted versus independent response to past climatic fluctuations. *Am. Nat.*, **155** (6), 755-768.
- Swofford, D. (2000) *PAUP: Phylogenetic analysis using parsimony. Version 4b3a*. Smithsonian Institution, Washington DC.
- Swofford, D.L. et Olsen, G. (1990) Phylogeny reconstruction. In : Molecular systematics. Hillis, D.M. et Moritz, C. (eds). Sinauer Ass., Sunderland, Massachusetts, pp. 411-501.
- Taberlet, P. (1998) Biodiversity at the intraspecific level: The comparative phylogeographic approach. *J. Biotech.*, **64**, 91-100.
- Taberlet, P., Fumagalli, L., Wust Saucy, A.G. et Cosson, J.F. (1998) Comparative phylogeography and post-glacial colonization routes in Europe. *Mol. Ecol.*, **7** (4), 453-464.
- Tajima, F. et Nei, M. (1984) Estimation of evolutionary distance between nucleotide sequences. *Mol. Biol. Evol.*, **1**, 269-285.
- Takezaki, N., Rzhetsky, A. et Nei, M. (1995) Phylogenetic test of the molecular clock and linearized trees. *Mol. Biol. Evol.*, **12**, 823-833.
- Tamura, K. (1992) Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G+C-content biases. *Mol. Biol. Evol.*, **9**, 678-687.
- Tamura, K. et Nei, M. (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.*, **10**, 512-526.
- Templeton, A.R. (1983) Convergent evolution and non parametric inferences from restriction fragment and DNA sequence data. In : Statistical analysis of DNA sequence data. Weir, B. (ed.). Marcel Dekker, New York, pp. 151-179.
- Templeton, A.R., Boerwinckle, E. et Sing, C.F. (1987) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. I. Basic theory and an analysis of alcohol dehydrogenase activity in drosophila. *Genetics*, **117**, 343-351.
- Templeton, A.R., Crandall, K.A. et Sing, C.F. (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics*, **132**, 619-633.
- Templeton, A.R. et Sing, C.F. (1993) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. IV. Nested analysis with cladogram uncertainty and recombination. *Genetics*, **134**, 659-669.
- Templeton, A.R. (1998) Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. *Mol. Ecol.*, **7**, 381-397.

- Thomas, H. (1985) The Early and Middle Miocene land connection of the Afro-Arabian plate and Asia: A major event for hominoid dispersal. *In* : *Ancestors: The Hard Evidence*. Delson, E. (ed.), A. R. Liss, New York, pp. 42-50.
- Thompson, J.D., Higgins, D.G. et Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, **22**, 4673-4680.
- Tong, H. et Jaeger, J.-J. (1993) Muroid rodents from the middle Miocene Fort Ternan locality (Kenya) and their contribution to the phylogeny of muroids. *Paleontographica*, **229**, 51-73.
- Tuffley, C. et Steel, M. (1997) Links between maximum likelihood and maximum parsimony under a simple model of site substitution. *Bull. Math. Bio.*, **59**, 581-607.
- Tutin, C. et White, L. (1999) The recent evolutionary past of primate communities: Likely environmental impacts during the past three millennia. *In* : *Primate Communities*. Fleagle, J.G., Janson, C. et Reed, K.E. (eds), Cambridge University Press, Cambridge, pp. 220-236.
- Underwood, E., Itoua, I., Olson, D., Dinerstein, E., Loucks, C. et Wettengel, W. (1998) Terrestrial Ecoregions of Africa. WWF-US Conservation Science Program.
- Uzell, T. et Corbin, K.W. (1971) Fitting discrete probability distribution to evolutionary events. *Science*, **172**, 1089-1096.
- Van Andel, T.H. et Tzedakis, P.C. (1996). Palaeolithic landscapes of Europe and environs: 150,000-25,000 years ago: an overview. *Quater. Sci. Rev.*, **15**, 481-500.
- van de Koppel, J. et Prins, H.H.T. (1998) The importance of herbivore interactions for the dynamics of African savanna woodlands: an hypothesis. *J. Trop. Ecol.*, **14**, 565-576.
- van der Kuyl, A., Kuiken, C.L. Dekker, J.T. et Goudsmit, J. (1995) Phylogeny of African monkeys based upon mitochondrial 12S rRNA sequences. *J. Mol. Evol.*, **40**, 173-180.
- van Der Kuyl, A.C., Dekker, J.T. et Goudsmit, J. (2000a) Primate genus *Miopithecus*: evidence for the existence of species and subspecies of dwarf guenons based on cellular and endogenous viral sequences. *Mol. Phylogenet. Evol.*, **14** (3), 403-413.
- van der Kuyl, A.C., Van Gennep, D.R.O., Dekker, J.T. et Goudsmit, J. (2000b) Routine DNA analysis based on 12S rRNA gene sequencing as a tool in the management of captive primates. *J. Med. Primatol.*, **29** (5), 309-317.
- Van der Straeten, E. (1984) Etude biométrique des genres *Dephomys* et *Stochomys* avec quelques notes taxonomiques (Mammalia, Muridae). *Rev. Zool. Afr.*, **98**, 771-798.
- Van der Straeten, E. et Verheyen, W.N. (1978) Taxonomical notes on the West-African *Myomys* with the description of *Myomys derooi* (Mammalia-Muridae). *Z. Säugetierkunde*, **43**, 31-41.
- Van Neer, W. (1990) Les faunes de vertébrés quaternaires en Afrique Centrale. *In* : *Paysages quaternaires de l'Afrique centrale atlantique*. Lanfranchi, R. et Schwartz, D. (eds), ORSTOM, Paris, pp. 195-220.
- van Noten, F., Cahen, D. et de Maret, P. (1980) L'Afrique centrale. *In* : *Histoire générale de l'Afrique*, vol. 2, "Afrique ancienne". UNESCO, Paris, pp. 673-693.
- Van Rompaey, R.S.A.R. (1993) Forest gradients in West Africa: a spatial gradient analysis. Doctoral thesis, Wageningen Agricultural Univ.
- Van Rompaey, R.S.A.R. (1994) Locating rain forest refugia in Liberia and western Côte d'Ivoire using biogeographic evidence from mainly endemic Caesalpiniaceae species: coastal versus sub-montane position. 14th Congress Association Etude Taxonomie et Flore Afrique Tropicale. Univ. Wageningen, p. 108.
- Van Valen, L. (1973) A new evolutionary law. *Evol. Theory*, **1**, 1-33.

- van Veller, M.G.P., Kornet, D.J. et Zandee, M. (2000) Methods in vicariance biogeography: Assessment of the implementations of assumptions 0, 1, and 2. *Cladistics*, **16**, 319-345.
- Vanzolini, P.E. et Williams, E.E. (1981) The vanishing refuge: a mechanism for ecogeographic speciation. *Pap. Avulsos. Zool.*, **34**, 251-255.
- Verheyen, W.N. (1962) *Contribution à la craniologie comparée des primates. Les genres Colobus Illiger 1811 et Cercopithecus Linné 1758*. Ann. Mus. Afr. Centr., Tervuren, Belgique.
- Verheyen, W.N., Colyn, M. et Hulselmans, J. (1996) Re-evaluation of the *Lophuromys nudicaudus* Heller, 1911 species-complex with a description of a new species from Zaïre (Muridae - Rodentia). *Bull. Inst. R. Sci. Nat. Belg. - Biologie*, **66**, 241-273.
- Verheyen, W.N., Hulselmans, J., Colyn, M. et Hutterer, R. (1997) Systematics and zoogeography of the small mammal fauna of Cameroun: description of two new *Lophuromys* (Rodentia: muridae) endemic to Mount Cameroun and Mount Oku. *Bull. Inst. R. Sci. Nat. Belg. - Biologie*, **57**, 163-186.
- Verheyen, W., Dierckx, T. et Hulselmans, J. (2000) The brushed-furred rats of Angola and southern Congo: description of a new taxon of the *Lophuromys sikapusi* species complex. *Bull. Inst. R. Sci. Nat. Belg. - Biologie*, **70**, 253-267.
- Via, S. (2001) Sympatric speciation in animals: the ugly duckling grows up. *T.R.E.E.*, **16**, 381-390.
- Viegas-Péquignot, E., Dutrillaux, B., Prod'Homme, M. et Petter, F. (1983) Chromosomal phylogeny of Muridae : a study of 10 genera. *Cytogenet. Cell Genet.*, **35**, 269-278.
- Vincens, A., Schwartz, D., Bertaux, J., Hilaire, E. et de Namur, C. (1998) Late Holocene climatic changes in Western Africa inferred from pollen from Lake Sinnda, Southern Congo. *Quaternary Research*, **50**, 34-45.
- Vitt, L.J., Caldwell, J.P., Zani, P.A. et Titus, T.A. (1997) The role of habitat shift in the evolution of lizard morphology: evidence from tropical *Tropidurus*. *Proc. Natl. Acad. Sci. U.S.A.*, **94**, 3828-3832.
- Voelker, G. (1999) Dispersal, vicariance, and clocks: Historical biogeography and speciation in a cosmopolitan passerine genus (*Anthus* : motacillidae). *Evolution*, **53**, 1536-1552.
- Vrba, E.S. (1985) Environment and evolution: alternative causes of the temporal distribution of evolutionary events. *S. Afr. J. Sci.*, **81**, 229-236.
- Vrba, E.S. (1995) The fossil record of African antelopes (Mammalia, Bovidae) in relation to human evolution and paleoclimate. In : Paleoclimate and evolution with emphasis on Human origins. Vrba, E.S. , Denton, G.H., Partridge, T.C. et Burckle, L.H. (eds), Yale University, Haven & London, pp. 385-424.
- Walker, S.E., Jolly, C.J. et Oates, J.F. (1988) Electrophoretic evidence for the evolutionary position of *Cercopithecus erythrogaster* and *Cercopithecus erythrotis*. *Folia Primatol.*, **51**, 220-226.
- Wallace, A.R. (1852) On the monkeys of the Amazon. *Proc. Zool. Soc. London*, **20**, 107-110.
- Wallace, A.R. (1876) *The Geographical Distribution of Animals*. Hafner, New York.
- Walsh, P.S., Metzger, D.A. et Higuchi, R. (1991) Chelex 100 as a medium for simple extraction of DNA for PCR based typing from forensic material. *Biotechniques*, **10**, 506-513.
- Welsh, J. et McClelland, M. (1991) Genomic fingerprints produced by PCR with consensus tRNA gene primers. *Nucl. Acid. Res.*, **19** (4), 861.
- Wenzel, J.W. (1992) Behavioural homology and phylogeny. *Annu. Rev. Ecol. Syst.*, **23**, 361-381.
- Wesselman, H.B. (1985) Fossil micromammals as indicators of climate change about 2.4 million years ago in the Omo Valley. Ethiopia. *South Afr. J. Sci.*, **81**, 260-261.
- Wheeler, W. (1996) Optimization alignment the end of multiple alignment in phylogenetics? *Cladistics*, **12**, 1-9.

- Wheeler, W.C. et Honeycutt, R.L. (1988) Paired sequence differences in ribosomal RNAs: evolutionary and phylogenetic implications. *Mol. Biol. Evol.*, **5** (1), 90-96.
- Wheeler, W.C. et Gladstein, D.S (1994) MALIGN: a multiple sequence alignment program. *J. Hered.*, **85**, 417.
- White, F. (1983) *The vegetation of Africa: a descriptive memoir to accompany the UNESCO/AETFAT/UNSO vegetation map of Africa*. UNESCO, Paris, France.
- White, F. (1993) Refuge theory, ice-age aridity and the history of tropical biotas: an essay in plant geography. *Fragm. Flor. Geobot.*, **2**, 385-409.
- White, L. (1992) Vegetation history and logging disturbance: effects on rain forest mammals in the Lope reserve, Gabon (with special emphasis on elephants and apes), PhD Thesis, University of Edimburgh.
- White, M.J.D. (1968) Models of speciation. *Science*, **159**, 1065-1070.
- Wieczorek, A.M., Drewes, R.C. et Channing, A. (2000) Biogeography and evolutionary history of *Hyperolius* species: application of molecular phylogeny. *J. Biogeog.*, **27**, 1231-1243.
- Wiley, E.O. (1987) Methods in vicariance biogeography. In : Systematics and Evolution: A Matter of Diversity. Hovenkamp, P. (ed.), Institute of Systematic Botany, Utrecht, pp. 283-306.
- Wiley, E.O. (1988) Vicariance biogeography. *Annu. Rev. Ecol. Syst.*, **19**, 513-542.
- Williams, S.E. et Pearson, R.G. (1997) Historical rainforest contractions, localized extinctions and patterns of vertebrate endemism in the rainforests of Australia's wet tropics. *Proc. R. Soc. Lond. Ser. B*, **264**, 709-716.
- Wilson, J.W. (1974) Analytical zoogeography of North American mammals. *Evolution*, **28**, 124-140.
- Wilson, K.H. (1995) Molecular biology as a tool for taxonomy. *Clinical Infectious Diseases*, **20**, 117-121.
- Winnepeninckx, B., Backeljau, T. et Dewachter, R. (1994) Small rRibosomal subunit RNA and the phylogeny of Mollusca. *Nautilus*, **108** (suppl.), 98-110.
- Wolsan, M. et Hutterer, R. (1998). Appendix : a list of the living species of shrews. In : Evolution of Shrews. Wojcik, J.M. et Wolsan, M. (eds), Mammal Research Institute, Polish Academy of Science, Bialowieza, pp. 425-448.
- Wolsan, M. et Wójcik, J.M. (1998) *Evolution of Shrews*. Mammal Research Institute, Polish Academy of Sciences, Bialowieza.
- Yang, Z. (1996) Phylogenetic analysis using parsimony and likelihood methods. *J. Mol. Evol.*, **42**, 294-307.
- Yang, Z., Goldman, N. et Friday, A.E. (1994) Comparison of models for nucleotide substitution used in maximum likelihood phylogenetic estimation. *Mol. Biol. Evol.*, **11**, 316-324.
- Yemane, K., Bonnefille, R. et Faure, H. (1985) Palaeoclimatic and tectonic implications of Neogene microflora from the Northwestern Ethiopian highlands. *Nature*, **318**, 653-656.
- Zandee, M. et Roos, M. (1987) Component compatibility in historical biogeography. *Cladistics*, **3** (4), 305-332.
- Zink, R.M. (1996) Comparative phylogeography in North American birds. *Evolution*, **50**, 308-317.
- Zink, R.M. et Slowinski, J.B. (1995) Evidence from molecular systematics for decreased avian diversification in the Pleistocene epoch. *Proc. Nat. Acad. Sci. USA*, **92**, 5832-5835.
- Zink, R.M., Blackwell Rago, R.C. et Ronquist, F. (2000) The shifting roles of dispersal and vicariance in biogeography. *Proc. R. Soc. London, Ser. B*, **267**, 497-503.

GLOSSAIRE

(les mots soulignés renvoient aux définitions de ce glossaire)

- **adaptation** : modifications évolutives conduisant à une meilleure survie et efficacité de reproduction, ou caractère morphologique, physiologique, développemental ou comportemental augmentant la survie ou le succès reproducteur d'un organisme.
- **allopatrie** : séparation spatiale, disjonction des aires de distribution.
- **apomorphie** : identité entre caractères homologues d'origine phylétique récente (= trait partagé dérivé).
- **barrière biogéographique** : obstruction physique à la dispersion d'un organisme ou au flux génique entre populations (exemple : une montagne, une rivière) ou limite climatique ou écologique empêchant la dispersion ou l'établissement d'une espèce (exemple : variations de température ou de pluviosité, limite forêt/savane).
- **biogéographie** : étude des distributions géographiques des organismes, en fonction de leur habitat (biogéographie écologique) et des facteurs historiques et biologiques qui les ont produits (biogéographie historique).
- **B.P.** (before present) : avant le temps présent.
- **cladogramme** : arbre phylogénétique qui décrit l'histoire évolutive supposée d'un taxon, sans indication du taux de changements entre lignées.
- **classification phylogénétique** : classification qui reflète les relations phylogénétiques entre taxons. Pour certains, tous les taxons doivent être strictement monophylétiques, ce qui présente un risque de création artificielle d'espèces par simple soucis d'homogénéisation des variations phylogénétiques. Pour d'autres, il peut y avoir paraphylie : si un taxon diverge fortement par rapport à l'ancêtre commun, il peut avoir un rang différent des autres taxons avec lesquels il se groupe à un nœud donné.
- **endémique** : caractéristique d'une espèce indigène dont la distribution est limitée à une certaine région géographique.
- **espèce** : la définition de l'espèce au sens biologique n'étant pas toujours applicable, plusieurs concepts d'espèces ont été définis :
 - *Concept biologique* : groupe d'individus potentiellement ou effectivement inter-féconds ayant atteint des trajectoires évolutives indépendantes par évolution de mécanismes d'isolement reproducteur ;
 - *Concept basé sur la reconnaissance du partenaire sexuel* : groupe d'individus partageant un mécanisme de fécondation commun via la reconnaissance du partenaire sexuel ;
 - *Concept de cohésion* : population la plus étendue d'individus pouvant assurer une cohésion phénotypique via des mécanismes de cohésion intrinsèques génétiques ou écologiques ;
 - *Concept phylogénétique* : plus petit groupe d'individu identifiable par une combinaison unique d'états de caractères pour lequel ancêtre et descendants forment une même lignée ;
 - *Concept généalogique* : groupe exclusif d'organismes historiquement monophylétique qui sont plus proches entre eux que de tout autre organisme situé en dehors du groupe ;

- *Concept évolutif* : lignée de population ancestrales et descendantes qui maintient son identité par rapport aux autres lignées et possède ses propres tendances évolutives et historiques ;
- *Concept de groupement génotypique* : groupe d'individus entre lesquels il n'existe pas ou peu d'intermédiaires.
- **évolution** : modification de caractères héritable dans une population, au cours du temps.
- **indel** : insertion ou délétion de nucléotides dans une séquence nucléotidique.
- **Holocène** : époque géologique du quaternaire correspondant aux derniers 10.000 ans B.P.
- **homologie** : similarité héritée d'un ancêtre commun.
- **homoplasie** : identité entre caractères homologues résultant d'une convergence et non d'un héritage par rapport à un ancêtre commun.
- **horloge moléculaire** : hypothèse qui suppose que les mutations ponctuelles se produisent à un intervalle de temps suffisamment régulier pour permettre la datation de dichotomies phylogénétiques, donc qu'il existe une relation directe entre taux de divergence moléculaire et temps écoulé depuis un événement de séparation entre lignées.
- **MA** : million d'années.
- **Miocène** : époque géologique du tertiaire comprise entre environ 26 et 5 MA B.P.
- **monophylétique** : se dit d'un groupe de taxons dérivés d'un seul et même ancêtre commun qui sont plus proches entre eux que de tout autre taxon situé en dehors du groupe, c'est-à-dire ayant un ancêtre commun qui n'est commun qu'à eux.
- **phylogénie** : histoire évolutive d'un groupe ou d'une lignée.
- **phylogéographie** : c'est l'étude de groupes monophylétiques dans le temps et dans l'espace. Elle recherche les principes et processus qui gouvernent la distribution géographique des lignées, au niveau intra et inter-spécifique.
- **phylogramme** : arbre phylogénétique dont les longueurs de branches indiquent les distances évolutives estimées entre lignées.
- **Pléistocène** : époque géologique du quaternaire comprise entre environ 1,8 MA et 10.000 ans B.P.
- **Pliocène** : époque géologique du tertiaire comprise entre environ 5 et 1,8 MA B.P.
- **polyphylétique** : se dit d'un groupe de taxons qui dérivent de plusieurs ancêtres distincts.
- **radiation adaptative** : diversification évolutive d'un taxon en plusieurs formes ayant des caractéristiques écologiques et des modes de vie diversifiés, généralement sur une courte période de temps.
- **refuge** : région géographique ayant des limites déterminées qui aurait permis la survie des individus

qu'elle a abrités durant des périodes de perturbations environnementales.

- **région faunique** : région caractérisée par une faune distincte.
- **spéciation** : formation de nouvelles espèces. Différents degrés de spéciation peuvent être distingués :
 - **sous-espèces** : discernables selon des critères morphologiques
 - **semi-espèces** : suffisamment différentes pour être considérées comme des espèces, mais pouvant s'hybrider
 - **allo-espèces** : lorsque des espèces pouvant s'hybrider s'excluent mutuellement et sont alors parapatricques
 - **espèce** : lorsque la reproduction est possible et donne une descendance fertile
 - **super-espèce** : complexe de semi-espèces ou d'allo-espèces
 - **groupe-espèce** : groupe monophylétique d'espèces étroitement apparentées mais ne pouvant s'hybrider. Elles peuvent donc être sympatriques.
- **substitution** : mutation ponctuelle conduisant au remplacement d'une base par une autre dans une molécule d'ADN.
- **sympatrie** : recouvrement des aires de distribution.
- **symplesiomorphie** : identité entre caractères homologues d'origine phylétique ancienne (= trait partagé ancestral).
- **syntopie** : caractéristique des taxons occupant le même habitat.
- **taxon** : groupe taxinomique de n'importe quel rang.
- **transition** : substitution d'un nucléotide purique par un autre nucléotide purique (échange entre A et G), ou d'un nucléotide pyrimidique par un autre nucléotide pyrimidique (échange entre C et T).
- **transversion** : substitution d'un nucléotide purique (A ou G) par un nucléotide pyrimidique (C ou T), ou *vice versa*.
- **zone de contact secondaire** : zone où deux espèces, semi-espèces ou sous-espèces se rencontrent après une période de séparation, avec intergradation ou recouvrement limité.

ANNEXE 1 : ECHELLE DES TEMPS GEOLOGIQUES

Ere	Période	Epoque	Limite inférieure (en MA)
Cénozoïque	Quaternaire	Holocène	0,01
		Pléistocène	1.8
	Tertiaire	Pliocène	5
	- Néogène	Miocène	26
		Oligocène	37
	- Paléogène	Eocène	53
		Paléocène	65
Mésozoïque	Crétacé		136
(= Secondaire)	Jurassique		190
	Trias		225

ANNEXE 2 : ECHANTILLONS DEMANDES A L'EXTERIEUR

Nous tenons à remercier les chercheurs et directeurs de Parcs Zoologiques qui nous ont aidé en nous fournissant du matériel biologique (biopsies ou poils).

B. Beer (Laboratory of Molecular Microbiology, Rockville, USA) : 1 échantillon primate

F. Catzeflis (CNRS Montpellier, France) : 12 échantillons soit 8 musaraignes et 4 rongeurs

R. Fons (CNRS Banyuls, France) : 1 échantillon musaraigne

P. Formenty (OMS, Abidjan, Côte d'Ivoire) : 4 échantillons soit 3 rongeurs et 1 primate

R. Hutterer (Museum A. Koenig, Bonn, Allemagne) : 13 échantillons musaraignes

J.C. Kerbis Peterhans (Field Museum of Natural History, Chicago, USA) : 4 échantillons musaraignes

H. Leirs (RUCA, Anvers, Belgique) : 3 échantillons rongeurs

J.-M. Lermould et P. Moisson (Zoo Mulhouse, France) : 2 échantillons primates

M. Ruedi (IZEA, Lausanne, Suisse) : 6 échantillons musaraignes

D. Schittler (Carnegie Museum of Natural History, Pittsburgh, USA) : 8 échantillons soit 6 rongeurs et 2 musaraignes

N. et P. Thivillon. (Espace zoologique de Saint Martin la Plaine, France) : 2 échantillons primates

W. Verheyen (RUCA, Anvers, Belgique) : 26 échantillons rongeurs

P. Vogel (IZEA, Lausanne, Suisse) : 1 échantillon musaraigne

ANNEXE 3 : PROTOCOLES D'EXTRACTION, AMPLIFICATION ET SEQUENÇAGE

EXTRACTION

A. Méthode Phénol-Chloroforme (Meyer)

1) Préparer le **tampon d'extraction** (100 ml): 1 ml NaCl 5 M

10 ml SDS 20%

86 ml H₂O

1 ml Tris 1 M

2 ml EDTA 0.5 M

puis la **solution d'extraction** (par éch.):

0,5 ml tampon extraction

20 µl protéinase K (20 mg/ml H₂O)

10 µl DTT 1 M (154,2 mg Dithiothréitol/ml H₂O)

2) Préparer les échantillons:

Vider chaque tube sur du papier absorbant et prélever un fragment de 1-2 mm³.

Mettre dans 1 ml H₂O, écraser un peu et laisser ½ h à température ambiante dans l'agitateur rotatif, puis pipeter la phase aqueuse et l'éliminer.

3) Suspendre dans 500 µl de solution d'extraction → au moins 1 h à 37°C avec agitation, jusqu'à dissolution complète.

4) Extraction:

- Extraction avec 0,5 ml de phénol (PCI 25:24:1) pour faire précipiter les sucres et les lipides dans le phénol et débris cellulaires à l'interphase.

Agiter 5 min, puis centrifuger 5 min à 6.000 tours.

Reprendre le surnageant et recommencer 2 fois.

- Elimination du phénol avec 0,5 ml de chloroforme:isoamylalcool 24:1.

Même protocole, à recommencer 2 fois.

5) Reprendre le surnageant avec 1 ml d'éthanol 100%.

Placer à -20°C, une nuit minimum.

6) Centrifuger 30 min à 4°C et 14.000 rpm, pour faire précipiter l'ADN.

Vider d'un trait la phase liquide. Ajouter 500 µl d'éthanol à 70% et recommencer.

Vider l'alcool d'un trait.

7) Sécher sous la pompe à eau, pendant au moins 2 heures.

8) Diluer dans 200 µl d'eau et agiter.

9) Placer soit au frigo à 4°C, soit au bain-marie à 65°C pour homogénéiser la solution avant PCR.

NB: Toutes les manipulations se font avec des gants, du matériel et de l'eau filtrée stériles.

B. Méthode CTAB

- 1) Préparer le **tampon extraction** (100 ml) :
 - 1 ml NaCl 5 M
 - 10 ml SDS 20%
 - 86 ml H₂O
 - 1 ml Tris 1 M
 - 2 ml EDTA 0.5 M
- 2) Préparer les échantillons comme pour la méthode Phénol-chloroforme.
- 3) Suspendre dans 100 µl de solution d'extraction.
Ecraser avec un piston stérile → au moins 1 h à 56°, avec agitation.
La solution doit devenir visqueuse et se clarifier.
- 4) Extraction au chloroforme:isoamylalcool 24:1 (1 volume), pour éliminer les débris cellulaires.
Agiter 20 min, puis centrifuger 5 min à 6.000 tours.
- 5) Reprendre le surnageant avec 2 volumes d'éthanol 100%.

Placer à -20°C, une nuit minimum.

- 6) Centrifuger 3 min à 6.000 tours, pour faire précipiter l'ADN.
Pipeter l'éthanol ou vider d'un trait la phase liquide.
Laver au moins 3 fois à l'éthanol 70%.
- 7) Sécher à l'étuve à 56°C, pendant au moins 1 heure.
- 8) Diluer dans 100 à 1000 µl d'eau et agiter.
- 9) Placer soit au frigo à 4°C, soit à l'étuve 56°C avec agitation pour PCR.

C. Méthode Chelex

- 1) Préparer les échantillons comme ci-dessus, mais prélever au plus 1 mm³.
- 2) Suspendre dans 500 µl de solution Chelex à 5% (75 mg pour 1.5 ml H₂O stérile).
Placer au moins 30 min à 56°C, avec agitation.
- 3) Porter à 95°C pendant 10 min.
- 4) Vortexer puis centrifuger 3 min à 12.000 rpm. C'est prêt!

AMPLIFICATION DE L'ADN PAR PCR

A. Amplification

1) Préparer les **solutions pour PCR** (1 pour chaque séquence) :

(Nb échantillons + 1 témoin) x 2,5 µl dNTP (1/50e)
2,5 µl Taq buffer 10x
5 µl amorce 1
5 µl amorce 2
8 µl H₂O stérile
0,13 µl Taq polymérase.

NB: Dégeler les produits dans la main. Ne sortir la TaqPol qu'au dernier moment.

2) Homogénéiser la solution et en mettre 20 µl dans chaque tube.

3) Ajouter 2 µl d'échantillon, puis 1 goutte d'huile minérale par tube.

Centrifuger brièvement (QR).

4) Lancer la PCR (94°C: 60», --°C: 60», 72°C: 1'30», puis retour à 4°C).

Après quelques heures...

B. Vérification des résultats

1) Préparer le gel d'électrophorèse (à 1.4%) : 2,1 g d'agarose

150 ml TAE

10 µl EtBr (très nocif)

Faire fondre au micro-onde avant d'ajouter le bromure d'ethidium (EtBr).

Couler le (ou les) gel(s).

2) Mélanger 2 µl de gel-loading solution avec 5 µl de chaque échantillon + avec le ladder.

3) Quand le gel a pris, ôter le scotch et le placer dans la cuve.

Remplir la cuve de TAE 1x.

Charger 6 µl par puits et faire migrer 40 min, à 80V et 150W.

4) Lire le gel sur la table UV.

Si tout va bien: faire une seconde amplification.

Sinon, retour à la case départ.

PURIFICATION

1) Réunir le produit des 2 PCR, compléter à 100 µl avec H₂O, puis QR.

Placer 1 filtre par échantillon dans le plateau supérieur et boucher les puits vacants.

2) Remplir les puits avec: 20 µl H₂O

200 µl solution tampon

50 µl résine préalablement agitée

100 µl échantillon en faisant attention à ne pas pipeter l'huile.

Attendre 3 min, puis fixer 1 min à 0,25 bars.

3) Mettre 400 µl de solution à base d'éthanol dans chaque puits. Eluer 1 min à 0,25 bars.

Recommencer et éluer 5 min. Vérifier que les puits sont bien secs.

4) Ajouter 40 µl d'eau stérile par puits. Attendre 5 min et mettre en place le plateau de récupération contenant un tube de collecte par échantillon.

Eluer 1 min à 0,25 bars, puis quelques secondes à 0,5 bars.

5) Stocker au froid.

SEQUENCAGE (ALF)

A - Amplification avec une amorce marquée et les ddNTP :

1) Diluer chaque amorce, en multipliant les quantités qui suivent par le nombre d'échantillons concernés:

1 µl d'amorce diluée au 1/15^e

13 µl H₂O stérile

Homogénéiser et déposer 3,4 µl au fond de chaque tube.

2) Ajouter 2,6 µl de l'échantillon concerné. Tapoter pour faire descendre au fond du tube.

3) Ajouter une goutte d'huile minérale dans chaque tube.

4) Ajouter 2 µl de l'un des 4 mix contenant l'enzyme et les di-nucléotides (A, C, G, ou T), QR.

4) Lancer l'amplification (94°C: 36», --°C: 36», 72°C:1'20», puis 5 min à 72°C et retour à 4°C).

Un peu plus tard... : Arrêter la réaction avec 5 µl de solution "Stop" par tube, QR.

B - Séquençage :

1) Nettoyer soigneusement les plaques du séquenceur à l'eau distillée, plusieurs fois.
Rincer à l'éthanol 70% pour enlever les traces grasses. Rincer encore et bien essuyer.

2) Préparer la solution de **Bind-Silane** :
2 ml ethanol 100%
7,5 µl Bind-Silane (très nocif)
0,5 ml acide acétique 10%

Travailler sous la hotte. Répartir immédiatement à l'aide d'une pipette pasteur sur les 5 cm supérieurs de chaque plaque, afin de consolider les puits.

Frotter et rincer à l'eau, puis à l'éthanol. Bien sécher.

3) Assembler le cadre.

4) Préparer le **gel de poly-acrylamide** :
25.2 g d'urée à diluer dans H₂O
12 ml de Ready-Mix IEF (30% acrylamide)
7.2 ml de TBE 10x (→ TBE 1,2%)
H₂O qsp 60 ml

Tout mélanger dans un becher 80 ml, sous la hotte (toxique avant polymérisation).

Diluer l'urée à l'aide de l'agitateur magnétique, puis ajuster à 60 ml avec H₂O.

5) Filtrer sur pompe à eau dans un buchner.

Remuer pour éliminer les bulles d'air. Verser doucement dans une bouteille en plastique.

6) Ajouter les agents polymérisants :
52.5 µl de TEMED
210 µl de persulfate d'ammonium 10 %

Rouler la bouteille. Couler le gel, et éliminer les bulles. Placer le peigne.

7) Diluer le tampon de migration **TBE** à 0.6 % (120 ml de TBE 10x, H₂O qsp 2 l)

8) Préparer le séquenceur ALF :

Enduire le boudin de la cuve supérieure de Silicone.

Placer la cuve inférieure, puis le gel. Connecter les tuyaux d'arrivée d'eau, et le contact électrique. Centrer.

Fermer les taquets et les visser. Fermer la protection laser.

Vérifier les conditions de migration (500 min, 45°C...). Contrôler l'eau, la lumière et le laser.

Remplir le bac supérieur avec le tampon TBE et ôter le peigne.

Rincer les puits plusieurs fois à l'aide de la seringue, pour éliminer les cristaux d'urée.

9) Charger les échantillons :

Dénaturer les échantillons (Stoppés) 5 min à 95°C, les placer dans de la glace.

Déposer au plus 7 µl d'échantillon par puits.

10) Fermer l'appareil. Lancer le Séquençage.

Résultats le lendemain.

ANNEXE 4 : AMORCES UTILISEES

ADN mitochondrial

ARN-r 16s

16Sar (L), 5'-CGCCTGTTTAACAAAAACAT-3' (Palumbi *et al.*, 1991)

16Sbr (H), 5'-CCGGTTTGAACCTCAGATCATGT-3' (Palumbi *et al.*, 1991)

16S-Hm, 5'-AGATCACGTAGGACTTTAAT-3' (nouvelle amorce → 1^{ère} moitié de l'ARNr 16s)

16S-Hc, 5'-TTATACTCCGAGGTCAC-3' (nouvelle amorce → 1^{er} quart de l'ARNr 16s)

Températures - PCR : 46 puis 49°C

- Séquençage : 49 puis 52°C (brin L) et 55°C (brin H)

Cytochrome b

L14724, 5'-CGAAGCTTGATATGAAAAACCAT-3' (Irwin, 1991)

L14723, 5'-ACCAATGACATGAAAAATCATCGTT-3' (C. Fadda, comm. pers., pour *Acomys*)

L14841m, 5'-GCACCATCAAATATTTTCATCATGATGAAA-3' (nouvelle amorce adaptée du L14841 de Kocher *et al.*, 1989, utilisé par Fumagalli *et al.*, 1999 sur *Sorex* → fragment d'environ 1000 pb)

L15408, 5'-ATAGACAAATCCCATTCCA-34 (C. Fadda, comm. pers., pour *Acomys*)

Cyt-Hm, 5'-TGGCTTTGTCTACTGAGAAGCCYCCTCA -3' (nouvelle amorce → 1^{ère} moitié du cyt b)

H15553, 5'-AGGCAAATAGGAAATATCATTCTGGT-3' (C. Fadda, comm. pers., pour *Acomys*)

H15915, 5'-AACTGCAGTCATCTCCGGTTTACAAGAC-3' (Irwin, 1991 → cyt b complet, utilisé par Fumagalli *et al.*, 1999 sur *Sorex*)

H15915b, 5'-TCTCCATTTCTGGTTTACAAGAC-3' (C. Fadda, comm. pers., pour *Acomys*)

Températures - PCR : 50°C

- Séquençage : 52°C

D-loop

L15926, 5'-TCAAAGCTTACACCAGTCTTGTAAC-3' (Kocher *et al.*, 1989 ; utilisé par Steward et Baker, 1994 sur des musaraignes)

DloopLm, 5'-CCCCACCATCAGCACCCAAAGC-3' (nouvelle amorce → 1^{er} tiers Dloop)

H16498, 5'-CCTGAAGTAGGAACCAGATG-3' (Kocher *et al.*, 1989 xx)

Températures - PCR : 49°C

- Séquençage : 52°C (brin L) et 55°C (brin H)

ARN-r 12s

12Sai 5'-AACTAGGATTAGATACCCTATTAT-3' (Kessing *et al.*, 1989, modifié du L1091 de Kocher *et al.*, 1989)

12Sbi 5'-AAGAGCGACGGGCGATGTGT-3' (Kessing *et al.*, 1989, modifié du H1478 de Kocher *et al.*, 1989)

Températures - PCR : 49°C

- Séquençage : 49°C (brin L) et 55°C (brin H)

ADN nucléaire

ARN-r 18s

4F18S (L), 5'- CTGGTTGATYCTGCCAGT -3' (Winne penninckx *et al.*, 1994)

1R18S (F), 5'- ATTACCGCGGCTGCTGGC -3' (Winne penninckx *et al.*, 1994)

Températures - PCR : 50°C et touch-down PCR (55 à 48°C)

- Séquençage : 52°C

ITS1

ITS1L, 5'-TCCGTAGGTGAACCTGCGGAAGGAT-3' (I. Schöen, comm. pers.)

ITS4F, 5'-TCCTCCGCTTATTGATATGC-3' (I. Schöen, comm. pers.)

Températures - PCR : 50°C et 46°C pendant 2 min

ANNEXE 5 : CLEF D'INTERPRETATION POUR L'ANALYSE DES GROUPEMENTS EMBOITES (16/02/2001)

Inference Key for the Nested Haplotype Tree Analysis of Geographical Distances

Start with haplotypes nested within a 1-step clade:

1. Are there any significant values for D_c , D_n , or I-T within the clade?
 - NO - the null hypothesis of no geographical association of haplotypes cannot be rejected (either panmixia in sexual populations, extensive dispersal in non-sexual populations, small sample size, or inadequate geographical sampling). Move on to another clade at the same or higher level.
 - YES - Go to step 2.
2. Is at least one of the following conditions satisfied?
 - a. The D_c 's for some tips are significantly small and the D_c 's for the interiors are significantly large or non-significant.
 - b. The D_c 's for tips are significantly small or non-significant and the D_c 's for some but *not* all of the interiors are significantly small.
 - c. The I-T D_c is significantly large.
 - NO - Go to step 11.
 - YES - Go to step 3.
 - Tip/Interior Status Cannot be Determined - **Inconclusive Outcome.**
3. Are any D_n and/or I-T D_n values significantly reversed from the D_c values, and/or do one or more tip clades show significantly large D_n 's or interior clades significantly small D_n 's or I-T significantly small D_n with the corresponding D_c values being non-significant?
 - NO - Go to step 4.
 - YES - Go to step 5.
4. Do the clades (or 2 or more subsets of them) with restricted geographical distributions have ranges that are completely or mostly non-overlapping with the other clades in the nested group (particularly interiors), and does the pattern of restricted ranges represent a break or reversal from lower level trends within the nested series (applicable to higher-level clades only)?
 - NO - **Restricted Gene Flow with Isolation by Distance (Restricted Dispersal by Distance in Non-sexual species).** This inference is strengthened if the clades with restricted distributions are found in diverse locations, if the union of their ranges roughly corresponds to the range of one or more clades (usually interiors) within the same nested group (applicable only to nesting clades with many clade members or to the highest level clades regardless of number), and if the D_c values increase and become more geographically widespread with increasing clade level within a nested series (applicable to lower level clades only).
 - YES - Go to step 9.
5. Do the clades (or 2 or more subsets of them) with restricted geographical distributions have ranges that are completely or mostly non-overlapping with the other clades in the nested group (particularly interiors), and does the pattern of restricted ranges represent a break or reversal from lower level trends within the nested series (applicable to higher-level clades only)?
 - NO - Go to step 6.
 - YES - Go to step 15.

6. Do clades (or haplotypes within them) with significant reversals or significant D_n values without significant D_c values define two or more geographically concordant subsets, or are they geographically concordant with other haplotypes/clades showing similar distance patterns?

- No - Go to step 7.
- YES - Go to step 13.
- **TOO FEW CLADES (< 2) TO DETERMINE CONCORDANCE - Insufficient Genetic Resolution to Discriminate between Range Expansion/Colonization and Restricted Dispersal/Gene Flow** - Proceed to step 7 to determine if the geographical sampling is sufficient to discriminate between short versus long distance movement.

7. Are the clades with significantly large D_n 's (or tip clades in general when D_n for I-T is significantly small) separated from the other clades by intermediate geographical areas that were sampled?

- NO - Go to step 8.
- YES - **Restricted Gene Flow/Dispersal but with some Long Distance Dispersal.**

8. Is the species absent in the non-sampled areas?

- NO - **Sampling Design Inadequate to Discriminate between Isolation by Distance (Short Distance Movements) versus Long Distance Dispersal**
- YES - **Restricted Gene Flow/Dispersal but with some Long Distance Dispersal over Intermediate Areas not Occupied by the Species.**

9. Are the different geographical clade ranges identified in step 4 separated by areas that have not been sampled?

- NO - **Past Fragmentation.** (If inferred at a high clade level, additional confirmation occurs if the clades displaying restricted by at least partially non-overlapping distributions are mutationally connected to one another by a larger than average number of steps.)
- YES - Go to step 10.

10. Is the species absent in the non-sampled areas?

- NO - **Geographical Sampling Scheme Inadequate to Discriminate Between Fragmentation and Isolation By Distance.**
- YES - **Allopatric Fragmentation.** (If inferred at a high clade level, additional confirmation occurs if the clades displaying restricted by at least partially nonoverlapping distributions are mutationally connected to one another by a larger than average number of steps.)

11. Is at least one of the following conditions satisfied?

- a. The D_c value(s) for some tip clade(s) is/are significantly large.
 - b. The D_c value(s) for all interior(s) is/are significantly small.
 - c. The I-T D_c is significantly small.
- NO - Go to step 17
 - YES - **Range Expansion**, go to step 12.

12. Are the D_n and/or I-T D_n values significantly reversed from the D_c values?

- NO - **Contiguous Range Expansion.**
- YES - Go to step 13.

13. Are the clades with significantly large D_n 's (or tip clades in general when D_n for I-T is significantly small) separated from the geographical center of the other clades by intermediate

geographical areas that were sampled?

- NO - Go to step 14.
- YES - **Long Distance Colonization.**

14. Is the species absent in the non-sampled areas?

- NO - **Sampling Design Inadequate to Discriminate between Contiguous Range Expansion and Long Distance Colonization.**
- YES - **Long Distance Colonization.**

15. Are the different geographical clade ranges identified in step 5 separated by areas that have not been sampled?

- NO - **Past Fragmentation.** (If inferred at a high clade level, additional confirmation occurs if the clades displaying restricted by at least partially non-overlapping distributions are mutationally connected to one another by a larger than average number of steps.)
- YES - Go to step 16.

16. Is the species absent in the non-sampled areas?

- NO - Go to step 18.
- YES - **Allopatric Fragmentation.** (If inferred at a high clade level, additional confirmation occurs if the clades displaying restricted by at least partially nonoverlapping distributions are mutationally connected to one another by a larger than average number of steps.)

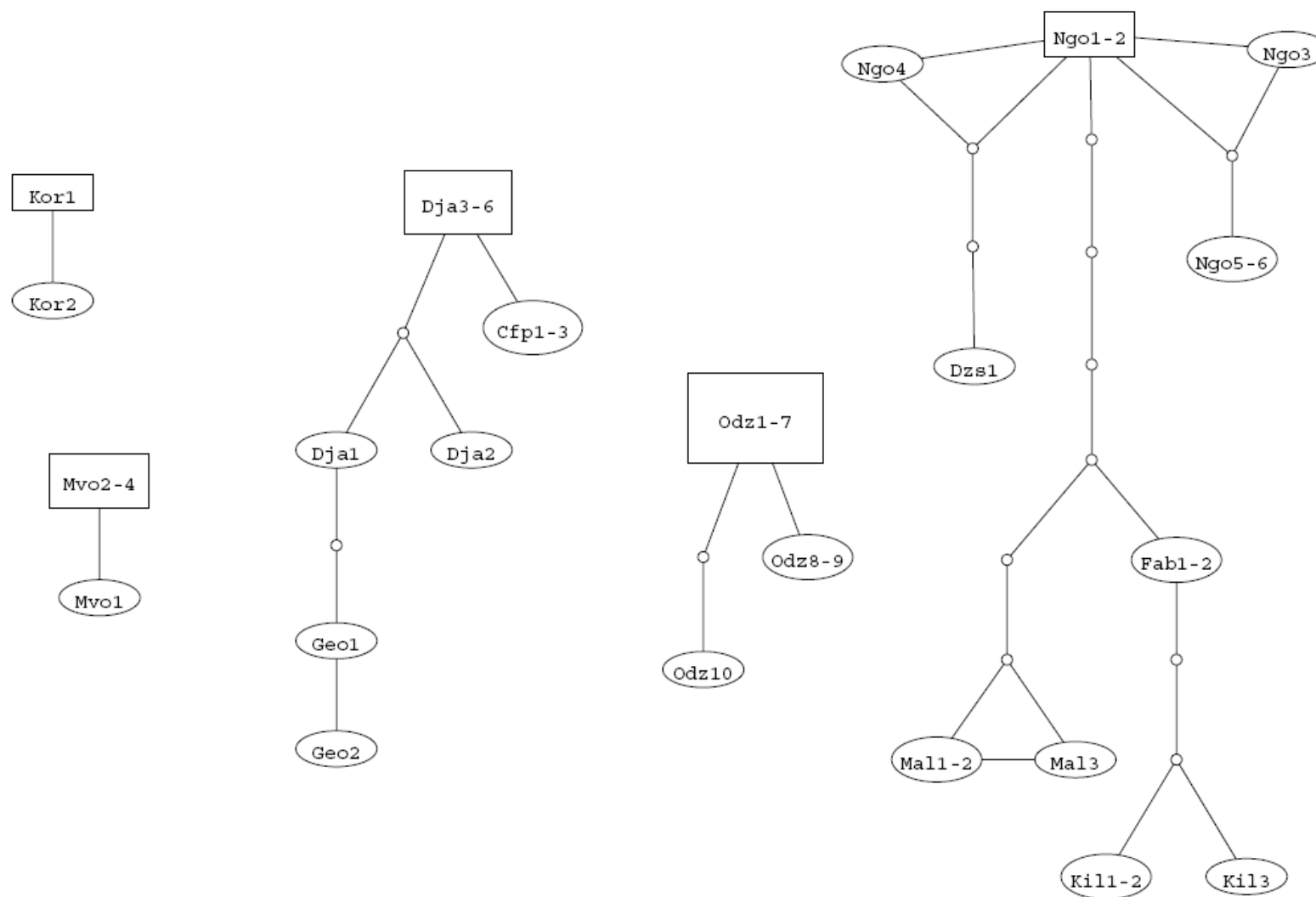
17. Are the D_n values for tip or some (but not all) interior clades significantly small, or the D_n for one or more interior clades significantly large, or is the I-T D_n value significantly large.

- NO - **Inconclusive Outcome.**
- YES - Go to step 4.

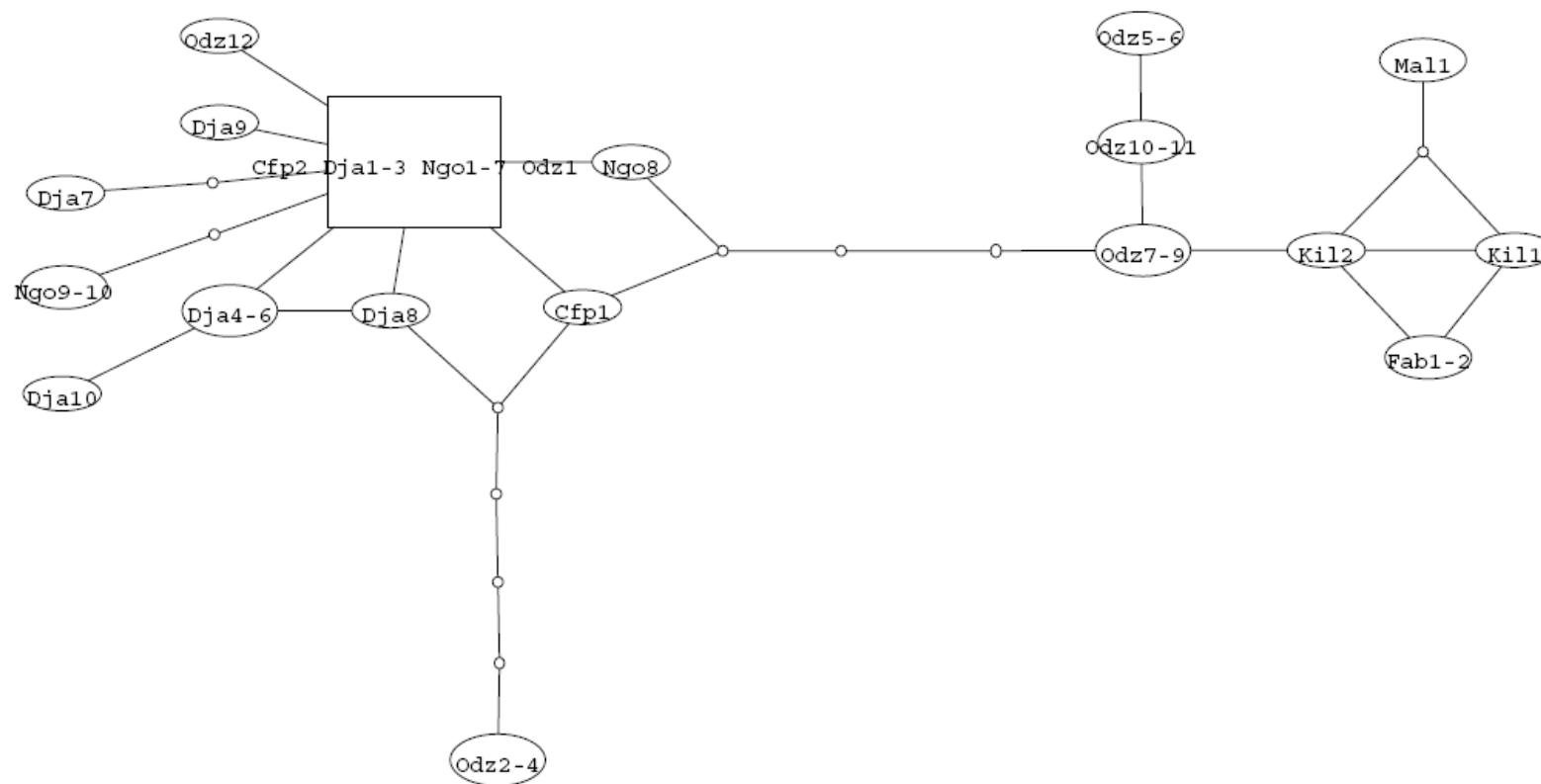
18. Are the clades found in the different geographical locations separated by a branch length with a larger than average number of mutational steps.

- NO - **Geographical Sampling Scheme Inadequate to Discriminate Between Fragmentation, Range Expansion, and Isolation By Distance.**
- YES - **Geographical Sampling Scheme Inadequate to Discriminate Between Fragmentation and Isolation By Distance.**

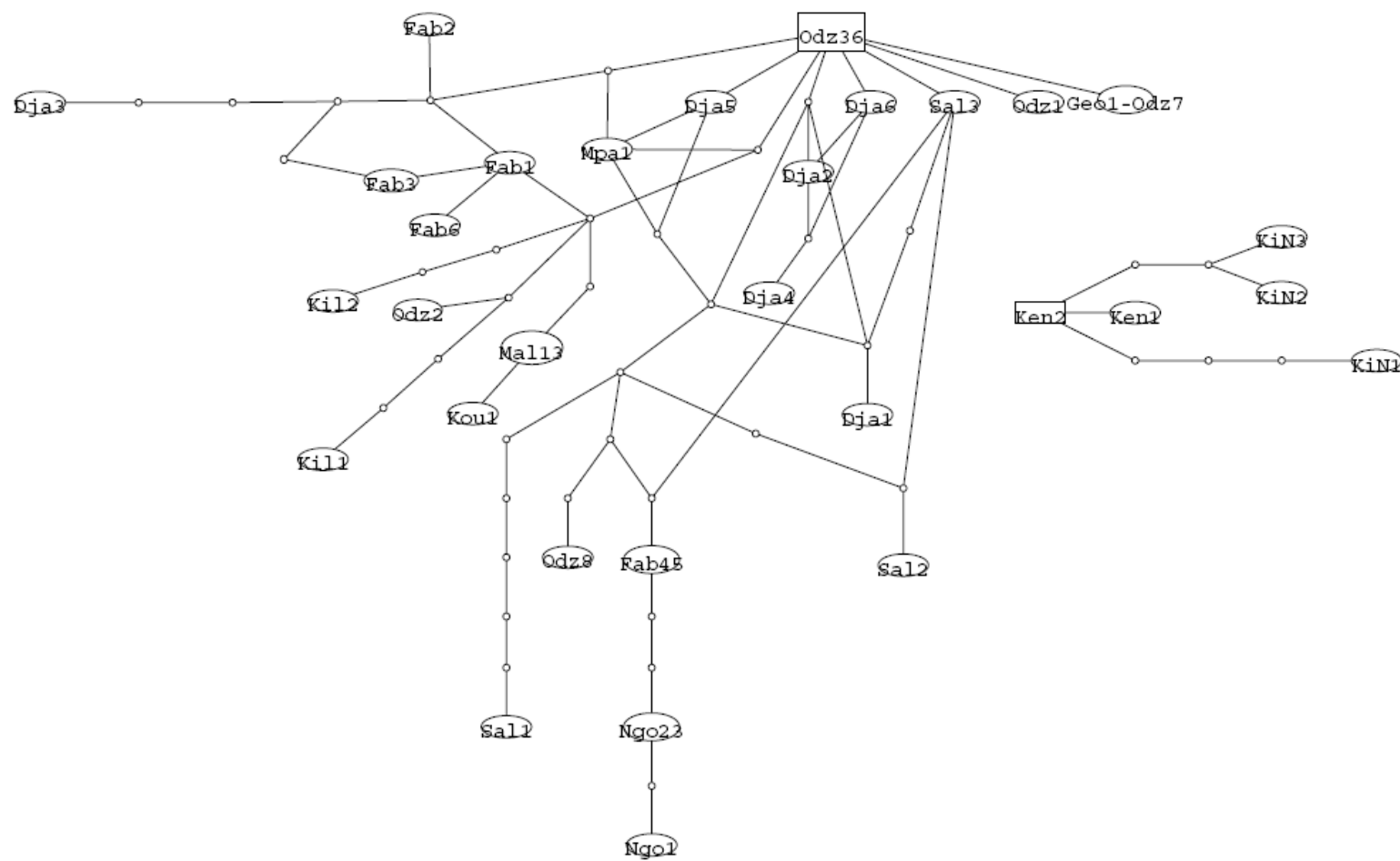
ANNEXE 6 : RESEAUX PARSIMONIEUX D'HAPLOTYPES OBTENUS POUR *S. JOHNSTONI*



ANNEXE 7 : RESEAU PARSIMONIEUX D'HAPLOTYPES OBTENU POUR *S. OLLULA*



ANNEXE 8 : RESEAUX PARSIMONIEUX D'HAPLOTYPES OBTENUS POUR *H. STELLA*



**ANNEXE 9 : RESEAUX PARSIMONIEUX D'HAPLOTYPES OBTENUS POUR
*S. LONGICAUDATUS***

