

MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE

**ÉCOLE PRATIQUE DES HAUTES ÉTUDES**  
**Sciences de la Vie et de la Terre**

**MÉMOIRE**

**Présenté**

**Par**

**Yannick Chaval**

Pour l'obtention du diplôme de l'École Pratique des Hautes Études

**Taxonomie intégrative de la tribu des Rattini  
(Rodentia, Muridae) en Asie du Sud-Est**

soutenu le 25/04/2014 devant le jury suivant :

<b>Dr Nicolas Navarro</b>	Maître de Conférences, EPHE, Dijon	<b>Président</b>
<b>Dr Julien Claude</b>	Maître de Conférences, UM2, Montpellier	<b>Tuteur scientifique</b>
<b>Dr Claudine Montgelard</b>	Maître de Conférences, EPHE, Montpellier	<b>Tutrice pédagogique</b>
<b>Pr Stéphane Aulagnier</b>	Professeur, UPS, Toulouse	<b>Rapporteur</b>
<b>Dr Violaine Nicolas</b>	Maître de Conférences, MNHN, Paris	<b>Examinatrice</b>
<b>Dr Jean-François Cosson</b>	Directeur de recherche, INRA, Montpellier	<b>Examineur</b>

**Mémoire préparé sous la direction de :**

*Julien Claude*

**Laboratoire :** UMR 5554 ISEM

**Directeur :** Agnès Mignot

**et de**

*Claudine Montgelard*

**Laboratoire :** UMR 5175 Cefe

**Directeur :** Philippe Jarne

**EPHE (Sciences de la Vie et de la Terre)**



# Remerciements

Tout d'abord un grand merci à Goy de l'université de Kasetart, à Nid et Tong de l'université de Maha Sarakham et à leur armée d'étudiants qui ont patiemment et méticuleusement préparé les crânes de rongeurs sur lesquels j'ai travaillé. Sans eux et leur labeur efficace mon travail n'aurait pas existé.

Je tiens ensuite à remercier mes collaborateurs les plus proches. Jean-François Cosson qui m'a recruté (ce qui était déjà un bon point) et m'a poussé à la réalisation de ce travail dans le cadre d'un cursus de diplôme EPHE. Tu m'as permis de découvrir de nouveaux horizons tant scientifiques que géographiques et culturels. Ces richesses font aujourd'hui partie de mon bagage professionnel et humain. Je t'en remercie infiniment et te souhaite un rétablissement rapide pour un retour en forme. Je remercie également Nathalie Charbonnel. Tu as été pour moi une encadrante administrative sans faille, pleinement impliquée dans ton rôle et toujours prête au dialogue. Tu m'as rapidement érigé en partenaire de tes expérimentations de terrain me déléguant l'encadrement des étudiants et m'invitant à participer à la mise au point des protocoles, je te remercie de la confiance que tu m'as témoigné. Je remercie Maxime Galan, moléculiste infatigable, débordant d'idées, de dynamisme et de compétences. Que d'aventures avons-nous partagées tous les quatre du Jura aux Ardennes, à chaque fois ce n'était pas tout Phrae mais ça nous a permis de boire des coups à Loei (Herbreteau, com. Pers.). Ces moments professionnels mais aussi les gros à côté (soirées, concerts, repas...) partagés avec vous ont été de grands moments de vie. Je ne regrette rien de ce temps et vous ouvre toutes grandes les portes de ma nouvelle demeure.

Un très grand merci également à Serge Morand, porteur des projets **CERoPath** et aujourd'hui **BiodivHeathSEA**, à la fois bulldozer de la science capable de monter des collaborations innombrables et fécondes, homme de terrain infatigable, homme de sciences incroyable, machine à écrire haut débit....bref une sorte de Sen Yai. Sacré Serge, tu as toujours un coup d'avance sur tout le monde et même quand on connaît ta botte secrète : « la projection dans le futur de demain », c'est quand même impressionnant. Sans toi rien n'aurait été possible et l'appellation porteur de projet trouve en toi son incarnation. Tu m'as traité sur un pied d'égalité et tu as su me faire prendre confiance en moi, me lançant des défis sans me les imposer et m'emmenant toujours plus haut. Que de bons moments partagés autour d'un ragout de rats ou à déguster des œufs embryonnés. Je te souhaite une longue carrière dans cette Asie du Sud-Est qui nous a tous deux conquis. Saches que pour toi aussi ma porte restera ouverte tant au niveau professionnel que pour partager de bons moments et de bons plats ensemble.

Merci à Vincent Herbreteau qui m'a tout appris des rongeurs asiatiques, m'a fait découvrir la culture et la vie thaïlandaise. Par ta connaissance du terrain, par ta maîtrise de la langue et des us et coutumes locaux, tu as été le trait d'union des cultures qui a permis la réussite de nos projets là-bas. Je n'oublierai jamais, et pour cause, nos réveils laotiens en fanfare au son du cochon que l'on égorge. Tu étais la seule personne avec qui je pouvais partager le charme subtil de ce réveil matin biologique car il en fallait plus pour troubler le sommeil de Serge. Là encore ma porte reste ouverte si tu passes par Toulouse un de ces jours et finalement .....si on le faisait ce bouquin ?

Un grand merci également à Marie Pagès, j'ai aimé ta rigueur sans faille. Ton travail de qualité a construit des bases solides sur lesquelles j'ai pu m'appuyer. Tu as joué pleinement le jeu

du partage et de la collaboration et c'était un vrai bonheur de travailler avec toi. Je te souhaite de trouver l'emploi que tu mérites, le plus rapidement possible. Une fois de plus la ville rose en ma personne t'accueillera à bras ouvert.

Merci aux personnes qui m'ont gentiment offert de relire les différentes parties du manuscrit : Nathalie, Gauthier, Manue, Guillaume, Marie, Jeff et bien sûr mes encadrants Claudine et Julien. Leur remarques ont été essentielles pour en améliorer la qualité.

Merci donc à Julien Claude pour la relecture du manuscrit mais également et surtout pour avoir pris le temps d'enseigner à cet étudiant attardé la poésie des statistiques lorsqu'elles sont mises en musique avec R. Ce ne fût pas une sinécure mais avec un peu d'énervement on finit par faire des choses avec ce truc. J'espère que le lecteur trouvera ces quelques productions à son goût.

Merci à Claudine Mongélard pour son encadrement et la relecture, tes conseils sur le manuscrit m'ont bien aidé.

Un grand merci particulier à Tonk, scientifique brillant actuellement en thèse à Liverpool et qui ira loin, c'est certain. Au-delà du réel plaisir que j'ai eu à travailler avec lui et à partager des moments en compagnie de sa petite famille, il a été lui aussi une porte d'entrée dans la culture thaïlandaise. Tu es d'une bienveillance extraordinaire lorsqu'on porte sur toi un regard français mais, il faut bien l'avouer sans intention de nuire à ton prestige, tout à fait dans les normes thaïlandaises. Je remercie également M, moins connu que son homonyme français, peut-être aussi moins bon chanteur mais très certainement beaucoup plus efficace sur le terrain. Nous resterons tout trois en contact car il en va ainsi des amitiés.

Merci aux chauffeurs et techniciens khmers, Oun le séducteur et Vibol l'indécis et son célèbre « may be » et bien sûr merci à Monsieur Balli pour le grand frisson sur la route. Merci à Kone, chauffeur du Nafri laotien et son fameux « safety first » prononcé alors qu'il bouclait sa ceinture tout en faisant demi-tour à la sortie d'un virage sans aucune visibilité. Ton entêtement à nous trainer au temple pour sauver notre âme m'a touché et mon cœur est léger aujourd'hui de stopper ma participation au génocide murin. Merci à nos chauffeurs thaïlandais : Joe, Kay, Toto, Som pot, To toujours prêts à rendre service.

Merci à Alice Latine pour le prêt des crânes de *Leopoldamys neilli*, à Aude Verwilghen pour le prêt des crânes de *Rattus tiomanicus*, à Madougou Garba pour le prêt des crânes de *R. rattus*, à Vincent Herbreteau une nouvelle fois mais cette fois pour le prêt des crânes collectés durant sa thèse. Merci à John Peter Yougote plus connu au Laos sous le nom de Dr Onco, porteur du projet ANR **Roboviroses à Hantavirus** de nous avoir associé au travail de terrain et introduit dans la formidable diversité murine du Sud-Est Asiatique.

Merci à Gauthier Dobigny pour la crise de rires lors de l'incendie (sans gravité) de l'appartement du docteur Onco. Merci également pour l'accueil au Niger et pour les discussions riches ou délirantes.

Merci aux copains du CBGP, que je ne peux pas tous citer mais il suffit de regarder l'organigramme de l'unité pour en avoir un aperçu à peine augmenté. Parmi ceux-ci je remercie particulièrement ceux que j'ai le plus fréquenté en dehors des jours ouvrables à savoir dans l'ordre alphabétique Anne, Armelle, Carine, Manue, Max, Nathalie (la grande et la petite), Jeff, Réjane,

Sylvain , Ainsi que mes thésards préférés Karine, Julie, Manu, Cécile ce fût un plaisir d'assister vos manips de terrain et Bertrand avec qui je n'ai pas fait de terrain mais qui est un bon gars quand même.

Merci à la formation permanente Inra et au CBGP de m'avoir permis de financer cette formation et à l'EPHE pour la qualité de la plupart de ses enseignements.



**ÉCOLE PRATIQUE DES HAUTES ÉTUDES**  
**SCIENCES DE LA VIE ET DE LA TERRE**

**Taxonomie intégrative de la tribu des Rattini  
(Rodentia, Muridae) en Asie du Sud-Est**

***Yannick Chaval***

**Soutenu le 25/04/2014**

**RÉSUMÉ**

L'Asie du Sud-Est est un point chaud de la biodiversité et, dans le même temps, cette région est un foyer actif d'émergence ou de réémergence de zoonoses (SRAS, grippe aviaire, leptospirose, roboviroses etc.). Le développement économique sans précédent et la croissance démographique humaine que connaît la zone biogéographique indochinoise sont un moteur de changements considérables des écosystèmes. Dans le contexte du changement global, il est primordial de comprendre comment ces modifications environnementales modifient les interactions entre les communautés d'hôtes et leurs communautés de parasites/pathogènes afin d'appréhender les transformations du paysage épidémiologique. Certaines espèces, qui sont des réservoirs de zoonoses constituent, de par leur mode de vie et leur statut de gibier des modèles particulièrement intéressants à étudier dans ce cadre. C'est le cas, pour la zone étudiée, des rongeurs et plus particulièrement des représentants de deux tribus de la sous-famille des Murinae : les Rattini et les Murini. La première partie de ce mémoire porte sur ces deux tribus, particulièrement diversifiés en Asie du Sud-Est. L'objectif a été d'étudier la faisabilité d'un outil d'identification web-service innovant basé sur la morphométrie. Cet outil viendra compléter des réalisations qui lui sont antérieures à savoir : un guide d'identification de terrain et un premier outil web-service basé sur l'identification moléculaire. Ces productions formeront un ensemble cohérent d'outils d'identification permettant d'offrir à la communauté des personnes désireuses d'étudier les rongeurs de la zone des accès multiples et adaptés à une identification de qualité. Pour la réalisation de cet outil d'identification morphométrique, nous avons tout d'abord étudié quels sont les facteurs intra et inter-spécifiques qui influencent la variation morphologique des espèces de rongeurs. Cette étude, au travers d'une approche de systématique intégrative a ainsi contribué à la stabilisation de la taxonomie des Rattini. L'étude morphométrique est entièrement basée sur un jeu de données de six cent soixante individus qui appartiennent à dix-sept espèces. L'ensemble des individus a fait au préalable l'objet d'une identification moléculaire réalisée à l'aide d'un gène mitochondrial. Dans le cadre de la création du web-service morphométrique, les sources d'erreurs relatives à la comparaison de jeux de données produits par des opérateurs différents sont identifiées et encadrées car elles pourraient compromettre l'identification des individus. Des seuils d'assignation sont également définis pour permettre à l'outil de gérer la confrontation avec des espèces qui ne sont pas représentées dans le jeu de données de référence. Dans un second temps nous avons souhaité replacer la disparité morphologique observée chez les espèces actuelles dans un contexte évolutif. L'objectif était de comprendre les patrons macro-évolutifs qui ont potentiellement gouverné la mise en place de la variation morphologique et ainsi d'évaluer le rôle de l'environnement sur celle-ci.

**MOTS-CLÉS : *Rattini, Murini, taxonomie intégrative, code barre moléculaire, analyse Procrustes, identification, web-service.***





# **Table des matières.**

**Abréviations**

**Liste des figures**

**Liste des tableaux**

## **Introduction**

<b>1. Contexte général</b>	2
<i>1.1. Maladies zoonotiques et communautés rongeurs/parasites en Asie du Sud-Est</i>	2
<i>1.2. Caractéristiques géographiques des sites d'étude</i>	3
<b>2. Objectifs de mon travail dans le cadre des programmes de recherche</b>	5
<b>3. Outils d'identifications</b>	5
<i>3.1. Le guide de terrain des murinés d'Asie du Sud-Est</i>	5
<i>3.2. L'outil web service d'identification moléculaire RodentSEA</i>	6
<i>3.3. Fonctionnement de l'outil RodentSEA</i>	6
<i>3.4. Le Museum virtuel accessible depuis RodentSEA</i>	8
<b>4. Limites des travaux réalisés en amont du présent travail</b>	9

## **Première partie : identifications morphométriques**

<b>A- Méthodes de morphométrie utilisées dans ce travail</b>	12
<b>1. Morphométrie traditionnelle et morphométrie géométrique</b>	12
<i>1.1. La morphométrie basée sur les mesures brutes</i>	
<i>1.2. Forme, taille et conformation</i>	13
<i>1.3. L'analyse en logshape-ratio</i>	13

<i>1.4. La morphométrie géométrique</i>	13
<i>1.4.1. Le principe de l'analyse Procrustes</i>	14
<i>1.4.2. Les types de points homologues</i>	15
<i>1.4.3. Les trois étapes de la superposition Procrustes</i>	15
<i>1.4.4. Les résultats de la superposition Procrustes</i>	16
<b>2. Sources d'erreurs lors de la production des données : estimation et correction</b>	17
<i>2.1. La précision</i>	17
<i>2.2. L'erreur de mesure stochastique</i>	18
<i>2.2.1. Influence de la préparation des spécimens</i>	18
<i>2.2.2. Positionnement et éclairage de l'objet et lors de la prise de vue</i>	19
<i>2.2.3. Protocole d'estimation de l'erreur de mesure intra-utilisateur</i>	19
<i>2.3. L'erreur de mesure systématique</i>	20
<i>2.3.1. Influence du positionnement de l'appareil photographique</i>	20
<i>2.3.2. Effet 'appareil photographique'</i>	21
<i>2.3.4. Correction de la déformation d'image</i>	21
<i>2.3.5. Effet 'opérateurs'</i>	22
<b>3. Traitement des données manquantes en morphométrie</b>	22
<i>3.1. Pour les mesures linéaires</i>	22
<i>3.2. Pour les points homologues</i>	22
<b>B. Application de l'étude morphométrique aux rongeurs du Sud-Est Asiatique</b>	23
<b>1. Echantillonnage, sexage et critère d'âge des rongeurs référence</b>	23
<i>1.1. L'échantillonnage des rongeurs</i>	23

1.2. Détermination des classes d'âge et du sexe	23
<b>2. Protocoles permettant la réduction de l'erreur de mesure lors de l'étude de crânes de rongeurs</b>	25
2.1. Protocole pour la préparation des crânes	25
2.2. Positionnement de l'appareil photographique et des sources lumineuses	26
2.3. Positionnement des pièces ostéologiques lors des prises de vue	26
<b>3. variables morphométriques utilisées dans cette analyse</b>	27
3.1. La prise de mesures externes	27
3.2. La prise de données de morphométrie géométrique	28
3.2.1. Préparation des prises de vue à la digitalisation des points homologues	28
3.2.2. Digitalisation des points homologues	28
<b>C- Résultats du traitement des données morphométriques</b>	30
<b>1. Données morphométriques et statistiques descriptives</b>	30
<b>2. Etude de la taille</b>	30
2.1. Facteurs influençant la variation de taille corporelle des adultes	30
2.2. Facteurs influençant la variation de taille crânienne des adultes	31
2.3. Comparaison de la taille corporelle et de la taille crânienne	32
<b>3. Description de la variabilité morphologique à l'aide de l'Analyse en Composantes Principale (ACP)</b>	32
3.1. L'analyse en composante principale sur les données brutes	33
3.2. L'analyse en composante principale sur les données logshape-ratio	33
3.3. Analyse en composante principale sur les données de morphométrie géométrique du crâne en vue ventrale	34
3.4. L'analyse en composante principale sur les données de morphométrie	35

<i>géométrie du crâne en vue dorsale</i>	
<i>3.5. Analyse en composante principale sur les données de morphométrie géométrique de la mâchoire en vue externe</i>	36
<i>3.6 L'analyse en composante principale sur les données de morphométrie géométrique de la mâchoire en vue interne</i>	37
<i>3.7. Discussion sur les variables utilisées dans ces ACP</i>	38
<b><i>4. Les mitotypes structurent-ils significativement la variation morphologique lorsque celle-ci correspond aux données brutes, aux logshape-ratio ou aux résidus Procrustes ?</i></b>	38
<i>4.1. Effets de l'âge, du sexe et de l'appartenance à un mitotype sur la forme et la conformation du corps des animaux</i>	39
<i>4.2. Effet de l'âge, du sexe et de l'appartenance à un mitotype sur la conformation et la forme du crâne en vue ventrale</i>	39
<b><i>5. Comparaison du pouvoir discriminant des mesures brutes, des logshape-ratios, des variables de conformation et de forme crâniennes</i></b>	40
<i>5.1. Méthode de comparaison</i>	40
<i>5.2. Discrimination des espèces sur la base des mesures corporelles</i>	41
<i>5.3. Discrimination des espèces sur la base des analyses Procrustes réalisées sur les différentes parties du crâne</i>	41
<i>5.4. Discussion</i>	43
<b><i>6. Optimisation du pouvoir discriminant des prises de vues crâniennes et simplification des conformations</i></b>	44
<i>6.1. Choix du nombre de composantes de conformation de l'ACP pour optimiser la fonction discriminante</i>	44
<i>6.2. Pouvoir discriminant des vues et simplification des points homologues</i>	44
<i>6.2.1. Méthodes</i>	44
<i>6.2.2. Résultats</i>	45
<i>6.2.3. Discussion</i>	45

<b>7- Erreur de mesure lors de l'acquisition des données de morphométrie géométrique et leur influence sur la capacité à discriminer les espèces</b>	47
7.1. Erreur intra-opérateur de taille en fonction de la structure et de la vue étudiée	48
7.2. Erreur intra-opérateur sur la conformation en fonction de la structure et de la vue étudiée	48
7.3. Influence sur la capacité à discriminer les espèces	49
<b>8. L'effet opérateurs multiples</b>	49
8.1. Effet de la différence d'appréciation du placement des points homologues par l'utilisateur et l'opérateur de référence	49
8.1.1. Méthodes	50
8.1.2. Résultats	51
8.1.3. Discussion	51
8.2. Effet de l'erreur opérateurs sur l'assignation des crânes des individus test	52
8.2.1. Méthode	52
8.2.2. Résultats	53
8.2.3. Discussion	53
8.2.4. Conclusion préliminaire sur la faisabilité de l'outil d'identification morphométrique	53
8.3. Représentation de l'erreur opérateurs	54
8.3.1. Méthode	54
8.3.2. Résultats	54
8.3.3. Discussion	55
8.4. Réduction de l'effet inter-opérateur	56
8.4.1. Méthode	56
8.4.2. Résultats	56

8.4.3. Discussion	57
8.4.4. Conclusion	57
<b>D- Mise en place de l'outil d'identification morphométrique</b>	<b>57</b>
<b>1. Test de l'outil d'identification</b>	<b>58</b>
1.1. Méthode	58
1.2. Résultats	58
1.3. Discussion	60
<b>2. Schéma de principe de l'outil</b>	<b>61</b>
2.1. Premier module : interface de correction de la lentille et de traitement de l'image	61
2.2. Deuxième module : entraînement de l'utilisateur (test de répétabilité)	61
2.3. Deuxième module : digitalisation d'un individu inconnu (test de reproductibilité)	62
2.4. Deuxième module : assignation de l'individu inconnu	62
2.5. Troisième modules : affichage du résultat	63
2.6. Choix de la structure étudiée	63
<b>3. Discussions : Limites et améliorations possibles de l'outil, conseils de mise en ligne</b>	<b>64</b>
3.1. Limites liées à la conformation moyenne	64
3.2. Limites de présentation du résultat	64
3.3. Limites liées à l'échantillonnage partiel des espèces de la zone d'étude	64
3.4. Limites liées à la reproductibilité des données	65
3.4.1. Les solutions avancées dans la littérature	65
3.4.2. La solution proposée ici et ses limites	66
3.5. Perspectives	66

## **Deuxième partie : Complémentarité des approches moléculaires et morphométriques :**

### **délimitation d'espèces et appréciation du rôle de l'environnement sur la variabilité morphologique interspécifique.**

<b>1. Une approche intégrative pour définir les limites d'espèces : le cas des mitotypes <i>R. tanezumi</i> et <i>R. phylogénétique R3</i> sensu Pagès (2010, 2013)</b>	<b>70</b>
<i>1.1. Présentation de l'article: « distordance cytonucléaire parmi les rats noirs du sud-est asiatique (complexe d'espèce <i>Rattus rattus</i> ) », Pagès et al. (2013)</i>	70
<i>1.2. Inférence de l'influence du mitotype sur la morphologie externe au sein du genre <i>Rattus</i> par l'utilisation des logshape-ratio (travail inclus dans l'analyse de Pagès et al., 2013)</i>	72
<i>1.3. Variation de la conformation crânienne en vue ventrale entre les deux clades mitochondriaux <i>R. tanezumi</i> et <i>Rattus R3</i> (analyse non publiée)</i>	74
<i>1.4. Scénarios évolutifs permettant la réconciliation des données mitochondriales et nucléaires</i>	76
<i>1.4.1. Trie imparfait des lignées ancestrales (ILS)</i>	76
<i>1.4.2. L'introgession mitochondriale</i>	77
<b>2. Facteurs influençant la variabilité morphologique</b>	<b>78</b>
<i>2.1. Données utilisées dans cette étude</i>	78
<i>2.1.1. Données morphométriques</i>	78
<i>2.1.2. Données moléculaires</i>	79
<i>2.1.3. La variable environnementale : le milieu de vie des espèces</i>	79
<i>2.2. Représentation de l'espace des conformations des espèces</i>	80
<i>2.3. Représentations conjointes de l'évolution des conformations de crâne, du milieu de vie et de la phylogénie</i>	80

2.3.1. <i>La reconstruction des caractères ancestraux</i>	81
2.3.2. <i>Représentation</i>	82
2.3.3. <i>Rôle du milieu de vie sur la morphologie des crânes</i>	82
2.5. <i>Corrélation des matrices morphologiques et moléculaires</i>	86
2.6. <i>Scénarios macro-évolutifs et évolution des conformations et de la taille</i>	88
2.6.1. <i>Description des modèles testés</i>	89
2.6.2. <i>La méthode de Pagel, signification et évolution des paramètres <math>\lambda</math>, <math>\delta</math>, <math>\kappa</math></i>	90
2.6.3. <i>Sélection du modèle qui correspond le mieux aux données observées</i>	91
2.7. <i>Effet du facteur environnemental sur la morphologie en tenant compte de l'information phylogénétique</i>	94
2.8. <i>Evolution des morphologies crâniennes</i>	97
2.9. <i>Comparaison des conformations actuelles et implications pour l'identification des espèces</i>	99
2.9.1. <i>Différences entre Rattus tanezumi et Rattus rattus</i>	99
2.9.2. <i>Différences entre Rattus tanezumi et Rattus tiomanicus</i>	100
2.9.3. <i>Différences entre Rattus rattus et Rattus tiomanicus</i>	100
2.9.4. <i>Différences entre Rattus tanezumi et Rattus andamanensis</i>	100
2.9.5. <i>Différences entre Rattus tanezumi et Rattus argentiventer</i>	101
2.9.6. <i>Différences entre Rattus tanezumi et Rattus exulans</i>	101
2.9.7. <i>Différences entre Rattus tanezumi et Rattus nitidus</i>	101
2.9.8. <i>Différences entre Rattus sakeratensis et Rattus argentiventer</i>	101
2.9.9. <i>Différences entre Bandicota savilei et B. indica</i>	101
2.9.10. <i>Différences entre Berylmys berdmorei et B. bowersi</i>	101
2.9.11. <i>Différences entre Niviventer hincei et N. fulvescens</i>	102



<i>2.9.12. Différences entre Leopoldamys sabanus et L. edwardsi</i>	102
<i>2.9.13. Différences entre L. neilli et L. edwardsi.</i>	102
<i>2.9.14. Différences entre L. edwardsi et M. surifer.</i>	102
<i>2.9.15. Différences entre les espèces du genre Mus</i>	102
<i>2.9.17. Différences entre Mus cookii et M. cervicolor</i>	102
<i>2.9.18. Différences entre M. cookii et M. caroli</i>	103
<i>2.9.19. Différences entre M. cervicolor et M. caroli</i>	103

**Conclusion générale** 105

**Bibliographie** 108

**Annexes**

<b>Annexe I</b>	117
<b>Annexe II</b>	133
<b>Annexe III</b>	134
<b>Annexe IV</b>	136
<b>Annexe V</b>	140
<b>Annexe VI</b>	141
<b>Annexe VII</b>	144
<b>Annexe VIII</b>	167
<b>Annexe IX</b>	168

**Glossaire** 169

## Abréviations et Acronymes

**ACP:** Analyse en Composante Principale.

**ALD:** Analyse Linéaire Discriminante.

**ANR :** Agence National de la Recherche.

**ANOVA:** ANalysis Of Variance.

**CERoPath:** « Community Ecology of Rodents and their Pathogen in South-East Asia ». Ecologie des communautés de rongeurs et de leur pathogènes en Asie du Sud-Est. Effets des changements de biodiversité et leurs impacts sur l'écologie de la santé.

**COI :** gène du cytochrome oxydase I.

**Cyt b :** gène du cytochrome b.

**IC99 % :** Intervalle de Confiance à 99 %.

**IRBP:** premier exon du gène codant pour la protéine de l'inter photorécepteur rétinoïde.

**IDH :** l'Indice de Développement Humain.

**K2P :** Kimura's two parameters.

**MANOVA :** Multivariate ANalysis Of Variance.

**Opr:** Opérateur de référence

**Roboviroses:** « Rodent Borne Viruses ».

**TPS :** Tin Plate Spline.

## Liste des figures (au verso des pages indiquées).

### *Première partie*

<b>Figure 1</b> : sites échantillonnés par les programmes ANR Roboviroses et ANR CERoPath.	22
<b>Figure 2</b> : les trois classes d'âges retenues dans cette étude en fonction du développement de la troisième molaire et de la maturité sexuelle.	23
<b>Figure 3</b> : positionnement de l'appareil	25
<b>Figure 4</b> : vue générale du montage de l'appareil et de l'éclairage.	25
<b>Figure 5</b> : positionnement du crâne en vue dorsale	26
<b>Figure 6</b> : positionnement du crâne en vue ventrale	26
<b>Figure 7</b> : les cinq mesures prises sur le terrain pour l'identification des rongeurs	27
<b>Figure 8</b> : Les photos sont transformées en négatif pour faire ressortir les sutures pariétales.	27
<b>Figure 9</b> : configurations de points homologues étudiés.	28
<b>Figure 10</b> : les groupes représentés par les lettres bleues sont constitués des mitotypes dont la différence de taille corporelle n'est pas significative au sein des groupes.	30
<b>Figure 11</b> : Droite de régression entre la mesure des tailles centroïdes des mitotypes en abscisse, des tailles corporelles en ordonnée et des valeurs moyennes des différents mitotypes pour ces traits.	31
<b>Figure 12</b> : représentation des composantes principales dans le plan des mesures corporelles brutes. Les axes sont les composantes principales 1 et 2 et le cercle est celui des corrélations.	32
<b>Figure 13</b> : analyse en composante principale sur les données corporelles brutes. Les ellipses contenant 90 % des données sont représentées.	32
<b>Figure 14</b> : représentation des logshape-ratio dans le plan des variables de l'ACP. Les axes sont les composantes principales 1 et 2 et le cercle est celui des corrélations.	33
<b>Figure 15</b> : représentation des ellipses contenant 90 % des données corporelles transformées en logshape-ratio dans le plan principal de l'ACP.	33
<b>Figure 16</b> : représentation des ellipses regroupant 90 % des individus de chaque mitotype dans le plan principal de l'espace des conformations crâniennes en face ventrale.	34
<b>Figure 17</b> : représentation des ellipses regroupant 90 % des individus de chaque mitotype en fonction de la taille centroïde des individus représentée en abscisse et la conformation des crânes en vue ventrale représentée ordonnées par la première composante de conformation.	34
<b>Figure 18</b> : représentation des ellipses regroupant 90 % des individus de chaque mitotype dans le plan principal de l'espace des conformations crâniennes en face dorsale.	35
<b>Figure 19</b> : représentation des ellipses regroupant 90 % des individus de chaque mitotype dans le plan principal de l'espace des conformations de la mâchoire externe.	36
<b>Figure 20</b> : représentation des ellipses regroupant 90 % des individus de chaque mitotype dans le plan principal de l'espace des conformations de la mâchoire interne.	36
<b>Figure 21</b> : Représentation dans le plan discriminant principal des ellipses regroupant 90 % des individus de chaque mitotype en vue ventrale.	39
<b>Figure 22</b> : Variation du pouvoir discriminant de l'ALD en fonction du nombre de composantes de conformation conservées sur l'ACP.	43
<b>Figure 23</b> : lecture d'un graphique synthétisant le pouvoir discriminant maximum (point vert) d'une conformation en fonction du nombre de composantes de conformations	44

associées (en noir sous le point vert).	
<b>Figure 24</b> : comparaison du pouvoir discriminant des structures, des vues et des conformations de points homologues.	44-45
<b>Figure 25</b> : Simplification de la conformation de la vue ventrale du crâne.	46
<b>Figure 26</b> : composition de l'échantillonnage pour tester l'effet utilisateur sur les vues ventrales	52
<b>Figure 27</b> : représentation de la variation de la conformation moyenne des espèces du groupe test à l'aide d'espaces vectoriels.	53
<b>Figure 28</b> : représentation de la variation de la conformation moyenne après « entraînement » de l'utilisateur sur les points 20 et 37 (a) et sur l'ensemble des points (b).	54
<b>Figure 29a et b</b> : gestion d'individus appartenant à un mitotype inconnu de la base de données de référence	57
<b>Figure 30</b> : schéma de principe de l'outil d'identification morphologique.	60
<b>Figure 31</b> : sortie dynamique et interactive de l'outil d'identification pour la vue ventrale du crâne.	62

## *Deuxième partie*

<b>Figure 1</b> : Evolution significative de la taille relative de certains traits dans la conformation corporelle des animaux.	22
<b>Figure 2</b> : Distribution géographique des monotypes échantillonnés dans l'article de Pagès et collaborateurs (2013).	23
<b>Figure 3</b> : L'histoire évolutive des gènes ne reflète pas toujours l'histoire des espèces.	75
<b>Figure 4</b> : Milieux fréquentés par chaque espèce d'après 1192 captures réalisées durant le programme CERoPath et dont la localisation est d'une précision inférieure ou égale à 100 mètres.	78
<b>Figure 5</b> : Représentation de la transformation d'une conformation de crâne en face ventrale d'une espèce.	79
<b>Figure 6</b> : Reconstruction en face ventrale de l'ancêtre commun des Rattini.	80
<b>Figure 7</b> : La zone grisée de la matrice (diagonale) en haut à droite représente la variance totale entre les feuilles et la racine sous l'hypothèse brownienne.	81
<b>Figure 8</b> : Représentation du phylo-éco-morpho espace de la tribu des Rattini en Asie du Sud-Est.	82
<b>Figure 9</b> : déformation de l'arbre phylogénétique en fonction des paramètres Lambda, Delta, Kappa	88
<b>Figure 10</b> : disparité morphologique observée et disparité morphologique attendue sous l'hypothèse d'une évolution suivant un mouvement Brownien.	94
<b>Figure 11</b> : La conformation de l'ancêtre commun est représentée en gris clair et la conformation cible en pointillés.	96
<b>Figure 12</b> : Evolution inter-spécifiques des conformations en vue ventrale des crânes du genre <i>Rattus</i> .	98
<b>Figure 13</b> : Critères de différenciation entre <i>R. tanezumi</i> et <i>R. rattus</i> à la même échelle.	99
<b>Figure 14</b> : Critères de différenciation entre <i>R. tanezumi</i> et <i>R. andamanensis</i> à la même échelle.	99
<b>Figure 15</b> : Evolution inter-spécifiques des conformations en vue ventrale des crânes de la tribu des Rattini.	101

## *Annexe I*

<b>Figure 1 :</b> Evolution significative de la taille relative de certains traits dans la conformation corporelle des animaux.	22
<b>Figure 2 :</b> richesse spécifique mondiale de la tribu des Rattini	120
<b>Figure 3 :</b> Arbre des relations phylogénétiques de la tribu des Rattini indochinois	122
<b>Figure 4 :</b> Distribution actuelle de la tribu des Rattini, Hydromyini et Murini dans les différentes sous-régions de la zone indo-australienne (Fabre et al., 2013).	127
<b>Figure 5 :</b> Limite des terres au dernier maximum glaciaire (19-23Ka).	128
<b>Figure 6 :</b> Arbre phylogénétique du genre <i>Mus</i>	130

## *Annexe II*

<b>Figure 1 :</b> Les étapes de la superposition Procrustes.	133
--	-----

## *Annexe IV*

Dénomination des points homologues	136
------------------------------------	-----

## *Annexe V*

Représentation des mesures et du poids des animaux adultes échantillonnés durant le programme CERoPath.	140
---	-----

## *Annexe VI*

Représentation du plan principal de l'ACP pour les données corporelles brutes	141
---	-----

## *Annexe VII*

<b>Figure 1:</b>	160
<b>a-</b> différences nucléotidiques entre le genre <i>Bandicota</i> et différents haplotypes	
<b>b-</b> photographie d'un gel d'agarose montrant l'amplification différenciée de parties du gène <i>Cyt b</i>	
<b>Figure 2:</b>	160
<b>a-</b> différences nucléotidiques entre le genre <i>Bandicota</i> et différents haplotypes	
<b>b-</b> photographie d'un gel d'agarose montrant l'amplification différenciée de parties du gène <i>Cyt b</i>	
<b>Figure 3:</b> identification a) et représentation b) des mesures corporelles diagnostiques	161
<b>Figure 4:</b> configuration de points homologues utilisée sur la mâchoire externe a), b) représentation du PC1 en fonction de la taille centroïde ; c) résultats de l'analyse discriminante	162
<b>Figure 5:</b> représentation du dimorphisme de taille chez <i>B.indica</i> et <i>B.savilei</i>	163
<b>Figure 6 :</b> transformation de la mâchoire en fonction de l'espèce ainsi que du sexe et de l'âge chez <i>B.savilei</i> .	164

## *Annexe IX*

Exemples de phylomorpha espaces reconstruits à partir de plusieurs méthodes d'estimation des caractères ancestraux	161
--	-----

## Liste des tableaux (au verso des pages indiquées).

### *Première partie*

<b>Tableau 1:</b> jeu de données morphométrique de référence de l'outil d'identification.	22
<b>Tableau 2:</b> résultats de l'analyse de variance de la taille corporelle des individus adultes.	29
<b>Tableau 3:</b> résultats de l'analyse de variance de la taille centroïde du crâne des individus adultes.	30
<b>Tableau 4:</b> Table d'analyse de variance de la taille centroïde du crâne en vue palatine des adultes en fonction du mitotype des animaux, de leur sexe et de l'interaction entre le mitotype et le sexe (se reporter au Tableau 3 pour la signification des colonnes).	31
<b>Tableau 5:</b> résultats de l'analyse de variance multivariée réalisée sur les cinq composantes principales de l'ACP des données brutes corporelles.	
<b>Tableau 6:</b> résultats de l'analyse de variance multivariée réalisée sur quatre composantes principales de l'ACP pour les logshape-ratio des données corporelles.	38
<b>Tableau 7:</b> résultats de l'analyse de variance multivariée réalisée sur les composantes principale de l'ACP pour les résidus Procrustes de la vue ventrale du crâne.	38
<b>Tableau 8:</b> résultats de l'analyse de covariance multivariée (Mancova type II) réalisée sur les composantes principale de l'ACP pour les résidus Procrustes de la vue ventrale du crâne.	38
<b>Tableau 9:</b> résultats de l'analyse de variance multivariée (Manova type II) réalisée sur les composantes principale de l'ACP de conformation et la taille centroïde de la vue ventrale du crâne.	38
<b>Tableau 10:</b> résultats des tests de robustesse de l'analyse discriminante pour les données brutes et les logshape-ratio des mesures corporelles ainsi que pour les différentes vues des différentes structures ostéologiques étudiées.	40-41
<b>Tableau 11:</b> résultat de l'identification des individus test en face dorsale et ventrale.	52
<b>Tableau 12:</b> synthèse et comparaison de l'erreur produite par l'opérateur de référence avec la variation de celle produite par l'utilisateur au cours de son entraînement.	55

### *Deuxième partie*

<b>Tableau 1:</b> résultats de l'Analyse de variance de la taille des crânes en vue palatine des mitotypes <i>Rattus tanezumi</i> et R3.	74
<b>Tableau 2:</b> résultats de l'analyse de variance multivariée sur les composantes principales de conformation des crânes en vue palatine des mitotypes <i>Rattus tanezumi</i> et R3.	74
<b>Tableau 3:</b> échantillonnage morphologique utilisé pour étudier l'évolution des conformations en vue ventrale de la tribu des Rattini en Asie du Sud Est.	77
<b>Tableau 4:</b> résultats de l'analyse de l'influence du milieu de vie des espèces sur la variance de la taille des crânes en vue ventrale.	83
<b>Tableau 5:</b> résultats de l'analyse de l'influence du milieu de vie des espèces sur la variance des 4 premières composantes de conformation des crânes en vue ventrale.	83
<b>Tableau 6:</b> Sélection du meilleur modèle macro évolutif pour expliquer l'évolution de la taille et des conformations (sur la CP1 et CP2) sur le crâne en vue ventrale.	92
<b>Tableau 7:</b> résultats de l'analyse de l'influence du milieu de vie des espèces sur la variance de la taille des crânes en vue ventrale en tenant compte des relations phylogénétiques entre espèces.	93
<b>Tableau 8:</b> résultats de l'analyse de l'influence du milieu de vie des espèces sur la variance des 4 premières composantes de conformation des crânes en vue ventrale en tenant compte des relations phylogénétiques entre espèces.	93

## *Annexe I*

<b>Tableau 1:</b> correspondance entre la dénomination des clades moléculaires de la figure 3 et les noms utilisés pour les nommer dans l'article de Pagès et al. (2010) et dans les parties de ce mémoire.	<b>120</b>
<b>Tableau 2:</b> Distance de Kimura à deux paramètres par paires (délétion par paires) de différentes lignées phylogénétiques (Chaval, non publié) définies à partir de l'analyse du gène mitochondrial Cyt B pour le genre <i>Niviventer</i> .	<b>124</b>

## *Annexe VI*

<b>Tableau 1:</b> nombre d'individus étudiés par espèces	<b>164</b>
<b>Tableau 2:</b> nombre de <i>Bandicota</i> étudiés par localités	<b>165</b>
<b>Tableau 3:</b> analyse de la variance de taille des individus adultes en fonction de l'espèce et du sexe	<b>165</b>
<b>Tableau 4:</b> analyse de covariance de la conformation de la mâchoire en fonction du sexe, de l'espèce et de la taille	<b>166</b>
<b>Tableau 5:</b> individus par espèce et par localité ayant subi une identification moléculaire	<b>166</b>









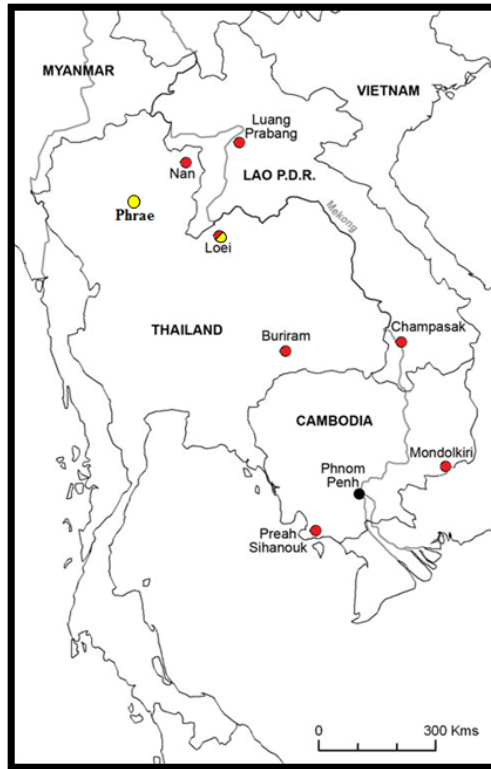
## ***1. Contexte général***

### ***1.1- Maladies zoonotiques et communautés rongeurs/parasites en Asie du Sud-Est***

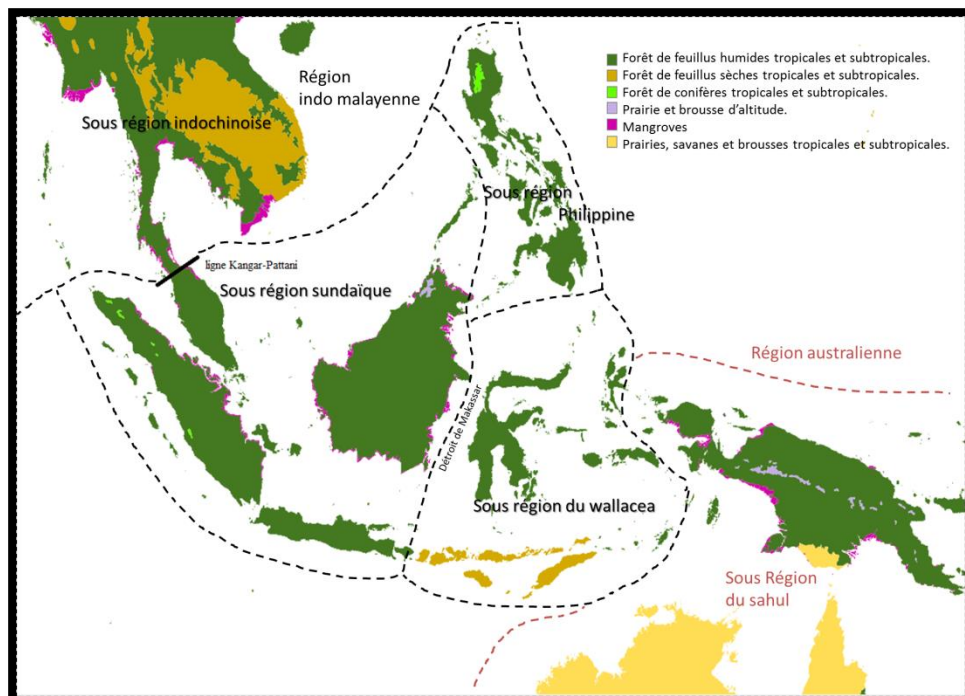
L'Asie du Sud-Est est l'une des zones de notre planète qui accueille la plus grande richesse biologique (Myers et al., 2000). Cette région abrite donc potentiellement un grand nombre d'espèces réservoir de maladies infectieuses. Ce travail s'inscrit dans le cadre de programmes de recherches scientifiques initiés en Asie du Sud-Est depuis 2006 (ANR **Roboviroses** à Hantavirus ("Biodiversité des roboviroses à hantavirus en Europe et en Asie du sud-est" dir. J.P. Hugot) et ANR **CERoPath** ("Community Ecology of Rodents and their Pathogens in South-east Asia" dir. S. Morand). L'objectif de ces projets est de comprendre la structuration des communautés de rongeurs et de leurs communautés de parasites et de pathogènes dans le contexte des changements globaux qui affectent les milieux naturels en Asie du Sud-Est. En effet, l'ordre des Rongeurs, qui comprend 2277 espèces valides, est le plus diversifié au sein de la classe des Mammifères (Carleton et Musser, 2005). Ce taxon à lui seul représente environ 42% de la diversité spécifique mammalienne totale (Wilson et Reeder, 2005) et est un groupe impliqué dans la transmission des zoonoses (Kruse et al., 2004). Les rongeurs sont en effet connus pour être le réservoir d'au moins soixante maladies zoonotiques (Meerburg et al., 2009).

L'homme et les rongeurs entretiennent des relations particulièrement étroites dans la mesure où un certain nombre d'espèces de rongeurs ont adopté un mode de vie commensal qui se traduit par l'occupation et l'utilisation des espaces agricoles, des moyens de transport de marchandises ainsi que des villes et des habitations. De plus en Asie du Sud-Est la pratique de la chasse aux rongeurs champêtres et forestiers constitue un apport protéique non négligeable pour les populations rurales lorsque les grands ongulés deviennent rares. Dans certaines régions, ils font également l'objet d'élevage et de vente à grande échelle sur les marchés et figurent même à la carte des restaurants. La promiscuité et le contact avec les fluides corporels lors de la préparation des animaux pour la consommation favorisent la transmission à l'homme de pathogènes portés par les rongeurs. D'autre part le mode de vie commensal a permis aux espèces s'y étant adaptées d'étendre leur aire de répartition à l'échelle de plusieurs continents et parfois à la planète entière (Schwarz, 1960 ; Aplin et al., 2011 ; Suzuki et al. ; 2013). Cette expansion de l'aire géographique des hôtes s'est parfois accompagnée de celle de leurs parasites et pathogènes favorisant l'émergence de pandémies à des endroits où elles étaient jusqu'alors inconnues. On peut citer à titre d'exemple le cas de la peste véhiculée par le rat noir, *Rattus rattus* Linnaeus, 1758, le long des voies maritimes (Gauthier-Clerc et Thomas, 2010).

Les changements globaux incluent à la fois des changements biotiques (mouvement des espèces animales et végétales et de leurs pathogènes) et abiotiques (changement climatique, changement d'usage des sols, pollutions chimiques) dont l'homme est la composante principale (Lorius et Carpentier, 2011). L'Asie du Sud-Est est un territoire en mutation rapide dont le développement économique fulgurant porte des atteintes importantes et sans précédent à l'environnement (Fahn, 2003 ; Sodhi et al., 2004). Les perturbations anthropiques bouleversent les relations au sein des communautés de rongeurs, affectent les relations hôtes/parasites/pathogènes et modifient le paysage épidémiologique (Thomas et al., 2009). En conséquence l'Asie



**Figure 1 :** les cercles jaunes représentent les sites échantillonnés lors de l'ANR Roboviroses à Hantavirus et les cercles jaunes lors de l'ANR CERoPath. Le cercle bicolore signifie que le site a été échantillonné lors des deux programmes.



**Figure 2 :** Régions, sous-régions biogéographiques D'après Corbet et Hill, 1992) et biomes actuels dans le foyer de diversification de la tribu des Rattini. (WWF du 06/01/2014).

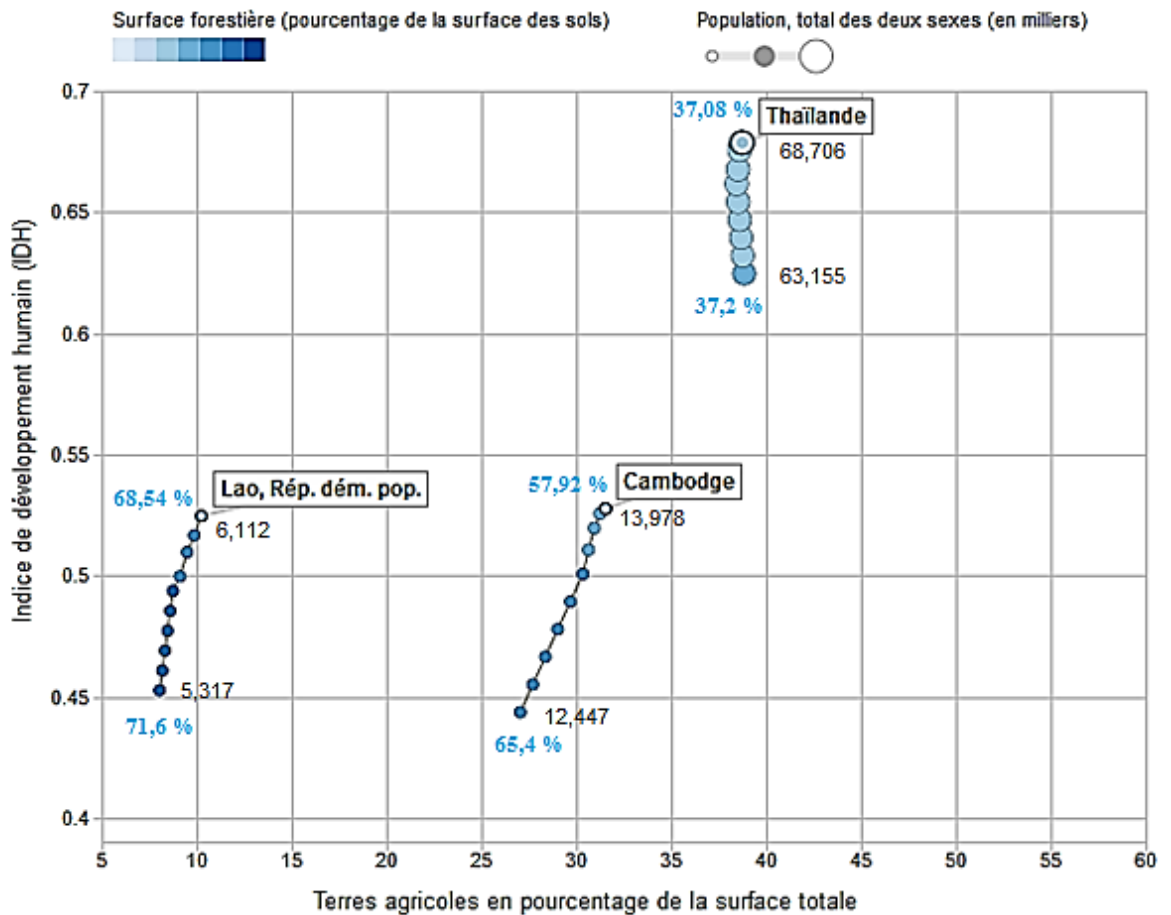
du Sud-Est est considérée aujourd'hui comme un point chaud de l'émergence des maladies infectieuses (Jones et al., 2008 ; Coker et al., 2011). La modification des communautés d'hôtes, de parasites et pathogènes associés, est en effet susceptible de conduire à l'émergence ou à la réémergence de zoonoses. Ainsi 75 % des maladies qui ont émergé dans les deux dernières décades ont pour origine la faune sauvage (Woolhouse, 2002). Les zoonoses constituent donc un sérieux défi pour la santé publique (Daszak et al. 2000). La finalité des programmes de recherche est, à plus longue échéance, de comprendre les facteurs qui sont responsables des phénomènes d'émergence ou de réémergence de zoonoses liées aux rongeurs afin d'œuvrer à la prévention des risques sanitaires qu'ils engendrent.

## *1.2. Caractéristiques des sites d'étude*

Les projets Robovirose à Hantavirus et CERoPath concernent la sous-région indochinoise de la région biogéographique Orientale (Marshall, 1977a). Cette dernière reprend dans les grandes lignes les limites de la zone indo-malayenne (Corbet et Hill, 1992). Les deux programmes couvrent trois pays : la Thaïlande, le Laos et le Cambodge (Figure 1). Différentes collaborations offrent également un accès à des échantillons de rongeurs de la sous-région sundaique (Figure 2). La sous-région indochinoise inclut la partie sub-tropicale de l'Asie du Sud-Est. Elle se caractérise par des volumes de précipitations compris entre 1000 et 2000 mm par an. Au sud de la Thaïlande, la ligne Kangar-Pattani (Figure 2), marque la transition avec la zone sundaique, au climat tropical qui recueille 2000 à 4000 mm d'eau par an. Les températures moyennes oscillent entre des valeurs inférieures à 25 °C au nord de la zone indochinoise à 29 °C pour le nord de la zone sundaique. Le climat du sud au nord est caractérisé par des saisonnalités de plus en plus marquées. Si l'on ne considère que la partie nord de la Thaïlande, qui est la seule partie de ce pays à avoir été échantillonnée par le programme CERoPath (Figure 1), on peut considérer que les trois pays étudiés présentent une situation climatique comparable et peuvent donc, à cet égard, être considérés comme des répliques spatiales. Le climat de la zone étudiée se caractérise par un régime de mousson empreint d'une saisonnalité marquée. L'année se divise en trois saisons : la saison froide (novembre-janvier) ; la saison chaude (février-juin) et la saison des pluies (juillet-octobre).

Les milieux naturels rencontrés sont principalement forestiers, entrecoupés de zones plus ouvertes couvertes de savanes arborées. Celles-ci se maintiennent naturellement par l'action des grands herbivores, du feu omniprésent en saison sèche et par le faible pouvoir colonisateur de la forêt en climat sub-tropical. Les différents types forestiers rencontrés sont très corrélés au volume et au régime des précipitations ainsi qu'à l'altitude (Gardner et al., 2000). Du sud au nord, on observe un gradient caractérisé par des types forestiers au feuillage de plus en plus caduc, la perte des feuilles étant en effet une réponse aux sécheresses provoquées par la diminution et le changement du régime des précipitations. Ainsi les forêts à feuillage persistant et mixte sont majoritaires dans la partie sundaique alors que les forêts sèches sempervirentes, mixte semi-décidue et finalement sèche caducifoliée à **diptérocarpes** se rencontrent plus au nord, dans la zone indochinoise (Marshall, 1977a, Gardner et al., 2000).

Au niveau du développement humain et de son impact sur l'environnement, les trois pays concernés ont été sélectionnés pour le projet CERoPath car ils affichent des profils économiques et environnementaux fort différents. La Figure 3 en présente une



Data from Rapport sur le développement humain 2013, Programme des Nations Unies pour le développement -

**\*IDH : Indice de développement humain**

Indice composite mesurant le niveau moyen atteint dans trois dimensions essentielles du développement humain : santé et longévité, accès à l'éducation et niveau de vie décent (voir <http://hdr.undp.org/fr/content/l'indice-du-developpement-humain-idh>).

**Figure 3 :** Comparaison, entre 2000 et 2009 (cercles) de l'évolution socio-environnementale de trois pays d'Asie du Sud-Est intégrant, le pourcentage de terres agricoles (abscisse), les surfaces forestières (coloration des cercles et chiffres en bleu), la population (taille des cercles et chiffres noir en millions d'habitants) et de l'IDH\*(ordonnée). Explorateur des données publiques de Google (<http://www.google.com/publicdata/explore?ds=kthk374hkr6tr> )

synthèse partielle. Les trois pays sont classés dans la catégorie des pays émergents. En terme de géographie et de peuplement humain, la Thaïlande arrive au premier rang avec environ 69 millions d'habitants (+8 % pour la période considérée) pour une surface de 513 120 km<sup>2</sup>. Le Laos est quasiment deux fois plus petit, il s'étend sur 236 800 km<sup>2</sup>. Sa population est plus faible, avec 6 millions d'habitants seulement mais une progression plus forte (13 % sur la même période). Le Cambodge, quant à lui, est le plus petit des trois pays, avec une surface de 181 035 km<sup>2</sup>, mais compte une population plus de deux fois plus importante que celle du Laos (14 millions d'habitants). La croissance de la population est de 11 % pour la période considérée et est intermédiaire par rapport aux deux autres pays. Le niveau de développement de ces trois pays est également contrasté. En effet, d'après le Programme des Nations Unies pour le Développement (données 2009), la Thaïlande a un indice de développement humain (**IDH** voir Figure3) nettement plus élevé que ceux du Cambodge et du Laos, qui sont tous les deux quasiment équivalents. L'évolution de l'IDH pour la période 2000/2009 est cependant plus lente en Thaïlande qu'au Cambodge et au Laos (Figure 3). En Thaïlande, les surfaces agricoles et forestières estimées en pourcentage de la surface totale sont relativement stables pour cette période. Cette stabilité est à porter au crédit de la loi de 1989, ambitieuse mais partiellement respectée, qui banni toute coupe commerciale de forêt naturelle ainsi que le changement d'usage des sols forestiers dans le pays (Usher, 2009). Les surfaces agricoles occupent un peu moins de 40 % de la surface totale et la surface forestière 37 %. Cependant la densité de la population rurale en Thaïlande avoisine 131 habitants/km<sup>2</sup> (CIA, 2012). Cette densité est élevée en comparaison des deux autres pays étudiés. Le Cambodge est un pays en mutation profonde qui connaît une expansion rapide de ses surfaces agricoles. Celles-ci sont passées de 27 à 31,5 % de la surface totale entre 2000 et 2009, et ce, au dépend de sa surface forestière (65,4 à 57,9 % sur la même période). La densité de population rurale est moins élevée qu'en Thaïlande mais reste importante avec 83 habitants/km<sup>2</sup>. Enfin le Laos, pays plus boisé et moins peuplé, connaît également une dynamique de déforestation (la forêt couvrait 71,6 % en 2000 et 68,5 % en 2009) au profit des terres agricoles (dont la couverture évolue de 8 à 10,2 % de la surface totale sur la même période). Cependant, d'après les statistiques officielles, la déforestation au Laos reste plus modérée qu'au Cambodge et la densité de population rurale est nettement inférieure à celle de ses deux pays voisins, avec 28 habitants/km<sup>2</sup>. Les transformations environnementales subies par les sites d'étude où les rongeurs ont été échantillonnés ont été intégrées dans le cadre du programme CERoPath, sur une période de vingt ans, et ce, au travers de l'analyse de séries d'images haute résolution (Dupuy et al., 2012).

Pour résumé, les localités échantillonnées dans les trois pays durant les programmes de recherche Robovirose à Hantavirus et CERoPath (Figure 1) affichent donc un profil climatique comparable mais avec un taux de pression humaine croissant du Laos à la Thaïlande avec une situation intermédiaire au Cambodge. Ils constituent de ce fait un cadre idéal pour les programmes de recherche dont l'objectif est de documenter la transformation du paysage épidémiologique et des risques associés à l'émergence des zoonoses liées aux rongeurs en fonction des changements du milieu naturel engendrés par les pressions humaines.





## ***2. Objectifs de mon travail dans le cadre des programmes de recherche***

La description des communautés de rongeurs ne peut se faire que sur la base d'une systématique claire et robuste à l'aide d'outils d'identification fiables et adaptés (Chaval et al., 2010). Dans le cadre des travaux initiés en Asie du Sud-Est par mon équipe de recherche et dans la mouvance des outils développés au sein de mon unité pour l'identification par barre-code moléculaire des organismes d'intérêt, mes missions étaient :

- d'apporter une expertise en matière d'identification des rongeurs Murinés pouvant servir de référence fiable aux travaux d'assemblage des communautés d'hôtes et de parasites;
- de fournir des outils d'identification spécifiques fiables et automatisés permettant d'offrir à des utilisateurs non spécialistes un diagnostic fiable tout en intégrant les dernières avancées taxonomiques.

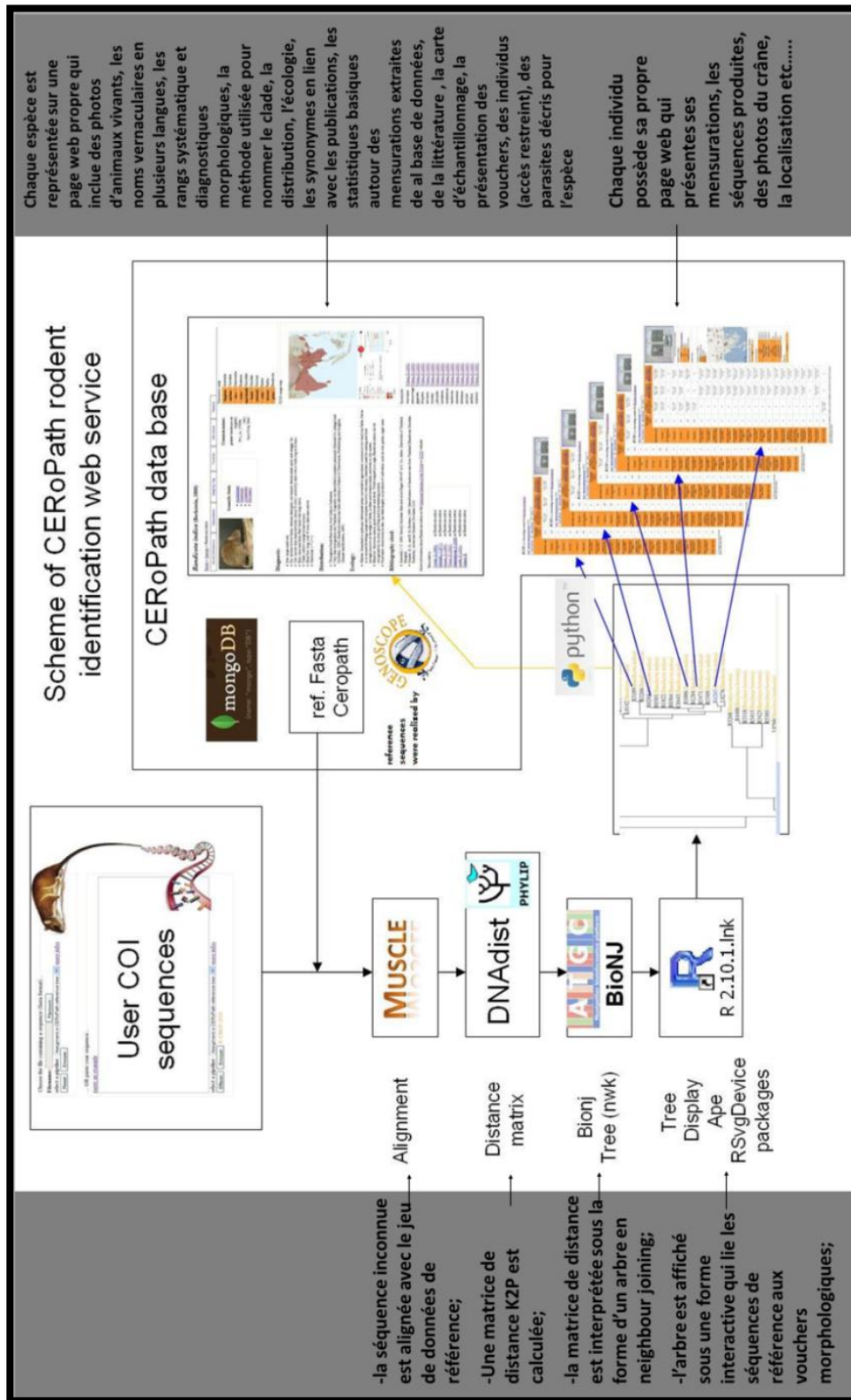
Le cadre méthodologique choisi pour accomplir ces objectifs a été celui de la systématique intégrative (Denys et al., 2003 ; Chaval et al., 2010) intégrant de ce fait différentes techniques allant de la biologie moléculaire à la morphologie et la morphométrie. Cette étude a pour finalité d'aider à la révision taxonomique et permettre le développement d'outils d'identification des rongeurs d'Asie du Sud-Est. Dans ce cadre général, j'ai ainsi pu développer un guide de terrain des murinés d'Asie du Sud-Est accessible en ligne (voir paragraphe 3.1), un outil web service d'identification moléculaire (RodentSEA ; paragraphe 3.2 et 3.3) ainsi qu'un muséum virtuel (centre d'information sur les rongeurs et les parasites ; voir paragraphe 3.4).

## ***3. Outils d'identifications***

La révision de la systématique des murinés conduite au cours des programmes de recherche m'a permis de mettre au point des outils adaptés à la fois à l'identification de terrain ([http://www.ceropath.org/references/rodent\\_field\\_guide](http://www.ceropath.org/references/rodent_field_guide) et voir paragraphe 3.1) et au laboratoire (voir paragraphe 3.2). L'ensemble des échantillons qui ont fait l'objet des études morphométriques présentées dans ce mémoire ont été identifiés via ces outils. L'annexe I présente l'état actuel des connaissances sur la systématique des taxons étudiés, la systématique suivie dans ce travail, ainsi que l'histoire biogéographique des tribus auxquelles ils appartiennent.

### ***3.1. Le guide de terrain des murinés d'Asie du Sud-Est***

J'ai réalisé un guide actualisé pour l'identification de terrain ([http://www.ceropath.org/references/rodent\\_field\\_guide](http://www.ceropath.org/references/rodent_field_guide)). Ce guide permet de répondre aux besoins propres de nos recherches mais aussi d'appuyer le travail de nos partenaires en Asie. Il correspond à une synthèse de la littérature disponible et comprend les données de morphométrie externe collectées durant nos programmes de recherche. La validité de cet ouvrage a été étendue à une zone géographique plus vaste que les trois pays considérés dans le présent travail. Il couvre l'Asie du Sud-Est continentale afin de répondre aux besoins de nouveaux programmes de recherche et de préparer l'édition d'un livre de synthèse sur les rongeurs de la zone. Il permet d'avoir une vision complète et à jour de la taxonomie des Murinés dans un format richement illustré et directement utilisable sur le terrain.



**Figure 4:** Schéma de fonctionnement de l'outil web service d'identification RODENTSEA et de son muséum virtuel associé.

### 3.2. L'outil web service d'identification moléculaire RodentSEA

Une seconde réalisation en réponse aux objectifs qui m'étaient fixés est l'outil d'identification RodentSEA. J'ai imaginé et développé cet outil en amont du travail réalisé durant mon cursus à l'EPHE. Sa réalisation est le fruit d'une collaboration avec un informaticien, Nicolas Clairon. Il a été mis en ligne en 2010 sous la forme d'un outil web-service. RodentSEA ([http://www.ceropath.org/barcoding\\_tool/rodentsea](http://www.ceropath.org/barcoding_tool/rodentsea), Figure 4) est un outil évolutif dédié à l'identification des rongeurs sauvages de la sous-région biogéographique indochinoise. Il est aisément adaptable à d'autres groupes de mammifères et/ou régions géographiques. Il permet à un utilisateur d'identifier spécifiquement un animal en soumettant en ligne sa séquence du gène mitochondrial COI. La comparaison de cette information avec une banque de données de séquences de référence permet de positionner cet individu parmi les 28 espèces Murines les plus communes de la zone (voir plus haut). Les individus à partir desquels les séquences de référence ont été produites font l'objet d'un travail de révision systématique intégratif, récent et toujours en cours (Pagès et al., 2010, 2013 ; Chaval et al., 2010, en prép. ; Bandenhorst et al., 2009, 2011, 2012, Latinne 2012a, Latinne et al., 2013b). Ainsi, les informations fournies par l'investigation du statut spécifique des taxons au travers de divers champs disciplinaires de la biologie sont recoupées et actualisées afin d'offrir en permanence à l'utilisateur l'expertise taxonomique la plus fiable possible. Un muséum virtuel ([http://www.ceropath.org/virtual\\_museum](http://www.ceropath.org/virtual_museum)) est lié à cet outil et permet un accès à la base de données (RDBSEA) du programme CERoPath. Ce muséum est divisé en deux centres d'informations. Le premier est dédié à la présentation de l'ensemble des données 'rongeurs' (<http://www.ceropath.org/rdbsea/species>) relatives aux espèces et individus collectés (6496 individus à ce jour) et parmi lesquelles figurent les référents morphologiques des séquences diagnostiques de l'outil d'identification. Le second présente les parasites (<http://www.ceropath.org/rdbsea/parasites>) portés par les rongeurs et dont certains sont susceptibles d'infecter l'Homme.

### 3.3. Fonctionnement de l'outil RodentSEA

Le schéma de fonctionnement de l'outil est présenté dans la figure 4.

RodentSEA est constitué :

- d'une interface permettant la soumission de séquences;
- d'une suite de procédures (pipeline) permettant de réaliser toutes les étapes de l'assignation d'une séquence inconnue à un jeu de séquences de référence;
- d'une interface interactive d'affichage des résultats qui prend la forme d'un phénogramme de distances génétiques et assure un accès aux données concernant les espèces et les individus de référence ;
- d'une base et d'une banque de données qui permettent, par l'intermédiaire d'un code en langage Python, l'affichage d'informations relatives aux espèces de rongeurs et de parasites ainsi que sur les individus constituant l'échantillonnage de référence de l'outil d'identification.

L'interface de soumission se présente sous la forme d'une fenêtre permettant à l'utilisateur de coller une ou plusieurs séquences du gène COI dans un format standard, le format fasta. Elle permet également le téléchargement des séquences à partir d'un fichier au même format. Une fois cette étape réalisée, l'utilisateur peut



choisir un jeu de données de référence. Ces jeux de données sont des fichiers de séquences au format fasta générés à partir d'individus ayant fait l'objet d'une expertise taxonomique. Par défaut, le fichier utilisé par l'outil est celui qui contient l'échantillonnage de référence du programme de recherche CERoPath. Il est constitué de séquences de référence provenant de l'ensemble des sites échantillonnés pour tenir compte de la variabilité intra-spécifique des espèces ainsi que des séquences produites par Pagès et al. (2010). Il est augmenté de séquences de rongeurs fournis par des partenaires et dont l'identification a été confirmée avant d'être intégrées à la base de données de référence. Un second 'pipeline' permet d'intégrer des séquences produites lors d'un travail réalisé à plus large échelle, du Sri Lanka aux îles Marquises (Robins et al., 2007). Ce dernier ne concerne que le genre *Rattus* dont l'identification morphologique est souvent délicate sur le terrain. Une fois les séquences soumises à RodentSEA, celles-ci sont insérées dans le 'pipeline' d'identification. La première étape de traitement est l'alignement des séquences utilisateurs à l'aide du programme MUSCLE (Edgar, 2004). L'alignement des séquences est le pré-requis à toute analyse phylogénétique pour s'assurer que les caractères comparés soient bien des homologies et non des analogies. L'alignement obtenu est ensuite aligné sur le jeu de données de référence qui a fait l'objet d'un pré-alignement. Cet alignement par bloc permet un gain de temps important. Une matrice de distance K2P (Kimura, 1980) est alors calculée à l'aide du logiciel DNAdist version 3.67 (Felsenstein, 2007). Cette matrice de distance est traduite en phénogramme par le logiciel Bionj (Gascuel, 1997) qui est affiché à l'aide du logiciel R (R Development Core Team, 2009). Des liens dynamiques permettent l'accès à diverses informations (voir chapitre suivant).

La technique d'identification phylogénétique incrémentée dans RodentSEA est basée sur l'utilisation du COI en tant que code-barre moléculaire des espèces (Hebert et al., 2003a, 2003b). Suite aux réflexions sur ses inconvénients et ses limites (Moritz et Cicero, 2005 ; Meyer et Paulay, 2005), la méthode d'identification moléculaire mise au point par Moritz et al. (1987) repose désormais sur un cadre méthodologique rigoureux. Le cadre proposé par Hebert a en effet évolué devant l'impossibilité d'identifier certaines parties de l'arbre du vivant avec un gène unique (Seberg et Petersen, 2009) et la faiblesse de la méthode en alpha taxonomie (Will et al., 2005). Ainsi d'autres gènes, qu'ils soient chloroplastiques (Rbcl, Matk) ou nucléaires (ITS) sont aujourd'hui utilisés (Kress, 2005 ; Kress et Erickson, 2008). Cependant pour les rongeurs murinés dont l'appartenance à une espèce peut être définie par d'autres critères morphologiques ou cytogénétiques clairement définis, le gène historique du code-barres moléculaire (COI) donne généralement des résultats d'assignation satisfaisants (Robins et al., 2007 ; Pagès et al., 2010 ; Lu et al., 2012). Toutefois, on sait aujourd'hui que des phénomènes tel que l'introgession ou le polymorphisme ancestral (voir partie 2 paragraphe 1.4) peuvent obscurcir le signal et l'arbre du COI peut alors ne pas correspondre à celui des espèces. Aussi le code-barres moléculaire devrait se limiter, lorsque d'autres informations sont disponibles pour parvenir à une expertise taxonomique intégrative (Chaval et al., 2010), à l'identification de taxons et non à la délimitation des espèces.

Lors de la création de l'outil RodentSEA, l'affichage du résultat sous la forme d'une représentation en arbre phylogénétique a été privilégié pour deux raisons. D'une part son interprétation est plus intuitive pour des utilisateurs sans compétence en biologie moléculaire et d'autre part elle est bien adaptée à un jeu de données réduit en nombre d'espèces.

Une seconde méthode qui est régulièrement utilisée pour l'identification de taxons, est fournie par l'algorithme Blastn (Basic Local Alignment Search Tool for



nucleotides, Altschul et al., 1997). Blastn est un algorithme heuristique ce qui lui confère une grande rapidité. Le principe est de découper la séquence inconnue en éléments de base d'une longueur de trois nucléotides et d'assembler ces éléments sur les séquences contenues dans les bases de données de référence (BOLD, NCBI). L'algorithme crée alors une matrice de score (coût à la similarité/dissimilarité). Le résultat est affiché sous la forme d'un taux de recouvrement et d'un pourcentage de similarité de la séquence inconnue aux séquences de la base de données de référence. Un score de 100% d'identité correspond à l'identité totale. Cette méthode rapide est bien adaptée à des jeux de données importants, elle est cependant plus difficile à lire par un public novice en biologie moléculaire.

### 3.4. *Le musée virtuel accessible depuis RodentSEA*

Le Museum virtuel ([http://www.ceropath.org/virtual\\_museum](http://www.ceropath.org/virtual_museum)) connecté à RodentSEA est un outil qui permet, pour les rongeurs, d'assurer le lien entre l'outil d'identification et les référents morphologiques des séquences de référence. Ce lien autorise l'utilisateur à porter un regard critique sur le résultat qu'il obtient en utilisant l'outil d'identification. Chaque individu de référence possède sa page propre de présentation (voir à titre d'exemple <http://www.ceropath.org/rdbsea/individual/r4265>) où sont répertoriés :

- l'âge, le sexe, les mensurations externes et crâniennes de l'individu ainsi que les mensurations de la littérature pour l'espèce à laquelle il appartient ;
- sa séquence consensus de référence du gène COI et éventuellement des séquences consensus d'autres gènes ainsi que les chromatogrammes des séquences. Ces données sont téléchargeables ;
- toute autre information relative à l'individu et fournie aux formats image et texte. Il est ainsi possible de visualiser les photos du crâne, des mâchoires, le caryotype de l'individu etc...

Deux autres pages sont également générées pour chaque individu. Elles sont accessibles sous la forme d'onglets depuis la page principale. L'une présente les informations géographiques et écologiques liées à la capture de l'individu, l'autre liste les échantillons biologiques prélevés sur l'animal et référence leur localisation dans les collections, le responsable de celle-ci de même que le mode de conservation des divers échantillons.

Un second lien renvoie vers les informations concernant l'espèce à laquelle appartient l'individu de référence sélectionné (fiche espèce). L'outil dans sa forme actuelle est capable de générer une page pour chaque espèce de mammifères décrite dans l'ouvrage de référence Wilson et Reeder (2005) mais seules les espèces de rongeurs étudiées dans nos programmes de recherche sont aujourd'hui documentées et accessibles. Une page espèce présente typiquement des photos de ses représentants vivants, son nom vernaculaire traduit en plusieurs langues, sa systématique, sa carte de répartition fournie par l'UICN ou par nos soins, les noms des espèces mis en synonymie (Musser et Carleton, 2005), une fiche de description de l'espèce, une liste de publications s'y rapportant. D'autre part, un menu nommé 'disciplines scientifiques' offre un accès vers des données de morphologie, de cytogénétique, de muséologie, de phylogénie ou de tout autre champ disciplinaire renseigné pour l'espèce concernée dans la banque de données. Une section 'mensurations' permet de





visualiser les statistiques basiques calculées à partir des individus adultes identifiés moléculairement comme appartenant à cette espèce et référencés dans la base de données RDBSEA. Ces mesures peuvent être comparées avec celles des principales publications décrivant l'espèce. Une section échantillonnage recense la localisation et les données écologiques relatives à chaque capture d'individus appartenant à l'espèce. Une section dresse la liste des individus de référence de l'outil d'identification RodentSEA et permet un accès à leur page spécifique (page individu). Enfin une partie restreinte aux membres du projet donne un accès à un moteur d'interrogation de la base de données ainsi qu'aux pages web de l'ensemble des individus.

#### ***4. Limites des travaux réalisés en amont du présent travail***

Concernant les programmes de recherche Robovirose à Hantavirus et CERoPath, les travaux réalisés s'appuient jusqu'alors sur la base de l'identification de terrain et de la délimitation d'espèces sur la base de caractères moléculaires et cytogénétiques. La morphologie des espèces et leurs mensurations sont prises en compte sur le terrain mais sans analyses permettant l'utilisation des outils modernes de la morphométrie. Cependant, l'incertitude sur la systématique des tribus étudiées et le souhait d'œuvrer dans une démarche de systématique intégrative rendent impérieuses la réalisation de telles analyses morphométriques. D'autre part RodentSEA est adapté à l'identification de tissus frais contenant de l'ADN non dégradé, cependant la plupart des collections de référence mammalogiques à ce jour sont constituées sur la seule base de matériel ostéologique. Il serait donc pertinent de pouvoir actualiser ces dernières à l'aide d'un jeu de données étalonnée à la fois sur des bases moléculaires et morphométriques.

L'objectif du travail réalisé durant le diplôme EPHE était double :

- 1- Etudier la faisabilité de la mise au point d'un outil d'identification complémentaire de RodentSEA utilisant la forme des crânes comme variable d'assignation à des groupes de référence définis sur des bases moléculaires (voir Partie 1);
- 2- Renforcer l'expertise taxonomique de RodentSEA par le croisement des approches moléculaire, cytogénétique, écologique, morphologique et morphométrique (voir Partie 2 chapitre 1);

La production d'un jeu de données morphométriques conséquent pour la réalisation de ces deux objectifs et congruent avec les travaux phylogénétiques de Pagès et al. (2010) ont permis d'étudier l'évolution des conformations de crânes au sein de la tribu des Rattini au regard de l'histoire évolutive des espèces qui la compose. Il a ainsi été possible de tester la part de l'adaptation à l'environnement et de la dérive dans cette évolution (voir Partie 2 chapitre 2)







Cette partie a pour objectif de décrire l'analyse morphométrique réalisée pour étudier la faisabilité de la mise en place d'un système d'identification semi-automatisée des espèces de rongeurs collectées en Asie du Sud-Est. L'identification des espèces a été réalisée en prenant comme base la trame systématique présentée en Annexe I. Elle intègre l'identification de terrain, des résultats de cytogénétique et la délimitation moléculaire des espèces. La morphométrie est un outil qui permet de reconnaître des entités taxonomiques. Cependant, l'étude des formes est susceptible de poser des problèmes de délimitation des taxons dans certains cas (Mayr et Ashlock, 1969). En outre, certaines espèces, dont l'isolement reproducteur est pourtant avéré, peinent à être discriminées morphologiquement. Ces espèces sont alors dites **cryptiques** (Mayr, 1963 ; voir Dobigny et al. 2002 pour un exemple chez les rongeurs ouest africains du genre *Taterillus*). De plus, la **plasticité phénotypique** des organismes, peut augmenter la **variance** des traits morphologiques. Par ailleurs, chez les mammifères, la taille et les proportions relatives des éléments du squelette d'un individu sont souvent variables au cours de son **ontogenèse** et en fonction de son sexe. Ces patrons de variation peuvent différer entre espèces ce qui rend plus délicate l'interprétation de la comparaison d'individus d'espèces, d'âges et de sexes différents lorsqu'elle est réalisée par des techniques de morphométrie traditionnelle à des fins de diagnose spécifique. Pour contourner ces difficultés, le systématicien se place parfois dans un cadre d'étude regroupant des individus de même âge et parfois de même sexe pour étudier les espèces (Marcus, 1990), ce qui, en retour, peut limiter la portée des études.

La quantification de l'anatomie par la morphométrie et l'étude des états de caractères discrets restent les outils historiques de la description, de la délimitation et de l'identification des espèces (Adams et al., 2004). Ainsi, la majeure partie de la biodiversité actuelle ou passée, conservée dans les collections des muséums d'histoire naturelle, sont souvent décrites sur les seuls caractères anatomiques. De plus, il est notable que la morphologie est un marqueur spécifique totalement indépendant de l'approche moléculaire portant sur des marqueurs mitochondriaux ou supposés neutres. Son étude est donc, dans la démarche de systématique intégrative que nous souhaitons mener, un très bon complément de la description moléculaire des taxons (voir Partie 2). De nombreuses méthodes d'analyse pour le traitement des données morphologiques sont aujourd'hui disponibles. Celles-ci vont de la représentation et de l'analyse des mesures brutes aux approches permettant de dissocier les informations de taille et de conformation contenues dans la forme des organismes. Ces méthodes ont bénéficié d'importants développements statistiques et techniques au cours du vingtième siècle (Pearson, 1901 ; Hotelling, 1933 ; Fisher, 1934; Gower, 1975 ; Bookstein, 1998). Enfin, l'avènement depuis une trentaine d'années des méthodes de morphométrie géométrique invite aujourd'hui, à la lumière des données moléculaires, à évaluer les méthodes de cette discipline pour l'identification des espèces collectées durant nos projets de recherche et la possibilité qu'elles pourraient offrir pour la mise en place d'outils d'identification web service innovants pour rendre notre expertise taxonomique accessible au plus grand nombre.

Dans cette partie, les méthodes de morphométrie qui ont été utilisées et leurs erreurs associées sont présentées (**section A**). Leur application spécifique à l'étude morphométrique des rongeurs du sud-est asiatique est développée et des protocoles permettant de minimiser l'erreur produite lors de l'acquisition des données sont fournis (**section B**). L'analyse des données morphométriques et l'impact de l'erreur sur l'assignation des individus est présentée à la **section C**. Enfin la mise en place de



l'outil lui-même, l'analyse de ses limites ainsi que les perspectives d'évolution sont détaillées à la **section D**.

## ***A- Méthodes de morphométrie utilisées dans ce travail***

### ***A1. Morphométrie traditionnelle et morphométrie géométrique***

La morphométrie traditionnelle décrit et compare la forme des objets biologiques à l'aide de mesures linéaires, de rapports de mesures ou encore de mesures d'angles. La morphométrie géométrique est basée sur des points homologues. Elle résume, à l'aide d'une configuration de points la géométrie de la morphologie des objets biologiques et cette information est préservée tout au long de son analyse (Adams et al., 2004).

#### ***A1.1. La morphométrie basée sur les mesures brutes***

La morphométrie basée sur des mesures non transformées peut être suffisamment informative pour distinguer des espèces très différentes (Marcus, 1990). Par exemple, dans notre cas, la mesure de la plus grande longueur du crâne d'un animal de l'espèce *Bandicota indica* (5 cm en moyenne) et celle d'un animal du genre *Mus* (2 cm en moyenne) sont fort différentes et permettent d'assigner un individu inconnu à l'un ou l'autre de ces taxons. Pour des espèces de tailles plus proches comme *B. indica* et *B. savilei*, la distinction ne pourra plus se faire que pour des individus adultes car leurs gammes de variation respectives sont très chevauchantes. Dans ce cas, des analyses bivariées peuvent parfois aider à assigner les animaux avec un pouvoir discriminant satisfaisant pour nombre d'applications. Ainsi les individus appartenant aux espèces *B. indica* et *B. savilei* peuvent être discriminés par le rapport de la longueur de leur corps sur celle de leur patte arrière (Aplin et al., 2003b, voir également Chaval et al., 2010 et Chaval et al., en prép). Cependant pour des espèces morphologiquement plus proches encore, comme celles du complexe *Rattus* (Musser et Carleton, 2005), la variance liée à la plasticité phénotypique, associée par exemple à l'âge et au sexe des individus ou encore à des conditions environnementales différenciées entre les populations est souvent suffisante pour rendre l'assignation d'individus à un groupe plus difficile à partir de quelques mesures seulement.

D'une façon plus générale, l'ensemble des mesures réalisées sur un objet biologique capture une partie de l'information relative à sa forme, qui est influencée par des facteurs multiples. C'est la raison pour laquelle, dans un but de discrimination, les morphométriciens utilisent souvent des techniques de transformation sur un ensemble de variables plus conséquent afin que certaines associations de variables soient les moins sensibles possibles aux sources de variation intra-spécifiques.





## A1.2. *Forme, taille et conformation*

La forme est une variable complexe qui peut être décomposée en deux variables qui sont parfois dépendantes: la taille (« size » en anglais) et la conformation (« shape »). La distinction entre forme et conformation est importante car un abus de langage répandu en français consiste à confondre la forme (« form » = shape + size) et la conformation. Lorsque la conformation se modifie avec la taille on parle alors de relations d'**allométrie**. Lorsque la conformation est indépendante de la taille on parle d'**isométrie**. L'influence de cette allométrie peut être réduite par des **procédures comme celle de Burnaby** (1966). Cette approche libre d'allométrie semble bénéfique en termes d'identification lorsque la taille interagit avec un autre facteur affectant la conformation des animaux (Claude, 2013, mais voir paragraphe C6.2.2 et C6.2.3).

## A1.3. *L'analyse en « log shape-ratio »*

Contrairement à l'analyse des données brutes, la transformation des données en log shape-ratios (Mosimann, 1970; Claude, 2013) permet d'étudier la forme d'un objet biologique en dissociant la variance associée à la taille de celle associée à la conformation de l'objet. Pour cela, un estimateur synthétique de taille doit être créé afin de transformer les variables de forme en variables de conformation.

Afin d'estimer la taille des individus, la moyenne géométrique des traits étudiés est calculée pour chaque individu :

$$\bar{x} = \sqrt[n]{x_1 * x_2 * \dots * x_n} = \sqrt[n]{\prod_{i=1}^n x_i}$$

$\bar{x}$  est la moyenne géométrique d'un nombre n de traits de valeur  $x_i$ .

Chaque trait mesuré est une variable de forme. Les variables de conformation sont obtenues en divisant chaque trait par l'estimateur synthétique de taille. On le nomme rapport de conformation ou « shape-ratio » en anglais. Les relations entre ces shape-ratios peuvent ne pas être linéaires, aussi travaille-t-on souvent sur les logarithmes des shape-ratios afin de linéariser leurs relations. Ces variables possèdent cependant certains inconvénients pour leur représentation. Elles n'ont pas de métrique directement interprétable (Marcus, 1990) puisqu'il s'agit de transformations par les logarithmes. De plus, leur représentation une à une ou multivariée ne permet pas de visualiser simplement les composantes de variation de la morphologie des animaux car les relations géométriques entre les variables ne sont pas conservées lors des analyses (Adams et al., 2004).

## A1.4. *La morphométrie géométrique*

D'autres méthodes de morphométrie appréhendant la morphologie des animaux indépendamment de leur taille (Bookstein, 1981, 1982) permettent de pallier



les inconvénients des log shape-ratio. Il s'agit notamment des approches de morphométrie géométrique qui analysent une configuration de points repères qui sont à priori considérés comme résumant la forme de l'objet biologique étudié. Dans ce travail, il sera essentiellement question de la méthode de superposition Procrustes qui dissocie la forme d'un objet en ses composantes de taille et de conformation (voir Adams et al., 2004). D'autres méthodes existent, comme les techniques d'analyses de contour (analyses de Fourier, analyses de Bézier). Bien que ces méthodes soient couramment utilisées pour décrire les contours de dents de rongeurs (Renaud et al., 1999 ; Navarro et al., 2004 ; Claude, 2013) ou les hémi-mandibules (Michaux et al., 2007), elles ne seront pas considérées dans ce travail car elles requièrent une étude plus complexe de l'erreur et de la sélection des variables de forme que celle des techniques Procrustes. De plus, la méthode de superposition Procrustes autorise l'analyse de structures plus complexes et une meilleure prise en compte de l'homologie des caractères pour les structures étudiées (voir paragraphe A1.4.1 et A1.4.2).

#### *A1.4.1. Le principe de l'analyse Procrustes*

La superposition Procrustes a pour objectif premier de rendre possible la comparaison de deux objets. L'exemple le plus didactique est celui de la comparaison de triangles (Annexe IIa), qui diffèrent par leur taille, leur position et leur orientation. Pour réaliser cette comparaison, l'information de forme doit être résumée à l'aide de points repères (Annexe IIa). Dans le cadre d'objets biologiques ces points sont souvent dit 'homologues', c'est-à-dire qu'ils doivent être positionnés, chez deux individus d'espèces différentes, à des positions structurellement ou ontogéniquement identiques. La notion d' **homologie** est ici relâchée par rapport à sa signification biologique où le caractère homologue de filiation (Lecointre et Le Guyader, 2009) est un caractère partagé par les descendants qui l'héritent de leur ancêtre commun. Ces points homologues peuvent être catégorisés en fonction de la manière dont ils sont reconnus anatomiquement (voir paragraphe suivant). Dans l'exemple de triangles, trois points homologues peuvent être positionnés (Annexe IIb), un à chaque sommet pour résumer l'information de forme. Dans ce cas, l'ensemble de l'information de forme est ainsi capturée puisqu'il est possible de reconstruire exactement l'objet à partir des trois sommets. Pour des structures plus complexes, telles que des crânes de rongeurs, le choix des points homologues est d'une grande importance car la puissance de l'analyse dépendra de l'information capturée par la position relative des points les uns par rapport aux autres (la configuration). Ces points sont définis dans l'espace cartésien directement à l'aide de coordonnées, ce qui permet de les localiser les uns par rapport aux autres dans un plan. Une troisième coordonnée peut être enregistrée afin de travailler sur des volumes mais son acquisition nécessite l'utilisation d'outil plus onéreux et engendre un temps de saisie qui peut être beaucoup plus long (cas par exemple de la technologie « ct-scan »). Cette deuxième approche moins économique financièrement et en terme de temps, n'est pas accessible au grand public actuellement, et il n'en sera donc pas question dans ce mémoire dont la motivation est de mettre les résultats à la disposition du plus grand nombre. Dans ce travail, les coordonnées des points sont numérisées à partir de photographies numériques donc en deux dimensions.



La position relative des points homologues a la propriété de rester inchangée après des transformations telles que la translation, la mise à l'échelle et la rotation. Ce sont ces propriétés qui vont être utilisées lors de la superposition Procrustes en vue de la comparaison de deux formes. Une fois la superposition réalisée, les comparaisons d'individus se font au travers d'analyses statistiques plus classiques, restant néanmoins dédiées à l'analyse des formes.

#### *A1.4.2. Les types de points homologues*

Bookstein (1991) a défini trois types de points homologues, suivant l'information biologique qu'ils portent. Les points homologues de type 1 sont positionnés sur des juxtapositions discrètes de tissus (croisement de sutures entre des plaques osseuses) et sont les seuls points homologues qui représentent de réelles homologues structurales. Les points de type 2 sont localisés sur des maximums de courbures. Ils correspondent à des similarités développementales entre individus mais leur valeur en termes d'homologies est plus faible que les précédents. Enfin, les points homologues de type 3 sont nommés pseudo-points homologues ou points homologues flous par Bookstein car ils résultent d'une construction géométrique à partir des deux types précédents. Pour Bookstein (1991) les points homologues de type 1 doivent être considérés en priorité par rapport aux deux autres types. Cependant, lors de la description des formes biologiques, certaines parties peuvent être sous-échantillonnées du fait de l'absence de juxtaposition de tissus dans certaines zones. Dans ce cas, les points homologues de type 2 et 3 s'avèrent utiles pour combler ce manque flagrant d'information (Claude, 2008). Ainsi, leur utilisation est de plus en plus admise en biologie.

#### *A1.4.3. Les trois étapes de la superposition Procrustes*

La superposition Procrustes se déroule en trois étapes : la translation, la mise à l'échelle et la rotation.

Avant de procéder à la translation, les coordonnées du centre de gravité des formes à comparer sont calculées de la manière suivante (Annexe IIc) :

Aire d'un polygone irrégulier :

$$A = \frac{1}{2} \sum_{i=0}^{n-1} (x_i y_{i+1} - x_{i+1} y_i)$$

Centre de gravité d'un polygone irrégulier :

$$x_g = \frac{1}{6A} \sum_{i=0}^{n-1} (x_i + x_{i+1}) * (x_i y_{i+1} - x_{i+1} y_i)$$

$$y_g = \frac{1}{6A} \sum_{i=0}^{n-1} (y_i + y_{i+1}) * (x_i y_{i+1} - x_{i+1} y_i)$$

La translation fait correspondre les centres de gravité de tous les objets à superposer avec l'origine (Annexe II d). Cette transformation consomme deux degrés



de liberté, car elle rend deux dimensions nulles à l'espace de variation de ces nouveaux objets (celle du barycentre des formes). Pour la réaliser, on retranche les coordonnées du centre de gravité à l'ensemble des coordonnées des points de chaque configuration :

$$(X, Y) = (x - x_g, y - y_g)$$

La mise à l'échelle va permettre d'extraire l'information de taille à partir de la forme de l'objet étudié et de transformer toutes les configurations de points de sorte qu'elles aient la même taille. On appelle alors ces configurations mises à la même échelle des pré-conformations (Annexe IIe). Cette transformation définit l'espace de pré-conformation et fait perdre également un degré de liberté (une dimension dans l'espace de variation des objets).

Pour appliquer cette transformation, une variable synthétique de taille, appelée la taille centroïde est calculée (Bookstein, 1986) : elle correspond à la racine carrée de la somme des carrés des distances entre chaque point et l'origine (i.e. le centre de gravité).

$$S = \sqrt{(x_1 - x_g)^2 + (y_1 - y_g)^2 + \dots + (x_n - x_g)^2 + (y_n - y_g)^2}$$

Chaque coordonnée est alors divisée par cette valeur :

$$(X, Y) = \left( \frac{x}{S}, \frac{y}{S} \right)$$

La dernière transformation (Annexe IIif) consiste à faire correspondre les pré-conformations de points homologues entre elles. Pour cela on fixe une référence, généralement la première pré-conformation et on lui superpose par rotation la seconde de façon à minimiser la somme des écarts aux carrés entre les points homologues. Lorsque l'on veut comparer plus de deux formes, une conformation moyenne, aussi appelée consensus est calculée : il s'agit de la moyenne arithmétique des coordonnées de chaque point pour l'ensemble du jeu de données. Les pré-conformations sont alors superposées sur le consensus (voir Dryden et Mardia, 1998), qui devient la référence, à l'aide d'un ajustement aux moindres carrés de façon à minimiser les résidus Procrustes, c'est-à-dire la somme des carrés des distances entre les coordonnées des points homologues et celles du consensus (Gower, 1975 ; Rohlf et Slice, 1990). Cette superposition est appelée la superposition Procrustes généralisée partielle. Cette transformation définit l'espace des conformations de Kendall (Kendall, 1984, 1989) et fait à nouveau perdre une dimension. Les superpositions Procrustes réalisées dans cette étude sont des pGLS (pour « partial Generalized Least-Squares Procrustes Superimposition») qui permettent de conserver l'indépendance des variables de taille et de conformation contrairement aux superpositions Procrustes non partielles (GLS) qui autorisent l'ajustement par la taille des conformations lors de la superposition.

#### *A1.4.4. Les résultats de la superposition Procrustes*

Pour résumer, la superposition Procrustes décompose la forme des objets en :

- une variable synthétique de taille, la taille centroïde ;





- une conformation définie par les coordonnées des points homologues superposés.

Elle produit :

- une conformation consensus qui est la moyenne arithmétique des coordonnées de chaque point;
- les résidus Procrustes qui correspondent à la différence entre les coordonnées en abscisse et en ordonnée de chaque conformation avec le consensus. Ce sont les variables de conformation initiale à partir desquelles est défini l'espace des conformations.

Malgré la réduction dimensionnelle de la variation des formes, les statistiques classiques ne peuvent pas directement s'appliquer sur les résidus Procrustes, car la superposition induit une métrique non euclidienne (Kendall, 1984, 1989). Pour contourner ce problème, les coordonnées des objets de l'espace des formes (ou espace de Kendall, Kendall, 1984, 1989) sont projetées dans un espace euclidien qui est tangent à celui-ci. C'est donc la projection des conformations dans l'espace euclidien qui sera étudiée. Afin de minimiser les distorsions lors de la projection, le consensus est généralement pris pour référence et placé au point tangent pour devenir le centre de l'espace des conformations. Cette projection ne peut s'opérer sans distorsion excessive que lorsque la variation entre objets est relativement faible par rapport à toutes les variations possibles. Il semble tout à fait raisonnable de supposer cette condition remplie dans ce travail, car la variation de conformation des objets biologiques est systématiquement limitée par de très nombreuses contraintes architecturales, développementales et évolutives.

## ***A2. Sources d'erreurs lors de la production des données : estimation et correction***

L'estimation de l'erreur de mesure est un point central à évaluer avant de pouvoir échanger des données entre morphométriciens ou produire un outil d'identification utilisable par des opérateurs divers (voir Dujardin et al., 2010). Une réflexion sur son appréciation et sa réduction est donc un travail préalable à toute production de jeux de données morphométriques. L'erreur de mesure se caractérise par un écart à la valeur réelle de la quantité mesurée (Rabinovich, 2005). Cet écart peut provenir de plusieurs facteurs et peut être divisé en trois composantes : la précision, l'erreur stochastique et l'erreur systématique (Arnqvist et Martenssen, 1998 ; Rabinovich, 2005).

### ***A2.1. La précision***

La précision, pour un opérateur donné, est influencée par les outils de mesures qu'il utilise. Ainsi pour des données corporelles, les réglets ont une précision de l'ordre du millimètre et les pieds à coulisse du dixième de millimètre (Fink, 1990). Pour les données de morphométrie géométrique collectées à partir de photographies, la précision est déterminée par la résolution de l'appareil photographique et la capacité de l'opérateur à placer les points de repère sur les points homologues. Les progrès techniques réalisés par les fabricants d'appareils photos font que le matériel disponible



sur le marché possède aujourd'hui une très bonne résolution, même pour des objets de la taille de crânes de rongeurs. La précision du placement des points homologues peut être influencée par la qualité de la préparation des crânes qui ne doivent plus comporter de restes de tissus ou de tendons. Pour les points situés à la périphérie de la structure, la qualité du contraste entre le sujet et le fond de la photographie influence également la capacité de l'opérateur à positionner précisément les points homologues. Des recommandations concernant la préparation des crânes (paragraphe B2.1) et les prises de vue (B.2.2 et B.2.3) sont fournies dans ce document.

## *A2.2. L'erreur de mesure stochastique*

Les sources d'erreurs stochastiques surviennent de façon ubiquiste (Arnqvist et Martensson, 1998) entre plusieurs individus mesurés et influencent la **reproductibilité** des mesures. Elles agissent sur la dispersion des données autour de leur valeur centrale. De ce fait, elles peuvent avoir un impact non négligeable sur la puissance des tests statistiques en augmentant artificiellement la variance observée. Elles peuvent, en outre, accroître l'erreur de type II qui est la probabilité de conserver l'hypothèse H0 (la variance observée du trait mesuré entre deux individus n'est pas lié à l'effet testé) alors que celle-ci est fausse.

Les sources principales d'erreur stochastiques pour un opérateur et un matériel optique unique sont :

- les conditions de préparation des spécimens qui influencent la forme d'un objet biologique d'un individu à un autre (augmentation de la variance inter-individus). Cette erreur peut devenir systématique si la variation induite est toujours la même d'un individu à l'autre;
- le positionnement de l'objet qui peut influencer la **répétabilité** du placement des points homologues (erreur intra-utilisateur);
- l'éclairage de l'objet qui peut influencer la **reproductibilité** et gêner l'utilisateur dans le placement des points homologues.

### *A2.2.1. Influence de la préparation des spécimens*

La préparation des objets biologiques étudiés (par exemple le nettoyage des os, la préparation histologique, la coloration, le montage entre lame et lamelle) est une source non négligeable possible d'erreur stochastique de mesure (Arnqvist et Martensson, 1998, Rabinovich, 2005, Dujardin et al., 2010). Elle influence la variance inter-individuelle de façon artificielle et risque d'interférer avec les facteurs que nous cherchons à mesurer et notamment d'altérer notre capacité à assigner un individu à un groupe morphologique. En effet, des différences dans le temps de cuisson des tissus, dans le degré de compression lors du montage, dans la torsion des organismes peuvent engendrer une variation de la conformation non négligeable. Ceci est particulièrement vrai pour les organismes à corps mou mais s'applique également aux squelettes de



rongeurs (fragilisation des structures connectives (sutures), mouvement des bulles tympaniques, contraction des plaques osseuses etc..). L'erreur produite lors de cette étape est généralement irréversible étant donné qu'un crâne déformé ne pourra pas être préparé une seconde fois. Il est par ailleurs impossible d'estimer cette erreur en valeur absolue, mais il est possible de la réduire au maximum en préparant les objets biologiques étudiés dans des conditions standard. Le protocole de préparation des crânes utilisé pour cette étude est présenté au paragraphe B2.1 (voir également Herbreteau et al., 2011).

#### *A2.2.2. Positionnement et éclairage de l'objet et lors de la prise de vue*

Contrairement à l'effet de préparation de l'échantillon, l'effet sur la variance intra-opérateur (parfois nommée intra-individuelle) du placement de l'objet et des points homologues peut être estimé dans des conditions standard (appareil photo fixé) et par un seul opérateur sur chaque objet étudié. Le protocole d'estimation de cette erreur est donné au paragraphe suivant. Le présent travail a été réalisé en deux dimensions car l'objectif est d'étudier la faisabilité d'un outil d'identification morphologique accessible au plus grand nombre sous la forme d'un web-service. L'utilisateur doit être en mesure de réaliser par lui-même les prises de vue et de pouvoir les télécharger sur un site internet dédié. Ces prises de vue doivent faire l'objet d'un grand soin pour pouvoir être comparables au jeu de données de référence constitué durant ce travail. De manière générale et pour tout type d'objet biologique, celui-ci doit être positionné au centre de la prise de vue et sa taille ne doit pas occuper plus d'un tiers de la surface de celle-ci. La déformation de l'image est en effet amplifiée sur le pourtour de la lentille (Fink, 1990). Pour l'étude spécifique des crânes de rongeurs, un protocole a été établi afin de réduire l'erreur liée au positionnement du crâne lors de la prise de vue. Il sera présenté au paragraphe B2.3.

L'éclairage, s'il varie d'une séance de prises de vues photographiques à l'autre peut influencer la reproductibilité des données. En effet, l'ombre portée de l'objet sur le support peut atténuer le contraste de la structure étudiée et ainsi engendrer une variation du placement des points par l'utilisateur. La source de lumière doit donc être multiple pour éviter les ombres portées de l'objet sur le support et garantir ainsi le meilleur contraste. Les sources lumineuses utilisées dans ce travail et leur positionnement seront explicitées dans le paragraphe B2.2.

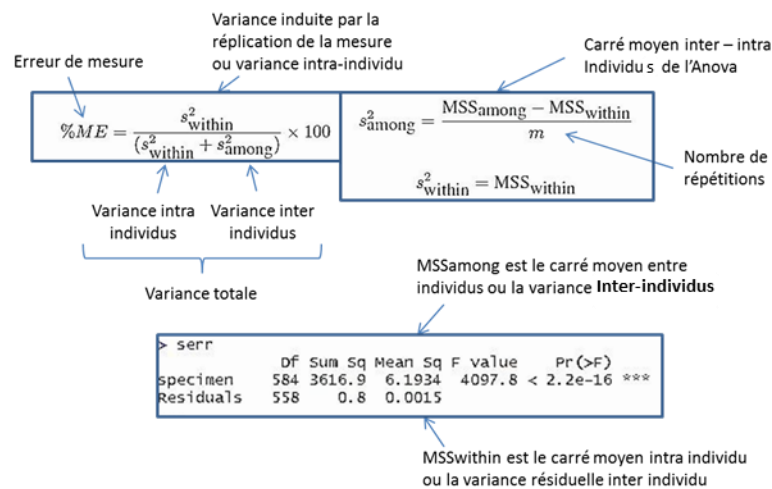
#### *A2.2.3. Protocole d'estimation de l'erreur de mesure intra-utilisateur*

Chaque pièce ostéologique étudiée a été photographiée et digitalisée deux fois de façon indépendante (crâne positionné *de novo* entre les deux prises de vues) afin de pouvoir calculer l'erreur commise lors du placement de l'objet et de la digitalisation des points homologues (reproductibilité).

L'erreur de mesure lors de la réalisation d'une mesure linéaire ou engendrée par le placement de l'objet lors de la prise de vue et le positionnement des points homologues peut être estimée en suivant le protocole de Yezerinac (1992). Dans ce protocole, l'erreur intra-opérateur ou erreur de mesure (ME) est généralement exprimée en pourcentage de variance par rapport à la variance totale.



Le schéma ci-dessous synthétise le calcul de l'erreur de mesure d'après le protocole de Yezerinac (1992) :



Yezerinac, S. M., Loogheed, S. C. and Handford, P. 1992. Measurement error and morphometric studies: statistical power and observer experience. *Systematic Biology* 41: 471-482.

Pour l'erreur de conformation, une approche de **MANOVA** Procrustes, qui est une adaptation multivariée du protocole de Yezerinac a été suivie ici (Klingenberg and McIntyre, 1998).

### A2.3. L'erreur de mesure systématique

L'erreur systématique existe pour un opérateur donné mais elle est d'autant plus importante à prendre en compte que l'on souhaite comparer les données produites par deux opérateurs différents. Dans la suite de ce travail, l'opérateur ayant produit le jeu de données de référence sera appelé l'Opérateur de référence (**Opr**) tandis que l'opérateur qui interroge l'outil d'identification morphométrique sera appelé l'utilisateur. Cette nouvelle source d'erreur de mesure de nature systématique est liée principalement :

- au positionnement de l'appareil photo lors de la prise de vue ;
- à la différence de matériel optique et de la distance focale utilisés lors de la constitution des jeux de données de référence et par l'utilisateur ;
- l'estimation propre à chaque opérateur du positionnement des points homologues.

#### A2.3.1. Influence du positionnement de l'appareil photographique

La réduction de dimensionnalité (i.e, le passage de trois à deux dimensions) lors de la prise de vue photographique peut influencer l'erreur intra-utilisateur en l'absence d'une parfaite orthogonalité du plan principal de l'objet photographié et de l'axe dans lequel est pris la photographie. Le positionnement de l'objet photographié et du matériel optique utilisé pour réaliser la prise de vue a donc une grande importance. Cette variance peut être réduite en suivant des règles précises.





Ainsi les méthodes que nous avons employées pour le placement de l'appareil photographique d'une part, et des pièces ostéologiques d'autre part sont présentées respectivement aux paragraphes B2.2 et B2.3

#### *A2.3.2. Effet 'appareil photographique'*

Tout appareil photographique engendre une distorsion qui est liée à la distance focale employée lors de la prise de vue ainsi qu'à la nature même des lentilles utilisées par le fabricant. Les lentilles photographiques convergentes sont de forme convexe et les déformations sont généralement plus importantes à la périphérie de celles-ci qu'en leur centre. C'est la raison pour laquelle, en suivant les recommandations de placement de l'objet (paragraphe B2.3), l'effet de la lentille ne devrait pas engendrer de déformation. Cependant, le cas échéant, la distorsion sphérique pourra être corrigée par l'outil d'identification pour un certain nombre d'appareils photographiques (paragraphe suivant et Annexe III).

#### *A2.3.4. Correction de la déformation de l'image*

La différence de lentille est une part non négligeable de la variance induite lors de la prise de vue. L'outil d'identification devra donc être capable de corriger les défauts des appareils des utilisateurs.

La distorsion sphérique, qui est le principal défaut engendré par les lentilles concaves des appareils photos, peut être décrite par l'équation :

$$r_{src} = (a \times r_{dest}^3 + b \times r_{dest}^2 + c \times r_{dest} + d) \times r_{dest}$$

$$\text{avec } d = 1 - (a + b + c)$$

Où  $r_{src}$  est la distance entre un pixel et le centre de l'image source et  $r_{dest}$  est la distance correspondante sur l'image corrigée. Le paramètre  $d$  est un paramètre de mise à l'échelle de la photographie. Si l'échelle des deux photos est la même ( $d=1$ ) et que  $a=b=c=0$ , alors l'image n'est pas modifiée. Les paramètres  $a$ ,  $b$  et  $c$  sont des paramètres de distorsion de l'image. Des distances négatives pour ces paramètres éloignent les points du centre alors que des valeurs positives les en rapprochent. Le paramètre  $a$  n'affecte que les points les plus à l'extérieur de l'image. Le paramètre  $b$  affecte l'image de façon plus uniforme. Le paramètre  $c$  contrôle des formes de distorsion plus complexes. Les trois paramètres  $a$ ,  $b$ ,  $c$  sont généralement inférieurs à 1.

Afin de corriger ces aberrations optiques, une base de données libre contenant les caractéristiques d'un certain nombre d'appareils photographiques a été téléchargée (PTLensDB\_06-02-08 :

<http://sourceforge.net/projects/hugin/files/PTLens%20Database/>). Par ailleurs le logiciel en ligne de commande (et donc utilisable dans un outil web-service)



Imagemagick v.6.7.3.3, permet de corriger les défauts de toutes les lentilles listées dans l'Annexe III.

#### *A2.3.5. Effet 'opérateurs'*

L'utilisateur et l'Opr ont chacun une façon qui leur est propre de placer les points homologues. Aussi la variance inter-opérateurs peut-elle être importante et parfois même dépasser la variance inter-individuelle. Ce dernier facteur d'erreur a été avancé comme étant un frein à la possibilité d'échanger des jeux de données de morphométrie géométrique concernant des espèces à faible variance interspécifique (Dujardin et al., 2010). Cependant, il est également montré que des opérateurs expérimentés produisent moins d'erreur inter-opérateurs que des opérateurs novices (Rasmussen et al., 2001). D'autre part, les effets inter-opérateurs sont mesurables (voir chapitre C8). Partant de ces constatations, il est possible d'émettre l'hypothèse que s'il est possible d'entraîner l'utilisateur à réduire la variance inter-opérateur à un niveau de variance inférieure à celle de l'effet mesuré, alors il sera possible de proposer un outil capable de comparer deux jeux de données produits par deux opérateurs différents.

### ***A3. Traitement des données manquantes en morphométrie***

#### *A3.1. Pour les mesures linéaires*

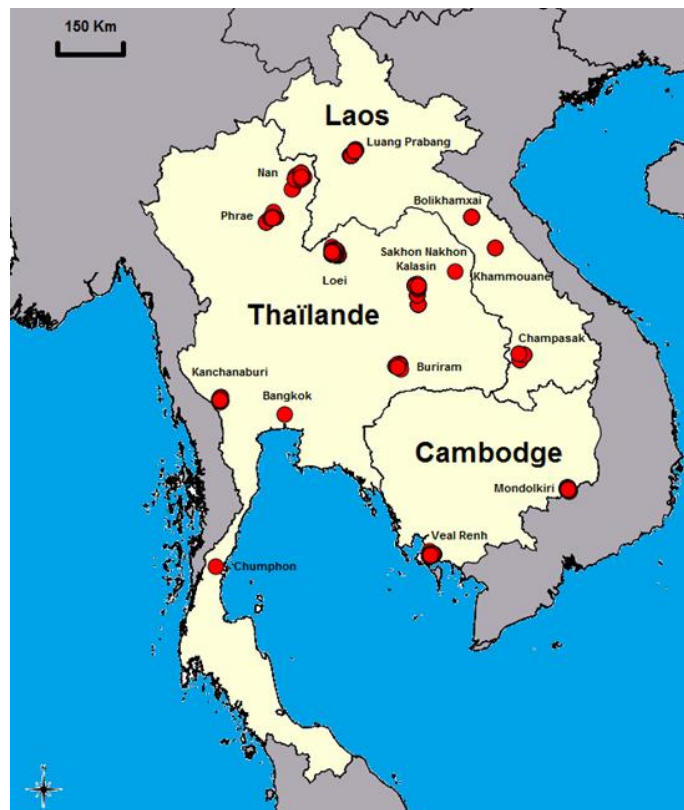
Lors du travail de terrain, un certain nombre de mesures peuvent ne pas avoir été enregistrées. Ceci peut provenir d'un oubli des expérimentateurs (1 % du jeu de données) ou de pièces anatomiques manquantes et donc non mesurables (ex. classiquement la queue chez les murinés : 11 % du jeu de données). Dans un souci de précision et de manière conservative, tous les individus où plus d'une donnée était manquante n'ont pas été inclus dans notre analyse. Pour les individus caractérisés par une seule donnée manquante, la valeur de celle-ci a été estimée par extrapolation à l'aide d'une régression linéaire décrivant les relations de ce trait avec la taille des animaux appartenant au même mitotype et au même sexe que l'animal considéré. Pour chaque trait possédant des données manquantes, une taille synthétique (la moyenne géométrique des traits), qui n'inclut pas ce trait, a été calculée (voir chapitre C2).

#### *A3.2. Pour les points homologues*

Lors de la digitalisation, chaque point homologue manquant (structure détruite) doit être noté (ID et numéro du point). Le fichier TPS est ensuite modifié en recherchant les données manquantes et en remplaçant celle-ci par « Na » (pour « non available »). Plusieurs fonctions peuvent ensuite être utilisées. Pour les objets possédant un axe de symétrie, une solution est de positionner le point homologue manquant par projection orthogonale de son symétrique par rapport à l'axe de symétrie de la structure. La fonction `Osymm` sous R réalise cette opération (<http://life.bio.sunysb.edu/morph/>). Une autre possibilité sur les structures ne possédant pas d'axe de symétrie est d'utiliser les grilles de déformation de Arcy Thompson (Thomson, 1917 ; voir partie 2 paragraphe 2.2). Le principe est de

**Tableau 1** : Nombre d'animaux par mitotype, par âge et par sexe utilisés pour le jeu de données morphométriques de référence de l'outil d'identification.

Espèces/sexe	Adultes		Sub Adultes			Jeunes		Total	
	Inconnu	F	M	Inconnu	F	M	F		M
<i>Bandicota indica</i>		15	15		11	7	2	3	53
<i>Bandicota savilei</i>		12	12		6	8	3	1	42
<i>Berylmys berdmorei</i>		18	12		3	4	1	2	40
<i>Berylmys bowersi</i>		13	10		6	6		3	38
<i>Leopoldamys edwardsi</i>		5	14	1	6	3			29
<i>Leopoldamys neilli</i>		4	23	1	6	6			40
<i>Maxomys surifer</i>		10	10		6	10	2	1	39
<i>Mus caroli</i>		4	15		4	11			34
<i>Mus cervicolor</i>		8	20		7	5			40
<i>Mus cookii</i>		8	12		10	7			37
<i>Niviventer huang</i>		13	23		6	2			44
<i>Rattus argentiventer</i>		10	11		10	9			40
<i>Rattus exulans</i>		11	9		5	6	2	2	35
<i>Rattus sakeratensis</i>		11	11		6	5	2	3	38
<i>Rattus nitidus</i>		3	2		5	3	1		14
<i>Rattus norvegicus</i>	1	8	3		5	7		1	25
<i>Rattus phylogenetic R3</i>		8	5		9	5	3	4	34
<i>Rattus tanezumi</i>		6	10		6	6	5	5	38
<b>Total</b>	<b>1</b>	<b>167</b>	<b>217</b>	<b>2</b>	<b>117</b>	<b>110</b>	<b>21</b>	<b>25</b>	<b>660</b>



**Figure 1** : Localités échantillonnées en Thaïlande, au Laos et au Cambodge pour constituer le jeu de données de référence morphométrique pour l'outil d'identification.

considérer la structure possédant une donnée manquante comme la structure cible de la déformation et de prendre comme référence une structure appartenant à un animal de la même espèce et de taille centroïde proche. Les points homologues existants servent à créer la grille de la conformation d'origine qui est ensuite déformée vers la conformation cible. Il est alors possible d'estimer les coordonnées du point homologue manquant. La fonction `estimate.missing{geomorph}` du logiciel R (R core team, 2012) réalise cette opération. Lorsque plusieurs points homologues sont manquants sur une même structure, il est déconseillé d'utiliser l'individu. Ainsi, dans ce travail quand plus d'un point était absent sur les héli-mandibules, ou un point et son symétrique sur les crânes, la pièce ostéologique concernée a été systématiquement écartée.

## ***B. Application de l'étude morphométrique aux rongeurs du Sud-Est asiatique***

Les rongeurs (voir tableau 1 p.22 ) ont été échantillonnés dans trois pays : Thaïlande, Cambodge et Laos. Pour chaque animal, cinq mensurations externes ont été prises, les crânes ont été préparés, photographiés, et des configurations de points homologues ont été digitalisées sur deux structures morphologiques (héli-mandibules et crânes) et pour deux vues : l'héli-mandibules en vue latérale interne et externe, le crâne en vue palatine (aussi appelée vue ventrale) et dorsale. Les données corporelles et les différentes configurations de points homologues ont ensuite été analysées afin de déterminer si l'une d'entre elles permettait une meilleure discrimination des différentes espèces de rongeurs étudiées.

### ***B1. Echantillonnage, sexage et critère d'âge des rongeurs de référence***

#### ***B1.1. L'échantillonnage des rongeurs***

L'échantillonnage constitué dans cette étude doit répondre à plusieurs contraintes. L'outil d'identification morphologique doit être capable d'identifier des animaux inconnus quels que soient leur sexe, leur âge et leur provenance géographique (dans la limite de l'échantillonnage de référence). En effet, l'ensemble de ces facteurs est susceptible d'influencer la morphologie externe (Claude, 2013) ou crânienne (Monteiro et al., 2003).

Afin d'avoir un échantillonnage représentatif morphologiquement du mitotype étudié et pour que la distribution des données mesurées se rapproche d'une loi normale, l'effectif de chaque mitotype a été porté à un minimum de trente individus à de rares exceptions près (tableau 1). De plus, notre échantillonnage a tenu compte de la possible variabilité géographique de la morphologie en incluant des individus capturés dans un maximum de sites distincts pour chaque espèce (figure 1). Les espèces avec des effectifs considérés comme trop faibles (*Leopoldamys sabanus*,



Figure 42: Different stages of the third molar (M3) growth (Photo: Chaval Y.)

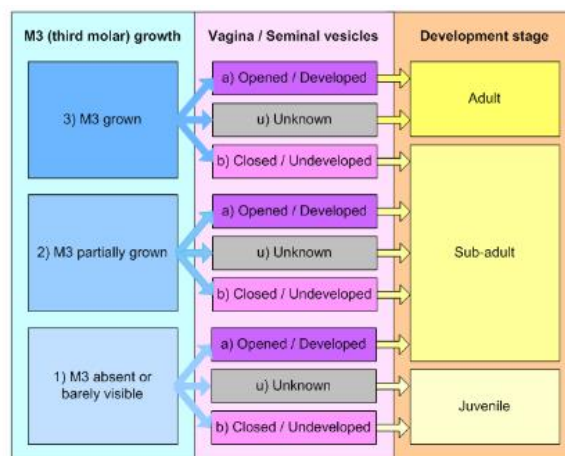


Figure 43: Decision tree to determine the maturity of a rodent, by considering the third molar growth and sexual organs externally

**Figure 2 :** les trois classes d'âges retenues dans cette étude en fonction du développement de la troisième molaire inférieure et de la maturité sexuelle (extrait de Herbreteau et al., 2011)

*Niviventer fulvescens*, *N. tenaster*, *Saxatilomys paulinae* (Musser et al., 2005b), *Chiropodomys gliroides* Blyth, 1856, *Hapalomys delacouri* Thomas, 1927, *Rattus andamanensis* Blyth, 1860, *Rhizomys pruinosus* Blyth, 1851, *Cannomys badius* Hodgson, 1841) ou les espèces échantillonnées dans des zones géographiques différentes de notre zone d'intérêt prioritaire (*Maxomys rajah*, *M. whiteheadi*, *Rattus tiomanicus* Miller, 1900, *Sundamys muelleri* Jentink, 1879 provenant de Sumatra et *Rattus rattus* Linnaeus, 1758 provenant du Niger) ont été utilisées pour tester la réaction de l'outil lorsqu'on lui présente un individu appartenant à un taxon qui n'appartient pas aux espèces de référence mais n'ont pas été inclus dans le jeu de données de référence. Celui-ci comprend donc dix-huit mitotypes appartenant aux tribus des Rattini et des Murini *sensu* Lecompte et al. (2008): *Bandicota indica*, *B. savilei*, *Berylmys berdmorei*, *B. bowersi*, *Leopoldamys edwardsi*, *L. neilli*, *Maxomys surifer*, *Mus caroli*, *M. cervicolor*, *M. cookii*, *Niviventer huang*, *Rattus argentiventer* Robinson and Kloss, 1916, *R. exulans* Peale, 1848, *R. sakeratensis*, *R. nitidus* Hodgson, 1845, *R. norvegicus* Berkenhout, 1769, *Rattus R3* (Pagès et al., 2010), *R. tanezumi* (voir le tableau 1 de l'Annexe I). Le sexe-ratio a été équilibré dans la mesure du possible, afin de limiter les effets d'un éventuel dimorphisme sexuel (tableau 1), comme souvent observé chez les murinés (Davis, 1983 ; Renaud, 2005 ; Siahsarvie et al., 2012 ; Claude, 2013). De même les individus ont été choisis de façon à ce que toutes les classes d'âge soient représentées.

### *B1.2. Détermination des classes d'âge et du sexe*

Chez les mammifères la maturité sexuelle est généralement atteinte avant le plein développement morphologique. Ceci signifie que la morphologie ne se stabilise pas avec la maturité sexuelle mais qu'elle continue d'évoluer. En conséquence, la maturité sexuelle n'est pas un critère suffisant pour classer efficacement les morphologies en fonction de l'âge (Herbretau et al., 2011). Cette classification des individus a donc été complétée à l'aide d'un critère morphologique porté par la mâchoire inférieure, à savoir le développement de la troisième molaire inférieure (M3) qui apparaît plus tardivement dans la dentition des rongeurs. Le critère d'âge a été établi en croisant le critère de maturité sexuelle et celui de développement de la M3 (figure 2). Afin de conserver des effectifs suffisants pour chaque classe d'âge, nous n'en avons distingué que trois.

Mâles et femelles peuvent être distingués en observant les organes sexuels. Pour les adultes sexuellement matures, la tâche est aisée : les testicules des mâles sont en position scrotale et masquent l'anus ; les mamelles des femelles sont visibles et la vulve est ouverte. Chez les jeunes cependant la tâche est plus ardue : les testicules des mâles sont en position abdominale et donc difficiles à détecter ; les mamelles des femelles sont invisibles ou difficiles à détecter dans l'épaisseur de la fourrure ventrale. Seul un examen de la distance entre l'anus et la papille génitale peut renseigner sur le genre des animaux. En effet, cette distance est comparativement plus grande chez les



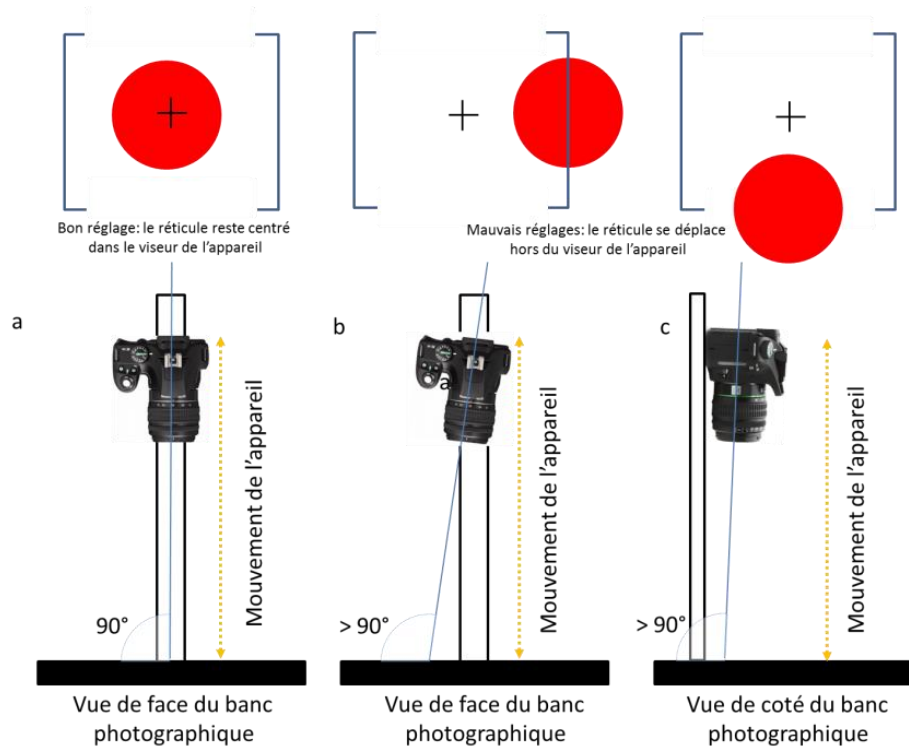


mâles que chez les femelles. Durant la dissection, la maturité des femelles peut être estimée par la présence ou l'absence d'embryons, ainsi que par la largeur de l'utérus qui est plus importante chez les individus après gestation (ce critère reste délicat à interpréter car très variable au cours du cycle menstruel). Chez les mâles la maturité sexuelle se détermine par la position scrotale des testicules et par le développement des vésicules séminales. La figure 2 résume les choix effectués dans ce travail (voir également Herbreteau et al., 2011)

## ***B2. Protocoles permettant la réduction de l'erreur de mesure lors de l'étude de crânes de rongeurs***

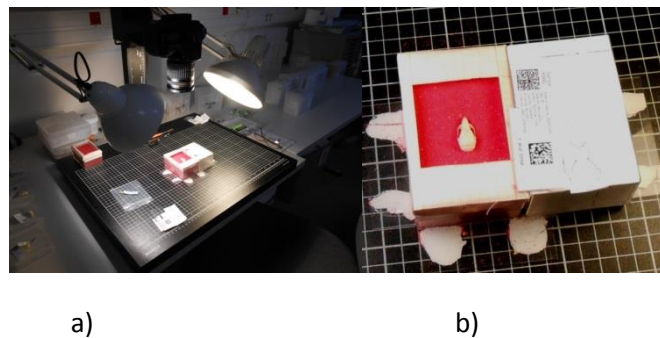
### ***B2.1. Protocole pour la préparation des crânes***

Lors de la dissection sur le terrain, le crâne est séparé du corps de l'individu à l'aide de ciseaux, en prenant soin de couper à bonne distance du foramen magnum de façon à ne pas endommager le crâne. Celui-ci est ensuite identifié à l'aide d'une marque indélébile fixée à l'aide d'un fil de nylon passé dans la trachée de façon à ce qu'elle ne puisse pas s'arracher. Le crâne est placé dans de l'éthanol à 96° pour être conservé jusqu'à sa préparation. Durant cette étude des marques plastiques imprimées en relief (Dymo ©) ont été utilisées et ont donné entière satisfaction. L'opérateur doit cependant veiller à ne pas laisser un fil de nylon trop long pour éviter que les fils de divers individus ne s'emmêlent entre eux si les crânes sont stockés collectivement dans une touque. Lors de la préparation du crâne, celui-ci est égoutté et placé dans un tube Falcon 50 ml résistant à la chaleur. Le tube est alors rempli d'eau et plongé dans l'eau bouillante pour 15 à 30 minutes suivant la taille de l'animal. Le crâne doit être cuit au minimum de façon à éviter toute déformation notamment au niveau des bulles tympaniques. L'ébullition doit être contrôlée pour que l'eau reste à peine frémissante afin que les tissus mous restent tendres et faciles à ôter par la suite. Une fois que les tissus mous entourant le crâne sont cuits, on retire le crâne de l'eau chaude et on le plonge dans un bain d'eau froide. On débarrasse le crâne de sa peau, de sa chair et des yeux en prenant soin de ne pas casser les arcades zygomatiques. Le cerveau est retiré à l'aide d'un poinçon ou d'un micro-écouvillon et en injectant de l'eau sous pression à l'aide d'une seringue. La mâchoire inférieure est séparée délicatement du crâne en prenant soin de ne pas perdre de dents qui peuvent s'être détachées durant la cuisson. L'opérateur doit ensuite enlever tous les tissus restant sur le crâne à l'aide d'une brosse à dent et éventuellement d'un peu de produit vaisselle pour les parties grasses. Une fois propre le crâne peut être blanchi à l'eau oxygénée diluée (37 %). Une fois qu'il est complètement sec, l'identifiant de l'individu est inscrit sur le crâne à l'aide d'encre indélébile, puis il est placé dans une boîte dans laquelle on prend soin de placer la marque originale apposée sur le terrain (traçabilité) afin de pouvoir corriger d'éventuelles erreurs commises lors du marquage du crâne.



**Figure 3 :** (a) bon réglage : lorsque l'on fait monter et descendre l'appareil, le réticule rouge reste à l'intérieur du viseur de l'appareil. (b et c) mauvais réglages : l'appareil photo n'est pas perpendiculaire au plan du banc photographique.

**Positionnement de l'objet biologique lors de la prise de vue.**



**Figure 4 :** (a) vue générale du montage de l'appareil et de l'éclairage. (b) crâne en vue dorsale positionné sur un lit de sable de couleur contrastée avec celle du crâne. Le papier millimétré entourant le crâne est utilisé pour donner l'échelle.

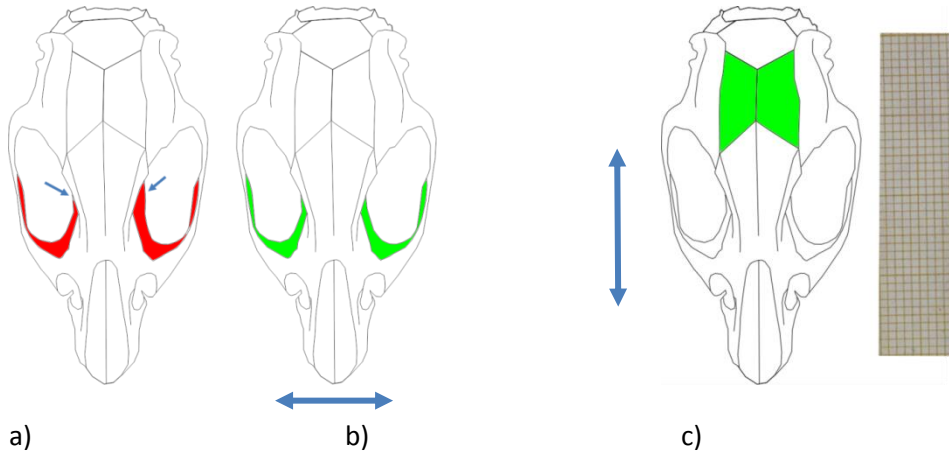
### *B2.2. Positionnement de l'appareil photographique et des sources lumineuses*

La première opération est de placer la lentille de l'appareil photographique parfaitement parallèle au plan de travail (figure 3a). Pour cela, l'appareil est monté sur un banc photographique dont l'horizontalité est testée avec un niveau à bulle. Un réticule de couleur est placée à l'endroit où se trouve l'objet à photographier. Le réticule est placé dans le viseur de l'appareil et celui-ci est élevé à l'aide de la crémaillère. Si le réticule sort du viseur de l'appareil, cela signifie que l'appareil n'est pas parfaitement positionné (figure 3b, c). Une fois l'appareil parfaitement perpendiculaire, la distance à l'objet a été fixée à 48 cm pour réaliser l'ensemble des prises de vue (figure 3a) des espèces de taille importante. Cette distance a été réduite à 30 cm pour les crânes de petite taille (genre *Mus* et *R. exulans*) de façon à ce que l'objet occupe une place assez importante (mais voir paragraphe A2.2.2) de la surface de l'image) ce qui facilite par la suite la digitalisation des points homologues. Un déclenchement légèrement retardé (2 secondes) par rapport à la pression sur le déclencheur a été programmé de façon à éviter tout tremblement de l'appareil lors du déclenchement. Dans cette étude, l'image a été enregistrée à l'aide d'un appareil photo de marque Pentax et de modèle K200D muni d'une lentille Pentax DA 16-45mm F4 ED AL. La distance focale était de 45 mm, ce qui correspond approximativement à la vision humaine. Cette distance focale doit être respectée si possible afin d'éviter les déformations de l'image (mais voir paragraphe A2.3.4). En effet un réglage macro, avec une distance focale de 17 mm par exemple produit une forte déformation de l'image (<http://www.imagemagick.org/Usage/lens/>, voir section « Examples of Lens Correction » de ce lien). Cependant si une distance focale de 45 mm ne peut être utilisée par l'opérateur une correction pourra être apportée (voir paragraphe A2.3.4). Deux sources lumineuses (Halogen eco 70 W, Megaman® SPIRAL 2700 K 20 W) ont été placées de part et d'autre de la pièce ostéologique à photographier de façon à éliminer les ombres portées (voir figure 4a).

### *B2.3. Positionnement des pièces ostéologiques lors des prises de vue*

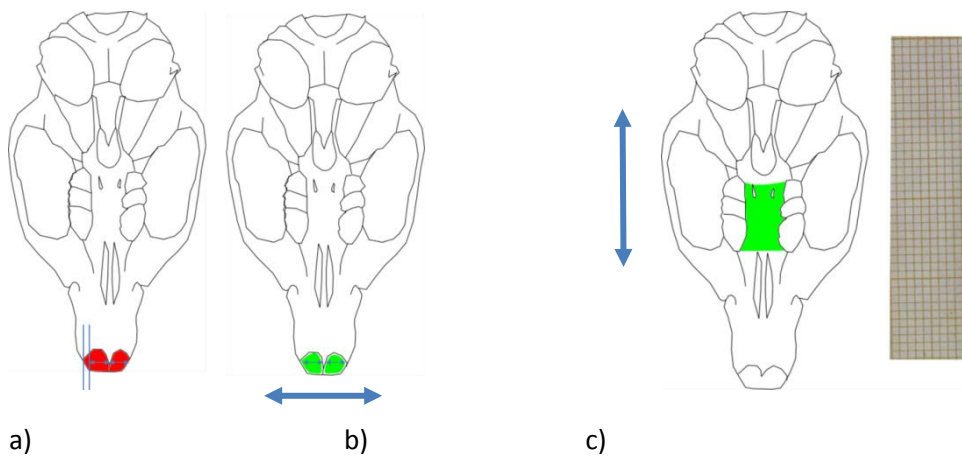
Pour cette étude, les crânes et les mandibules ont été positionnés sur un lit de sable (figure 4), de façon à permettre un positionnement du plan principal de la pièce ostéologique parfaitement parallèle à celui de la lentille. Les pièces ostéologiques sont systématiquement positionnées au centre du dispositif de façon à éviter les déformations liées à la lentille de l'appareil photographique (voir paragraphe A2.2.2). Un sable de couleur fortement contrastée avec celle du crâne, ici de couleur rouge, a été choisi pour faciliter le placement ultérieur des points homologues du pourtour de la pièce osseuse. Pour les héli-mandibules, que la vue soit interne ou externe, le positionnement ne pose pas de problème particulier car la structure est relativement plate : l'héli-mandibule est déposée sur le lit de sable parfaitement plat. Pour obtenir un lit de sable horizontal, la boîte contenant celui-ci est délicatement secouée avant chaque prise de vue pour en égaliser la surface. Une fois l'héli-mandibule positionnée de manière approximative, sa position est rectifiée à l'aide d'un poinçon de façon à ce que son plan principal soit parfaitement parallèle au papier millimétré donnant l'échelle. Pour le crâne en vue dorsale, la position du crâne est ajustée sur le lit de sable (figure 5a) de façon à ce que la partie interne de l'orbite apparaisse parfaitement

**Positionnement du crâne en vue dorsale sur le lit de sable.**



**Figure 5 :** Le crâne est déposé sur un lit de sable parfaitement plat (a). La position du crâne est ensuite rectifiée de façon à ce que la partie visible à l'intérieur de l'arcade zygomatique soit parfaitement égale de chaque côté (b). Les plaques pariétales sont alors orientées de façon à former un plan parallèle au papier millimétré (c).

**Positionnement du crâne en vue palatale sur le lit de sable.**



**Figure 6 :** Le crâne est déposé sur le lit de sable parfaitement plat (a). La position du crâne est ensuite rectifiée de façon à ce que l'espace inter incisives soit parfaitement perpendiculaire au plan d'observation (b). Le palais est alors orienté de façon à être dans le plan du papier millimétré (c).

égale des deux côtés du crâne (figure 5a,b). Le crâne est ensuite ajusté dans la troisième dimension en positionnant les os pariétaux dans le plan du papier millimétré (figure 5c). Pour la vue ventrale la démarche est similaire. On vérifie tout d'abord que les incisives soient parfaitement droites (figure 6a,b) puis on positionne le palais dans le plan du papier millimétré (fig 6c).

### ***B3. Variables morphométriques utilisées dans cette analyse***

#### ***B3.1. La prise de mesures externes***

Les mesures corporelles de terrain doivent être prises avec le plus grand soin d'autant plus qu'elles portent sur des parties molles du corps des individus sur lesquelles les instruments sont difficiles à positionner. Dans ce travail, cinq mesures ont été enregistrées pour aider à identifier les rongeurs sur le terrain (Aplin et al., 2003a ; Herbreteau et al., 2011). Ces mesures standard couplées à l'examen morphologique permettent l'identification des animaux adultes dans de nombreux cas mais peuvent s'avérer moins discriminantes dans le cas des jeunes individus et pour certains groupes d'espèces.

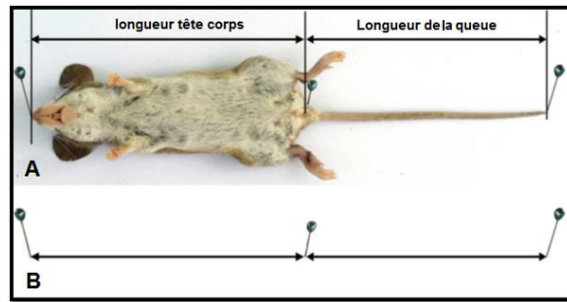
Les mesures doivent être prises immédiatement après l'euthanasie de l'animal afin d'éviter toute raideur cadavérique qui rend difficile l'enregistrement de mesures précises.

**La longueur tête corps et celle de la queue** (figure 7a et 7b page suivante) sont prises de façon concomitante, pour plus de précision car elles possèdent un point de repère en commun. L'animal est positionné sur la planche de dissection, parfaitement allongé sur le dos, la queue dans l'axe du corps, la tête aplatie sur la surface de travail. Trois épingles sont placées, au bout du museau, à l'anus (chez les mâles, le scrotum doit être soulevé pour visualiser celui-ci) et à l'apex de la queue. L'animal est ensuite retiré et les mesures sont prises à l'aide d'un réglet (précision 1 mm). La longueur tête - corps est mesurée entre les deux premières épingles et celle de la queue entre la seconde et la troisième. Cette dernière longueur n'inclut pas celle d'un éventuel pinceau terminal de la queue dont la mesure peut être notée à part si elle contient une valeur informative pour l'identification.

**La longueur de la patte postérieure** (figure 7c) est mesurée du talon à l'orteil le plus long à l'aide d'un réglet et ne comprend pas la griffe (cette dernière mesure peut être notée à part si elle revêt un caractère diagnostique dans certaines régions).

**L'oreille** (figure 7d) est mesurée de la base du conduit auditif à la limite externe du pavillon de façon à enregistrer la plus grande longueur mesurable. Elle s'effectue à l'aide d'un réglet.

**La longueur du crâne** (figure 7e) est mesurée de la crête occipitale à l'extrémité du rostre nasal à l'aide d'un pied à coulisse (précision 0,1 mm). Cette mesure bien



C



D



E

**Figure 7:** les cinq mesures prises sur le terrain pour l'identification des rongeurs (Auffray et al., 2011)

Photo du crâne du rongeur R5574 en vue dorsale.



**Figure 8:** Les photos sont transformées en négatif pour faire ressortir les sutures pariétales. Cette photo montre les codes-barres dont la lecture permet le renommage automatisé de la photo.

qu'informative en terme d'identification de terrain, est très imprécise car elle est effectuée sur des crânes non préparés et donc entourés de parties molles (muscles et cartilages). Elle s'avère cependant très utile et assez précise lorsqu'elle est ensuite recoupée avec la plus grande longueur du crâne prise sur des crânes préparés et permet notamment de déceler d'éventuels échanges de crânes lors de leur préparation.

Les mesures, qu'elles soient corporelles, crâniennes ou d'une autre nature, devraient être prises au moins deux fois par chaque opérateur pour pouvoir estimer le taux d'erreur associé à chacune d'elles (voir le paragraphe A2.2.3). Il serait alors possible de savoir si les jeux de données produits lors d'études distinctes sont comparables. Cependant la multiplication des mesures externes n'a malheureusement pas été réalisée lors du travail de terrain qui a permis la constitution de notre jeu de données et durant lequel plusieurs opérateurs ont réalisé les mesures.

### *B3.2. La prise des données de morphométrie géométrique*

#### *B3.2.1. Préparation des prises de vue à la digitalisation des points homologues*

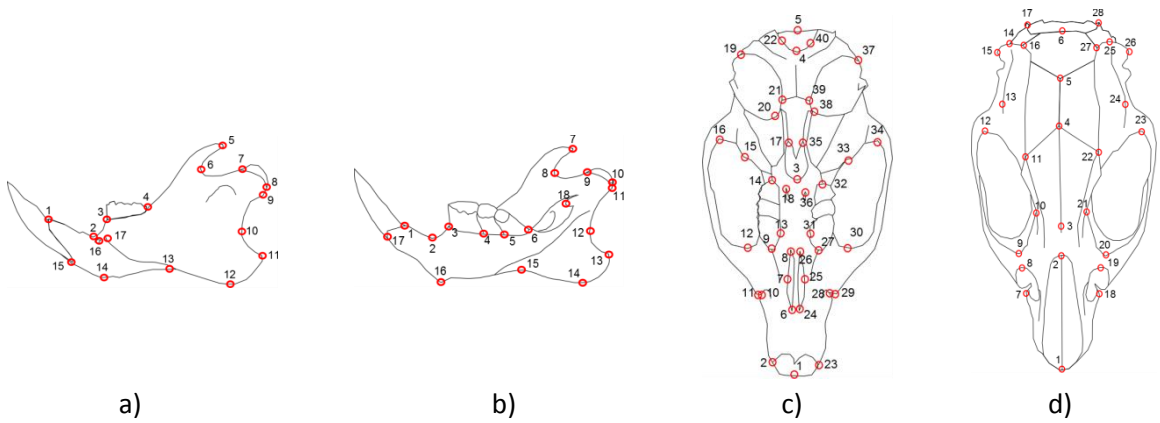
Les prises de vues peuvent être retravaillées pour faciliter le positionnement et augmenter la précision des points homologues (Rohlf, 1990). Pour cela, un logiciel libre en ligne de commande, Imagemagick 6.7.3.3 (<http://www.imagemagick.org/script/index.php>) a été utilisé pour pouvoir directement retravailler les photos à partir des scripts R et ainsi inclure le logiciel dans l'outil d'identification. Les images ont toutes été repositionnées dans le même sens pour chaque structure et les hémi-mandibules gauches ont fait l'objet d'une réflexion horizontale afin d'être transformées en « mâchoires droites » pour faciliter leur digitalisation. Les crânes en vue dorsale ont été transformés en image négative, ce qui facilite le placement des points homologues sur les sutures osseuses. Toutes les photos ont été renommées de façon automatique afin de réduire le temps passé à cette tâche ainsi que les erreurs qui pourraient découler d'un travail manuel. Pour cela, un ensemble d'utilitaires libres libdmtx 0.7.3-1 (<http://libdmtx.git.sourceforge.net/git/gitweb.cgi?p=libdmtx/libdmtx;a=summary>) a été utilisé. Ces logiciels libres en lignes de commande permettent la création et la lecture de datamatrix (code barre à deux dimensions) dans lesquels ont été codés les identifiants uniques de chaque individu. Des datamatrix ont également été produits pour coder chaque prise de vue (type de vue et numéro de la prise de vue). Ces informations ont été combinées lors des prises de vue et utilisées pour renommer les photos. A titre d'exemple la lecture du datamatrix (figure 8) contenant le numéro d'individu « R5574 » et de celui portant l'information « deuxième prise de vue dorsale » sur la photo, permet de renommer automatiquement la photographie :

R5574~2\_skull\_dorsal.jpeg

#### *B3.2.2. Digitalisation des points homologues*

La digitalisation des points homologues de l'échantillonnage de référence a été réalisée à l'aide du logiciel TPSdig2 (Rohlf, 2013). Ce logiciel permet pour chaque individu digitalisé de générer une entrée comportant le nombre de points homologues,

**Configurations de points homologues étudiées.**



**Figure 9 :** Pour la dénomination du placement des points voir annexe IV.



les coordonnées (x y) de chaque point, un identifiant unique et le nom de l'image digitalisée. Les configurations de points homologues étudiés dans ce travail en fonction des vues des pièces ostéologiques considérées sont présentées à la figure 9 La définition des points anatomiques est présentée en Annexe IV.

Les configurations de points portées par les hémi-mandibules en vue externe comprennent 17 points homologues (figure 9a). Ces points sont distribués principalement sur le contour de la mâchoire à l'exception de deux d'entre eux, le point 17 qui se trouve à l'extrémité de la crête massétérique du ramus et le point 16 qui se trouve sur le foramen mentonnier. Ces points sont de type 2 (10/17) et de type 1 (7/17).

Les configurations de points portées par les hémi-mandibules en vue internes comprennent 18 points homologues (figure 9b). Ces points sont distribués principalement sur le contour de la mâchoire à l'exception de quatre d'entre eux, distribués de part et d'autre des molaires et au niveau du foramen mentonnier. Les points de cette configuration sont de type 2 (10/18) et de type 1 (8/18).

Pour les deux vues précédentes la digitalisation du jeu de données a été partielle car il n'est constitué que des adultes et des sub-adultes de chaque espèce et n'inclut pas les individus du genre *Mus* ni ceux de l'espèce *L. neilli*.

Les configurations de points portées par les crânes en vue dorsale comprennent 28 points dont 22 sont symétriques deux à deux par rapport à l'axe de symétrie du crâne (figure 9c). Ces points ont été choisis de façon à recouvrir un ensemble de mesures précédemment utilisées pour caractériser les espèces de rongeurs en Asie du Sud-Est (voir Musser, 2006) et ont été complétés par des points portés par des sutures osseuses. Les points représentant la conformation crânienne sont de type 2 (18/28) et de type 1 (10/28).

Les crânes en vue ventrale comprennent le plus de points homologues avec 40 points dont 18 sont symétriques deux à deux par rapport à l'axe de symétrie du crâne (figure 9d). Ici encore, les points ont été choisis de façon à recouvrir un ensemble de mesures utilisées pour caractériser les espèces de rongeurs en Asie du Sud-Est (voir Musser et al., 2006). Les points représentant la conformation crânienne sont de type 2 (33/40) et de type 1 (7/40).

Chaque vue ostéologique de chaque individu a fait l'objet de deux prises de vues indépendantes (i.e. l'objet a été positionné *de novo* à chaque prise de vue) et chacune d'entre elles a fait l'objet d'une digitalisation indépendante. Pour cela, les prises de vues ont été mélangées de façon aléatoire dans le fichier TPS (le fichier contenant les coordonnées des points et les identifiants des photos) afin que deux prises de vues d'une même structure pour un même individu ne soient pas digitalisées coup sur coup.

Par la suite, l'information contenue dans ces configurations de points homologues a été comparée (voir chapitre C5) pour estimer quelle était la conformation qui avait le pouvoir discriminant le plus important pour les mitotypes considérés. Le pouvoir d'assignation de l'analyse pour les conformations de points homologues les plus discriminantes a alors été optimisé en fonction du nombre de variables de conformation considérées et simplifiées (voir chapitre C6) pour servir de jeu de données de référence à l'outil d'identification.

**Tableau 2:** Moyenne des mesures corporelles par espèces en mm +/- écart type, les effectifs des individus adultes sans données manquantes sont entre parenthèses (NA = non disponible).

Espèces	Tête Corps	Queue	Patte/arriere	Oreille	Poids	Crâne
<i>Bandicota indica</i> (46)	267 +/- 3,3	233,5 +/- 3,4	51,2 +/- 0,6	30 +/- 0,3	514 +/- 20,5	61,7 +/- 0,7
<i>Bandicota savilei</i> (138)	200,1 +/- 1,7	158,7 +/- 1,4	34,4 +/- 0,2	23,6 +/- 0,1	207,4 +/- 5,4	49,6 +/- 0,4
<i>Berylmys bermorei</i> (27)	204,1 +/- 4	162,7 +/- 2,5	36 +/- 0,6	24 +/- 0,3	208,4 +/- 10,6	47,6 +/- 0,6
<i>Berylmys bowersi</i> (18)	255,8 +/- 3,3	268,8 +/- 3,9	51,7 +/- 0,7	30,9 +/- 0,3	378,7 +/- 14,5	59,9 +/- 0,8
<i>Chiropodomys gliroides</i> (2)	90,5 +/- 0	118,5 +/- 2,5	18,5 +/- 2,5	16,5 +/- 1,5	25,4 +/- 0,1	27,5 +/- 0,7
<i>Hapalomys delacouri</i> (1)	132,5 +/- NA	144 +/- NA	23,5 +/- NA	12 +/- NA	55,3 +/- NA	36,8 +/- NA
<i>Leopoldamys edwardsi</i> (22)	237,9 +/- 2,3	339,9 +/- 4,1	45,7 +/- 1	30,2 +/- 0,3	321,3 +/- 10,1	61,2 +/- 0,4
<i>Leopoldamys neilli</i> (13)	213 +/- 3,7	265,7 +/- 2,7	40,1 +/- 0,4	28,3 +/- 0,6	242,6 +/- 10,5	56,9 +/- 0,5
<i>Leopoldamys sabanus</i> (2)	235 +/- 5	367,5 +/- 17,5	45,5 +/- 0,5	30,5 +/- 2,5	310,5 +/- 27,5	58 +/- 2
<i>Maxomys surifer</i> (43)	182,2 +/- 1,8	175,6 +/- 1,9	37,7 +/- 0,3	24,7 +/- 0,3	144,1 +/- 4,8	46,5 +/- 0,4
<i>Mus caroli</i> (99)	80,3 +/- 0,6	78,5 +/- 0,5	17,1 +/- 0,2	13,1 +/- 0,1	14,2 +/- 0,3	22,6 +/- 0,2
<i>Mus cervicolor</i> (144)	83,7 +/- 0,6	61,4 +/- 0,4	15,1 +/- 0,1	13,2 +/- 0,1	16,7 +/- 0,4	22,8 +/- 0,1
<i>Mus cookie</i> (169)	91,6 +/- 0,5	79,7 +/- 0,5	18,2 +/- 0,1	14,8 +/- 0,1	21,7 +/- 0,3	25,4 +/- 0,1
<i>Mus fragilicauda</i> (1)	76,5 +/- NA	57 +/- NA	16 +/- NA	12,5 +/- NA	13,4 +/- NA	20,2 +/- NA
<i>Niviventer huang</i> (63)	142,6 +/- 1,4	163,7 +/- 1,7	27,5 +/- 0,2	20,6 +/- 0,1	74,2 +/- 2,1	38,2 +/- 0,3
<i>Niviventer fulvescens</i> (1)	97 +/- NA	140 +/- NA	26 +/- NA	20 +/- NA	40 +/- NA	36 +/- NA
<i>Rattus andamanensis</i> (5)	178,2 +/- 3,2	198,1 +/- 4	32,3 +/- 0,6	23,6 +/- 0,6	129,2 +/- 7,8	45,5 +/- 1,3
<i>Rattus argentiventer</i> (45)	174,6 +/- 3,2	167,2 +/- 2,7	33,9 +/- 0,3	20,6 +/- 0,2	132 +/- 7,1	42,9 +/- 0,5
<i>Rattus exulans</i> (117)	115,3 +/- 0,8	126,6 +/- 0,8	22,8 +/- 0,1	15,9 +/- 0,1	36,6 +/- 0,8	31,3 +/- 0,2
<i>Rattus nitidus</i> (6)	151,4 +/- 12,2	136,2 +/- 10,9	32,8 +/- 1,3	20,1 +/- 0,9	82 +/- 17,5	39,9 +/- 1,9
<i>Rattus norvegicus</i> (9)	196,4 +/- 8	174,6 +/- 7,1	41,5 +/- 0,8	19,8 +/- 0,5	174,9 +/- 29,3	46,8 +/- 1,3
<i>Rattus rattus</i> (7)	185 +/- 4,2	212,4 +/- 5	35,4 +/- 0,6	23,8 +/- 0,6	150,9 +/- 16,4	NA +/- NA
<i>Rattus sakeratensis</i> (33)	154,6 +/- 2,2	133,2 +/- 2,1	28,1 +/- 0,3	17,2 +/- 0,2	83,6 +/- 4,2	39,2 +/- 0,5
<i>Rattus tanezumi</i> (270)	171,4 +/- 0,9	172,2 +/- 1,1	31,8 +/- 0,1	22 +/- 0,1	117,1 +/- 1,9	42,7 +/- 0,2
<i>Rhizomys pruinus</i> (1)	335,5 +/- NA	133 +/- NA	49 +/- NA	23 +/- NA	1430 +/- NA	78,1 +/- NA
<i>Saxatilomys paulinae</i> (3)	147 +/- 2,3	161,3 +/- 2,7	29 +/- 0,5	24,7 +/- 0,2	80 +/- 6,7	40,5 +/- 1,4
<i>Suncus murinus</i> (2)	131,8 +/- 6,8	75,5 +/- 0	18,8 +/- 0,2	14,2 +/- 1,8	48,6 +/- 10,4	31,5 +/- 0,6

## ***C- Résultats du traitement des données morphométriques***

### ***C1. Données morphométriques et statistiques descriptives***

Un tableau rassemble la moyenne, l'**écart-type** et l'effectif pour chaque espèce échantillonnée (tableau 2) pour chaque mesure corporelle et le poids. En plus des taxons spécifiquement ciblés au cours de notre étude, l'échantillonnage contient une espèce de soricomorphe *Suncus murinus* Linnaeus, 1766 ainsi que quelques espèces de rongeurs n'appartenant pas aux tribus étudiées. Nous avons choisi de présenter (sans toutefois les commenter plus avant) leurs statistiques descriptives afin de mettre cette information à disposition et également afin de décrire la totalité de l'échantillonnage des programmes de recherche.

Une représentation simple de la dispersion des mesures et du poids par espèce peut être réalisée sous la forme de boîtes à moustaches (voir Annexe V pour la représentation des données de mesures externes récoltées durant le travail de terrain du programme CERoPath et concernant les individus adultes uniquement). Cette représentation permet de caractériser rapidement les proportions corporelles de chaque espèce en décrivant chacun des traits par une variable de position, la médiane et une représentation de la dispersion des données autour de cette valeur centrale par quartile (chaque quartile représente la valeur prise à 25 % et 75 % de la distribution de la population totale).

Les mesures corporelles sont presque systématiquement corrélées à la taille des individus (Bookstein *et al.*, 1985), il est donc plus prudent de ne comparer que des individus d'une même cohorte avec ce type de données.

### ***C2. Etude de la taille***

Il est souvent intéressant de considérer la taille indépendamment de la conformation. Des méthodes permettant de les séparer ont donc été utilisées. Pour les données corporelles, la moyenne géométrique de la taille a été considérée alors que la taille centroïde a été considérée pour les données crâniennes. L'effet des facteurs susceptibles d'influencer chacun de ces deux estimateurs (le sexe et l'appartenance à un mitotype pour une cohorte donnée) ainsi que les relations de taille qui unissent les espèces ont été inférées. Une visualisation de la relation entre ces deux estimateurs a ensuite été réalisée afin de les comparer. Le facteur localité n'a pas été étudié car l'échantillonnage étudié est trop restreint pour que les différentes modalités des autres facteurs ne soient représentées dans chaque site. Pour cette analyse, seuls les adultes ont été sélectionnés afin d'éviter les variations **ontogéniques** qui viendraient masquer les relations spécifiques qui nous intéressent. Le jeu de données utilisé (n = 375) est le jeu de données de référence de l'outil d'identification dont le détail est donné en tableau 1 et figure 1.

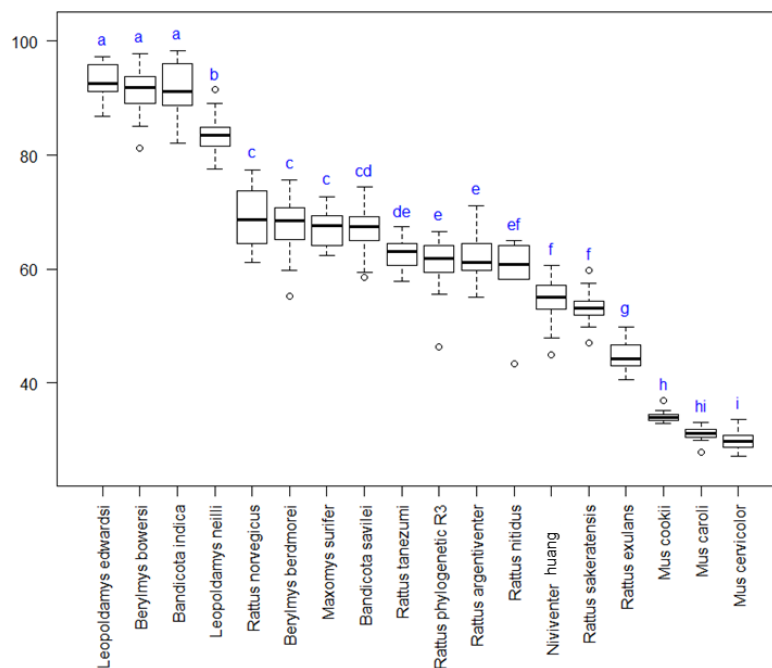
#### ***C2.1. Facteurs influençant la variation de taille corporelle des adultes***

Afin de comprendre quelles variables explicatives (parmi le sexe et le mitotype) peuvent traduire les patrons de variance observés dans le jeu de données, une analyse de variance (**Anova**) a été réalisée sur la taille corporelle des adultes (tableau 3). Quinze individus pour lesquels deux données mesurées étaient manquantes ont été

**Tableau 3:** Table d'analyse de variance de la taille corporelle des adultes en fonction du mitotype des animaux, de leur sexe et de l'interaction entre le mitotype et le sexe. **Df** est le degré de liberté, **Sum Sq** est la somme des carrés des écarts (noté **SCE**) entre les valeurs observées et leur moyenne, **Mean sq** est le carré moyen noté **CM** (avec **CM= SCE/Df**). **F** value est la valeur du test de Fisher-Snedecor, **F observé** =  $CM_{inter-modalités}/CM_{intra-modalité}$ . **F observé** est comparé au **F seuil** (lu dans la table de Fisher-Snedecor avec un risque à 5% par défaut). Si **F<sub>obs</sub> > F<sub>seuil</sub>** l'hypothèse nulle **H<sub>0</sub>** n'est rejetée pour le risque considéré. Dans le cas contraire **H<sub>0</sub>** est acceptée. **H<sub>0</sub>** est l'hypothèse nulle selon laquelle le facteur n'a pas d'effet sur la variable considérée. **Pr(>F)** est la probabilité d'obtenir une valeur égale ou supérieure à **Fobs** si **H<sub>0</sub>** était vrai.

	<b>Df</b>	<b>Sum Sq</b>	<b>Mean Sq</b>	<b>F value</b>	<b>Pr(&gt;F)</b>
mitotype	17	148754	8750	801,920	< 2e-16
sexe	1	485	485	44,490	1,04e-10
mitotype:sexe	17	382	22	2,061	0,00821
Residuals	339	3699	11		

**Représentation graphique de la distribution (médiane et quartiles de variance) des tailles composites (moyennes géométriques) des individus adultes par mitotype.**



**Figure 10:** les espèces dont la distribution des tailles composites ne sont pas significativement différentes sont rassemblées au sein des groupes matérialisés par des lettres de couleur bleue de a à i.

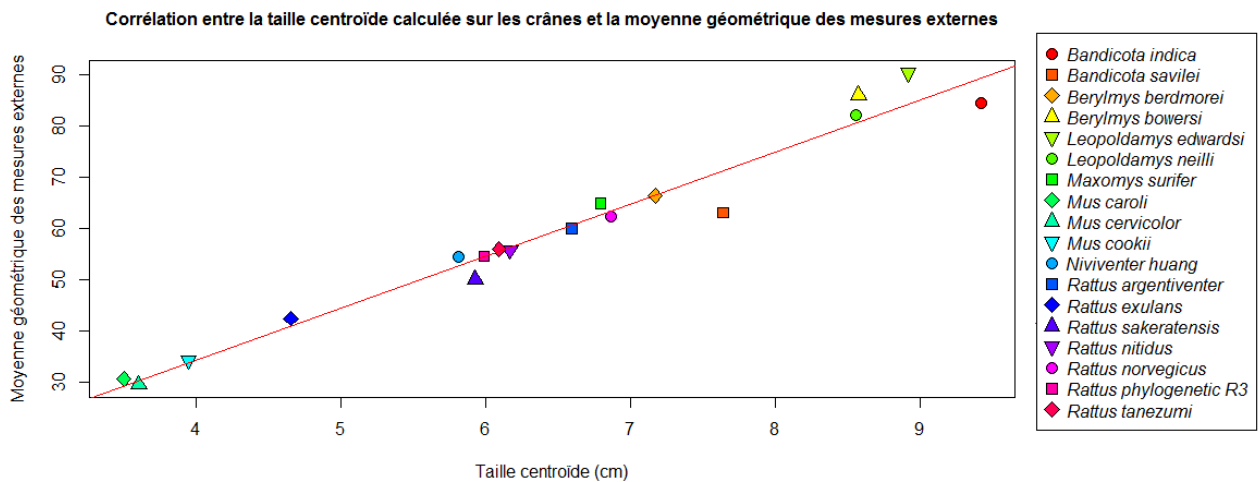
retirés des analyses. Les résultats montrent que la taille est expliquée par l'appartenance à un mitotype ( $F(17,339)= 801,92$  ;  $P < 2,210^{-16}$ ) et le sexe des animaux ( $F(1,339)= 44,49$  ;  $P < 1,0410^{-10}$ ) de manière significative et qu'il y a une interaction significative entre le sexe et le mitotype ( $F(17,339)= 2,061$  ;  $P < 0,00821$ ). Cette interaction signifie que la différence de taille entre les mâles et les femelles n'est pas toujours la même en fonction des mitotypes. Un test de permutation deux à deux de Tuckey permet de mettre en évidence la tendance qu'ont les mâles des différents mitotypes à avoir en moyenne une taille supérieure à celle des femelles, à l'exception des mitotypes *Leopoldamys neilli* et *Rattus nitidus* pour lesquels la moyenne de taille des femelles est légèrement supérieure à celle des mâles. Cependant, il ne s'agit là que d'une tendance et seuls les mâles de *B. indica* sont significativement plus gros que les femelles de la même espèce ( $P= 0,0124$ ). D'autre part le même test appliqué aux mitotypes nous permet de créer des groupes en fonction de la significativité statistique des différences de taille avec un risque de 5% (i.e. dans un groupe donné, les individus ne diffèrent pas significativement en taille, et les groupes sont ordonnés par taille décroissante). La figure 10 représente ces groupes nommés de « a » à « i » en fonction de la valeur de leur indicateur synthétique de taille. L'échantillonnage est constitué d'un groupe (a) de grosses espèces regroupant *L. edwardsi*, *B. indica* et *B. bowersi*, un groupe formé de *L. neilli* seul (b), quatre groupes (c, d, e, f) regroupant des animaux de taille moyenne : *R. norvegicus* (c), *B. bermorei* (c), *M. surifer* (c), *B. savilei* (cd), *R. tanezumi* (de), *R. argentiventer* (e), *Rattus* lignée R3 (e), *R. nitidus* (ef), *N. huang* (f), *R. sakeratensis* (f) puis un groupe (g) constitué de *R. exulans* et qui correspond aux petits animaux, et enfin deux groupes (h) et (i) de très petits animaux qui sont constitués des représentants de la tribu des Murini, à savoir *M. cookii* (h), *M. caroli* (hi) et *M. cervicolor* (i)

## C2.2. Facteurs influençant la variation de taille crânienne des adultes

Afin de comprendre quelles variables explicatives peuvent traduire les patrons de variance observés dans le jeu de données, une analyse de variance (Anova) a été réalisée sur la taille centroïde du crâne en vue palatine des adultes (tableau 4, page suivante). Par souci de comparaison, le jeu de données utilisé est le même qu'au paragraphe précédent de sorte que les mêmes individus ont été exclus. Les résultats montrent que la taille est expliquée par l'appartenance à un mitotype ( $F(17,339)= 699,514$  ;  $P < 2,210^{-16}$ ) et le sexe des animaux ( $F(1,339)= 19,857$  ;  $P < 1,1410^{-5}$ ) de manière significative et qu'il y a une interaction significative entre le sexe et le mitotype ( $F(17,339)= 1,922$  ;  $P < 0,0156$ ). Cela signifie que la différence de taille entre les mâles et les femelles n'est pas toujours à la faveur du même sexe d'un mitotype à l'autre. Un test de permutation deux à deux de Tuckey permet de nous rendre compte que les mâles des différents mitotypes ont une moyenne de taille généralement supérieure à celle des femelles à l'exception de *L. edwardsi*, *L. neilli*, *M. surifer* et *R. nitidus* et *M. cookii*. Chez ces espèces la moyenne de taille des femelles est légèrement supérieure à celle des mâles. Cependant cette tendance n'est jamais significative avec un risque de 5%. Le même test sur les mitotypes permet, comme cela avait été fait pour les mensurations externes, de créer des groupes de taille pour les différents mitotypes étudiés (données non montrées).

**Tableau 4:** Table d'analyse de variance de la taille centroïde du crâne en vue palatine des adultes en fonction du mitotype des animaux, de leur sexe et de l'interaction entre le mitotype et le sexe (se reporter au Tableau 3 pour la signification des colonnes).

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
mitotype	17	1458,8	85,81	699,514	< 2e-16
sexe	1	2,4	2,44	19,857	1,14e-05
mitotype:sexe	17	4,0	0,24	1,922	0,0156
Residuals	339	41,6	0,12		



**Figure 11:** Droite de régression entre la mesure des tailles centroïdes des mitotypes en abscisse, des tailles corporelles en ordonnée et des valeurs moyennes des différents mitotypes pour ces traits.

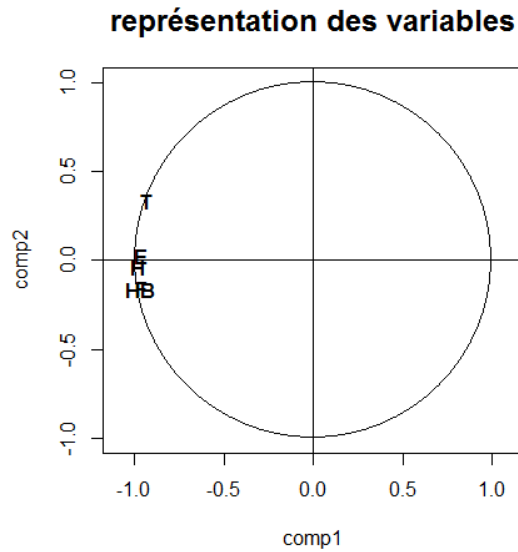
Les groupes sont nommés de « a » à « k » en fonction de la valeur de leur indicateur synthétique de taille. Ainsi, notre échantillonnage est constitué de trois groupes de grosses espèces parmi lesquelles le groupe (a) formé de *B. indica* seul, (b) constitué de *L. edwardsi* et *B. bowersi*, ce dernier mitotype appartenant également au groupe c avec *L. neilli*. Cinq groupes (d à h) représentent les espèces de taille moyenne groupant respectivement : *B. savilei* (d), *R. norvegicus* (de), *B. berdmorei* (ef), *M. surifer* (fg), *R. tanezumi* (g), *R. argentiventer* (g), *R3* (g), *R. nitidus* (gh). Enfin deux groupes (i) et (j) englobent deux petites espèces, *N. huang* (i) et *R. exulans* (j) tandis que le groupe (k) ne contient que de très petites espèces représentant les souris : *M. cookii* (k), *M. caroli* (k) et *M. cervicolor* (k). Deux entités supplémentaires sont définies ici par rapport aux groupes constitués à partir des données corporelles. Ainsi les crânes de *B. indica* sont significativement plus gros que ceux des autres mitotypes. Ceux de *L. neilli* (c) ne forment plus un groupe dont il est l'unique représentant mais se groupent avec les crânes de *B. bowersi* (bc) eux-mêmes groupés avec ceux de *L. edwardsi* (b), comme cela était déjà le cas lorsque les mensurations externes étaient considérées (voir au-dessus). Viennent ensuite les groupes (d, e, f, g, h) des espèces de taille moyenne mais dont se détache cette fois *N. huang*, alors que nous avons vu au paragraphe précédent que la taille corporelle de ce dernier le groupait avec *R. sakeratensis* et *R. nitidus*. Au contraire, *N. huang* forme désormais le groupe (i) dont il est l'unique représentant. C'est également le cas des crânes de *R. exulans* unique représentant du groupe (j), et qui s'isolait déjà avec la taille corporelle. Enfin, les représentants des Murini forment un groupe k unique alors qu'ils formaient deux groupes sur la base de la taille corporelle.

### *C2.3. Comparaison de la taille corporelle et de la taille crânienne*

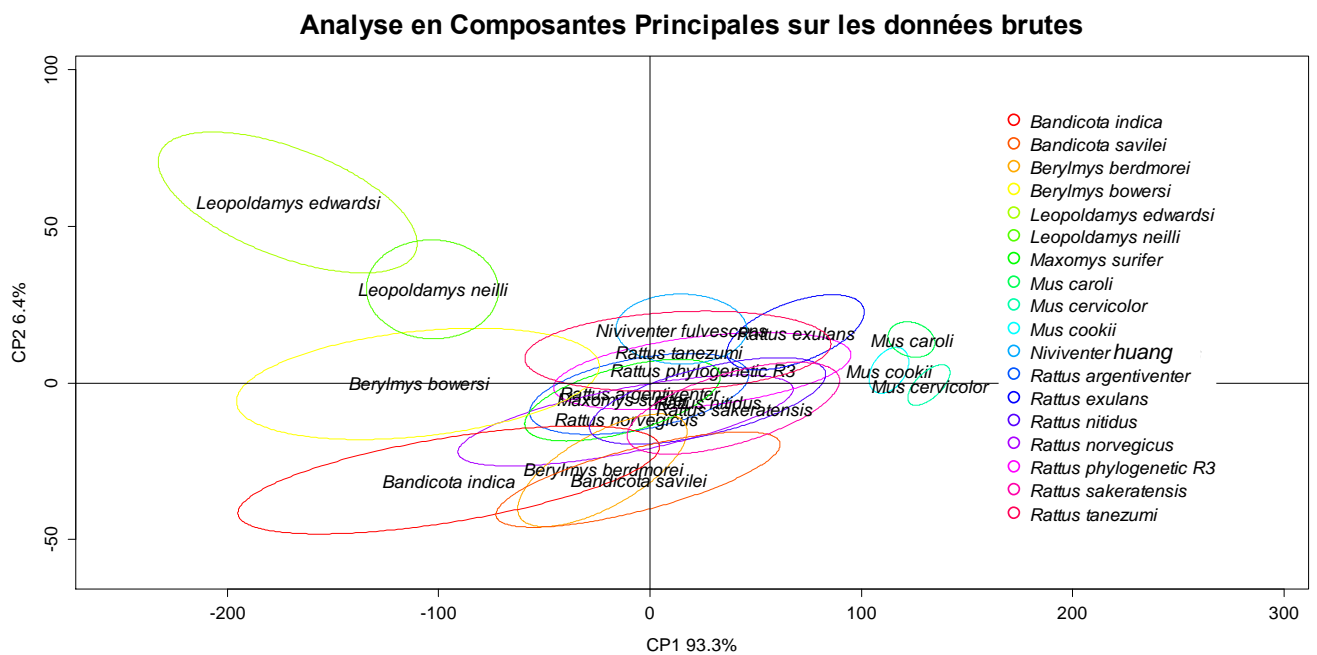
Assez logiquement, la taille crânienne et la taille corporelle sont très fortement corrélées ( $r_{\text{Pearson}} = 0,984$  ;  $t = 21,823$  ;  $df = 16$  ;  $p = 2,482 \cdot 10^{-13}$ ). Cependant, les données crâniennes semblent un peu plus discriminantes que les données corporelles puisqu'elles permettent de séparer les mitotypes en des groupes plus nombreux. La représentation de la distribution de la valeur moyenne de chaque mitotype autour de la droite de régression, de la taille corporelle sur la taille crânienne (figure 11) renseigne sur les proportions relatives des animaux. Les espèces qui sont au-dessus de cette droite présentent des crânes plus petits qu'attendus proportionnellement à la moyenne géométrique de leurs mesures corporelles, et ceux situés en dessous de la droite ont, à l'inverse, des crânes proportionnellement plus gros que leur taille corporelle ne le laisserait présager. Ainsi *B. indica*, *B. savilei* et, dans une moindre mesure, *R. sakeratensis* ont une tête proportionnellement plus grosse que ne prévoit leur taille corporelle, tandis que *L. edwardsi* et *B. bowersi* ont une taille centroïde proportionnellement plus petite que leur taille corporelle.

## ***C3. Description de la variabilité morphologique à l'aide de l'Analyse en Composantes Principale (ACP)***

Afin de caractériser la variabilité morphologique des différents jeux de données générés à partir des mêmes individus et de visualiser leur structure, différentes analyses en composantes principales (ACP) ont été réalisées.



**Figure 12:** représentation du plan des variables de l'ACP réalisée sur les données brutes. Les axes sont les composantes principales 1et 2 et le cercle est celui des corrélations.



**Figure 13:** Analyse en Composante principale sur les données corporelles brutes. Les ellipses contenant 90 % des données sont représentées. Le plan principal représente près de 100 % de l'inertie du nuage de points. La taille structure fortement le jeu de données sur la première composante principale.



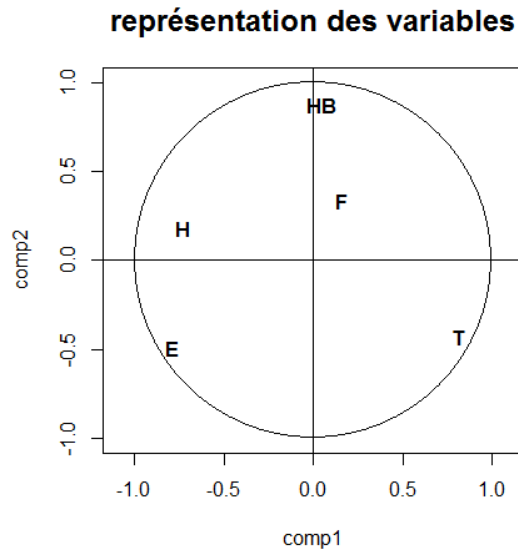
Elles permettent de mettre en évidence l'adéquation entre la systématique et la structuration des premières composantes de conformation et de forme (conformation plus taille). Dans un souci de lisibilité, seules les ellipses contenant 90 % des individus ont été représentées par la suite ; cependant, les représentations complètes sont disponibles en Annexe VI. Le jeu de données utilisé (n = 660) est le jeu de données de référence de l'outil d'identification complet dont le détail est donné en tableau 1 et figure 1

### *C3.1. L'analyse en composante principale sur les données linéaires brutes*

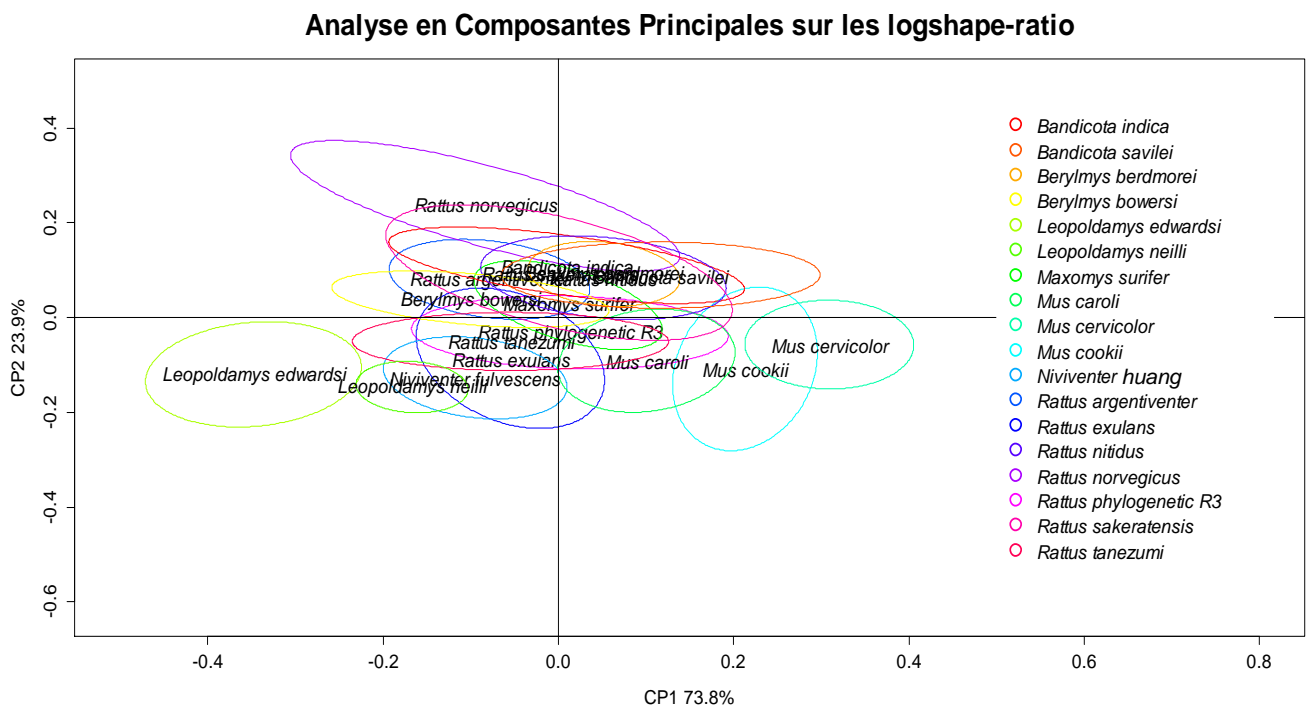
La représentation de la matrice de corrélation entre les composantes principales et les mesures corporelles (figure 12), met en évidence que la longueur de la queue (T) est la variable la moins corrélée avec les autres mesures et également la moins corrélée à la première composante. C'est aussi la variable qui contribue le plus à la seconde composante principale. L'ensemble des données brutes sont corrélées entre elles, notamment la longueur de la tête (notée H) et celle de l'oreille (E). Ces deux longueurs sont par ailleurs également très corrélées à la première composante principale. La longueur de la tête et du corps (HB) et celle de la patte arrière (F) sont très corrélées entre elles. Elles le sont également, mais dans moindre mesure que les variables H et E, avec la première composante principale. La représentation du plan principal de l'espace des individus, dont la variance au sein des mitotypes est résumée par les ellipses contenant 90% des individus dans le plan principal (figure 13), montre une forte structuration par la taille. Cette information de taille est portée par la première composante principale (Jolicoeur et Mosimann, 1960, Mosimann, 1970, Pimentel, 1979) qui capture à elle seule 93,3 % de l'inertie totale du nuage de point. On retrouve globalement le long de cette composante les groupes formés d'après les critères de tailles corporelles (figure 10). En effet, les mitotypes possédant les ordinations les plus positives sur la première composante principale appartiennent à la tribu des Murini qui forme le groupe des animaux de petite taille (figure 10). A l'inverse, le groupe des animaux de grande taille présente les ordinations les plus négatives. Enfin les groupes contenant les animaux de taille intermédiaire occupent une position centrale et se séparent assez mal les uns des autres. La seconde composante principale, qui porte 6,4 % de l'inertie totale, sépare les animaux avec une taille de queue proportionnellement plus longue (comme ceux du genre *Leopoldamys*) de ceux avec une queue proportionnellement plus courte (comme les individus du genre *Bandicota*). L'association des axes 1 et 2 nous permet d'individualiser des espèces morphologiquement très différentes comme celles du genre *Leopoldamys* et du genre *Mus*. Néanmoins, la majeure partie des taxons/mitotypes demeurent difficile à séparer sur la base de ces deux composantes.

### *C3.2. L'analyse en composantes principales sur les données log shape-ratio*

La représentation des corrélations entre les composantes principales et les log shape-ratio (figure 14) montre que les variables sont beaucoup moins corrélées entre elles, même si certaines d'entre elles le restent cependant (ex. la longueur tête corps (HB) et celle de la patte arrière (F)). Lorsque l'on regarde la contribution de chaque variable aux composantes principales, on s'aperçoit que c'est la longueur de la queue (T) et, de manière opposée, celles de la tête (H) et de l'oreille (E) qui contribuent le plus à la première composante, alors que la seconde est très influencée par la longueur HB et, dans une moindre mesure et de manière



**Figure 14 :** représentation du plan des variables de l'ACP réalisée sur les données transformées en log shape-ratio. Les axes sont les composantes principales 1et 2 et le cercle est celui des corrélations.



**Figure 15:** représentation des ellipses contenant 90 % des données corporelles transformées en log shape-ratio dans le plan principal de l'ACP.

opposée par la longueur de la queue et de l'oreille. La première composante principale capture ici 73,8 % de l'inertie du nuage de points (figure 15). L'ordination des mitotypes est toujours fonction de leur taille sur la première composante principale mais de manière beaucoup moins nette que lorsque l'on considère uniquement les données brutes (voir ci-dessus). Cela signifie que la morphologie externe des espèces est corrélée positivement avec leur taille ou, pour le dire d'une autre manière, que l'ordination des espèces est structurée par des relations d'allométrie inter-spécifiques. La seconde composante principale porte 24 % de l'inertie et sépare les animaux possédant un corps allongé, une patte arrière proportionnellement longue, une queue et des oreilles relativement plus courtes (ce qui est typiquement le cas de *R. norvegicus*), des mitotypes possédant au contraire des corps ramassés, des pattes courtes, des queues longues et des oreilles développées. Néanmoins, le plan principal fait apparaître de très nombreux recouvrements entre les ellipsoïdes des différentes entités taxonomiques et il est ainsi visuellement difficile de reconnaître une espèce par rapport aux valeurs prises sur les axes de ces ACP.

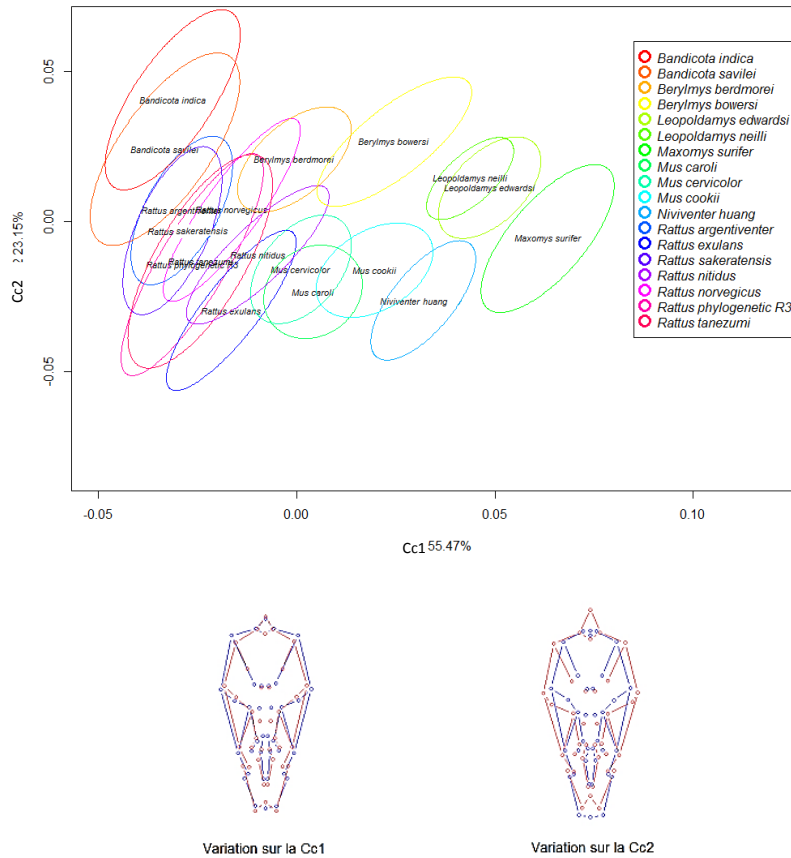
### *C3.3. Analyse en composantes principales sur les données de morphométrie géométrique du crâne en vue ventrale*

A chaque paire de coordonnées correspond l'ordination d'un individu sur les composantes principales de l'ACP. Il est possible de recalculer la conformation de points homologues correspondante à l'ordination de chaque individu. Les composantes principales définissent l'espace des conformations et sont alors nommées composantes de conformation (notées cc). Ainsi, la variation des conformations sur les composantes de conformation 1 et 2 (cc1 et cc2) ont été représentées (figure 16) ainsi que les ellipses à 90 % des mitotypes dans le premier plan principal.

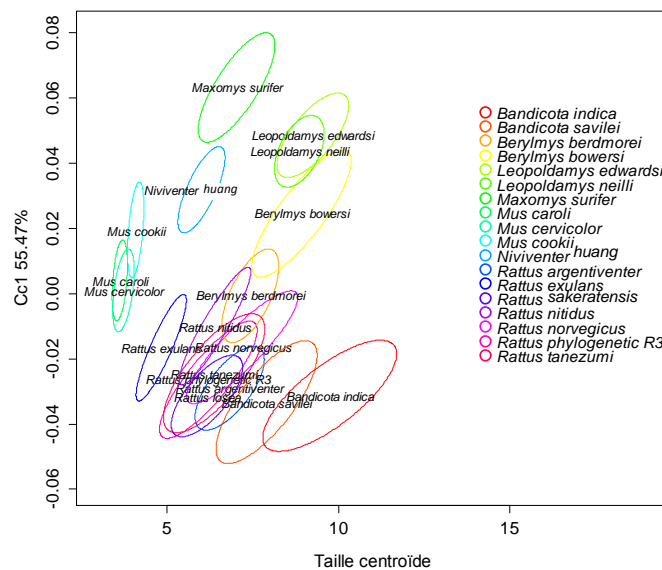
Les individus possédant des ordinations négatives sur l'axe 1 de l'espace des conformations ont des apophyses mastoïdes plus développées, des bulles tympaniques et un palais plus longs ainsi que des foramens palatins placés plus en arrière du palais que les animaux ayant des abscisses positives. De plus, ces animaux possèdent des foramens incisifs plus allongés vers l'arrière en direction de la M1, et une apophyse zygomatique plus avancée vers les incisives ; de fait, l'os prémaxillaire apparaît plus court et légèrement plus large chez ces individus. Les individus caractérisés par des valeurs négatives sur la seconde composante de conformation ont, par rapport à celles

qui ont des valeurs positives sur cette même composante, une boîte crânienne proportionnellement moins développée, un foramen magnum qui apparaît moins ouvert, un palais de même longueur mais plus fin et des foramens palatins postérieurs placés plus en arrière. L'os prémaxillaire est légèrement plus court et plus fin, tandis que les incisives sont projetées en arrière vers les foramens incisifs.

La différenciation des genres s'accroît lorsque l'on s'intéresse à la forme du crâne en fonction de la taille centroïde des individus (figure 17). Cette représentation fait apparaître clairement des différences entre les allométries ontogéniques intragroupe matérialisées ici par l'orientation du plus grand axe des ellipses. On constate que les Murini suivent une loi



**Figure 16 :** représentation des ellipses regroupant 90 % des individus de chaque mitotype dans le plan principal de l'espace des conformations crâniennes en face ventrale. La variation de conformation sur les axes est représentée en dessous, la conformation minimale en bleue et la conformation maximale en rouge



**Figure 17 :** Représentation des ellipses regroupant 90 % des individus de chaque mitotype en fonction de la taille centroïde des individus représentée en abscisse et la conformation des crânes en vue ventrale représentée ordonnées par la première composante de conformation.

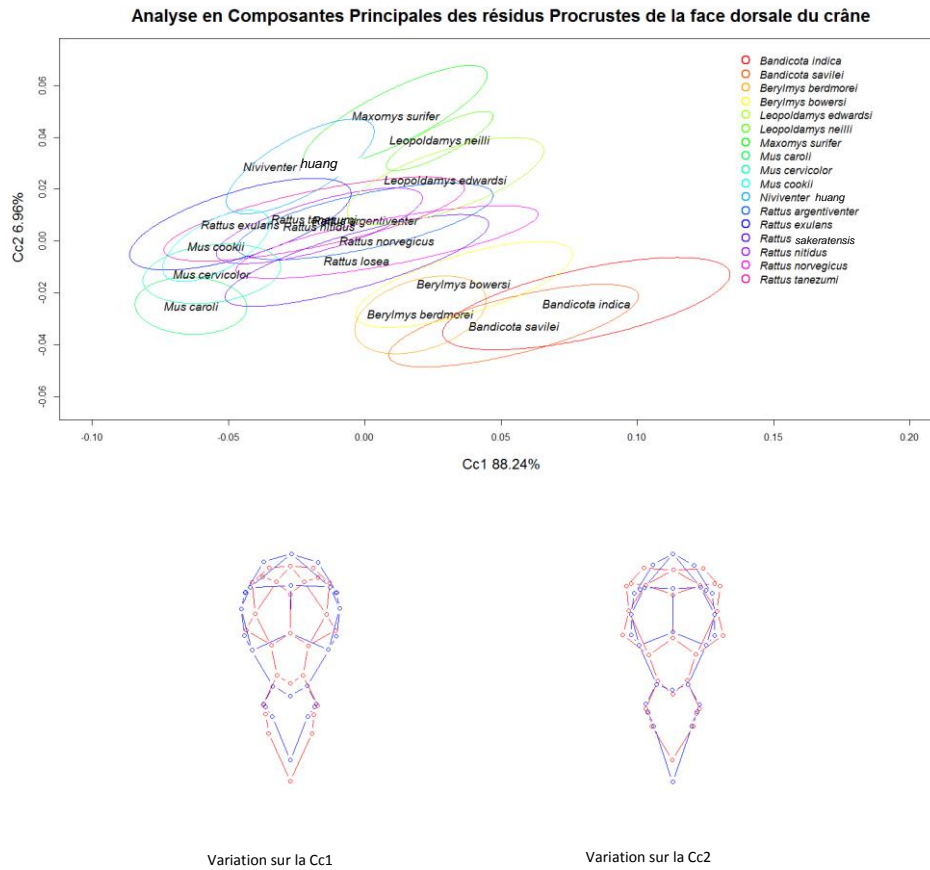
d'allométrie très différente des Rattini, ce qui suggère que la conformation des crânes de Murini varie en fonction de l'âge de manière différente de celle des crânes de Rattini. Pour les Murini, le changement de conformation en relation avec la taille se traduit par une diminution relative du volume des bulles tympaniques, un avancement du palais, des foramens palatins et foramens incisifs. Pour les Rattini, les variations de taille entre jeunes et adultes sont plus marquées et la variation de conformation se traduit principalement par un changement des proportions entre la boîte crânienne dont le volume augmente et la partie prémaxillaire qui se raccourcit. Cependant, on retrouve des traits communs avec les Murini, comme le palais et les foramens incisifs qui se raccourcissent durant l'ontogenèse et les foramens palatins qui avancent. Il existe également au sein de chaque tribu, et en particulier chez les Rattini, des différences d'allométries interspécifiques. Ces différences d'allométries ontogéniques pourraient être intéressantes en termes de systématique de nos rongeurs car elles constituent une part non négligeable des différences évolutives qui permettent de caractériser les espèces (Dujardin et al., 2010). Cependant elles le sont moins dans le cadre d'un outil d'identification où l'on cherche à assigner un individu à un nuage de points et non à dissocier deux nuages de points.

La représentation des ellipses fait par ailleurs apparaître que les animaux ne sont pas organisés par taille de crânes puisque le genre *Leopoldamys* est voisin de *N. huang* alors que leurs tailles centroïdes sont très différentes (figure 17). De plus on constate une organisation intéressante de l'espace des conformations par genre taxonomique sur le plan principal. Ainsi, les espèces du genre *Bandicota* ont des valeurs positives sur la seconde composante et négative sur la première. Les représentants du genre *Rattus* sont groupés avec une ordination proche de l'origine. Les trois espèces du genre *Mus* ont une ordination faiblement positive sur la première composante et faiblement négative sur la seconde. Les individus du genre *Berylmys* ont une ordination avec des valeurs positives sur la première composante. Cependant les valeurs de ces coordonnées sont supérieures chez *B. bowersi* par rapport à celles de *B. berdmorei*. Enfin les genres *Niviventer*, *Leopoldamys* et *Maxomys* possèdent les valeurs les plus fortement positives sur la première composante de conformation et des valeurs allant de faiblement négatives pour *N. huang* à légèrement positives pour les espèces du genre *Leopoldamys*. Il est à noter que les différences morphologiques décrites dans ce plan représentent 79 % de la variance et que 21 % restent donc à explorer sur les plans de plus haut rang.

#### *C3.4. L'analyse en composantes principales sur les données de morphométrie géométrique du crâne en vue dorsale*

La variation des conformations sur les composantes de conformation 1 et 2 (cc1 et cc2) a été représentée (figure 18) ainsi que le plan principal.

L'étude de la vue dorsale du crâne fait apparaître que les individus possédant des ordinations négatives pour la première composante de l'espace des conformations sont caractérisés par des rostres plus courts que ceux possédant des ordinations positives car la constriction inter-orbitale avance concomitamment au raccourcissement du rostre. La boîte



**Figure 18:** représentation des ellipses regroupant 90 % des individus de chaque mitotype dans le plan principal de l'espace des conformations crâniennes en face dorsale. La variation de conformation sur les axes est représentée en dessous, la conformation minimale en bleue et la conformation maximale en rouge.

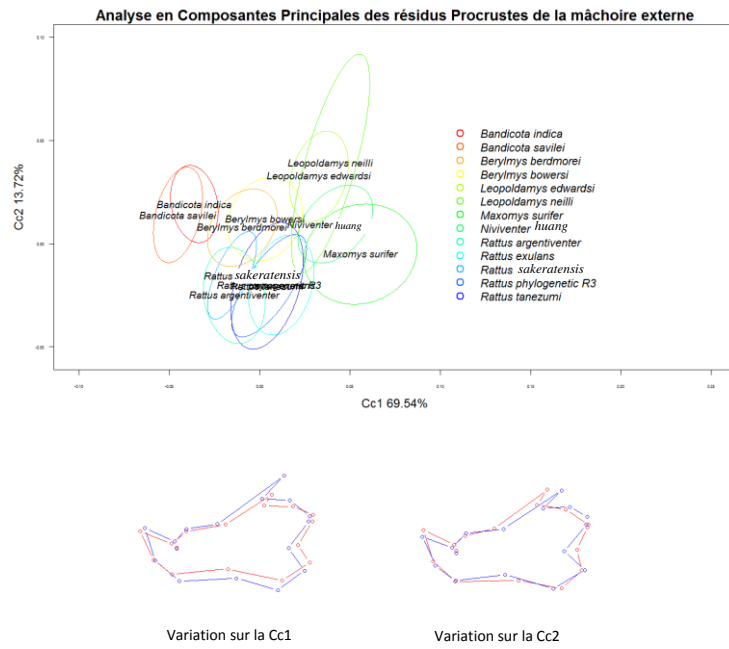
crânienne de ces individus est plus développée. Ceci est dû à un élargissement important de la partie postérieure de l'os frontal, de la plus grande largeur de la boîte crânienne, l'allongement vers l'arrière et un élargissement de l'os interpariétal ainsi qu'un recul des condyles occipitaux. Cette composante principale oppose des animaux aux morphologies très différentes comme les Murini (qui ont une morphologie proche de celle qui vient d'être décrite) et les *Bandicota* (qui ont une morphologie proche du maximum de cette composante). Sur la seconde composante, la décroissance des valeurs est caractérisée par un allongement du rostre qui est presque exclusivement dû à l'avancée de sa partie apicale. Le reste de la structure bouge peu, à l'exception d'un élargissement de la base de l'apophyse squamosale et des crêtes supra-occipitales. Enfin, l'os interpariétal est plus allongé vers l'arrière. Cette composante oppose des espèces comme *B. savilei* (qui possède la valeur minimale sur cette composante) et *M. surifer* (qui possède la valeur maximale). Cependant, la variance portée par cette composante est faible (7 % de la variance totale) et ces transformations représentent donc une très petite proportion de la variation de la morphologie du crâne en vue dorsale.

Les ellipses contenant 90% des individus pour chacun des mitotypes dans le premier plan nous montrent, comme pour la vue ventrale, un intéressant partage de l'espace des conformations entre les genres taxonomiques. Tout d'abord, les deux espèces du genre *Bandicota* possèdent des ordinations positives sur la première composante et négative sur la seconde. A proximité de ces deux espèces se trouvent les représentants du genre *Berylmys* dont l'ordination a des valeurs un peu inférieures sur la première composante de conformation et des valeurs équivalentes sur la seconde. Vient ensuite en position centrale, le genre *Rattus*, dont la plus petite espèce, *R. exulans*, se rapproche des individus du genre *Mus* aux ordinations négatives sur les deux composantes. Les représentants des genres *Maxomys*, *Niviventer* et *Leopoldamys* possèdent des valeurs d'ordination positives sur la seconde composante mais des valeurs négatives (comme pour *N. huang*), nulles (pour le genre *Maxomys*), ou positives (pour le genre *Leopoldamys*) sur la première composante. Les différences morphologiques décrites sur ce plan représentent 95 % de la variance totale, ce qui est bien plus que pour les conformations en vue ventrale.

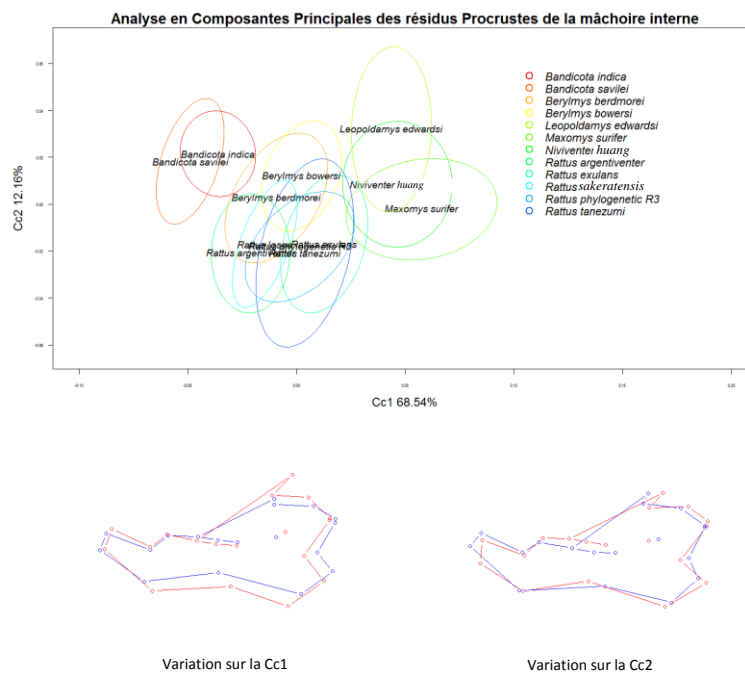
### *C3.5. Analyse en composantes principales sur les données de morphométrie géométrique des hémi-mandibules en vue externe*

La variation des conformations sur les composantes de conformation 1 et 2 (Cc1 et Cc2) ainsi que les projections des individus dans le plan principal ont été représentées sur la figure 19 (page suivante).

Les individus possédant des ordinations négatives pour l'axe 1 de l'espace des conformations ont une mâchoire globalement plus massive que les individus possédant une ordination positive. Cela se traduit par un léger épaissement de l'os mandibulaire, mais surtout par un développement plus important du ramus, et ce tant aux niveaux des apophyses angulaire que coronoïde. Cependant l'échancrure de la partie postérieure du ramus entre les apophyses angulaires et condyloïdes est plus prononcée. Le long de la seconde composante, les individus ayant des valeurs fortement négatives se caractérisent par une échancrure du



**Figure 19:** représentation des ellipses regroupant 90 % des individus de chaque mitotype dans le plan principal de l'espace des conformations de la mâchoire externe. La variation de conformation sur les axes est représentée en dessous, la conformation minimale en bleue et la conformation maximale en rouge.



**Figure 20:** représentation des ellipses regroupant 90 % des individus de chaque mitotype dans le plan principal de l'espace des conformations de la mâchoire interne. La variation de conformation sur les axes est représentée en dessous, la conformation minimale en bleue et la conformation maximale en rouge.



ramus également plus prononcée entre les apophyses angulaire et condyloïde, ainsi qu'une apophyse coronoïde située plus en arrière que les individus possédant des valeurs positives.

Les ellipses contenant 90% des individus dans le plan principal montrent que l'espace des conformations des mâchoires externe est structuré par les genres taxonomiques. Les espèces du genre *Bandicota* possèdent des ordinations positives sur la seconde composante et négative sur la première. A proximité de ces deux espèces se trouvent les individus du genre *Berylmys* avec des ordinations aux valeurs un peu supérieures sur la première composante de conformation et des valeurs équivalentes sur la seconde. Le genre *Rattus* est positionné autour de l'origine mais avec des valeurs plutôt négatives pour la seconde composante. Le genre *Maxomys* a des valeurs supérieures à celles du genre *Rattus* sur la première composante et des valeurs nulles pour la seconde. La variance de la conformation de la mâchoire du genre *Niviventer* chevauche de façon importante celle du genre *Maxomys*. Les espèces du genre *Leopoldamys* ont des valeurs positives sur la seconde composante. On constate que la variance de la mâchoire de *Leopoldamys neilli* est plus importante que celle de *Leopoldamys edwardsi*. Cependant, l'échantillonnage pour ce premier mitotype est très limité (trois individus seulement) et cette tendance est donc à relativiser. Les différences morphologiques décrites sur ce plan représentent 83 % de la variance totale.

### *C3.6 L'analyse en composante principale sur les données de morphométrie géométrique des héli-mandibules en vue interne*

La variation des conformations sur les composantes de conformation 1 et 2 ainsi que le plan principal ont été représentés sur la figure 20.

Les individus possédant des ordinations négatives pour l'axe 1 de l'espace des conformations ont une mâchoire globalement plus fine que les individus possédant une ordination positive. Ces espèces sont également celles qui présentaient un développement plus important de l'héli-mandibule en vue externe. Il semble donc que les mâchoires de ces animaux soient bombées vers l'intérieur. L'échancrure du ramus entre les apophyses angulaire et condyloïde est moins prononcée que pour les espèces possédant une ordination positive. Le long de la seconde composante, les individus ayant des valeurs fortement négatives se caractérisent par une échancrure moins prononcée du ramus entre les apophyses angulaire et condyloïde, et une apophyse coronoïde plus en avant que les individus possédant des valeurs positives. La mâchoire est amincie au niveau de la rangée de molaire et l'os mandibulaire et l'incisive qu'il porte est plus relevée.

Les ellipses contenant 90% des individus dans le plan principal montrent une ordination des animaux assez semblable à celle de la vue externe de la mâchoire. Les espèces du genre *Bandicota* ont une ordination qui correspond à des valeurs positives sur la seconde composante et négative sur la première. A proximité de ces deux espèces se trouvent les représentants du genre *Berylmys* avec des valeurs un peu supérieures sur la première composante de conformation et des valeurs équivalentes sur la seconde. Leur variance est très



chevauchante avec celles des espèces du genre *Rattus* dont les ordinations sont toujours positionnées autour de l'origine mais avec des valeurs négatives pour la seconde composante principale. Le représentant du genre *Maxomys* a des valeurs supérieures à celles du genre *Rattus* sur la première composante et des valeurs nulles pour la seconde. La variance de la conformation des héli-mandibules en vue interne du genre *Niviventer* chevauche largement celle du genre *Maxomys*, comme cela était déjà le cas en vue externe. Seule une espèce du genre *Leopoldamys* a été représentée ici et la variance de son héli-mandibule en vue interne est chevauchante avec celle du mitotype *N. huang* dans le premier plan principal. Les différences morphologiques décrites sur ce plan représentent 80 % de la variance totale.

### *C3.7. Discussion sur les variables utilisées dans ces ACP*

Lorsque les données corporelles brutes sont considérées, la taille qui est variable possédant la plus grande variance, influence fortement la première composante principale. Les mitotypes s'agencent le long de cette composante et s'isolent par groupes de genres aux tailles très différentes, comme le genre *Leopoldamys* et le genre *Mus* qui sont complètement séparées des genres *Rattus*, *Niviventer*, *Maxomys*, *Berylmys* et *Bandicota* (figure 13). Lorsque l'on retire la variable de taille, comme c'est le cas dans la représentation en log shape-ratios (figure 15), la séparation des genres devient encore moins nette car les proportions corporelles des animaux sont relativement homogènes au sein des tribus étudiées. Lorsque l'on s'intéresse aux crânes, le regroupement des espèces par genre apparaît plus évident, et ce quelle que soit la structure (crâne (figures 16, 18) ou héli-mandibule (figures 19, 20)) ou la vue étudiée (ventrale (figure 16), dorsale (figure 18), externe (figure 19) ou interne (figure 20)). Le signal morphologique porté par les conformations des héli-mandibules est assez semblable en vues externes et internes et l'ordination des espèces est assez similaire. Pour les vues du crâne, l'ordination des espèces diffèrent quelque peu suivant que l'on s'intéresse à la vue ventrale ou à la vue dorsale. Ainsi, les genres *Mus* et *Berylmys* n'occupent pas des ordinations semblables par rapport aux autres espèces suivant la vue du crâne à laquelle on s'intéresse. En termes d'identification cette constatation est intéressante car elle montre que l'étude des deux vues, si elle contient une part de redondance, contient également une part d'information propre à chacune d'elle.

### ***C4. Les mitotypes structurent-ils significativement la variation morphologique lorsque celle-ci correspond aux données brutes, aux log shape-ratio ou aux résidus Procrustes ?***

La part du facteur 'mitotype' sur la variance morphologique, décrite par les ACP, est testée à l'aide d'une analyse de variance multivariée sur les composantes non nulles de variation (en effet certaines composantes sont nulles à la suite des transformations du jeu de données par le biais des analyses de formes). Une Manova, utilisant les carrés de type II permet de connaître la significativité d'un effet tout en considérant les effets confondants des autres facteurs (Claude et al., 2004) tels que l'âge et le sexe.

**Tableau 5 :** Tableau de l'analyse de variance multivariée (Manova type II) réalisée sur les cinq composantes principales de l'ACP pour les données brutes corporelles. L'effet du mitotype, du sexe, de l'âge et l'interaction entre le sexe et l'âge sont testés (se reporter au Tableau 3 pour la signification des colonnes). **Pillai** (voir glossaire), **num df** et **den df** sont les degrés de libertés utilisés pour calculer F.

	<b>Df</b>	<b>Pillai</b>	<b>approx F</b>	<b>num Df</b>	<b>den Df</b>	<b>Pr(&gt;F)</b>
mitotype	17	2,86904	47,994	85	3030	< 2,2e-16
sexe	1	0,07172	9,303	5	602	1,509e-08
age	2	0,47827	37,904	10	1206	< 2,2e-16
mitotype:age	17	0,17208	1,271	85	3030	0,04971
Résidus	606					

**Tableau 6 :** Tableau de l'analyse de variance multivariée (Manova type II) réalisée sur les quatre premières composantes principale de l'ACP pour les données de log shape-ratios. L'effet du mitotype, du sexe, de l'âge et l'interaction entre le sexe et l'âge sont testés. (se reporter au Tableau 3 pour la signification des colonnes).

	<b>Df</b>	<b>Pillai</b>	<b>approx F</b>	<b>num Df</b>	<b>den Df</b>	<b>Pr(&gt;F)</b>
mitotype	17	2.05460	37.648	68	2424	< 2.2e-16
sexe	1	0.04110	6.461	4	603	4.287e-05
age	2	0.43959	42.538	8	1208	< 2.2e-16
mitotype:age	17	0.15624	1.449	68	2424	0.01026
Résidus	606					

**Tableau 7 :** Tableau de l'analyse de variance multivariée (Manova type II) réalisée sur les composantes principale de l'ACP pour les résidus Procrustes de la vue ventrale du crâne. L'effet du mitotype, du sexe, de l'âge sont testés.

	<b>Df</b>	<b>test stat</b>	<b>approx F</b>	<b>num Df</b>	<b>den Df</b>	<b>Pr(&gt;F)</b>
mitotype	17	8.8805	8.8435	1190	10098	<2e-16
sexe	1	0.2376	1.0613	140	1102	0.3069
age	2	0.8429	5.7346	140	1102	<2e-16
mitotype:age	28	3.8119	1.2990	1960	16156	5.662e-16
Résidus	619					

**Tableau 8:** Table de l'analyse de covariance multivariée (Mancova type II) réalisée sur les composantes principale de l'ACP pour les résidus Procrustes de la vue ventrale du crâne.

	<b>Df</b>	<b>Pillai</b>	<b>F</b>	<b>num Df</b>	<b>Den Df</b>	<b>Pr(&gt;F)</b>
Size	1	0.95119	166.48	70	598	< 2.2e-16
Résidus	667					

Les interactions entre l'âge et le sexe n'ont pas été testées car, pour beaucoup de mitotypes, les jeunes animaux ne sont pas toujours représentés pour les deux sexes dans notre échantillonnage (tableau 1).

#### *C4.1. Effets de l'âge, du sexe et de l'appartenance à un mitotype sur la conformation et la forme du corps des animaux*

Pour les données brutes, toutes les composantes principales de l'ACP sont considérées dans l'analyse de variance. L'âge, qui divise les animaux (paragraphe B1.2) en trois groupes (**Pillai** (2, 606)= 0,4783.  $P < 2,2 \times 10^{-16}$ ), leur sexe (**Pillai** (1, 606)=0,0717.  $P = 1,509 \times 10^{-8}$ ) et leur appartenance à un mitotype (**Pillai** (17, 606)=2,869.  $P < 2,2 \times 10^{-16}$ ) ont un effet significatif sur la morphologie (taille et forme) des animaux (tableau 5). Une interaction significative apparaît entre l'âge et le mitotype (**Pillai** (17, 606)= 0,1721.  $P = 0,0497$ ) montrant que les lois d'allométries ontogéniques diffèrent en fonction du mitotype considéré.

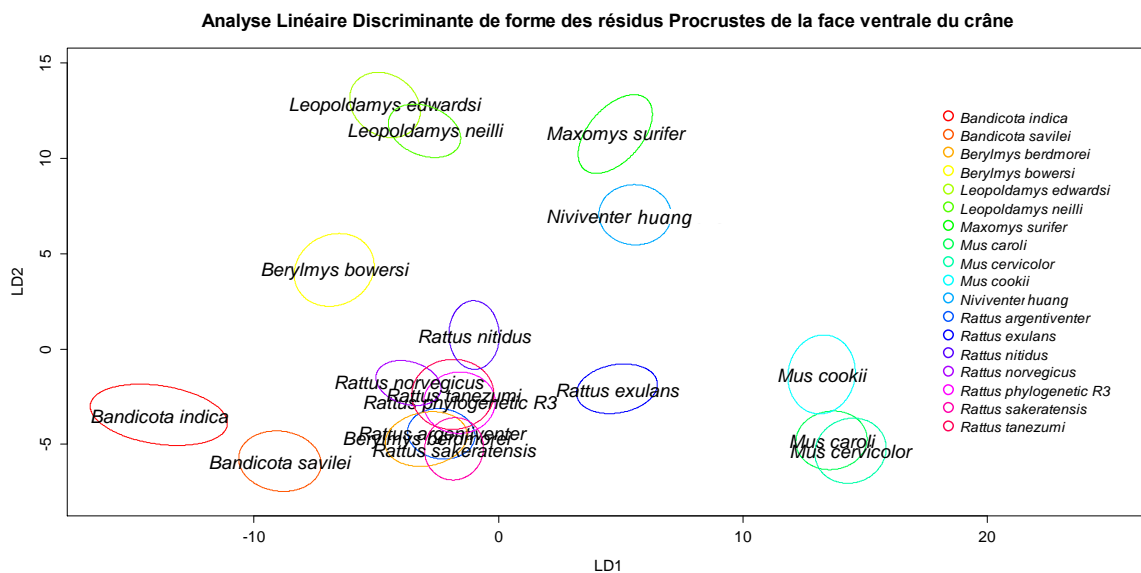
A la suite de la mise à l'échelle, l'information apportée par les log shape-ratios est redondante (Claude, 2013) et seules les quatre premières composantes principales de l'ACP ont donc été considérées dans l'analyse de variance. Les résultats (tableau 6) sont proches de ceux obtenus sur les données brutes. L'âge des animaux (**Pillai** (2, 606)= 0,4396.  $P < 2,2 \times 10^{-16}$ ), leur sexe (**Pillai** (1,606)=0,041 ;  $P = 4,287 \times 10^{-5}$ ) et leur appartenance à un mitotype (**Pillai** (17,606)= 2,0546 ;  $P < 2,2 \times 10^{-16}$ ) ont un effet significatif sur les proportions corporelles des animaux. Une interaction significative existe entre l'âge et le mitotype (**Pillai** (17,606)= 0,156 ;  $P = 0,010$ ) montrant que la morphologie externe ne varie pas de la même manière au cours de la croissance des animaux en fonction du mitotype auquel ils appartiennent. Le mitotype, l'âge et le sexe des animaux ont un effet hautement significatif sur la variance observée dans l'espace de l'ACP, que la taille soit prise en compte (données brutes) ou non (log shape-ratio).

#### *C4.2. Effet de l'âge, du sexe et de l'appartenance à un mitotype sur la conformation et la forme du crâne en vue ventrale*

L'analyse de variance de conformation de la vue ventrale du crâne en considérant l'ensemble des composantes de conformation de l'ACP montre que l'âge des animaux (**Pillai** (2, 619)= 0,843 ;  $P < 2,2 \times 10^{-16}$ ) et leur appartenance à un mitotype (**Pillai** (17,619)= 8,63 ;  $P < 2,2 \times 10^{-16}$ ) sont des facteurs affectant fortement la conformation du crâne (tableau 7). Cependant, il n'existe pas de dimorphisme sexuel de conformation significatif (**Pillai** (1, 619)= 0,238 ;  $P = 0,307$ ) alors qu'il existait un dimorphisme de taille pour cette vue (voir paragraphe C2.2). L'interaction entre âge et mitotype indique que la variation de conformation liée à l'âge diffère d'un mitotype à l'autre (**Pillai** (28, 619)= 3,8119 ;  $P = 5,662 \times 10^{-16}$ ). L'existence d'allométries peut être vérifiée en testant l'influence de la taille comme variable explicative de la conformation du crâne (tableau 8). Cette analyse montre que la conformation du crâne est expliquée de façon très hautement significative par la taille centroïde (**Pillai** (1, 667)=0,9512 ;  $P < 2,2 \times 10^{-16}$ ). Il existe donc bien des allométries ontogéniques au sein des tribus étudiées. Les résultats de l'analyse de variance de forme de la vue ventrale du crâne (la

**Tableau 9:** Tableau de l'analyse de variance multivariée (Manova type II) réalisée sur les composantes principale de l'ACP pour les résidus Procrustes et la taille centroïde (i.e. la forme) de la vue ventrale du crâne. L'effet du mitotype, du sexe, de l'âge sont testés.

	Df	test stat	approx F	num Df	den Df	Pr(>F)
mitotype	17	8.7787	8.9183	1207	10081	<2e-16
sexe	1	0.2369	1.0939	142	1156	0.2253
age	2	0.7845	5.2540	142	1156	<2e-16
Résidus	647					



**Figure 21 :** Représentation dans le plan discriminant principal des ellipses regroupant 90 % des individus de chaque mitotype en vue ventrale.

variable de forme est obtenue en ajoutant la taille centroïde aux données précédentes) sont en concordance avec les résultats précédents (tableau 9). L'âge des animaux (Pillai (2,647)= 0,784 ;  $P < 2,2 \times 10^{-16}$ ) et leur appartenance à un mitotype (Pillai (17, 647)= 8,779 ;  $P < 2,2 \times 10^{-16}$ ) sont des facteurs affectant fortement la forme des crânes alors que le dimorphisme sexuel n'a pas d'effet significatif sur la conformation du crâne en vue ventrale (Pillai (1, 647)= 0,237 ;  $P = 0,2253$ ).

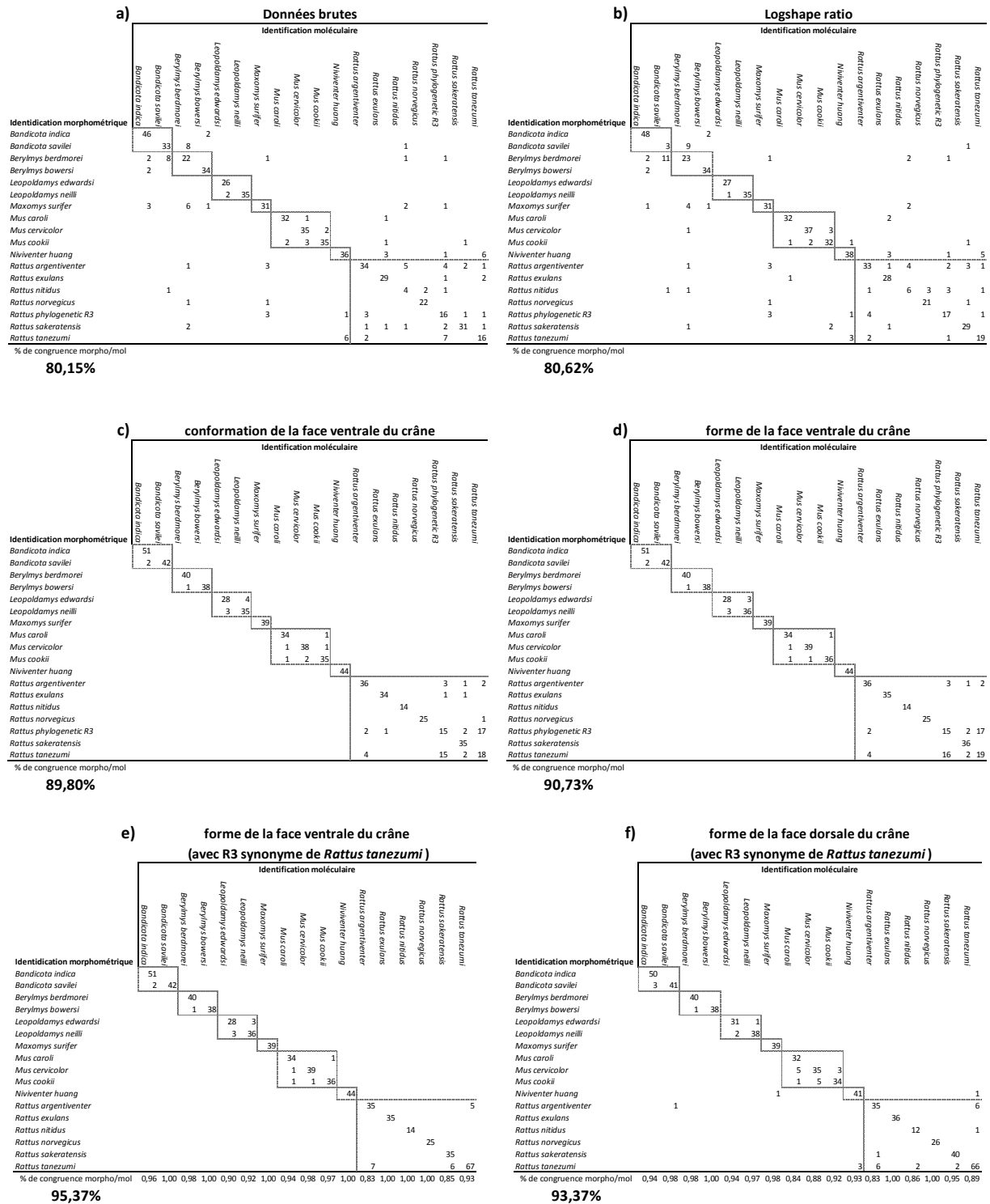
Ces résultats permettent de conclure que, dans le cadre de la mise en place d'un outil d'identification basé sur la morphométrie, nous devons prendre en considération le sexe des animaux pour la morphologie externe, alors que la variation du crâne en vue ventrale (forme et conformation) n'est pas sensible à ce facteur. Ces patrons de variations sont confirmés lorsque l'on étudie la face dorsale (données non montrées). Cette analyse n'a pas été conduite sur les héli-mandibules en raison du faible effectif échantillonné pour chacun des mitotypes (mais voir Annexe VII).

## ***C5. Comparaison du pouvoir discriminant des mesures brutes, des log shape-ratios, des variables de conformation et de forme crâniennes***

### *C5.1. Méthode de comparaison*

Les morphologies externes et crâniennes sont influencées par l'âge (i.e., la taille) et le mitotype des animaux. La morphologie externe l'est également par le sexe. L'objectif de ce nouveau chapitre est d'évaluer dans quelle mesure il est possible, à l'aide de fonctions mathématiques basées sur les traits morphologiques, de discriminer efficacement les mitotypes dans l'objectif de la construction d'un outil d'identification. L'analyse linéaire discriminante (ALD) est une technique d'analyse de données qui permet, à partir de la connaissance *a priori* de l'appartenance d'individus à des groupes, de calculer des fonctions discriminantes indépendantes entre elles et capables de retrouver au mieux ces groupes en fonction des variables à expliquer (Macleod et al., 2007). Pour cela l'analyse considère que la matrice de variance-covariance de  $k$  groupes peut être subdivisée en composantes de variance intra-groupe et inter-groupe. La variance intra-groupe est calculée comme étant la somme des écarts de chaque individu à la moyenne de son groupe pour toutes les variables, tandis que la variance inter-groupe correspond à la somme des écarts de chaque moyenne de groupe à la moyenne générale pour toutes les variables. L'analyse cherche alors à créer des fonctions discriminantes en minimisant la variance intra-groupe tout en maximisant la variance inter groupes (Macleod et al., 2007). A titre d'exemple, la figure 21 est une représentation des ellipses contenant 90 % des données crâniennes en face ventrale dans l'espace discriminant principal. La robustesse des fonctions discriminantes ainsi calculées est testée à l'aide d'un ré-échantillonnage au sein de l'échantillonnage par Jackknife (voir Marcus, 1990 ; Chiari et Claude, 2012 ; Claude, 2013). Cette technique de ré-échantillonnage exclut de façon itérative un individu du calcul des fonctions discriminantes pour évaluer comment les probabilités postérieures d'assignation de l'ensemble du jeu de données sont affectées par ce ré-échantillonnage. Cette méthode évite une démarche circulaire qui consiste à classer des animaux par une fonction discriminante qu'ils ont contribué à créer. Dans un premier temps,

**Tableau 10:** résultats de Jackknife présentant en colonne le classement des individus identifiés moléculairement à des groupes morphométriques en lignes. Les genres sont délimités par des traits pleins. Ce classement est effectué à partir des données mesurées brutes (a), des données mesurées transformée en log shape-ratio (b), des variables de conformation et de forme crâniennes en vue ventrale (c et d), des variables de forme en vue ventrale et dorsale du crâne avec *Rattus phylogenetic R3* considéré comme synonymes de *Rattus tanezumi* (e et f), de la forme de la face interne et externe de la mâchoire (g et h, page suivante), de la forme de la vue ventrale du crâne à partir des prises de vues (i, page suivante)





le pouvoir discriminant des données brutes et des données transformées en log shape-ratio ont été comparées pour les données corporelles. Dans un deuxième temps, cette comparaison a porté sur les données corporelles et crâniennes.

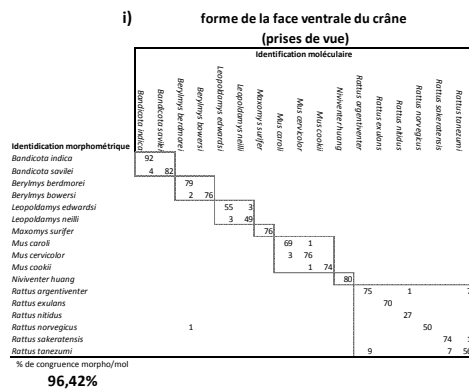
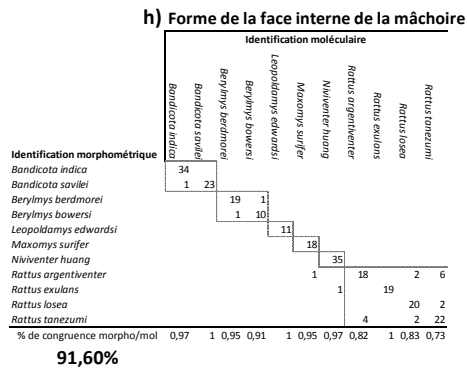
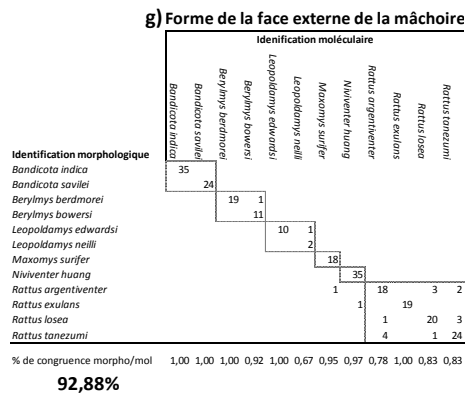
### *C5.2. Discrimination des espèces sur la base des mesures corporelles*

L'analyse linéaire discriminante appliquée sur le jeu de données brutes et sur les log shape-ratios donne des résultats comparables (tableau 10a,b). Les individus des mitotypes correspondant au genre *Bandicota* sont souvent confondus avec ceux du genre *Berylmys* et, plus rarement, avec ceux du genre *Maxomys*. Il est à noter que les individus identifiés moléculairement comme appartenant au mitotype *Bandicota savilei* sont très souvent confondus avec ceux identifiés en tant que *Berylmys berdmorei* lorsque l'on retire la variable de taille du jeu de données de référence (tableau 10b). De même, les souris, de taille très inférieure aux autres animaux étudiés, sont effectivement assignées au genre *Mus* lorsque la taille est prise en compte, mais aux genres *Mus* et *Rattus* lorsque la conformation est considérée seule. A l'opposé, les individus au mitotype *M. cervicolor* sont, eux, légèrement mieux assignés au sein du genre *Mus* lorsque la taille n'est pas prise en compte. De même, le mitotype *N. huang* est moins souvent assigné à *R. tanezumi* lorsque l'on ne tient pas compte de la taille des animaux. La taille peut donc à la fois être un facteur de nuisance et une variable améliorant l'assignation des individus à un mitotype. Pour les autres mitotypes, les résultats sont quasiment équivalents, que l'on considère la forme ou la conformation des animaux. Les erreurs de genre sont relativement nombreuses et lorsque l'on applique un test de ré-échantillonnage par Jackknife, le taux de congruence entre assignations moléculaire et morphométrique est de 80,2 % sur les données corporelles brutes, et de 80,6 % en utilisant les log shape-ratio.

### *C5.3. Discrimination des espèces sur la base des analyses Procrustes réalisées sur les différentes parties du crâne*

Lorsqu'on réalise l'analyse discriminante sur les différentes conformations des vues des hémi-mandibules et des vues crâniennes, on obtient un pouvoir discriminant entre mitotypes de 79,6% pour la vue interne de la mâchoire, 84,2% pour la vue externe de l'hémi-mandibule, 85,6% pour la vue dorsale du crâne et 89,8% pour la vue ventrale du crâne. La vue ventrale est par ailleurs la seule à assigner tous les individus à leur genre taxonomique défini moléculairement (tableau 10c). L'assignation s'améliore pour toutes les vues lorsque l'on ajoute la variable de taille à l'analyse discriminante (Dobigny et al., 2002; Claude, 2013). L'ajout de cette variable revient à considérer l'espace des formes lors de l'analyse (Mitteroecker et al., 2004, 2013). Un point dans cet espace représente une forme unique et les relations entre ces formes sont linéaires (Mitteroecker et al., 2013). L'espace est donc Euclidien sous deux conditions : 1) la variation des conformations observées doit rester faible. Cette condition est vérifiée dans le cadre de l'identification d'espèces aussi proches taxonomiquement que celles que nous étudions; 2) une transformation par le logarithme népérien de la taille centroïde doit être utilisée afin que les

**Tableau 10:** résultats de Jackknife présentant en colonne le classement des individus identifiés moléculairement à des groupes morphométriques en lignes. Les genres sont délimités en pointillés, Ce classement est effectué à partir de la forme de la face interne et externe de la mâchoire (g et h), de la forme de la vue ventrale du crâne à partir des prises de vues (i,)



transformations induites par une variation de conformation soient isotropes dans l'espace des formes.

Ainsi pour la vue ventrale, le pouvoir discriminant entre mitotypes passe à 90,73 % (tableau 10d). Des erreurs d'assignation intra-générique subsistent pourtant. L'assignation est particulièrement mauvaise pour les deux mitotypes *R. tanezumi* et *R3*. Les individus de ces deux groupes sont assignés pour moitié à l'un des mitotypes et pour moitié à l'autre indépendamment de leur identification moléculaire. Lorsque la forme du crâne est considérée dans l'analyse discriminante, 47 % des individus moléculairement identifiés comme appartenant au mitotype *R3* sont assignés au groupe morphométrique de même nom et 44 % au groupe des *R. tanezumi*. De même 50 % des individus identifiés moléculairement comme appartenant au mitotype *R. tanezumi* sont identifiés morphométriquement comme tel, tandis que 45 % le sont comme *R3*. Nous verrons dans la partie 2 qu'une étude morphométrique combinée avec une étude de marqueurs moléculaires neutres n'a pas permis non plus de dissocier ces deux mitotypes qui pourraient en fait ne correspondre qu'à un seul taxon (Pagès et al., 2013). Pour cette raison, dans la suite de la mise en place de l'outil d'identification, nous considérerons ces deux mitotypes comme des synonymes de *R. tanezumi*. Une fois ce regroupement opéré, le pouvoir discriminant testé par Jackknife, qui était de 89,8% et 90,7% respectivement pour la conformation et la forme, augmente pour atteindre 94,3% (données non montrées) pour les conformations et 95,4 % lorsque l'on considère la forme des crânes en vue ventrale (tableau 10e). L'augmentation du pouvoir discriminant lorsque l'on regroupe les deux mitotypes et que l'on ajoute la variable de taille dans l'analyse ne se limite pas à la vue ventrale du crâne. En effet, lorsque l'on considère la forme des mâchoires en vue interne et après avoir regroupé *R. tanezumi* et *R3*, le pouvoir discriminant passe à 91,6% (tableau 10f). Il passe à 92,9% pour la vue externe de la mâchoire et à 93,7% pour la forme de la face dorsale du crâne (tableau 10g, h). En revanche, la discrimination de certains mitotypes du genre *Rattus* se dégrade légèrement lorsque l'on considère *R. phylogenetic R3* comme synonyme de *R. tanezumi*. Ainsi, pour la vue ventrale, le taux d'incongruence entre l'identification morphométrique et l'identification moléculaire pour *R. argentiventer* passe de 14,3% à 16,7% et celle de *R. sakeratensis* de 12,2% à 14,6%. Cette augmentation est liée au fait que le regroupement de *R. tanezumi* et *R3* double l'effectif de *R. tanezumi*, ce qui donne plus de poids à ce groupe lors du calcul des fonctions discriminantes. Lorsque l'on ne considère que l'un des deux mitotypes dans l'analyse, le pourcentage d'assignation congruente de *R. argentiventer* et *R. sakeratensis* revient à son niveau d'origine mais le taux d'incongruence pour *R. tanezumi* (ou *R. phylogenetic R3* suivant le mitotype conservé) passe de 7 % à 16 %. Cet exemple montre l'importance d'opter pour un échantillonnage équilibré lorsque l'on mène une analyse discriminante.

On constate que les erreurs d'identification diffèrent suivant la structure et la vue considérée. L'étude réalisée sur les mâchoires portait sur un échantillonnage retreint (n = 250 individus). Par conséquent, l'assignation des individus à l'aide de la forme des hémimandibules, bien qu'elle apparaisse ici globalement moins performante qu'avec les vues crâniennes, ne doit pas être négligée (tableau 10e, f, g, h). Ainsi, pour le genre *Bandicota*, notre échantillonnage réduit de la vue externe de la mâchoire ne semble pas fournir d'identification

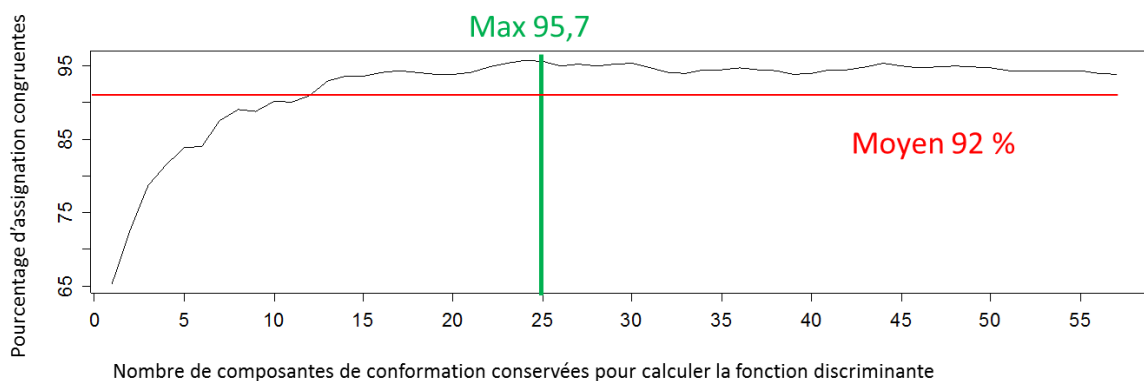


différente de celle basée sur l'ADN mitochondrial, alors que l'identification à partir de la vue ventrale du crâne classe deux individus appartenant aux mitotype *B. indica* au sein du mitotype *B. savilei*. Une étude comprenant un échantillonnage large d'hémi-mandibules droite et gauche en vue externe appartenant à 358 individus des deux mitotypes (Chaval, en prép, voir Annexe VII) a confirmé que la forme de cette structure était très pertinente pour la discrimination des deux mitotypes. Cependant les deux individus mal classés lorsque l'on considère les vues crâniennes et qui ne sont pas inclus dans l'échantillonnage utilisé ici pour la vue externe de la mâchoire, sont également mal classés lorsque l'on considère un échantillonnage plus étendu de la vue externe de la mâchoire (voir Annexe VIII pour une présentation des erreurs par prise de vue). La discrimination n'est donc pas améliorée mais le nombre plus limité de points repères à positionner sur l'hémi-mandibule par rapport aux faces crâniennes en font une structure intéressante dans le cadre de la mise en place d'un outil web-service, d'autant que les hémi-mandibules sont parfois le seul matériel disponible (comme dans l'étude des pelotes de réjection de rapaces nocturnes). Pour le genre *Berylmys* la vue externe de la mâchoire permet l'identification d'un individu mal assigné en vue crânienne. *L. edwardsi* et *L. neilli* sont relativement mal identifiés lorsque la vue ventrale est étudiée seule ; mais là encore, l'identification peut être améliorée à l'aide des conformations des mâchoires ou de la vue dorsale du crâne. Les genres *Niviventer* et *Maxomys* sont bien identifiés en face ventrale du crâne. Toutes les *Mus* sont bien identifiées lorsque l'on considère les différentes vues du crâne, la vue ventrale restant cependant la plus discriminante. Enfin, au sein du genre *Rattus*, un individu est toujours mal identifié quelles que soient la structure et la vue considérées. Trois individus sont mal identifiés sur la base de la forme de leur crâne et ne font pas partie de l'échantillonnage des animaux identifiés à l'aide des mâchoires alors que tous les autres individus appartenant à ce genre sont tous correctement assignés soit par la forme de leur crâne soit par celle de leurs mâchoires.

#### C5.4. Discussion

L'utilisation de la forme (i.e., taille et conformation) augmente le pouvoir discriminant de l'ensemble des vues étudiées. Ce résultat a déjà été mis en avant chez les rongeurs (Dobigny et al., 2002 ; Claude, 2013). Le regroupement de *R. tanezumi* et du clade *R3* permet également d'améliorer le taux d'assignation globale basé sur les différentes vues. Cependant, le pouvoir discriminant des vues pour certains mitotypes (i.e. *R. argentiventer* et *R. sakeratensis*) baisse alors. L'ensemble des individus des clades *R3* et *R. tanezumi* ont pourtant été conservés dans la suite de ce travail en considérant ces deux mitotypes comme une synonymie de *R. tanezumi*, ceci afin de conserver un pouvoir discriminant fort sur cette espèce commensale qui est invasive et régulièrement échantillonnée (Aplin et al., 2011 ; Lack et al., 2012). L'analyse des animaux mal identifiés (Annexe VIII) par l'une ou l'autre des vues a montré que les vues sont complémentaires les unes des autres. L'information apportée par l'étude de la forme des crânes et des mâchoires de rongeurs est plus discriminante que l'information portée par les données corporelles. Cependant ces dernières peuvent s'avérer pertinentes dans certains cas (voir Chaval et al., en prép., Annexe VII) car bien que moins précises que la morphologie crânienne pour l'assignation d'individus inconnus à un groupe,

**Représentation, pour une conformation donnée, de la variation du pouvoir discriminant en fonction du nombre de composantes de conformation considérées**



**Figure 22 :** Variation du pouvoir discriminant de l'ALD en fonction du nombre de composantes de conformation conservées sur l'ACP. Le pouvoir discriminant moyen est représenté par le trait rouge, le maximum en vert, la courbe représente le pourcentage d'individus assignés à leur identification moléculaire en ordonnées en fonction du nombre de composantes de conformation conservées pour l'ALD.

elles peuvent parfois atteindre un taux de congruence avec l'assignation moléculaire élevé. Elles sont par ailleurs beaucoup plus simples à mettre en œuvre sur le terrain.

## ***C6. Optimisation du pouvoir discriminant des prises de vues crâniennes et simplification des conformations***

Le pouvoir discriminant des fonctions discriminantes varie en fonction de :

- la vue considérée ;
- la configuration de points homologues sélectionnée pour représenter la conformation du crâne;
- du nombre de composantes principales considérées dans l'ACP pour calculer les fonctions discriminantes.

L'optimisation du pouvoir discriminant de l'analyse pour le jeu de données de référence doit tenir compte de ces trois conditions simultanément.

### ***C6.1. Choix du nombre de composantes de conformation de l'ACP pour optimiser la fonction discriminante***

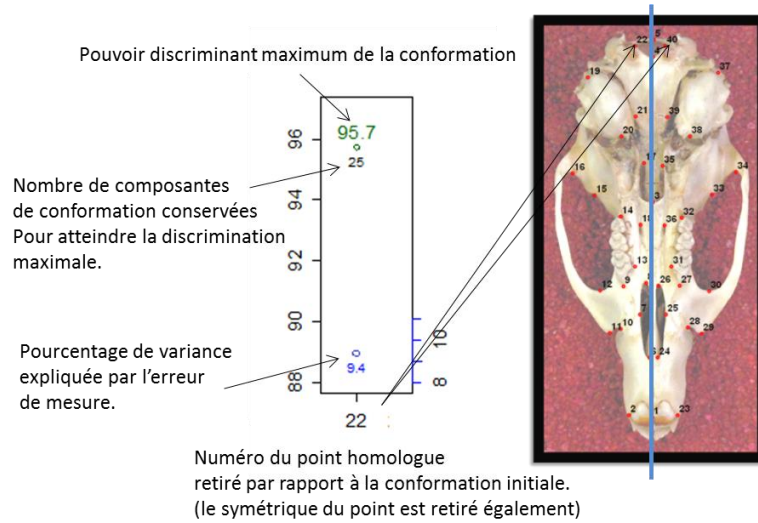
Pour réaliser une identification morphologique rigoureuse, il faut parvenir à un équilibre entre la taille de l'échantillon étudié et le nombre de variables de forme considérées (Chiari et Claude, 2012). Théoriquement l'assignation augmente avec le nombre de composantes considérées ; cependant elle atteint un plateau au-delà duquel, si de nouvelles variables sont considérées, l'assignation deviendra échantillon-dépendante et manquera donc de robustesse lors du ré-échantillonnage par Jackknife.

La stratégie mise en œuvre dans ce travail a donc été, pour une conformation donnée de calculer le pourcentage d'individus reclassés correctement par le test de validation croisée de l'analyse discriminante en fonction de leur identification moléculaire et du nombre de composantes de conformation considérées de façon à optimiser celui-ci. La figure 22 illustre ce travail d'optimisation pour une configuration de points homologues donnée, cette figure représente l'évolution du pouvoir discriminant en fonction du nombre de variables de conformations considérées dans l'analyse discriminante.

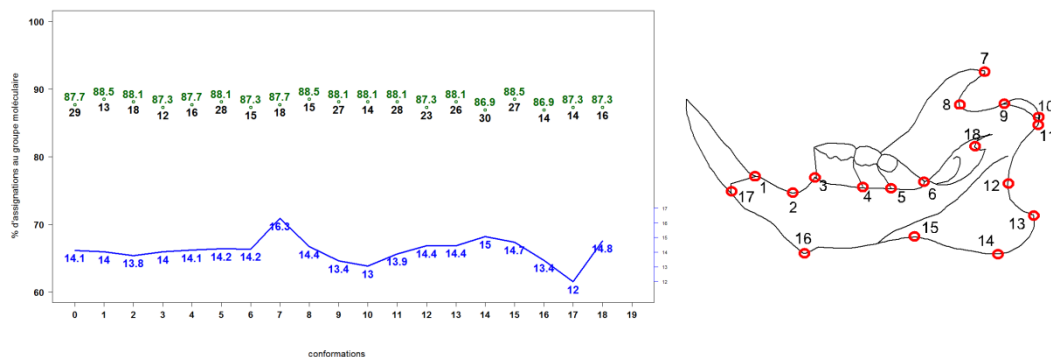
### ***C6.2. Pouvoir discriminant des vues et simplification des points homologues***

#### ***C6.2.1. Méthodes***

Les différentes structures et vues du crâne et des mâchoires sont complémentaires pour parvenir à la mise au point de l'outil d'identification. Pour une structure et une vue données, la configuration qui comporte le plus de points n'est pas toujours celle qui possède le meilleur pouvoir discriminant car certains points apportent plus d'incertitudes qu'ils n'apportent d'information discriminante. Ainsi, afin de mesurer la variation du pouvoir discriminant, les conformations de chaque structure et chaque vue ont été simplifiées en enlevant tour à tour un point homologue ou les deux points image lorsque la vue comprend un



**Figure 23 :** lecture d'un graphique synthétisant le pouvoir discriminant maximum (point vert) d'une conformation en fonction du nombre de composantes de conformations associées (en noir sous le point vert). Le point bleu correspond à l'erreur de conformation associée. L'abscisse représente la conformation considérée. La conformation 22 correspond à la configuration d'origine à laquelle on a ôté le point 22 et son symétrique (le point 40) par rapport à l'axe de symétrie du crâne matérialisé par le trait bleu.



a)

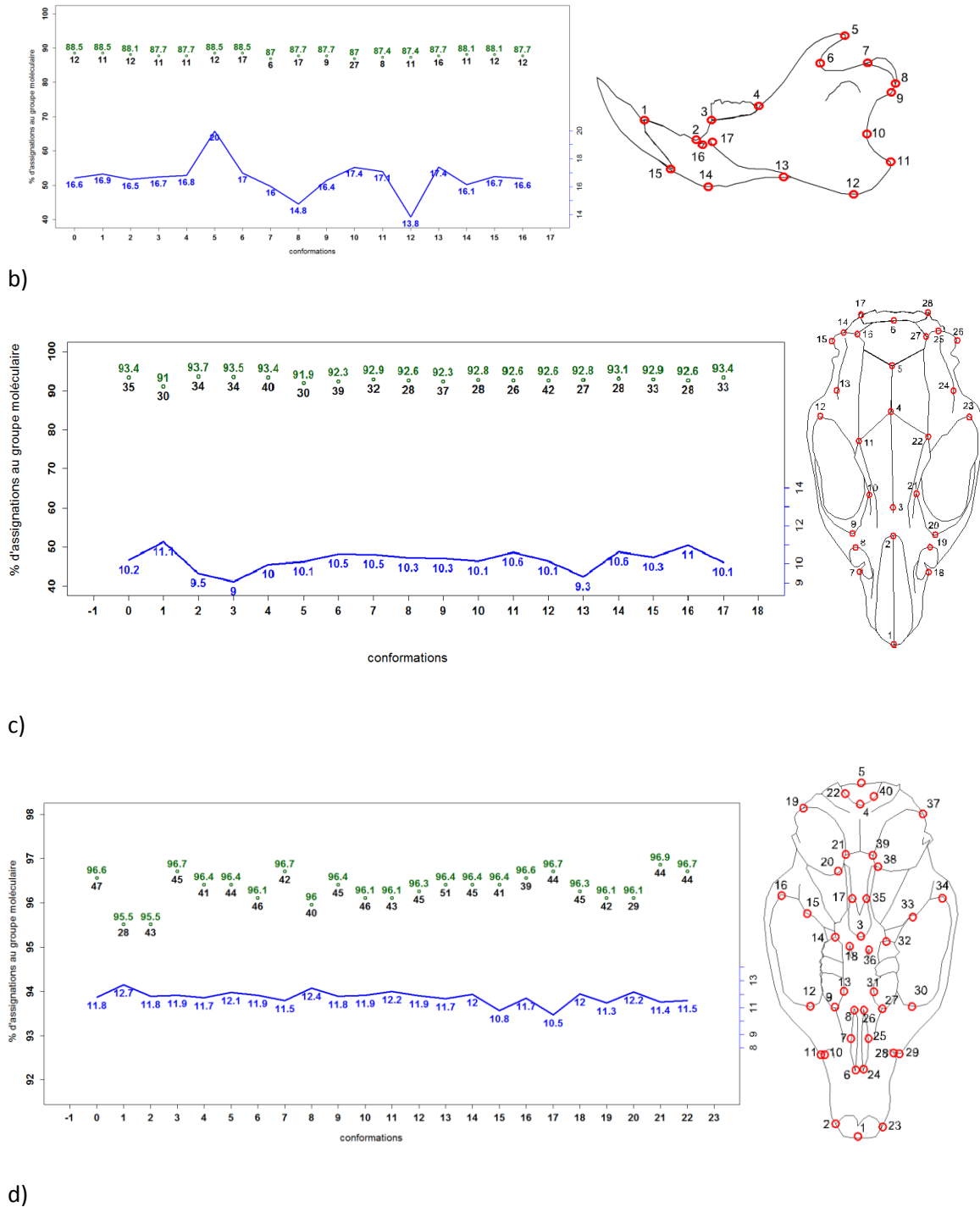
**Figure 24 :** Comparaison du pouvoir discriminant des structures, des vues et des conformations de points homologues (voir également page suivante). Le dessin à droite correspond à la conformation initiale pour l'hémi-mandibule en vue interne (a) positionnée au 0 de l'axe des abscisses. Les autres abscisses correspondent à cette conformation à laquelle on a enlevé le point homologue correspondant au chiffre en abscisse et à son symétrique pour les vues ventrales et dorsales du crâne. Le point vert correspond au pourcentage d'individus assignés au groupe défini par leur mitotype par l'analyse discriminante utilisant le nombre de composante de conformation de l'ACP noté en noir. La courbe bleue correspond à l'évolution de l'erreur de mesure associée en pourcentage de la variance totale.



axe de symétrie, ce qui est le cas des vues ventrales et dorsales du crâne. L'erreur de mesure associée à une configuration varie conjointement avec la simplification de celle-ci. Ainsi, la variation de l'erreur de mesure et celle du pouvoir discriminant peuvent être étudiées en fonction des simplifications de la configuration d'origine. Le choix des points homologues à retirer pour simplifier les configurations se fera en recherchant la configuration la plus simple et qui offre un pouvoir de discrimination maximal pour une erreur de mesure minimale. La figure 23 explicite la lecture des graphiques produits pour optimiser les configurations de points homologues. Sur cet exemple, le point 22 et son symétrique 40 sont retirés de la configuration d'origine ; pour cette nouvelle configuration, le pouvoir discriminant maximal en fonction du nombre de composantes de conformation optimum considéré est affiché, de même que l'erreur intra-utilisateur associée. Ainsi, chaque configuration d'origine ainsi que les configurations simplifiées qui en découlent ont fait l'objet d'une superposition Procrustes, les composantes de conformation ont été calculées à partir des résidus Procrustes et une série d'analyses discriminantes comprenant entre 2 et le maximum de composantes de conformation a été réalisée. Le pouvoir discriminant ainsi que la robustesse de l'assignation de chacune des analyses a été testé avec des procédures de Jackknife. Un graphique reprenant ces résultats est tracé pour chaque vue étudiée (figure 24). La comparaison de ces structures permet de choisir la conformation qui offre le meilleur résultat d'assignation des individus aux groupes pré-définis moléculairement.

#### *C6.2.2. Résultats*

La vue interne des héli-mandibules offrent un pouvoir discriminant assez limité (figure 24a). La conformation d'origine (à l'origine de l'axe des abscisses) comprenant 18 points homologues (l'erreur de mesure représente 14,1% de la variance totale), offre un pouvoir discriminant de 87,7% lorsque l'on conserve 29 composantes de conformations. Ce pouvoir discriminant s'améliore de 0,8% lorsqu'on enlève certains points, comme par exemple le point 1, placé à la jonction entre l'incisive et la mandibule, le point 8, à la base de l'apophyse coronoïde, ou encore le point 15, positionné au point le plus interne de la base du ramus. On note d'autre part que le point 7, positionné au sommet de l'apophyse coronoïde est un point dont le placement offre une bonne reproductibilité car lorsqu'on le retire, l'erreur de mesure augmente sensiblement. A l'opposé, la variance générée par l'erreur de mesure sur le placement du point 17, à la jonction de l'incisive et de la mandibule est importante comparativement à la variance totale. En effet, lorsqu'on retire ce point, l'erreur décroît de façon importante. L'espace de conformation est constitué de 32 composantes et le nombre de composantes à conserver pour le calcul de la meilleure fonction discriminante est très variable selon la configuration de points considérée. Celui-ci varie de 12 composantes pour la configuration d'origine à laquelle on a ôté le point 3, à 30 pour la configuration d'origine à laquelle est ôté le point 14. Lorsque l'on enlève tour à tour de la configuration d'origine les points homologues 1, 8 et 15 (figure 24a), on obtient les configurations de points qui conduisent à la meilleure discrimination du jeu de données pour cette structure. La configuration à laquelle on a retiré le point 1 est celle qui offre la meilleure reproductibilité. Si l'on poursuit la simplification en enlevant un second point à l'une de ces configurations,



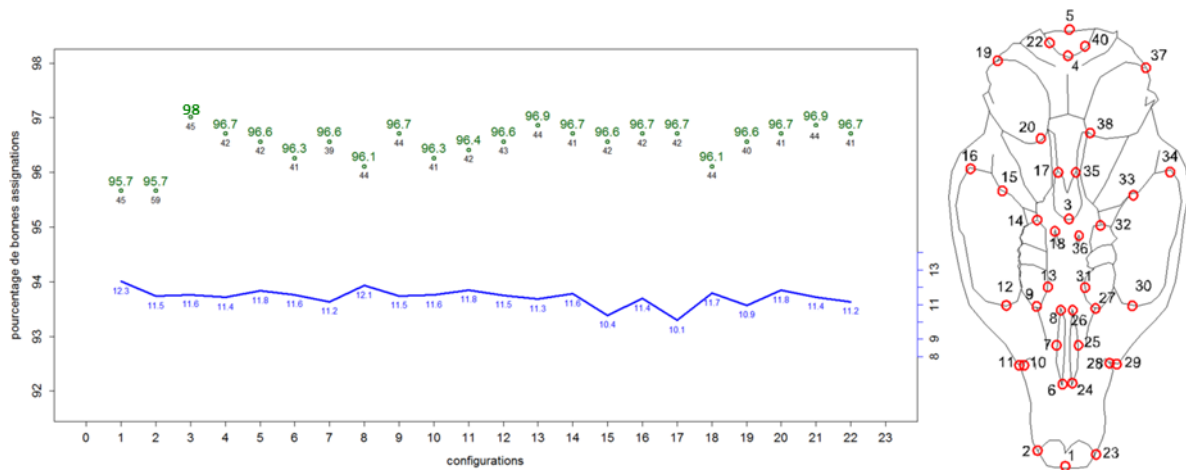
**Figure 24** : Comparaison du pouvoir discriminant des structures, des vues et des conformations de points homologues. Le dessin à droite correspond à la conformation initiale positionnée au 0 de l'axe des abscisses. b) face externe de l'hémi-mandibule, c) face dorsale du crâne, d) face ventrale du crâne. Les autres abscisses correspondent à cette conformation à laquelle on a enlevé le point homologue correspondant au chiffre en abscisse et à son symétrique pour les vues ventrales et dorsales du crâne. Le point vert correspond au pourcentage d'individus assignés au groupe défini par leur mitotype par l'analyse discriminante utilisant le nombre de composante de conformation de l'ACP noté en noir. La courbe bleue correspond à l'évolution de l'erreur de mesure associée en pourcentage de la variance totale.

toutes les configurations qui découlent ont un pouvoir discriminant moindre (données non montrées).

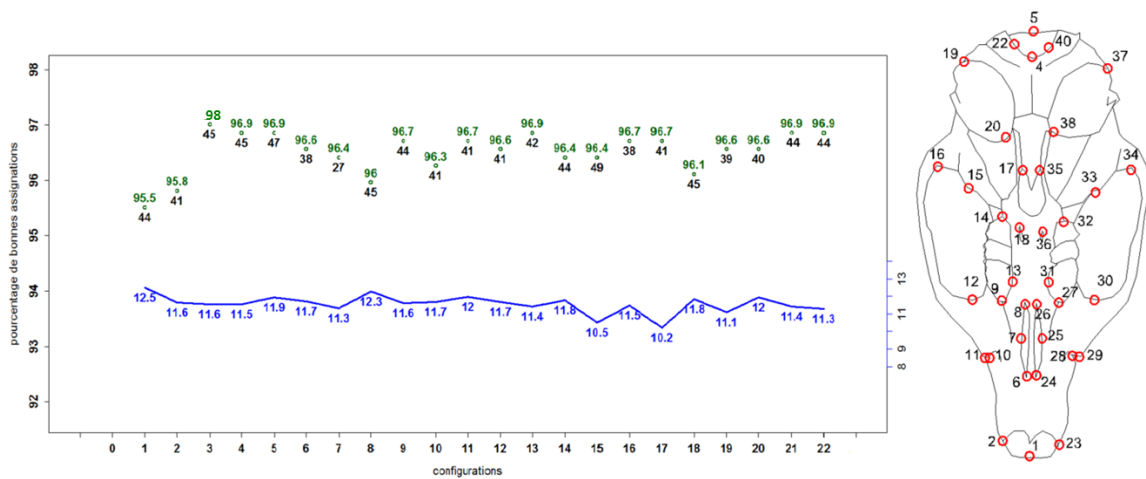
La vue externe des héli-mandibules qui comprend 17 points homologues (figure 24b), dans sa configuration d'origine, et offre un pouvoir discriminant équivalent aux meilleures configurations de points homologues de la vue interne : 88,5% des assignations morphologiques sont congruentes avec l'assignation moléculaire. Ce résultat est atteint lorsqu'on conserve 12 composantes de conformation comme variables à expliquer dans l'analyse discriminante. Ici encore, on constate que le placement du point situé au sommet de l'apophyse coronoidé (7) produit peu d'erreurs de mesure alors que le placement du point 12, à la base de l'apophyse angulaire, en produit beaucoup. L'espace des conformations de la vue externe de la mâchoire est constitué de 34 composantes de conformation. La variation du nombre de composantes à sélectionner pour calculer la meilleure fonction discriminante varie également de façon importante passant de 6 pour la configuration d'origine à laquelle on a enlevé le point antérieur de la surface articulaire du processus condilaire (7) à 27 pour la conformation où le point le plus antérieur de la partie postérieure du ramus (10) a été retiré. Cependant pour la majorité des conformations dérivées de la conformation initiale, l'information discriminante semble être davantage portée par les composantes de bas rang que pour les conformations dérivées de la face interne de la mâchoire. La configuration originale de la mâchoire en vue externe est la plus discriminante, et ne peut pas être simplifiée sans perdre de pouvoir discriminant (données non montrées, la figure 24b affiche des valeurs arrondies).

La conformation de points homologues portée par la vue dorsale du crâne (figure 24c), comprend 28 points dont 22 sont symétriques deux à deux. Dans sa configuration d'origine, cet ensemble de points homologues permet d'identifier 93,4% des individus des 17 espèces moléculaires étudiées lorsque l'on considère 35 composantes de conformation sur les 56 que comprend l'espace des conformations de cette vue. L'erreur de mesure associée à cette configuration est de 10,2%. Lorsque l'on enlève le point 3 ou dans une moindre mesure le point 13, cette erreur de mesure atteint son niveau minimal à savoir 9% de la variance totale lorsque le point 3 est ôté et 9,3% lorsque les points 13 et 24 le sont. Le pouvoir discriminant s'améliore légèrement (passant de 93,4% à 93,5%) lorsqu'on retire le point 3 de l'analyse (point qui sépare le rostre de la boîte crânienne) et décroît légèrement lorsque l'on retire les points 13 et 24 qui correspondent à l'intervalle entre les sutures symétriques des os pariétal et squamosal. La suppression du point homologue 2, positionné à la limite postérieure de l'os nasal, et la sélection de 34 variables de conformations améliore l'identification qui passe à 93,7% de congruence entre identifications moléculaire et morphologique (figure 24c). L'erreur intra-utilisateur passe de 10,2% pour la configuration d'origine à 9,5 % pour cette configuration. Une fois ce point ôté, aucune configuration de points obtenue en retirant un point supplémentaire n'améliore le pouvoir discriminant de cette vue (données non montrées).

La vue ventrale du crâne (figure 24d) offre le meilleur pouvoir discriminant. Sa configuration d'origine comporte 40 points homologues dont 18 sont symétriques deux à deux. Le pouvoir discriminant est de 96,6% lorsque l'on considère 47 composantes de conformation sur les 80 formant l'espace des conformations. L'erreur de mesure associée à



a)



b)

**Figure 25 :** Simplification de la conformation de la vue ventrale du crâne. Le dessin à droite correspond à la conformation initiale positionnée au 0 de l'axe des abscisses. Les autres abscisses correspondent à cette conformation à laquelle on a enlevé le point homologue correspondant au chiffre en abscisse et à son symétrique pour les vues ventrales et dorsales du crâne. Le point vert correspond au pourcentage d'individus assignés au groupe défini par leur mitotype par l'analyse discriminante utilisant le nombre de composante de conformation de l'ACP noté en noir. La courbe bleue correspond à l'évolution de l'erreur de mesure associée en pourcentage de la variance totale.

cette conformation représente 11,8% de la variance totale. On constate que les points homologues portés par les incisives (1, 2 et son symétrique 23) ont un poids important dans le pouvoir discriminant de l'analyse car lorsqu'on les retire, celui-ci chute de façon assez importante. C'est également le cas des points placés aux extrémités des foramens incisifs (6 et 8) et postérieurs (18 et son symétrique 36), des points placés à la jonction de l'apophyse mastoïde et de la bulle tympanique (19 et 37) ainsi que de ceux placés à l'intersection de l'os basiophénoïde et de la bulle tympanique (20 et 38). Ces quatre derniers points définissent le volume relatif des bulles tympaniques par rapport au reste du crâne. D'autres points homologues semblent de moindre importance pour la discrimination au sein de notre jeu de données. C'est le cas du point 3 placé à l'extrémité postérieure du palais, des points 7 et 25 qui définissent la plus grande largeur des foramens incisifs, et des points 21 et 39 qui définissent la largeur de la suture basioccipital-basiophénoïde. Pour la vue ventrale, la suppression du point 21 et de son symétrique 39 (figure 24d), qui se trouvent à la jonction des bulles tympaniques et de la suture entre l'os basioccipital et le basiophénoïde, permet d'améliorer l'assignation des individus à leur mitotype. Le pourcentage d'assignations passe alors de 96,6% pour la conformation d'origine à 96,9% lorsque l'on conserve 44 composantes de conformation. L'erreur de mesure, elle, passe alors de 11,8% à 11,4% de la variance totale pour cette configuration simplifiée. Une fois cette simplification effectuée, il apparaît que la suppression du point 3 (figure 25a) permet encore d'améliorer légèrement le pouvoir discriminant qui s'élève à 97% lorsque l'on conserve 45 composantes de conformation. L'erreur de mesure, quant à elle, remonte légèrement pour atteindre 11,6% de la variance totale. Si l'on poursuit la simplification de la conformation, le pouvoir discriminant se met alors à baisser (figure 25b). Le pouvoir d'assignation de la vue ventrale est proche de celui atteint par des outils équivalents décrits dans la littérature (Gaston et O'Neill, 2004). La procédure de Burnaby (1966) a été utilisée sur la face ventrale du crâne pour les espèces du genre *Bandicota* et *Leopoldamys* (voir Claude, 2013) afin de tenter d'améliorer la discrimination au sein de ces taxons mais celle-ci s'est avérée moins bonne lorsque l'allométrie a été supprimée.

### C6.2.3. Discussion

Une identification basée sur le croisement de l'information apportée par les différentes vues apparaît ici pertinente. Une identification générique sera d'abord donnée à partir de la vue ventrale car celle-ci exclut toute erreur de genre. De plus, l'erreur intra-utilisateur générée lors de la production de données à partir de cette vue est plus importante qu'elle ne l'est sur la vue dorsale mais elle reste cependant modérée. Les résultats d'assignation à l'aide de la face dorsale du crâne sont généralement moins bons que ceux de la face ventrale (tableau 10e,f). Cependant, les résultats obtenus à partir de la vue dorsale sont meilleurs pour le genre *Leopoldamys* et pour *R. sakeratensis*. L'erreur d'assignation est de 6% pour *L. edwardsi* en vue dorsale contre 9,7% en vue ventrale, et de 2,6% pour *L. neilli* contre 7,7% en vue ventrale. Pour *R. sakeratensis*, la vue dorsale assigne les individus avec un taux d'erreur de 4,8% contre 14,7% en vue ventrale. Ainsi, l'identification en face dorsale peut être complémentaire de l'identification en face ventrale. Cette démarche d'identification croisée pourra être poursuivie à l'aide de la face externe des héli-mandibules dont les résultats



préliminaires montrent un bon taux d'assignation au sein du genre *Bandicota* et *Leopoldamys*. Notre jeu de données « hémi-mandibules » devra être complété pour confirmer l'intérêt de cette partie du squelette en complément des vues crâniennes. Cependant l'établissement d'un outil web-service demande également la prise en compte du temps que devra allouer l'utilisateur à son utilisation (voir le paragraphe D3.4.2) ainsi que la gestion de l'erreur engendrée par la multiplicité des opérateurs lors de la comparaison des données. L'approche libre d'allométrie (Burnaby, 1966 ; Claude, 2013) n'a pas permis d'améliorer la discrimination au sein des genres *Bandicota* et *Leopoldamys* mais elle s'est avérée intéressante dans le cadre de la discrimination entre *R. exulans* et *R. tanezumi* (Claude, 2013). Ces deux espèces sont discriminées ici à 100% et ne nécessitent pas ce genre de correction.

### ***C7- Erreur de mesure lors de l'acquisition des données de morphométrie géométrique et leur influence sur la capacité à discriminer les espèces***

L'erreur de mesure sur les données de morphométrie corporelle n'a pu être estimée car les mesures n'ont été réalisées qu'une seule fois pour chaque individu. En revanche l'erreur de mesure sur les données de morphométrie géométrique a pu être évaluée pour les structures et vues ostéologiques étudiées ici. Elle a été estimée pour la taille et la conformation suivant le protocole de Yezerinac (1992), dont le principe a été décrit au paragraphe A2.2.3.

#### ***C7.1. Erreur intra-opérateur de taille en fonction de la structure et de la vue étudiée***

L'erreur de mesure sur la taille est toujours faible quelle que soit la structure étudiée. Ainsi la variance induite par l'erreur de mesure sur la mâchoire en vue externe est la plus forte mais cette dernière ne représente que 0,067 % de la variance de taille totale. Pour les autres prises de vue, cette erreur de mesure ne dépasse pas 0,064% pour la vue interne de la mâchoire, 0,042 % pour la vue ventrale du crâne et 0,019 % pour la vue dorsale du crâne.

#### ***C7.2. Erreur intra-opérateur sur la conformation en fonction de la structure et de la vue étudiée***

La variance induite par l'erreur de mesure sur l'estimation de la conformation est nettement plus importante que pour la taille. C'est un résultat assez général dans les études de morphométrie géométrique (voir Arnqvist et Martensson, 1998 ; Siahsharvie et al., 2012 ; Claude, 2013). On retrouve cependant un patron semblable entre les structures avec une erreur de mesure sur la mâchoire plus importante que sur le crâne quelles que soient les vues considérées. Ainsi, l'erreur sur la mâchoire externe représente 16,62 % de la variance totale, celle sur la mâchoire interne 12,52 %, celle sur la face ventrale du crâne 11,75 % et celle sur la face dorsale 10,18 %.

La face dorsale et la face ventrale du crâne sont donc les vues pour lesquelles la **reproductibilité** est la meilleure.





### *C7.3. Influence sur la capacité à discriminer les espèces*

Les résultats précédents ont été obtenus à partir des moyennes des deux digitalisations effectuées indépendamment sur chaque individu par l'Opr. Afin de tester la reproductibilité de l'assignation, l'analyse discriminante a été réalisée à partir des prises de vues pour tous les individus pour la face ventrale du crâne (tableau 10i). Cela signifie que chaque prise de vue est considérée comme un individu statistique indépendant de son réplicat, ce qui double la population étudiée par rapport à l'échantillonnage d'origine. On constate que l'assignation est alors légèrement meilleure que sur les individus biologiques (moyenne des réplicats) ce qui veut dire que l'erreur de mesure réalisée par un opérateur ne dégrade globalement pas le pouvoir discriminant de l'analyse. On note néanmoins qu'une prise de vue d'un animal identifié moléculairement comme appartenant au mitotype *B. berdmorei* est identifiée morphologiquement comme appartenant au taxon *R. norvegicus*, c'est-à-dire dans un genre différent, alors que l'autre prise de vue est correctement assignée à l'espèce moléculaire. Le travail sur la moyenne des prises de vue permet d'éviter ce type d'assignations erronées au niveau du genre. Il est donc préférable de travailler sur une forme moyenne par individu que sur une forme uniquement liée à une prise de vue, ceci afin d'éviter toute erreur d'assignation générique.

### **C8. L'effet opérateurs multiples**

L'effet « opérateurs multiples » est susceptible d'affecter l'assignation d'un individu inconnu au jeu de données de référence à deux niveaux. Le premier effet peut être induit par l'utilisation d'un matériel différent entre l'utilisateur et l'Opr. L'outil corrigera automatiquement cet effet (voir paragraphe D2.1) pour un certain nombre de matériels optiques (Annexe III). Le second est lié à l'appréciation du placement des points homologues par l'utilisateur qui peut être différent de celui de l'Opr. C'est ce dernier effet qui est le plus souvent mis en avant comme étant un frein à l'échange de données entre opérateurs en morphométrie géométrique (Dujardin et al., 2010).

#### *C8.1. Effet de la différence d'appréciation du placement des points homologues par l'utilisateur et l'opérateur de référence*

Les configurations de points homologues du crâne en vue dorsale et ventrale ont été simplifiées (chapitre C6) afin d'améliorer le pouvoir discriminant tout en minimisant l'erreur intra-utilisateur. Ainsi pour la vue ventrale, les points homologues 3, 21 et son symétrique 39 ont été retirés ; pour la vue dorsale le point 2 a été supprimé de la configuration d'origine. D'autre part, le nombre de composantes de conformation à conserver pour calculer les meilleures fonctions discriminantes à partir de ces conformations est déterminé. Enfin, l'erreur de mesure intra-opérateur est connue pour chacune d'entre elles. L'étape suivante dans le développement de l'outil web-service est l'évaluation des erreurs de mesures systématiques (voir paragraphe A2.3) induites par l'effet « opérateurs », et celle de leur influence sur la reproductibilité de la production de données.



### C8.1.1. Méthodes

Un échantillon de 31 individus adultes, avec un sexe ratio le plus proche de l'équilibre possible et appartenant au jeu de données de référence a été utilisé pour estimer la variance inter-opérateurs. Ces individus appartiennent à trois mitotypes, *B. indica* (11 individus), *M. surifer* (10 individus), *R. exulans* (10 individus). Ces espèces ont été choisies car elles correspondent à trois tailles d'espèces, *B. indica* représente les espèces de grande taille, *M. surifer* les espèces de taille moyenne et *R. exulans* les espèces de petite taille. Les crânes de ces 31 individus ont été photographiés deux fois en vue dorsale et ventrale par une personne novice avec l'appareil ayant servi à la constitution du jeu de données de référence (Pentax et de modèle K200D avec une lentille Pentax DA 16-45mm F4 ED AL). Le positionnement de l'appareil photo sur le banc photographique a été réalisé par l'utilisateur après qu'il ait été formé à cette technique par l'Opr. Des photos comportant les points homologues de chaque vue ont été fournies à l'utilisateur qui a ensuite été laissé seul pour effectuer la digitalisation sur les prises de vue. Les conformations obtenues par l'utilisateur et celles des mêmes individus appartenant au jeu de données de référence ont été comparées afin de mesurer l'évolution de l'erreur de mesure intra et inter-opérateur sur la taille centroïde et la conformation. L'erreur intra-individuelle (ou opérateur global) a d'abord été estimée à l'aide du protocole de Yezerinac (1992) en considérant les conformations digitalisées par l'utilisateur comme des répliqués des conformations de référence. Dans ce calcul, chaque individu a donc été photographié 4 fois pour chaque vue.

La variance totale est la somme des variances inter-individuelle et intra-individuelle. Cette dernière est équivalente à l'effet opérateur global qui peut être dissociée en variance intra et inter-opérateurs:

$$S^2_{\text{tot}} = S^2_{\text{inter-ind}} + S^2_{\text{inter-op}} + S^2_{\text{intra-op}}$$

Avec :

$S^2_{\text{intra-op}}$  la variance résiduelle ( $MSS_{\text{intra-op}}$ ) qui n'est expliquée ni par la variance inter-individuelle ni par la variance inter-opérateurs.

$$S^2_{\text{inter-op}} = \frac{MSS_{\text{inter-op}} - MSS_{\text{intra-op}}}{2}$$

Pour 2 opérateurs.

$$S^2_{\text{inter-ind}} = \frac{MSS_{\text{inter-ind}} - MSS_{\text{inter-op}}}{4}$$

Pour 2 répétitions par deux opérateurs.

La part de variance inter-individuelle est, en pourcentage de la variance totale:

$$ME_{\text{inter-ind}} = 100 * \frac{S^2_{\text{inter-ind}}}{S^2_{\text{tot}}} \quad (1)$$



La part de variance inter-opérateurs dans la variance totale est :

$$ME_{inter-op} = 100 * \frac{S^2_{inter-op}}{S^2_{tot}} \quad (2)$$

La part de variance intra-opérateur dans la variance totale est :

$$ME_{intra-op} = 100 * \frac{S^2_{intra-op}}{S^2_{tot}} \quad (3)$$

Pour plus d'informations se reporter à Palmer (1994).

### 8.1.2. Résultats

#### *Effet utilisateur sur la taille centroïde*

*Vue ventrale :*

La variance inter-individuelle (équation 1) représente 99,965% de la variance totale de taille. La variance se décompose en variance inter-opérateurs (2) qui représente 0,005% de la variance totale et en variance intra-opérateurs (3) qui représente 0,030% de la variance totale.

*Vue dorsale :*

La variance inter-individuelle (1) représente 99,81% de la variance totale de taille. La variance inter-opérateurs (2) représente 0,14% de la variance totale et la variance intra-opérateurs (3) représente 0,05% de la variance totale.

#### *Effet utilisateur sur la conformation*

*Vue ventrale*

La variance inter-individuelle (1) représente 66,95% de la variance totale de conformation. La variance inter-opérateurs (2) représente 25,21% de la variance totale et la variance intra-opérateurs (3) représente 7,84% de la variance totale.

*Vue dorsale*

La variance inter-individuelle (1) représente 78,6% de la variance totale de conformation. La variance inter-opérateurs (2) représente 10,6% de la variance totale et la variance intra-opérateurs (3) représente 10,8% de la variance totale.

### C8.1.3. Discussion

*Effet opérateur sur la taille centroïde*

L'augmentation de variance induite par la présence d'un second opérateur sur la taille centroïde n'est pas importante. En effet l'erreur opérateur globale ou erreur intra-individuelle représente 0,035% de la variance totale (somme des équations (2) et (3)) pour la face ventrale et 0,19% pour la face dorsale. Cette variance est inférieure à la variance



intra-individuelle induite par l'Opr (0,042%) pour la vue ventrale. Elle est cependant supérieure à celle-ci pour la vue dorsale (0,019%). Dans tous les cas ces variances sont très inférieures à la variance entre les individus et elles peuvent donc être considérées comme négligeables.

#### *Effet opérateurs sur la variance de conformation*

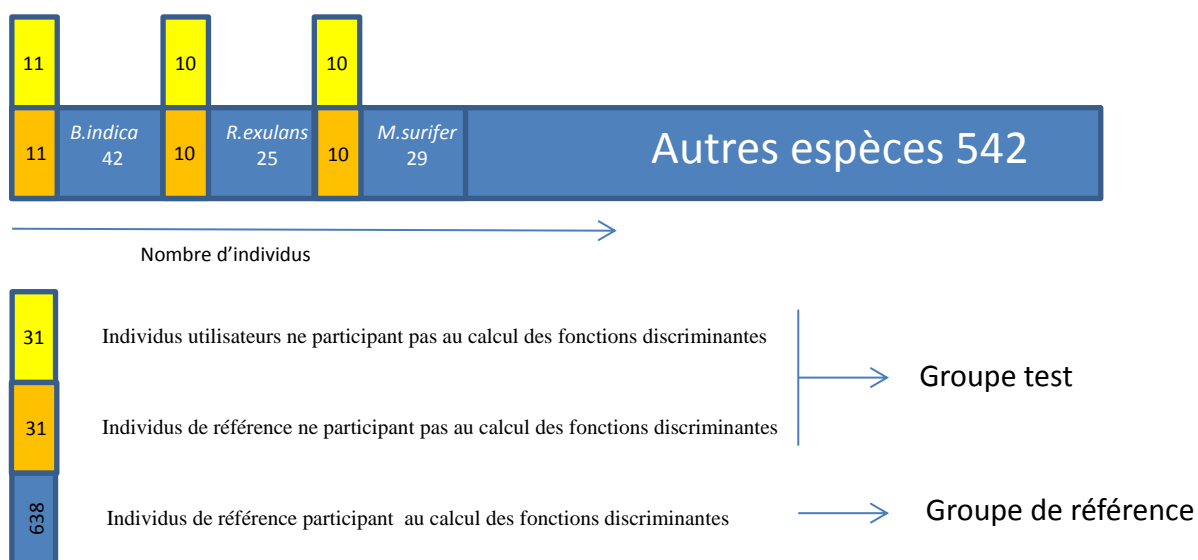
La variance induite par la présence d'un second opérateur pour l'estimation de la morphologie est plus importante qu'elle ne l'était pour l'estimation de la taille (Arnqvist et Martensson, 1998). La variance globale induite par les opérateurs en pourcentage de la variance globale est ici de 21,4% pour la face dorsale et de 33,1% pour la face ventrale. Cette variance est très supérieure à celle générée par l'Opr qui est de 10,2% pour la vue dorsale et de 11,7% pour la vue ventrale. Ce résultat montre que si les données générées par l'utilisateur pour la variance de taille pouvaient être considérées comme un réplikat des données Opr, la situation est différente en ce qui concerne la conformation du crâne. L'apparition d'un utilisateur augmente de façon importante la variance de conformation des crânes qui devient nettement plus importante que la variance générée par un seul opérateur (Arnqvist et Martensson, 1998, Dujardin et al., 2010). D'autre part, cette tendance est plus marquée pour la face ventrale que pour la face dorsale. Pour la vue ventrale, la plus grosse proportion de la variance induite par les opérateurs provient de la variance inter-opérateurs qui est de 25,2%, contre 7,8% pour la variance intra-opérateur. La part importante de la variance inter-opérateurs par rapport à la variance totale ne se retrouve pas pour la vue dorsale où les variances inter-opérateurs et intra-opérateurs sont très proches avec des valeurs respectivement de 10,6 et 10,8% de la variance totale. Cependant même lorsque l'on ne considère que la vue dorsale pour laquelle l'inflation de l'erreur reste modérée, la variance imputable aux opérateurs représente 21,4% de la variance totale contre 10,2% pour la variance intra-opérateur lorsque l'Opr est considéré seul. Cet effet est donc important même si la variance inter-individuelle reste toujours une part importante de la variance totale puisqu'elle représente 78,6% de la variance dans le cas de la vue dorsale et 66,9% dans le cas de la vue ventrale.

### *C8.2. Effet de l'erreur opérateurs sur l'assignation des crânes des individus test*

L'effet opérateur augmente la variance intra-individuelle de conformation de façon plus ou moins importante suivant la vue du crâne considérée. Il est donc important de tester la capacité de l'analyse à assigner les individus malgré l'inflation artificielle de cette variance.

#### *C8.2.1. Méthode*

Le jeu de données des conformations de crânes en vue dorsale et ventrale, incluant d'une part les individus de références de l'outil d'identification (digitalisés par l'Opr voir tableau 1) d'une part, et la fraction de cet échantillonnage digitalisé une seconde fois par l'utilisateur d'autre part, ont fait l'objet d'une superposition



**Figure 26 :** Composition de l'échantillonnage pour tester l'effet utilisateur sur les vues ventrales (l'échantillonnage pour les vues dorsales est très proche et ne diffère que de quelques individus supplémentaires dans le groupe de référence).

**Tableau 11 :** Assignation morphométrique par l'ALD des individus test en face ventrale. Les formes digitalisées par l'Opr (n= 31) et celles digitalisées par l'utilisateur (n= 31) sont assignées de manière congruente avec leur identification mitochondriale malgré l'effet « opérateurs » (le résultat est le même pour la face dorsale).

identification moléculaire	identification morphométrique	
	Utilisateur	Opr
<i>Bandicota indica</i>	22	
<i>Maxomys surifer</i>		20
<i>Rattus exulans</i>		20



Procrustes (voir la figure 26 pour une présentation de l'échantillonnage utilisé). Les 31 individus digitalisés à la fois par l'Opr et par l'utilisateur constituent le groupe test. Ils ont été retirés du jeu de données de référence après l'étape de superposition Procrustes. De cette façon, les individus appartenant à ce groupe sont projetés dans l'espace des conformations (ACP) comme des individus supplémentaires et ne prennent pas part au calcul des fonctions discriminantes qui déterminent, par la suite, leur assignation à une espèce. Autrement dit, c'est le groupe de référence (figure 26) et lui seul qui détermine comment sont ségrégués les groupes morphologiques en fonction de leur mitotype. Ainsi l'effet du groupe test sur l'analyse reste limité au calcul de la conformation moyenne lors de la superposition. L'identification moléculaire des individus du groupe test (digitalisés à la fois par l'Opr et par l'utilisateur) a été masqué à l'analyse. Ils forment donc un groupe d'identification inconnue.

#### *C8.2.2. Résultats*

Les individus du groupe de référence sont assignés de manière congruente à leur identification moléculaire dans 93,8% des cas lorsque l'on considère la face dorsale et dans 96,9% des cas lorsque que l'on considère la face ventrale. Quelle que soit la vue du crâne considérée, lorsque l'on projette les individus du groupe test dans l'espace discriminant et que l'on teste la robustesse de leur assignation par Jackknife, 100% des individus test, digitalisés par l'Opr ou par l'utilisateur, sont assignés à leur groupe moléculaire (tableau 11).

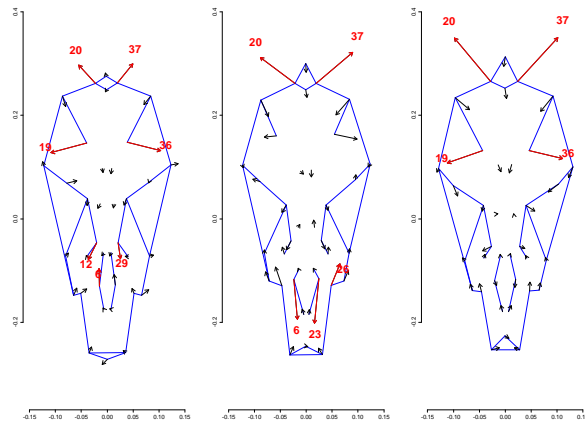
#### *C8.2.3. Discussion*

Le pouvoir discriminant de l'analyse sur le groupe de référence (jeu de données de référence moins le groupe test) est quasiment équivalent à celui sur le jeu de données de référence pour les vues dorsales (93,7%) et ventrales (97%). Le retrait des 31 individus du groupe test a eu peu d'influence sur le calcul des fonctions discriminantes. Les résultats indiquent que l'effet opérateurs multiples n'agit pas sur la capacité d'assignation de l'analyse. Les probabilités postérieures d'assignation des individus aux groupes mitotypiques sont généralement égales ou voisines de 1, que ce soit pour les individus digitalisés par l'Opr ou par ceux saisis par l'utilisateur. Il semble que les spécimens que nous étudions ici possèdent une variance inter-individuelle suffisamment importante pour que l'assignation reste robuste malgré l'augmentation de variance liée à la différence entre opérateurs. Ce résultat est très important car il montre que si la variance induite par l'erreur opérateurs peut être maintenue ou, le cas échéant ramenée à un niveau inférieur ou égal à 33% de la variance totale pour la face ventrale et à 21,4% pour la face dorsale, alors la mise en place d'un outil d'identification est possible.

#### *C8.2.4. Conclusion préliminaire sur la faisabilité de l'outil d'identification morphométrique*

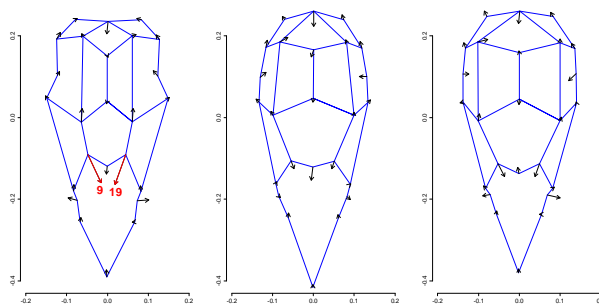
Les résultats obtenus sont très encourageants quant à la possibilité de la mise en place d'un outil d'identification sur ce jeu de données. L'analyse est en effet tolérante à une variance importante induite par l'utilisateur. Cette tolérance est lié au fait que les espèces du jeu de données de référence sont considérées comme occupant l'ensemble de l'espace des conformations. En réalité, d'autres espèces qui n'appartiennent pas à l'échantillonnage de

***Bandicota indica*   *Maxomys surifer*   *Rattus exulans***



a)

***Bandicota indica*   *Maxomys surifer*   *Rattus exulans***



b)

**Figure 27 :** représentation de la variation de la conformation moyenne des espèces du groupe test à l'aide d'espaces vectoriels. La forme moyenne des individus de référence est représentée en bleu pour chaque espèce. Les flèches rouges indiquent le mouvement moyen de chaque point homologue pour obtenir la conformation moyenne de l'utilisateur. La longueur de l'ensemble des vecteurs ainsi définis a été multipliée par trois pour plus de lisibilité. a) face ventrale, b) face dorsale

référence occupent également cet espace. Aussi l'étude du comportement de l'outil lors de la confrontation d'individus provenant d'espèces différentes de celles constituant le jeu de données de référence nous renseignera sur la tolérance qu'il faut préserver afin d'éviter que des morphologies inconnues ne soient assignées à tort à l'échantillonnage de référence (voir chapitre D1).

### *C8.3. Représentation de l'erreur opérateurs*

La variance intra-opérateurs de conformation représente généralement une part plus faible que la variance inter-opérateurs (voir paragraphe C8.1.2). Elle représente cependant une mesure de la **reproductibilité** de la création du jeu de données et sera contrôlée dans le cadre de l'outil d'identification (voir paragraphe D2.3).

L'erreur inter-opérateurs est portée par un ensemble de points homologues répartis sur l'ensemble de la structure étudiée. Il est important de comprendre la répartition de la variance induite par l'effet inter-opérateurs pour savoir si elle est générée régulièrement sur l'ensemble des points ou au contraire seulement sur certains d'entre eux. L'erreur inter-opérateurs est engendrée par la différence de placement des points homologues entre l'Opr et l'utilisateur. Cette différence peut être visualisée afin d'en apprécier la structuration.

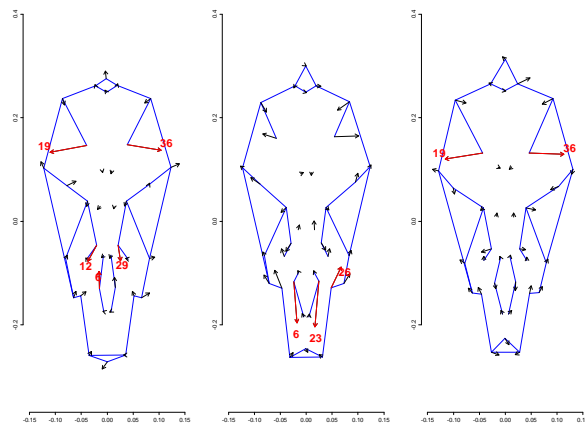
#### *C8.3.1. Méthode*

Pour les vues ventrales et dorsales du crâne des trois espèces test étudiées (*B. indica*, *M. surifer*, *R. exulans*), la variation du positionnement des points homologues a été estimée en comparant les conformations moyennes obtenues par les deux opérateurs. Cette comparaison est basée sur la représentation de vecteurs entre points homologues analogues d'un opérateur à l'autre (Claude, 2008). Ces vecteurs sont représentés avec un code couleur permettant de savoir directement si le positionnement des points est compris (vecteur noir) ou non (vecteur rouge) dans l'**IC<sub>99%</sub>** calculé à partir des individus de référence. Pour plus de lisibilité la longueur des vecteurs lors de l'affichage est multipliée par trois (figure 27 et figure 28).

#### *C8.3.2. Résultats*

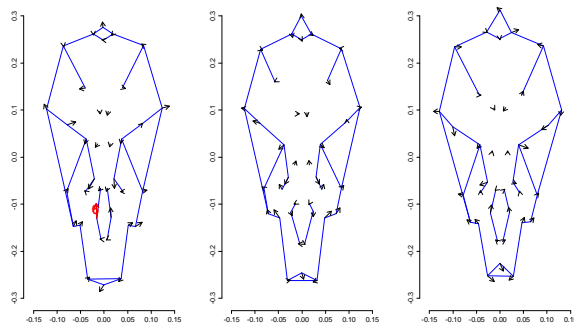
La conformation moyenne des individus digitalisés par l'Opr a été représentée en bleu pour la vue ventrale (figure 27a) et dorsale (figure 27b) de chacune des trois espèces. Les points 20 et 37 placés de part et d'autre du foramen magnum sur les renflements du condyle exoccipital par l'Opr ont été placés sur des structures différentes par l'utilisateur pour les trois espèces. C'est également le cas des points 19 et 36 chez *B. indica* et *R. exulans* qui sont placés à la base externe de l'os du tube eustachien ainsi que les points 6 et 23 chez *M. surifer* qui représentent la plus grande largeur des foramens incisifs. Enfin les points représentant la largeur la plus faible du palais à la M1 (12 et 29) sont placés différemment chez *B. indica*. Ces différences dans le placement des points expliquent certainement la valeur élevée de l'erreur inter-opérateurs observée pour la vue ventrale (i.e. 25,21%). D'autres différences entre opérateurs sont également constatées. C'est le cas du point 6 chez *B. indica* et du point 26 chez *M. surifer*. Cependant, ces écarts sont de moindre ampleur que les précédents et sont produits de façon asymétrique. Ceci suggère qu'elles ne proviennent pas d'un positionnement

*Bandicota indica* *Maxomys surifer* *Rattus exulans*



a)

*Bandicota indica* *Maxomys surifer* *Rattus exulans*



b)

**Figure 28** : représentation de la variation de la conformation moyenne après « entraînement » de l'utilisateur sur les points 20 et 37 (a) et sur l'ensemble des points (b).

différent des points sur la structure mais de la variabilité induite par la présence de deux opérateurs lors du placement des points. Pour la vue dorsale (figure 27b), les différences sont moins marquées et seuls deux points (9 et 19 qui mesurent la plus faible largeur de la constriction inter-orbitale) présentent chez *B. indica* une distance entre opérateurs supérieure à l'IC<sub>99%</sub> des données de référence. Le point le plus postérieur de l'os-interpariétal est également différemment placé par les deux opérateurs mais il entre, malgré tout, dans l'IC<sub>99%</sub> des données de référence de chacune des espèces.

### C8.3.3. Discussion

Les points 6 et 23 sont placés par l'Opr pour qu'ils forment la plus grande largeur des foramens incisifs. Celle-ci se situe en général dans la partie centrale de ces foramens. Cependant, chez *M. surifer*, les foramens palatins ont une forme élargie dans la partie postérieure tout à fait caractéristique. Pour cette raison, les points devraient, pour cette espèce et d'après la règle suivie par l'Opr, être positionnés plus en arrière vers la première molaire. L'utilisateur a, quant à lui, placé ces points chez *M. surifer* au centre des foramens comme il l'a fait pour les autres espèces. Cela explique la différence observée seulement pour cette espèce. Cette différence d'appréciation dans le positionnement des points homologues se retrouve pour les points 20, 37, 19 et 36. Ceci implique que le placement de ces points devra faire l'objet d'une description très précise dans la documentation de l'outil web-service. Lorsque les deux opérateurs ont la même appréciation de l'endroit où doit être placé un point homologue, alors la variance induite par la présence d'un second opérateur est beaucoup plus faible et entre généralement dans l'IC<sub>99%</sub> du placement des points par l'Opr. Ce cas de figure semble se rencontrer pour l'ensemble des autres points (vecteur noirs) à l'exception des points 12 et 29 chez *B. indica* ainsi que 6 et 23 pour *M. surifer*. Ces résultats sont en accord ceux de l'étude de Rasmussen et al. (2001) qui montrent que des mesures prises par des opérateurs expérimentés influencent moins l'assignation à un groupe pré-défini que des mesures prises par des opérateurs inexpérimentés. En effet, l'opérateur expérimenté, quel qu'il soit, positionne de façon plus adéquate l'instrument de mesure, ou ici le point homologue, que l'utilisateur novice. Le niveau des opérateurs est donc, logiquement, un point important à prendre en compte pour pouvoir comparer des données prises par deux opérateurs (Dujardin et al., 2010). D'autre part le mauvais placement de certains autres points (comme 12, 29, 6 et 23) sont espèce-dépendants, ce qui montre que le placement des points par l'utilisateur doit être vérifié pour chaque espèce.

Cependant il est difficile d'estimer quel pourcentage d'erreur inter-opérateurs est acceptable pour que l'identification reste fiable, et ce d'autant plus que toutes les espèces recensées dans la zone d'étude ne sont pas représentées dans notre jeu de données de référence. Dans le cadre de cette expérimentation, une variance inter-opérateurs importante pour la face ventrale n'entrave pas les identifications des individus lors de l'analyse discriminante. Cependant, l'origine de la majeure partie de cette variance était concentrée dans le placement inadéquat de quelques points homologues ; pourtant, l'analyse s'est montrée robuste à l'introduction de cette variance artificielle. Dans l'hypothèse d'une variation entre opérateurs, de même ampleur, mais induite par des points homologues différents, ou, alternativement, d'une variation répartie de façon plus uniforme sur l'ensemble des points

**Tableau 12:** synthèse des variances induites par l'opérateur de référence et par un utilisateur avant (données brutes) et après son entraînement. Une première correction effectuée par l'utilisateur porte sur les points placés le plus différemment de ceux placés par l'opérateur de référence (20 et 37). Une seconde correction porte sur l'ensemble des points qui n'entre pas dans l'IC<sub>99%</sub> des données produites par l'opérateur de référence. La variance de taille est toujours négligeable alors que la variance de conformation s'améliore lors de l'entraînement de l'utilisateur pour devenir comparable à la variance induite par l'opérateur de référence.

	opérateur référence	données brutes	utilisateur correction points 20, 37	correction de tous les points
<b>Taille</b>				
Intra-opérateurs	0,042	0,030	0,032	0,028
Inter-opérateurs		0,005	0,032	0,033
<b>Somme : effet opérateurs</b>	<b>0,042</b>	<b>0,035</b>	<b>0,064</b>	<b>0,062</b>
<b>Conformation</b>				
Intra-opérateurs	11,6	7,8	8,0	8,5
Inter-opérateurs		25,2	15,0	4,9
<b>Somme: effet opérateurs</b>	<b>11,6</b>	<b>33,0</b>	<b>23,0</b>	<b>13,4</b>

homologues, le résultat aurait pu être fort différent. L'étude du comportement de l'analyse à l'accroissement et à la répartition de l'erreur opérateurs pourrait se faire en simulant une augmentation de la variance progressive sur l'ensemble des points homologues, ou sur des binômes de points symétriques, afin de localiser la tolérance possible pour chacun des points par rapport à la variation des autres. Cependant, cette approche par simulation sort du cadre du présent travail, et devra être entreprise dans le cadre d'une autre étude pragmatique et empirique sur la réduction de l'effet opérateurs.

#### *C8.4. Réduction de l'effet inter-opérateurs*

L'outil doit être capable d'identifier des individus inconnus qui appartiennent à des espèces effectivement référencées dans le jeu de données de référence mais il doit aussi se montrer capable de gérer des individus d'espèces non représentées dans le jeu de données de référence. La gestion de cette contrainte (voir chapitre D1) peut rendre l'outil plus sensible aux effets opérateurs. L'erreur inter-opérateurs est l'effet « opérateurs » qui apporte le plus de variance par rapport à un opérateur unique (voir paragraphe C8.1.2). Il est donc utile de comprendre comment il est possible de le maîtriser. Pour cela l'outil doit être en mesure d'emmener d'une façon ou d'une autre l'utilisateur à positionner les points homologues de façon à réduire les variances inter-opérateurs à des niveaux les plus proches possibles de ceux de la variance intra-opérateur de l'Opr. Dans ce cas, l'effet utilisateur sera presque totalement supprimé.

##### *C8.4.1. Méthode*

La représentation des patrons de différence entre l'utilisateur et l'Opr produite précédemment a été montrée à l'utilisateur novice et il lui a été demandé de reprendre son jeu de données pour les points 20 et 37. Pour cela, des photographies montrant le placement des points homologues digitalisés par l'Opr lui ont été fournies. Une fois ce travail effectué, les patrons de différence entre Opr et utilisateur ont à nouveau été inférés (figure 28a) et les variances intra- et inter-utilisateurs ont été recalculées pour la conformation de la vue ventrale du crâne (tableau 12). Cette correction correspondrait à un premier niveau d'exigence de la part de l'outil d'identification avec une première réduction de la variance générée par les opérateurs. Il a ensuite été demandé à l'utilisateur de modifier l'ensemble des points dont la distance moyenne est supérieure à l'IC<sub>99%</sub> des données produites par l'Opr. Cette correction correspond à un second niveau d'exigence de l'outil d'identification (figure 28b).

##### *C8.4.2. Résultats*

La figure 28a montre que le positionnement des points 20 et 37 entre les deux opérateurs est devenu tout à fait comparable. Les variances inter-opérateurs et intra-opérateurs pour la taille représentent 0,032% de la variance totale (tableau 12). La variance inter-opérateurs de conformation représente 15% de la variance totale et la variance intra-opérateurs représente 8% de la variance totale. La variance inter-individuelle représente 77% de la variance totale de conformation.





La figure 28b montre que le positionnement des points entre les deux opérateurs est similaire excepté pour le point 6 chez *B. indica* dont les coordonnées sont toujours au-delà de l'IC<sub>99%</sub> des données produites par l'Opr. Pour la taille, la variance inter-opérateurs représente 0,033% de la variance totale et la variance intra-opérateurs 0,028% (tableau 12). La variance inter-opérateurs de conformation représente 4,9% de la variance totale et la variance intra-opérateurs en représente 8,5%. La variance inter-individuelle représente 86,6% de la variance totale de conformation.

#### *C8.4.3. Discussion*

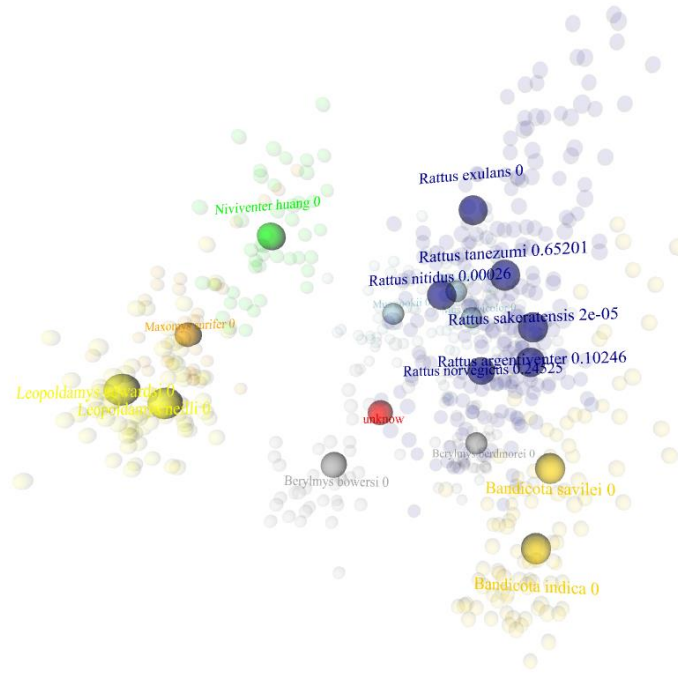
Le tableau 12 résume l'évolution des variances en fonction de l'entraînement de l'utilisateur. Pour la taille, l'effet intra-opérateur reste stable, ce qui montre que la répétabilité est bonne. Etrangement, la variance inter-opérateurs augmente au cours de l'entraînement de l'utilisateur. Ceci est vraisemblablement lié au fait que la variance opérateurs est exprimée comme un pourcentage de la variance totale : cette dernière diminue très fortement au cours de l'entraînement grâce à la réduction importante de la variance inter-opérateurs de conformation, ce qui a pour effet d'augmenter artificiellement la part de la variance inter-opérateurs dans la variance totale. Celle-ci reste cependant toujours très négligeable. Comme pour la taille, la variance intra-opérateur de conformation varie peu durant l'entraînement de l'utilisateur, représentant toujours environ 8% de la variance totale. La répétabilité ne varie donc pas. A l'opposé, la variance inter-opérateurs qui représente la part la plus importante de la variance induite dans le système par la présence de deux opérateurs, diminue fortement pour devenir finalement inférieure à la variance intra-opérateur. L'effet « opérateurs » pour la conformation varie ainsi de 33% à 13,4% entre le début et la fin de l'entraînement pour atteindre une valeur proche de l'erreur de mesure générée par l'Opr (11,6%). En fin d'entraînement, les données produites par l'utilisateur sont un quasi-réplicat de celles produites par l'Opr.

#### *C8.4.4. Conclusion*

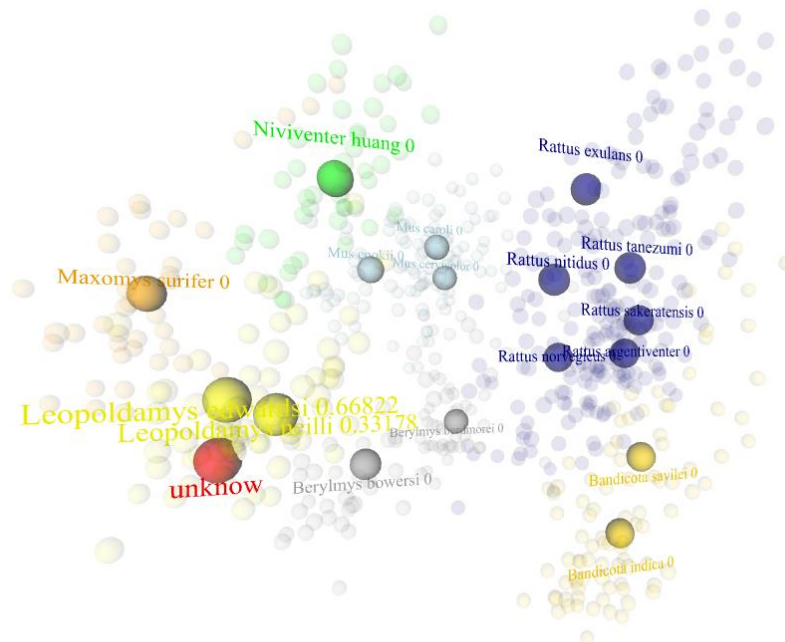
La variance de l'utilisateur peut être contrôlée et être ramenée à un niveau comparable à celle produite par l'opérateur de référence. Il sera donc possible de conformer les données produites par l'utilisateur aux exigences de l'outil d'identification web-service à condition de mettre en place des procédures d'accompagnement des utilisateurs qui soient claires, précises et didactiques.

### ***D- Mise en place de l'outil d'identification morphométrique***

La mise en place de l'outil se révèle possible techniquement dans la mesure où, comme nous venons de le démontrer, il est possible d'encadrer l'effet lié à la présence d'un second opérateur. Cependant, il reste à définir les limites de l'outil ainsi que la qualité des données qu'il faudra lui fournir pour permettre l'assignation d'un individu à une des entités de l'échantillonnage de référence.



1)



2)

**Figure 29<sub>a</sub>** : Vues dynamiques générées par l’outil lors de la soumission de différents mitotypes qui ne sont pas connus du jeu de données de référence. a) *R. rattus* et b) *L. sabanus* ne sont pas assignés car leur probabilités postérieures d’assignation sont plus faibles que la plus faible valeur des individus de référence identifiés morphologiquement dans ce groupe.

## ***D1. Test de l'outil d'identification***

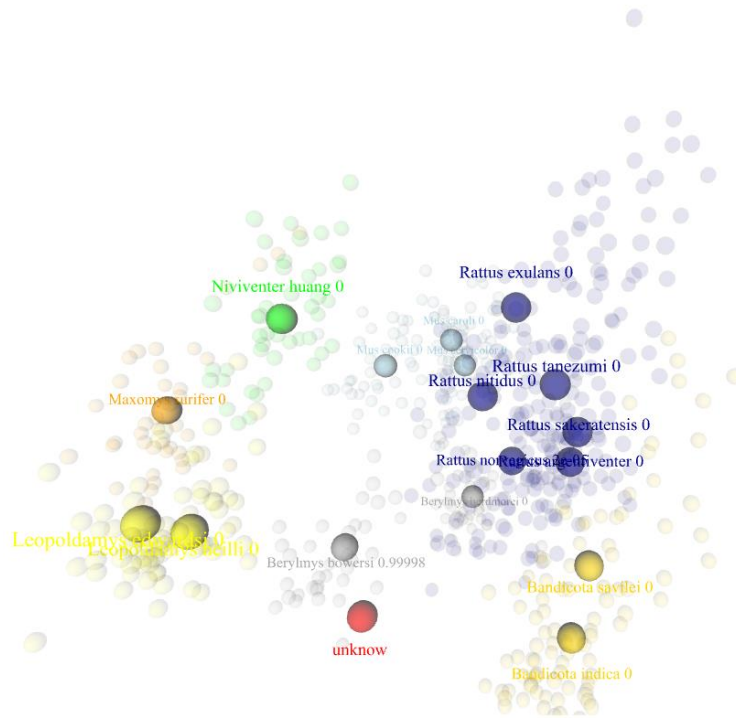
L'analyse discriminante a montré une bonne tolérance à l'erreur (même importante) induite par la présence de deux opérateurs. Cependant la dispersion artificielle des données liée à l'utilisateur ne sera acceptable que si l'outil se montre également robuste pour gérer des espèces qui ne sont pas présentes dans le jeu de données de référence.

### ***D1.1. Méthode***

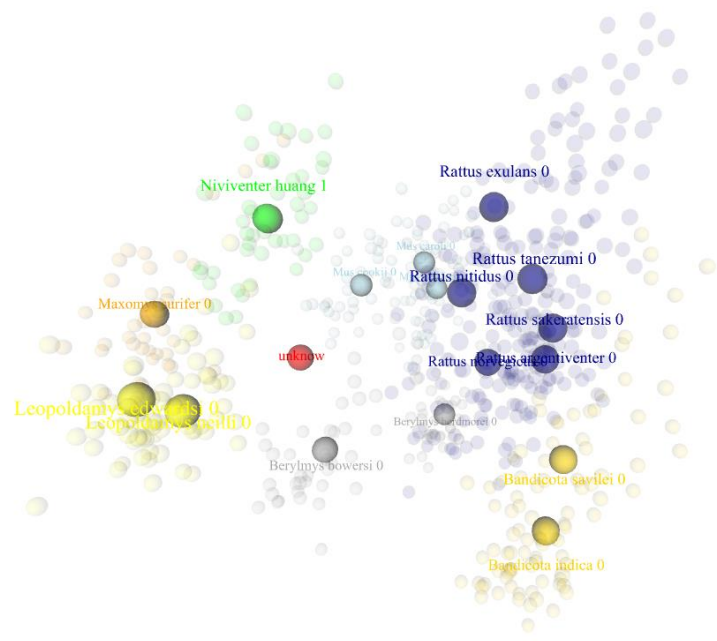
L'outil a été testé avec différents mitotypes n'appartenant pas aux jeux de données de référence, à savoir des individus des genres *Rattus* (*R. rattus*, *R. andamanensis*, *R. tiomanicus*), *Leopoldamys* (*L. sabanus*), *Niviventer* (*N. fulvescens*, *N. tenaster*), *Maxomys* (*M. rajah*, *M. hylomyoides*, *M. whiteheadi*) mais aussi *Saxatilomys paulinae*, *Cannomys badius*, *Sundamys muelleri*, *Micromys minutus*, *Rhizomys pruinosus*. Ce test a pour objectif de définir le meilleur paramétrage de l'outil d'identification et, par voie de conséquence, de fixer les exigences qu'il faut avoir en termes d'encadrement de l'effet opérateurs. Le paramétrage a pour but de permettre à l'outil de gérer des espèces non référencées pour que des individus appartenant à celles-ci ne soient pas assignés aux espèces du jeu de données de référence. Pour cela, nous avons fixé deux contraintes agissant l'une au niveau de la probabilité postérieure d'assignation, l'autre à celui de la distance de Mahalanobis. De plus, des individus appartenant à des espèces présentes dans le jeu de données de référence et provenant de la zone étudiée ont été identifiés à la suite de ce paramétrage pour vérifier que les conditions définies ne soient pas trop sévères et ne compromettent pas leur identification.

### ***D1.2. Résultats***

Il a tout d'abord fallu tenir compte de l'incapacité de l'analyse à assigner morphologiquement tous les individus d'un mitotype dans un même groupe moléculaire (erreurs d'identification). Ainsi, des individus peuvent être assignés à un groupe différent de leur groupe mitotypique avec une probabilité postérieure d'assignation faible. Un seuil de valeur minimale de probabilité postérieure d'assignation a donc été défini pour éviter qu'un individu inconnu ne soit assigné à un groupe avec une probabilité d'assignation trop faible, ce qui pourrait n'être que la signature d'une assignation erronée entre entités de référence. D'après ce seuil, pour être assigné effectivement à un groupe mitotypique de référence, un individu inconnu doit posséder une probabilité postérieure d'assignation supérieure ou égale à la plus petite probabilité postérieure des individus du jeu de données de référence assignés morphologiquement à ce groupe mitotypique. De cette façon les individus appartenant à des mitotypes inconnus du jeu de données de référence et qui sont assignés à un groupe par l'analyse discriminante avec une probabilité postérieure trop faible ne sont pas identifiés par l'outil et cette non-identification est explicite. La figure 29a est un exemple de l'application de ce seuil à des individus appartenant aux mitotypes correspondant à *R. rattus* (1) et *L. sabanus* (2). Ces mitotypes ne sont pas connus du jeu de données de référence. Sans paramétrage, l'analyse discriminante assigne ces individus respectivement aux groupes *R. tanezumi* et *L. edwardsi* avec des probabilités postérieures de 0,652 et 0,668. Or les probabilités postérieures les plus faibles du dernier individu de référence assigné à chaque



1)



2)

**Figure 29<sub>b</sub>** : même figure que la figure 29 mais pour des cas de figures différents. a) *Sundamys muelleri* et b) *Niviventer tenaster* ne sont pas assignés malgré une probabilité postérieure d'assignation élevée car la distance morphologique euclidienne qui les sépare de la conformation moyenne est plus élevée que celle de l'individu de référence assigné morphologiquement à ce groupe.

groupe sont de 0,789 et 0,990. En appliquant notre seuil, les nouveaux individus ne sont pas assignés par l'outil. Cette non assignation se traduit visuellement par le fait que la sphère rouge, qui matérialise l'individu inconnu, n'est pas liée à la forme moyenne du groupe (voir figure 31 à la page 62 verso, montrant le résultat d'une assignation) mais reste isolée dans l'espace des conformations (figure 29a).

D'autre part, afin de ne pas influencer les fonctions discriminantes, l'individu que l'utilisateur souhaite identifier est projeté dans l'espace discriminant. Si l'outil se borne à donner simplement les résultats bruts de l'analyse discriminante, alors les données de l'utilisateur sont associées au groupe connu qui possède la probabilité postérieure d'assignation la plus grande sans tenir compte de la distance de Mahalanobis (celle-ci sépare l'individu de la conformation moyenne du groupe). Dans ces conditions des individus d'espèces non représentées dans le jeu de données de référence peuvent être facilement assignés à des espèces de référence. Une limite à la distance de Mahalanobis séparant l'individu inconnu et la moyenne a donc été définie : celle-ci doit être inférieure à la distance maximale du dernier individu correctement assigné morphologiquement dans le groupe mitotypique considéré. La figure 29b représente cette situation pour deux mitotypes non référencés. L'un est un individu appartenant au mitotype *Sundamys muelleri* (1) et l'autre est un individu appartenant au mitotype *Niviventer tenaster* (2) espèce morphologiquement proche de *N. huang*. Si aucune contrainte n'est fixée, ces individus sont assignés respectivement aux mitotypes *B. bowersi* et *N. huang* avec des probabilités postérieures élevées, (respectivement 0,999 et 1). Pourtant, la distance de Mahalanobis les séparant de la conformation moyenne est supérieure à la distance maximale du groupe de référence, et leur assignation au groupe n'est donc pas réalisée par l'outil (pas de lien entre l'individu inconnu et la forme moyenne du groupe).

Grâce à l'une ou l'autre des valeurs seuil définies, l'outil n'a pas donné d'assignation pour l'ensemble des espèces absentes du jeu de données de référence excepté pour trois individus qui appartiennent respectivement aux mitotypes *M. rajah*, *M. hylomyoides*, *N. fulvescens*. Les deux premiers individus ont été identifiés comme appartenant à l'espèce *M. surifer* et le troisième à l'espèce *N. huang*. D'autre part, l'outil a assigné, pour divers mitotypes (*R. exulans*, *B. indica*, *B. savilei*, *B. bowersi*, *M. caroli*) deux individus avec succès. Les seuils fixés n'entravent donc pas l'identification d'individus nouveaux mais appartenant à des espèces référencées.

En revanche, la contrepartie des contraintes sur l'identification des individus inconnus est qu'elles abaissent très fortement la tolérance de l'outil à la variance induite par un utilisateur lorsque celui-ci soumet un individu appartenant à un mitotype de référence. En effet l'utilisateur novice, au travers de l'erreur qu'il génère, soumet une estimation approximative de la conformation du crâne. Les seuils que nous avons posés interdisent alors l'assignation de l'individu. Ainsi les données brutes fournies par un utilisateur novice qui apportent une variance importante et celles partiellement corrigées (voir chapitre C8.4) ne permettent plus l'assignation des individus soumis par l'utilisateur à leur groupe mitotypique. Seules des données produites par un utilisateur parfaitement entraîné (chapitre C8.4 et tableau 12) permettent une assignation des individus



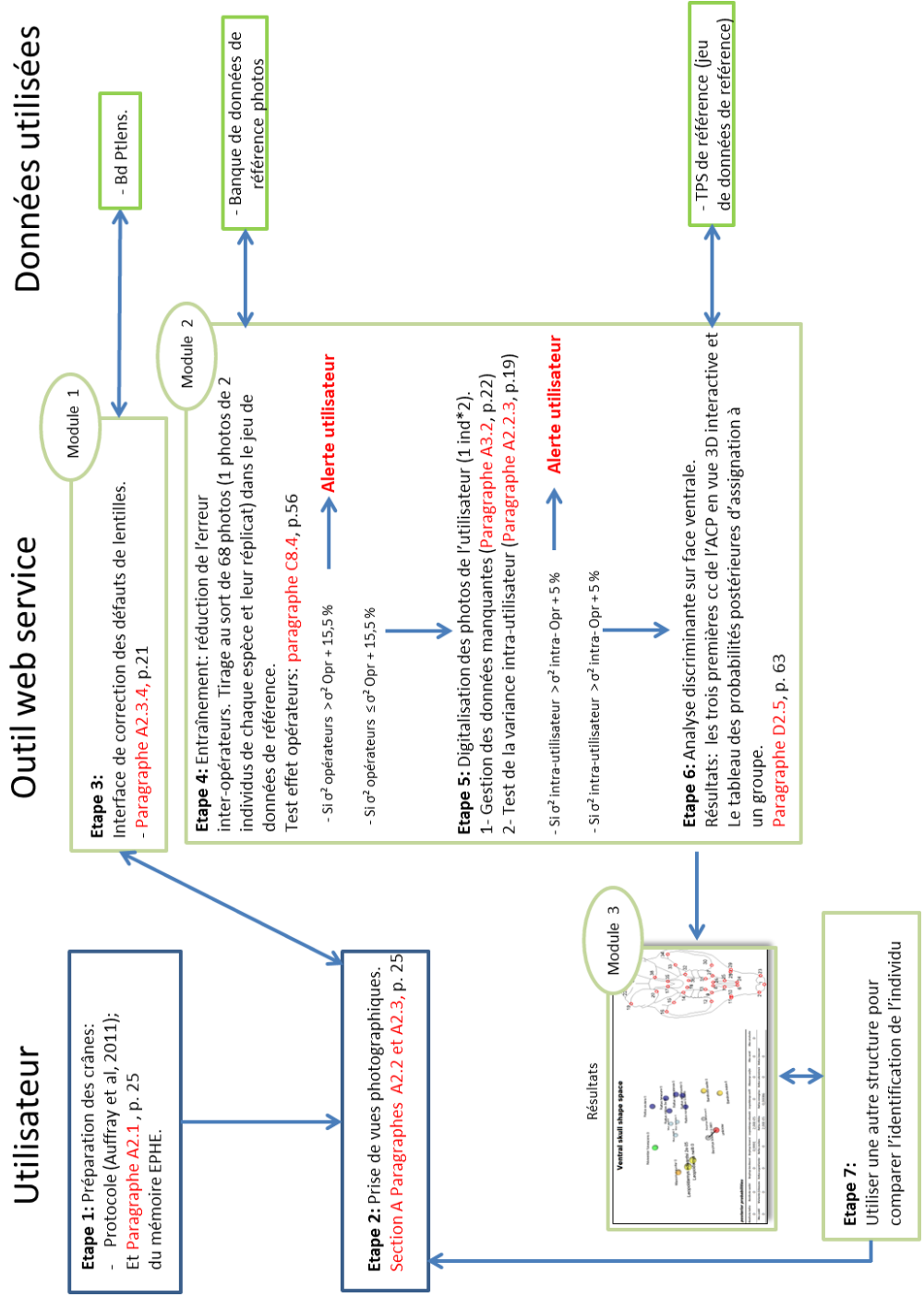
### D1.3. Discussion

Les seuils précédemment définis permettent de limiter les erreurs d'assignation d'espèces inconnues aux individus ayant à la fois une très forte probabilité d'assignation à un groupe de référence et une distance de Mahalanobis inférieure ou égale à la plus grande distance des individus correctement assignés morphologiquement à ce mitotype. Ces erreurs sont liées à des individus qui se trouvent dans les zones de recouvrement phénotypique des espèces et pour lesquels il serait, de toute façon, impossible de trancher à l'aide de la morphologie. C'est là la limite de l'outil d'identification morphologique. Cette limite ne pourrait être levée que via l'intégration de nouveaux éléments du squelette et relèverait probablement d'analyses au cas par cas, ce qui dépasse le cadre de ce mémoire, et de sa philosophie, à savoir la mise en œuvre d'un outil à la disposition du plus grand nombre. Cette limite ne s'applique par ailleurs probablement que pour un nombre très limité de situations réelles.

L'étude menée sur des mitotypes non représentés dans l'échantillonnage de référence montre que certains d'entre eux peuvent être confondus avec des espèces lui appartenant. C'est le cas d'un individu appartenant au mitotype *M. hylomyoides* qui provient de Sumatra et ne se rencontre pas sur la zone d'étude. C'est également le cas d'un individu appartenant à l'espèce *M. rajah*, espèce sondaïque qui se rencontre seulement dans l'isthme de Kra en Thaïlande. La confusion entre les deux espèces sera à signaler aux utilisateurs qui devront utiliser l'outil moléculaire pour distinguer ces espèces ([http://www.ceropath.org/barcoding\\_tool/rodentsea](http://www.ceropath.org/barcoding_tool/rodentsea)). *N. fulvescens* se rencontre lui aussi en Thaïlande mais il est inféodé aux massifs karstiques et n'a été échantillonné à ce jour que dans la province de Kanchanaburi à l'ouest du pays (Latinne, 2012a, Balakirev et al., 2011). Là encore l'utilisateur devra utiliser l'outil moléculaire pour distinguer ces espèces dans ces régions. Pour les mitotypes présents dans le jeu de données, les résultats préliminaires sont très encourageants. Néanmoins, des essais seront poursuivis pour tester l'outil avec des individus supplémentaires de mitotypes appartenant à l'échantillonnage de référence. L'outil s'est montré informatif concernant l'appartenance éventuelle d'un mitotype inconnu du jeu de données de référence à un genre connu de celui-ci. Ainsi si des individus appartenant à des mitotypes tels que *R. rattus*, *R. andamanensis* ou *R. tiomanicus* ne sont pas assignés par l'outil, on observe que les probabilités postérieures d'assignation non nulles à des mitotypes du genre *Rattus* (figure 29a) laissent à penser que l'animal est susceptible d'appartenir à ce genre. Ces informations phénétiques doivent cependant être prises avec une valeur relative car il est établi que les homoplasies (convergence, réversion) sur les caractères morphologiques sont courantes (Zelditch et al., 1995).

Enfin les seuils que nous avons déterminés et qui permettent d'éviter l'identification erronée d'espèces qui ne sont pas présentes dans le jeu de données de référence obligent l'utilisateur à étalonner sa prise de données sur celle de l'Opr s'il souhaite obtenir une assignation par l'outil d'identification.

# Schéma de principe de l'outil d'identification morphométrique



**Figure 30:** Schéma de principe de l'outil d'identification morphologique.



## ***D2. Schéma de principe de l'outil***

La figure 30 schématise le fonctionnement de l'outil qui est détaillé dans les paragraphes suivants. Chacun des trois modules correspond à une interface web programmée en langage Java Script qui permettra l'entrée de données et l'affichage de résultats. Les données de l'utilisateur seront alors traitées et comparées au jeu de données de référence à l'aide de scripts rédigés en langage R et développés dans le présent travail.

### ***D2.1. Premier module : interface de correction de la lentille et de traitement de l'image***

Les appareils photographiques corrigés par l'outil d'identification sont ceux dont les caractéristiques optiques sont contenues dans la base de données PTlens (Annexe III). L'utilisateur sera invité à télécharger les photos de(s) l'individu(s) à identifier au format .zip ou .rar puis à sélectionner les références de la lentille qu'il a utilisée pour faire les prises de vue. La distorsion cylindrique de l'image sera alors corrigée et un fichier compressé lui sera renvoyé. Ce module permettra également la modification de certains paramètres élémentaires des photos, tels que son orientation, sa taille, son contraste, sa lumière ou sa transformation en négatif. Ces transformations sont utiles pour augmenter la précision du placement des points homologues par l'utilisateur. D'après les tests réalisés lors du présent travail (données non montrées), l'effet de la lentille n'est pas une part prédominante de l'effet « opérateurs » si des appareils de type reflex sont utilisés. Cependant, l'ensemble des appareils du marché ne pouvant être testé, il est conseillé d'utiliser un matériel optique pouvant faire l'objet d'une correction.

### ***D2.2. Deuxième module : entraînement de l'utilisateur (test de répétabilité)***

Dans un deuxième temps l'utilisateur sera dirigé vers le module d'entraînement à la digitalisation. L'interface sera la même que celle utilisées pour la digitalisation et la soumission de points homologues placés sur les individus à identifier. L'utilisateur sera invité à comparer sa digitalisation à celle de l'opérateur de référence pour le calcul de l'effet « opérateurs ». Il est libre d'accepter ou de passer à l'étape suivante mais sera informé que, si tel est son choix, l'outil ne sera certainement pas capable d'assigner les individus qu'il propose à une espèce de référence. Si l'utilisateur accepte la comparaison, il se verra proposer un jeu de 68 photographies de crânes de référence en vue ventrale sur lesquelles il devra digitaliser une configuration de points homologues (la configuration de points simplifiés établie au paragraphe C 6.2.2 et visualisée à la figure 25b). Ce jeu de données sera composé des deux vues répliquées de deux individus appartenant à chacune des 17 espèces en présence. L'utilisateur pourra digitaliser les vues directement sur l'interface web ou télécharger un fichier compressé contenant les photos et un fichier au format **TPS** vierge référençant les prises de vues dans un ordre aléatoire. Ce fichier et les photos associées permettront à l'utilisateur de produire un jeu de données qui pourra être comparé à son répliquat dans le jeu de données de référence. Une fois le jeu de données soumis (fichier TPS rempli), l'outil calculera la **répétabilité** du placement des points homologues d'un opérateur à l'autre. Un tableau présentera l'effet « opérateur » (ligne en gras du tableau 12; pour la taille sur la



première ligne et pour la conformation sur la seconde) pour chaque opérateur en colonne (Opr et utilisateur). Si l'effet « opérateurs » est supérieur à l'effet Opr avec une tolérance de 15,5% (marge calculée en fonction des résultats de l'utilisateur entraîné au tableau 12), alors la valeur correspondante dans le tableau s'affichera en rouge et la représentation des patrons de différence entre l'utilisateur et l'Opr sera affichée afin de permettre à l'utilisateur d'identifier rapidement les points placés différemment. Il sera alors invité à corriger sa digitalisation et sera aidé en cela par la possibilité qui lui sera offerte de visualiser des photos digitalisées par l'Opr. Un nouveau jeu de données aléatoire et son fichier TPS vierge correspondant lui sera alors fourni pour réaliser une nouvelle évaluation de la digitalisation qu'il produit.

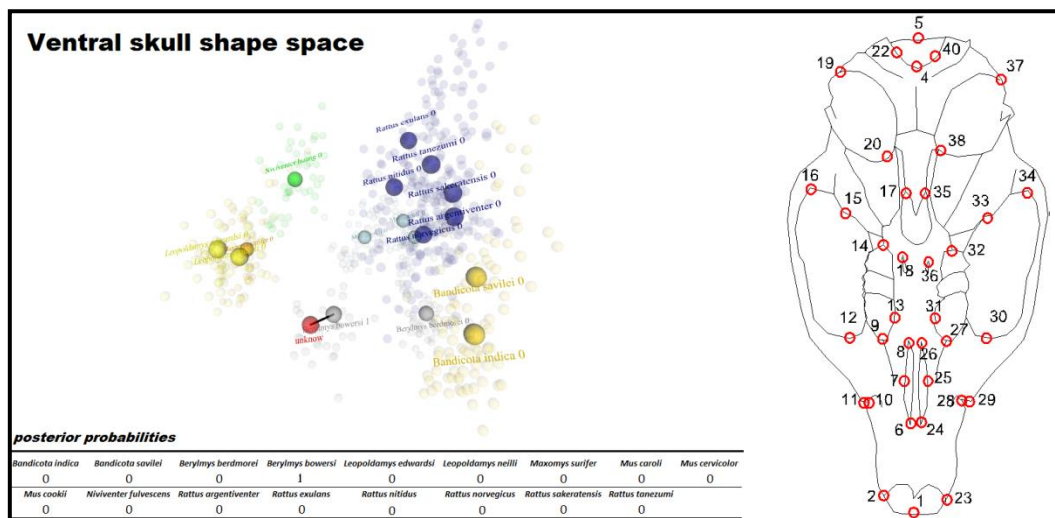
Si l'erreur « opérateurs » est limitée à un niveau proche de celle de l'Opr (toujours avec 15,5% de tolérance), alors l'utilisateur sera invité à soumettre ses propres individus l'outil d'identification.

### *D2.3. Deuxième module : digitalisation d'un individu inconnu (test de reproductibilité)*

Pour l'identification d'un individu inconnu, l'utilisateur sera invité à télécharger une première photographie (ayant été au préalable corrigée à l'étape 3 par le module 1) et à digitaliser la configuration de points correspondant à la vue choisie. Si des données sont manquantes, l'utilisateur devra être en mesure de le signaler par un symbole « NA » (pour « non available » en anglais) à l'emplacement supposé du point correspondant. Les configurations comprenant plus d'un point « NA » (soit deux coordonnées) seront refusées. Dans le cas contraire la donnée manquante sera calculée (chapitre A3). Une fois la première conformation soumise à l'outil, une seconde prise de vue du même individu sera demandée. Cette seconde prise de vue permettra de tester la **reproductibilité** de la production de données par l'utilisateur. La **reproductibilité** est une mesure de la **répétabilité** mais qui prend également en compte le placement du crâne et de l'appareil photographique. Le calcul de l'erreur intra-utilisateur pour la taille et la conformation sera réalisé pour la taille et la conformation. Si l'erreur de mesure est supérieure à l'erreur générée par l'Opr (avec un seuil arbitraire de 5%) alors l'utilisateur sera invité à relire les instructions de placement des crânes lors de la prise de vue et à recommencer l'étape de digitalisation avec de nouvelles photographies corrigées du crâne. L'utilisateur pourra choisir d'ignorer cette nouvelle alerte et soumettre l'individu à l'identification tel quel.

### *D2.4. Deuxième module : assignation de l'individu inconnu*

L'individu inconnu est projeté dans l'espace des conformations en multipliant ces coordonnées Procrustes par les vecteurs propres calculés à partir des individus de référence. Ces coordonnées dans l'espace discriminant sont calculées par l'intermédiaire des fonctions discriminantes calculées sans lui et la robustesse de l'assignation est testée par Jackknife. L'individu est alors assigné à un groupe en tenant compte des seuils fixés au chapitre D1, à savoir une probabilité postérieure minimale et une distance de Mahalanobis maximale au groupe d'affection.



**Figure 31** : Sortie dynamique et interactive de l'outil d'identification pour la vue ventrale du crâne. Le crâne inconnu (sphère rouge) est lié à son mitotype d'assignation, *B. bowersi*, par une ligne représentant la distance de Mahalanobis. Les probabilités postérieures d'assignation à chaque espèce, obtenues sont affichées en dessous. Chaque sphère, qui représente la conformation moyenne pour chaque espèce, est liée dynamiquement à la fiche espèce correspondante du site de Ceropath.

### D2.5. Troisième module : affichage du résultat

Le résultat est affiché sur le navigateur de l'utilisateur sous la forme d'une représentation tri-dimensionnelle des trois premières composantes de l'espace des conformations (figure 31).

L'animal inconnu est représenté par une sphère de couleur rouge. Les autres sphères de même taille, dont la couleur correspond aux genres taxonomique de l'échantillonnage de référence, représentent la forme moyenne de chaque espèce. Les individus de référence sont quant à eux représentés par des sphères plus petites avec une transparence plus marquée. La représentation des individus donne une idée de la dispersion des données autour de la valeur moyenne. Ces sphères (espèces ou individus) sont interactives et offrent des liens dynamiques vers les pages du muséum virtuel du site internet du programme CERoPath correspondant aux espèces (à titre d'exemple [http://www.ceropath.org/rdbsea/species/bandicota\\_indica](http://www.ceropath.org/rdbsea/species/bandicota_indica)) ou aux individus (<http://www.ceropath.org/rdbsea/individual/I0142>). La valeur de la probabilité postérieure d'assignation de l'individu inconnu à chaque groupe est affichée à la suite du nom de chaque espèce (figure 29 a et b et figure 31). Ces valeurs sont synthétisées par espèce dans un tableau qui s'affiche sous le graphique. Un dessin est positionné à droite de l'écran et présente la configuration de points homologues étudiée. Si l'animal est assigné à un groupe de référence, alors son assignation est matérialisée par une ligne solide qui le lie à son groupe d'affectation (figure 31). Dans l'exemple fourni par la figure 31, l'individu est assigné à l'espèce *B. bowersi* avec une probabilité postérieure d'assignation de 1. Dans le cas où l'individu n'est pas assigné par l'outil, la sphère représentant l'individu ne sera reliée à aucun groupe par une ligne solide comme c'est le cas pour les exemples représentés à la figure 29.

### D2.6. Choix de la structure étudiée

La vue ventrale du crâne est la seule à offrir une assignation au genre qui est sans équivoque. Aussi, cette vue sera la première utilisée pour l'identification afin de déterminer avec précision le genre de l'animal étudié. Si l'individu soumis à l'outil appartient à une lignée mitochondriale non contenue dans l'échantillonnage de référence ou si l'utilisateur ne souhaite pas conformer sa prise de données à celle de l'Opr, l'individu ne sera généralement pas assigné à un groupe mais les probabilités postérieures peuvent le renseigner sur son appartenance à un genre (voir le chapitre D1).

Cette identification pourra être complétée par une identification à partir de la face dorsale, notamment pour le genre *Leopoldamys* et par des identifications sur les hémimandibules externes pour le genre *Bandicota*, comme nous avons pu le voir plus haut. Dans ce cas, l'opérateur pourra reprendre toutes les étapes de l'identification et tester la répétabilité et la reproductibilité du placement des points homologues s'il le souhaite. Cependant, pour les hémimandibules, l'échantillonnage de référence n'est pas complètement finalisé et nous avons conscience qu'un effort d'échantillonnage supplémentaire devra être réalisé avant de proposer ce service.



### ***D3. Discussion : limites et améliorations possibles de l'outil, conseils de mise en ligne***

#### ***D3.1. Limites liées à la conformation moyenne***

Pour être comparées entre elles, les conformations ont été ici superposées à la conformation moyenne ou consensus. Aussi le jeu de données utilisateur influence-t-il le jeu de données de référence lors du calcul du consensus. Or, ce dernier est pris comme référence lors de la superposition et comme origine lors de la projection dans l'espace euclidien. Le jeu de données de référence comprend 660 individus, ce qui implique une influence très limitée de l'ajout d'une morphologie nouvelle sur le calcul de la forme moyenne. Au contraire, la soumission de nombreux individus simultanément pourrait influencer la conformation du consensus de façon non négligeable et, dans certains cas, modifier le résultat de l'analyse. Autrement dit il est souhaitable que l'analyse des crânes à identifier se fasse individu par individu, de façon à influencer la reconstruction du consensus au minimum. Cette contrainte n'interdit pas la soumission par l'utilisateur de jeux de données importants mais elle ralentira le traitement (voir paragraphe D3.4.2). Ce traitement pourra cependant être accéléré (voir paragraphe D3.5)

#### ***D3.2. Limites de présentation du résultat***

L'espace des formes (celui des individus dans l'ACP) a été choisi par défaut pour présenter les résultats plutôt que l'espace discriminant. En effet, dans l'espace des formes calculé à partir des variables de conformation et du logarithme népérien de la taille centroïde les relations qui lient les individus sont linéaires et les distances sont cartésiennes (Mitteroecker et al., 2013). Au contraire, l'espace discriminant est déformé, ce qui rend les distances de Mahalanobis plus difficiles à interpréter visuellement. Ainsi, dans l'espace discriminant, l'ordination d'un individu ne correspond pas directement à sa morphologie crânienne, alors que cela est le cas dans l'espace des conformations. De ce fait, l'affichage des résultats dans l'espace discriminant pourrait induire l'utilisateur en erreur sur la proximité relative des morphologies qu'il observe.

Toutefois la représentation des trois premières composantes de l'espace des formes (92% de la variance totale) peut emmener l'utilisateur à penser que certaines espèces proches morphologiquement, comme celles du genre *Rattus*, ne sont pas correctement discriminées par l'analyse. Le module d'affichage du résultat devra donc laisser l'opportunité à l'utilisateur de visualiser le résultat sur des composantes principales de plus haut rang ainsi que dans l'espace discriminant.

#### ***D3.3. Limites liées à l'échantillonnage partiel des espèces de la zone d'étude***

L'échantillonnage, toujours en cours, devrait apporter de nouvelles espèces censées (ou non) être présentes dans la zone, ainsi que de nouvelles informations sur la variabilité géographique de leur morphologie. Ces nouvelles données devront être intégrées à l'outil au





fur et à mesure de leur collecte. Bien entendu, un groupe de référence nouveau ne peut être ajouté sans influencer en profondeur l'ensemble de l'espace discriminant (Macleod et al., 2007). Aussi l'ajout d'individus devra-t-il s'accompagner d'une nouvelle démarche d'optimisation des configurations et du choix du nombre de composantes de conformation à conserver pour réaliser l'analyse discriminante. De même, de nouveaux seuils d'assignation devront être définis.

#### *D3.4. Limites liées à la reproductibilité des données*

La reproductibilité des données d'un opérateur à l'autre est le talon d'Achille de la morphométrie géométrique, ce qui a constitué jusqu'alors un frein fondamental au développement d'outils automatisés d'identification.

##### *D3.4.1. Les solutions avancées dans la littérature*

Des outils web services d'identification basés sur la morphométrie géométrique ont déjà vu le jour. La solution détaillée dans ce travail au problème de l'effet « opérateurs » propose une solution originale à la comparaison de jeux de données morphométriques constitués indépendamment (voir Dujardin et al., 2010 ; Arnqvist et Martensson, 1998 ; Rasmussen et al., 2003). Les outils d'identification basés sur la morphométrie géométrique développés jusqu'alors ont traité l'effet « opérateurs » par d'autres méthodes que celle proposée ici. Le système expert ApiClass (<http://apiclass.mnhn.fr/>, visité le 23/10/2013) permet d'identifier les lignées et les sous-espèces d'ouvrières de l'abeille domestique *Apis mellifera* à partir d'images d'ailes proposées par les utilisateurs. ApiClass effectue des mesures et renvoie l'identification la plus probable des ouvrières. Dans ce cas, un système de détection automatique des points homologues indépendants de tout effet « opérateurs » a été développé. Ces systèmes sont généralement basés sur la détection automatique des contours à partir des valeurs de gris des pixels de la photo (Rohlf, 1990 ; Tofilski, 2007 ; Palaniswamy et al., 2007 ; Claude 2008) ainsi que de la détection de la nervation des ailes par le tracé des contours et la linéarisation de ceux-ci par squelettisation (Rohlf, 1990). Ils sont adaptés à des structures fortement contrastées, typiquement des ailes d'insectes hyalines photographiées sur fond blanc et possédant une nervation noire (voir <http://drawwing.org/>). Pour l'étude des rongeurs, la mise en place de tels modules de placement automatique de points homologues paraît n'être encore qu'une perspective lointaine car les points homologues situés à l'intérieur des structures étudiées sont placés sur des fonds faiblement contrastés et la moindre imperfection dans la préparation du crâne (ex .morceau de cartilage ou de tendon restant attaché à l'os après préparation) nécessite l'expertise de l'œil humain pour pouvoir être prise en compte. Pour les hémi-mandibules qui ne présentent pas (ou peu) de point interne, une approche de détection automatique et d'analyse de contour pourrait toutefois être envisagée à l'avenir. Un autre outil d'identification basé sur la morphométrie géométrique, le package CLIC (Dujardin et al., 2010) est destiné à identifier les insectes d'intérêt en entomologie médicale. Il propose d'éliminer l'effet utilisateur en fournissant à celui-ci les photographies d'ailes d'insectes de référence précisément identifiées et une suite de logiciels pour l'assister dans l'analyse. L'utilisateur doit ensuite digitaliser ce jeu de photographies et y inclure ses propres échantillons. Cette méthode est radicale puisque l'effet inter-utilisateur est éliminé mais une



telle démarche apparaît difficilement envisageable pour des jeux de données importants tel que le nôtre ou la digitalisation de 1320 photographies par l'utilisateur (pour chacune des vues crâniennes et le double pour les vues des mâchoires) rendrait l'outil indigeste, voire totalement repoussant.

#### *D3.4.2. La solution proposée ici et ses limites*

Pour être véritablement opérationnel, utilisable et utilisé par le plus grand nombre, un outil web service se doit d'être rapide, précis et convivial. Notre outil est raisonnablement rapide puisqu'une identification s'effectue en 30 secondes une fois les données fournies à l'outil (mais voir paragraphe D3.5). Il possède une bonne précision avec 98% d'identifications congruentes avec l'identification réalisée à partir d'un gène mitochondrial. La présentation du résultat, interactive et en trois dimensions est conviviale et attractive. Elle permet en outre des liens dynamiques avec la présentation illustrée des espèces concernées ainsi que vers les photos des crânes de référence, ce qui permet à l'utilisateur de porter un regard critique sur son identification. Enfin, cet outil est innovant dans le domaine de la morphométrie car il permet de distinguer les deux sexes ainsi que toutes les classes d'âge des espèces étudiées, ce qui est rarement le cas pour ce genre d'outils.

Le défaut principal de cet outil est le temps d'entraînement que l'utilisateur doit consacrer à son utilisation et la rigueur qu'il doit observer pour parvenir à une identification fiable. A l'inverse de la collection d'images CLIC (Dujardin et al., 2010), l'ensemble de l'échantillonnage de référence n'est pas à digitaliser à nouveau par l'utilisateur. Néanmoins, pour s'assurer de la reproductibilité des données entre utilisateur et Opr, au minimum soixante-huit photos par vue étudiée seront à digitaliser. A titre indicatif, ce travail prend entre une demi-journée et une journée suivant le niveau de l'utilisateur. Mais cette opération peut être à répéter plusieurs fois si la variance introduite par l'utilisateur est trop importante et excède les contraintes imposées par l'outil pour parvenir à une identification fiable. L'utilisateur est toutefois laissé libre d'utiliser l'outil sans s'entraîner préalablement. Dans ce cas, il n'y aura certainement pas d'assignation de l'individu à un groupe de référence, mais les valeurs des probabilités postérieures d'assignation à chacun des groupes pourront malgré tout le guider pour savoir à quel genre taxonomique appartient le crâne de l'animal qu'il a soumis à l'outil. Ce mode « dégradé » d'utilisation de l'outil pourrait convenir à l'usage d'utilisateurs occasionnels. L'investissement en temps d'entraînement que suppose l'utilisation de l'outil pour parvenir à une assignation ne devrait pas cependant être un frein à l'utilisation de celui-ci par un public plus averti et possédant des jeux de données conséquent à identifier (ex. responsables de collections de Musée d'Histoire Naturelle, scientifiques et experts).

#### *D3.5. Perspectives*

Dans la version actuelle de l'outil, les données utilisateurs sont utilisées pour calculer la forme moyenne afin de minimiser les résidus Procrustes. Cette procédure force l'outil à effectuer la superposition de 661 configurations de crâne à chaque interrogation. Elle ne permet pas un temps de traitement inférieur à 30 secondes par échantillon. Le temps de traitement deviendra vite important si de nombreuses identifications sont requises. L'outil doit



donc évoluer pour être en mesure d'accepter l'identification simultanée d'un grand nombre d'individus. Une première évolution consisterait par exemple à utiliser la conformation moyenne de l'échantillonnage de référence comme consensus. La conformation utilisateur serait alors directement alignée sur celle-ci. La proximité des conformations étudiées rendrait possible cette modification sans que cela n'influence trop fortement les résidus Procrustes pour l'individu soumis par l'utilisateur au point d'en modifier l'assignation à un groupe de référence. Dans ce cas, la superposition du jeu de données de référence serait effectuée une fois pour toute. Ainsi l'outil superposerait la conformation de l'individu de l'utilisateur sur un bloc de conformation pré-superposé et la procédure serait quasiment instantanée. L'outil se révélerait alors capable de traiter l'identification de plusieurs milliers d'individus les uns après les autres en des temps très brefs. Le troisième module qui présente les résultats devra, dans le cas d'identifications multiples, présenter l'ensemble des résultats à l'aide d'identifiants uniques pour chaque individu soumis par l'utilisateur. Les individus seront projetés dans l'espace des conformations et un tableau des probabilités postérieures d'assignation pour chaque individu sera fourni avec un code couleur qui visualisera l'assignation de l'individu à un groupe de référence.

La réduction du temps d'entraînement de l'utilisateur est un point central dans l'amélioration de l'outil car elle conditionne son utilisation et la possibilité d'améliorer l'identification en collectant l'information sur plusieurs vues et plusieurs structures. La digitalisation de la vue ventrale du crâne de l'animal permet une assignation au genre non ambiguë. Si le temps d'entraînement de l'utilisateur était réduit, cette assignation pourrait, dans un second temps, être améliorée par l'utilisation de la vue dorsale du crâne. Cette démarche serait pertinente pour affiner l'identification au sein du genre *Leopoldamys* et pour l'espèce *R. sakeratensis* (voir paragraphe C5.3). Une autre amélioration possible, si le temps de digitalisation était réduit, serait d'assembler l'information portée par les deux faces du crâne et éventuellement celle des héli-mandibules (en tenant compte de la redondance apportée par des points communs sur les différentes vues) dans une matrice unique de caractères afin d'augmenter le pouvoir discriminant de l'analyse.

L'utilisation de la symétrie bilatérale du crâne permet de réduire le nombre de points à digitaliser par l'utilisateur ce qui permet un gain de temps non négligeable notamment lors de son entraînement. Des tests préliminaires effectués sur le jeu de données de référence montrent une très légère diminution du pouvoir discriminant (moins 0,2%) lorsque l'on symétrise les conformations. Cette diminution du pouvoir discriminant est liée à la taille de l'échantillonnage de référence. Elle devrait disparaître pour un échantillonnage plus étendu du jeu de données de référence. Cependant la faible dégradation du taux d'assignation est à mettre en relation avec le gain de temps réalisé qui est proche de 50% lors du placement des points.

Nos résultats préliminaires montrent que les héli-mandibules sont généralement moins informatives que les vues crâniennes. Elles sont cependant, dans certains cas la seule information disponible pour identifier les espèces. Cela est par exemple le cas lors de l'étude des restes osseux dans les pelotes de réjection des rapaces nocturnes.



Un autre développement possible serait l'automatisation de la prise de données sur les hémimandibules. L'extraction automatisée de l'information sur cette structure est en effet plus aisée que sur le crâne car elle est essentiellement localisée sur son contour. Des méthodes pour extraire automatiquement cette information existent mais font appel à des techniques d'analyse d'image non abordées dans le présent travail. La prise de données sur le contour du crâne pourrait également être partiellement automatisée à l'aide de pseudo-points homologues. Le temps de digitalisation serait alors réduit au placement des points homologues internes de la structure et l'erreur 'opérateurs' fortement diminuée.

Le second module de l'outil, qui permet l'entraînement de l'utilisateur pourrait également être l'objet du développement d'un outil propre. L'encadrement de l'effet « opérateurs » a en effet toute sa pertinence dans une plate-forme de morphométrie ou un muséum d'histoire naturelle dans lesquels les opérateurs peuvent être nombreux et les jeux de données disparates. En effet si chaque utilisateur est entraîné pour que son jeu de données soit comparable aux jeux de données produits par les opérateurs l'ayant précédé et que les données sont stockées dans une base de données dans laquelle chaque point homologue est identifié par un identifiant unique, alors il devient envisageable d'assembler des jeux de données originaux en fusionnant les jeux de données provenant d'opérateurs multiples dans l'objectif de réaliser des méta-analyses. La production de données en morphométrie géométrique, comme dans d'autres disciplines, est un travail de longue haleine et ce type d'outils permettant d'assurer leur standardisation pourrait se révéler extrêmement précieux pour accroître la productivité des plateaux techniques de morphométrie.







... ET ON POURRA NOTER UN  
CHANGEMENT MORPHOLOGIQUE  
IMPORTANT CHEZ LES ESPÈCES  
NICHAUT DANS LES ÉCO-  
SYSTÈMES PRAIRIAUX SOUMIS  
À LA FAUCHE ANNUELLE ...



*Crex crex*



*Vanellus vanellus*



*Dendix pardi*



La systématique est la discipline qui a pour but de découvrir, d'identifier, de classier et d'interpréter la diversité biologique. Sous ce terme générique de systématique, est inclu ici la taxonomie qui décrit, nomme et classe les taxons et la phylogénie qui étudie l'histoire évolutive des taxons. Ces deux disciplines ont un objectif commun qui est de parvenir à une classification stabilisée. Des progrès techniques importants ont été réalisés au cours de la fin du vingtième siècle pour l'étude des organismes que ce soit sur le plan moléculaire (systématique moléculaire et génomique), ou pour la description du phénotype (morphométrie géométrique). La systématique intégrative propose de croiser ces approches à la lumière des avancées conceptuelles (**cladistique**), pour affiner la classification du vivant et comprendre les différences entre organismes dans le cadre conceptuel de la biologie évolutive. L'intérêt d'une démarche de systématique intégrative sur les modèles rongeurs a été démontrée pour différents groupes africains, notamment pour les genres *Praomys*, *Myomys* et *Mastomys* (Denys et al., 2003 ; Lecompte et al., 2003, 2005a, 2005b). Cette méthode a servi de base pour la mise en place d'outils d'identification performants pour le genre *Mastomys* (Lecompte et al., 2005b). Lors de nos programmes de recherche en Asie du Sud-Est, centrés sur l'étude des interactions rongeurs/parasites/milieus, nous avons souhaité suivre une démarche équivalente (Chaval et al., 2010) afin de clarifier la systématique des espèces. L'approche cytotonomique s'est avérée peu informative à ce sujet et pour la zone étudiée (Badenhorst et al., 2009), alors que les études moléculaires et morphométriques (voir introduction et Partie 1) font apparaître les limites de l'utilisation des mitotypes pour délimiter les taxons. Dans cette partie, le statut taxonomique des mitotypes *R. tanezumi* et *R3* sera discuté en croisant les résultats moléculaires, ceux de morphométrie traditionnelle appliquée aux mesures externes (Pagès et al., 2013) et ceux de morphométrie géométrique sur la vue ventrale du crâne car elle est celle qui possède le pouvoir discriminant le plus élevé pour ce jeu de données (voir partie 1).

Dans un second temps, la diversité biologique observée sera interprétée, à l'échelle de la tribu des Rattini, en testant l'existence de corrélation entre l'évolution du phénotype et l'environnement en tenant compte de la phylogénie. Les phénotypes observés sont en effet influencés par l'histoire évolutive commune entre les taxons actuels et leur adaptation à leur environnement (Caumul et Polly, 2005). En considérant que l'histoire évolutive des taxons étudiés est fidèle à la phylogénie moléculaire de la tribu des Rattini établie par les travaux de Pagès et al. (2010) et en utilisant celle-ci pour reconstruire l'histoire des morphologies crâniennes d'un jeu de données comportant les mêmes taxons, nous pourrions tester le rôle de l'environnement sur la disparité morphologique observée chez les Rattini en Asie du Sud-Est.

## **1. Une approche intégrative pour définir les limites d'espèces : le cas des mitotypes *R. tanezumi* et *R3* sensu Pagès et al. (2010, 2013)**

### *1.1. Présentation de l'article: « discordance cytonucléaire parmi les rats noirs du sud-est asiatique (complexe d'espèce *Rattus rattus*) », Pagès et al. (2013)*

L'étude phylogénétique de la tribu des Rattini en Asie du Sud-Est basée sur le séquençage de gènes (Pagès et al., 2010), a permis d'identifier deux clades mitochondriaux pour l'espèce nommée sur le terrain *Rattus tanezumi*. L'un de ces clades (R2 dans les travaux de Pagès et al., 2010) a été attribué à *R. tanezumi* et le second a été nommé *R3*. La distance génétique (K2P) au niveau de l'ADN mitochondrial séparant les clades *R3* et *R. tanezumi* est relativement importante (5,72% pour la Cytochrome Oxydase I et 6,24% pour le Cytochrome b) et *R3* est le groupe frère de *R. sakeratensis* nommé R4 dans l'étude de Pagès et al. (Pagès et al.,



2010 ; voir la figure 3 de l'annexe I). La cytotaxonomie n'a pas permis de discriminer ces deux clades mitochondriaux. En effet, le caryotype réalisé sur la lignée R3 (individu R4009, données non publiées) entre dans la variabilité importante des caryotypes obtenus pour *R. tanezumi* (Badenhors et al., 2009) et ne permet pas de certifier un isolement reproductif potentiel entre les deux lignées (Dobigny, comm.pers.).

Il est alors possible d'expliquer la situation observée à l'aide de deux hypothèses.

La première hypothèse est que les deux clades représentent deux espèces valides, c'est-à-dire des **espèces biologiques** au sens de Mayr (1942). L'histoire évolutive des lignées mitochondriales serait alors en accord avec l'histoire évolutive des espèces. Dans ce cas R3, mise en évidence sur des bases moléculaires, devra être décrite sur des bases morphologiques pour répondre à la définition phylogénétique de l'espèce (Lecointre et Le Guyader, 2009). Cette description permettra aux taxonomistes de rapprocher ce clade mitochondrial de l'un des quatre-vingt synonymes de l'espèce *R. tanezumi* actuellement recensés dans la littérature (Musser et Carleton, 2005). Cependant, si un tel rapprochement s'avère impossible, une nouvelle espèce devra être décrite pour ce clade. Ce scénario est plausible car l'histoire évolutive des lignées mitochondriales est souvent représentative de l'histoire évolutive récente des espèces (Hebert et Gregory, 2005). De plus, Musser et Carleton (2005a, p. 1491) soulignent: «*More than one species may be present in our concept of R. tanezumi, but careful taxonomic revision using morphological, chromosomal and molecular characters will have to test such an hypothesis*»;

La seconde hypothèse est que les deux clades mitochondriaux, comme le suggère l'identification de terrain et l'approche cytotaxonomique, ne représentent qu'une seule espèce. Dans ce cas, l'espèce est **paraphylétique** sur la base des marqueurs mitochondriaux et il faudra essayer de comprendre les mécanismes évolutifs susceptibles d'expliquer la divergence constatée entre l'histoire évolutive des lignées mitochondriales et celle de l'espèce *R. tanezumi*. Ce scénario est également plausible car de tels mécanismes (*i.e.* introgression ou tri imparfait des lignées ancestrales) ont été documentés dans d'autres groupes taxonomiques (Moritz and Cicero, 2005 ; Dasmahapatra and Mallet, 2006).

Pour tester ces hypothèses, les deux voies d'investigation indépendantes proposées par Pagès et al. (2010) ont été conduites.

La première approche utilisée est basée sur la morphométrie externe traditionnelle et l'étude des rapports de proportions ou log shape-ratio (présentée dans la partie 1). Cette étude morphologique est présentée ici de façon détaillée. Elle sera complétée par une étude originale de morphométrie géométrique étudiant la forme des crânes en vue ventrale pour ces deux clades mitochondriaux. Ainsi si le facteur mitotype influence significativement la morphologie externe ou crânienne alors l'hypothèse de la présence de deux espèces distinctes sera renforcée. Dans le cas contraire l'hypothèse d'une incongruence entre l'histoire des gènes mitochondriaux et celle des espèces sera renforcée.

La seconde approche utilisée dans l'étude de Pagès et al. (2013) est basée sur l'analyse d'un gène nucléaire codant (le gène **IRBP**) et de marqueurs nucléaires neutres, très variables, les microsatellites. Cette analyse ne sera pas développée ici mais son objectif est de quantifier les flux de gènes nucléaires entre deux populations. Si les deux populations sont éloignées génétiquement, cela signifie que les flux de gènes sont interrompus. Cet isolement reproductif est un marqueur de la spéciation biologique.



## 1.2. Inférence de l'influence du mitotype sur la morphologie externe au sein du genre *Rattus* par l'utilisation des log shape-ratio (travail inclus dans l'analyse de Pagès et al., 2013)

### Matériel et méthode

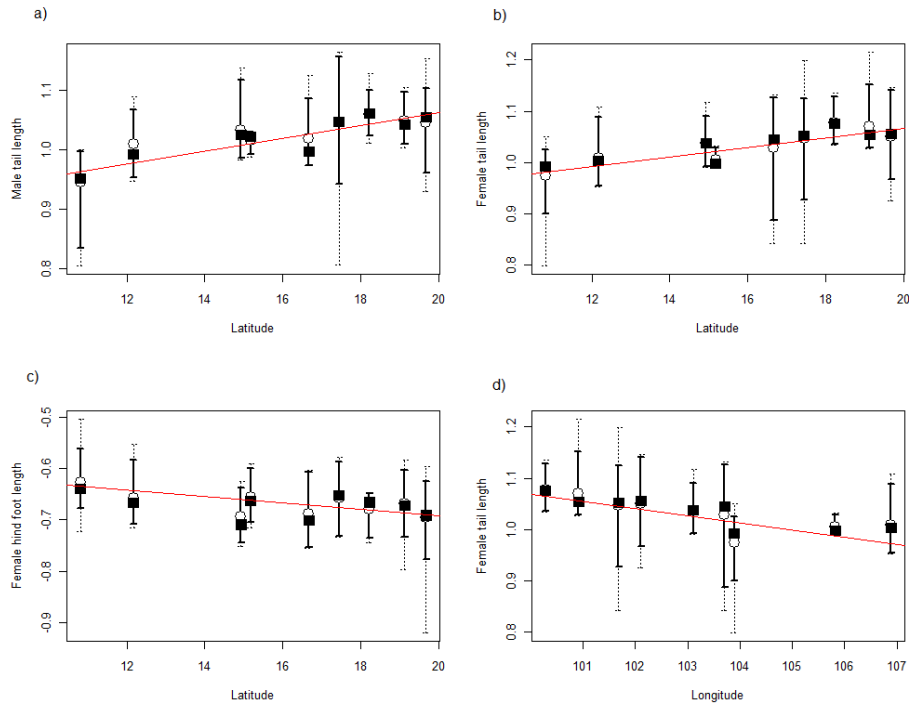
Les mensurations externes comprenant les longueurs : tête-corps, queue, patte arrière, oreille et tête (Voir Partie 1 paragraphe B3.1) d'un groupe d'individus identifiés à l'aide de leur séquence COI et de l'outil web service ([http://www.ceropath.org/barcoding\\_tool/rodentsea](http://www.ceropath.org/barcoding_tool/rodentsea)) et appartenant aux clades mitochondriaux *R. tanezumi* (n=164 dont 69 femelles et 95 mâles) et *R3* (n=164 dont 78 femelles et 86 mâles) ont été considérées pour réaliser cette analyse. Les animaux ont été sexés à partir des caractères sexuels primaires, les jeunes et les sub-adultes ont été exclus de l'étude (voir la Partie 1 paragraphe B1.2 pour la détermination de l'âge). La taille et la conformation ont été estimées avec la méthode des log shape-ratio (Mosimann, 1970 ; voir Partie 1 paragraphes C2.1, C3.2). Un estimateur de taille (la moyenne géométrique des variables) et de conformation (le rapport entre le logarithme de chaque variable et l'estimateur de taille) ont été calculés. Pour éviter les biais statistiques introduits par le calcul des log shape-ratio (voir Partie 1 paragraphe C3.2), une analyse en composantes principales a été réalisée et seules les composantes principales non nulles ont été retenues pour les analyses statistiques. Un modèle linéaire expliquant la taille par le mitotype (2 modalités), le sexe (2 modalités) et la localité (9 modalités) de provenance des animaux a été construit. L'échantillonnage n'étant pas équilibré entre les deux sexes au sein des espèces, les sommes des carrés de type II ont été utilisées dans cette analyse de variance pour éviter les problèmes artéfactuels de corrélation entre facteurs. Un test post-hoc de permutation de Tukey a été appliqué pour connaître l'origine de la variation de taille en fonction des facteurs significatifs. De même, pour la conformation, un modèle linéaire multivarié a été construit sur les quatre premières composantes de l'espace de conformation. L'effet du sexe, du mitotype et de la localité d'échantillonnage des individus sur la conformation de leur corps a été étudié à l'aide d'une analyse de variance multivariée. Comme ces comparaisons étaient multiples, une **correction de Bonferroni** a été appliquée. La variation géographique des log shape-ratio et de la taille a ensuite été explorée pour chaque sexe en utilisant des boîtes à moustaches avec la longitude ou la latitude en abscisse. Les relations entre conformation et taille avec la longitude et la latitude ont été estimées par une régression linéaire entre les conformations ou taille moyenne et les localités.

### Résultats

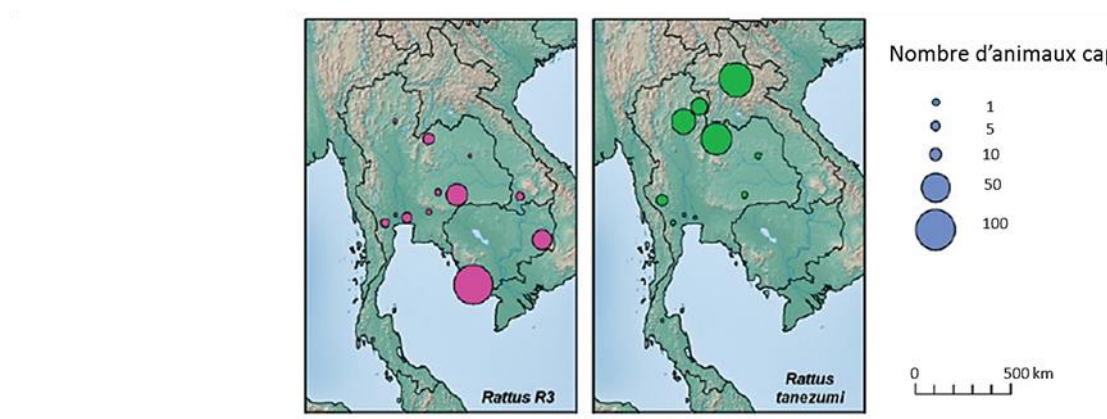
La taille des animaux varie de façon significative en fonction du sexe ( $F(1,303)=31,25$ ;  $p < 0,001$ ) et des localités ( $F(8, 303)= 2,56$ ,  $p= 0,010$ ) mais pas des mitotypes ( $F(1, 303)= 0,17$ ;  $p= 0,68$ ). De plus aucune interaction entre ces facteurs n'a pu être mise en évidence ce qui montre que la taille des deux sexes varie de la même façon quel que soit la localité ou le mitotype considéré. Le test post-hoc de Tukey montre que les mâles sont significativement plus gros que les femelles et que les individus de Kalasin (Thaïlande) sont significativement plus gros que ceux de Mondolkiri (Cambodge) et Luang Prabang (Laos). La longitude et la latitude n'expliquent pas significativement la taille des mâles ( $F(1,7) < 1,58$ ,  $p > 0,24$ ) ou des femelles ( $F(1,7) < 0,75$ ,  $p > 0,41$ ).

La conformation est significativement expliquée par le sexe (Pillai(1, 303)= 0,15,  $p < 0,001$ ) et par les localités (Pillai(8, 303)=0,35,  $p < 0,001$ ). Cependant l'appartenance à l'un ou l'autre des mitotypes n'explique pas significativement la variance de la conformation du

## Effet de la latitude et de la longitude sur la variation de traits morphologiques en fonction du sexe.



**Figure 1:** Evolution significative de la taille relative de certains traits dans la conformation corporelle des animaux (*R.tanezumi* et *R3*). Allongement de la longueur relative de la queue avec la latitude pour les mâles (a) et les femelles (b). Diminution de la longueur du pied podtérieur des femelles avec l'augmentation de la latitude (c) et diminution de la longueur de la queue des femelles avec l'augmentation de la longitude (d).



Extrait de: Pagès M, Bazin E, Galan M, Chaval Y, Claude J, et al. (2013) Cytonuclear discordance among Southeast Asian black rats (*Rattus rattus* complex). *Mol Ecol* 22: 1019-1034.

**Figure 2:** Distribution géographique des monotypes échantillonnés dans l'article de Pagès et collaborateurs (2013).



corps (Pillai(1, 303)= 0,021,  $p= 0,164$ ). De plus aucune interaction entre ces facteurs n'a été trouvée significative. Le corps a une forme plus allongée chez les mâles que chez les femelles ( $F(1,326)=31,2$ ,  $p<0,001$ ) mais les mâles possèdent des oreilles comparativement plus courtes ( $F(1,326)= 18,04$ ,  $p<0,001$ ). La comparaison deux à deux pour les mâles révèle que les animaux de Krong Preah Sihanouk (Cambodge) sont significativement différents ( $p< 0,0014$ ) de ceux des autres localités, excepté de ceux de Champasak (Laos). On retrouve le même patron chez les femelles qui ont des conformations significativement différentes ( $p< 0,0014$ ) à Krong Preah Sihanouk par rapport aux autres localités, à l'exception de Champasak et Mondolkiri. Les individus de Krong Preah Sihanouk ont des longueurs relatives de corps et d'oreille supérieures aux autres individus mais des longueurs de queue relativement plus courtes.

Les relations entre la conformation du corps des espèces et la géographie ont été étudiées en comparant les variations morphologiques en fonction de la longitude et la latitude. Pour les mâles, aucune relation significative n'a été observée entre les log shape-ratio et la longitude ( $\text{abs}(t(7)) < 1,88$ ;  $p>0,10$ ). La longueur relative de la queue des mâles augmente avec la latitude (pente= 0,009,  $t(7)=4,56$ ;  $p=0,003$ ) mais aucun autre rapport de forme n'est influencé par cette variable ( $\text{abs}(t(7)) < 2,03$ ;  $p>0,082$ ) (figure 1a). Pour les femelles le rapport de conformation de la longueur relative de la queue croit significativement (pente=0,010,  $t(7)=5,03$ ,  $p=0,002$ ) et la longueur relative du pied décroît significativement (pente= -0,005,  $t(7)= -2,42$ ,  $p= 0,046$ ) avec la latitude mais aucun autre rapport de conformation n'est significativement corrélé avec la latitude ( $\text{abs}(t(7)) < 1,39$ ,  $p> 0,21$ ) (figure 1b,c). D'autre part, pour les femelles, la longueur relative de la queue dans la proportion corporelle diminue significativement avec la longitude (pente= -0,012;  $t(7)= -3,50$ ;  $p=0,010$ ) (figure 1d), aucun rapport de forme n'est corrélé pour les femelles avec la longitude ( $\text{abs}(t(7)) < 2,19$ ,  $p> 0,064$ ).

### Discussion

Les individus ayant des mitotypes différents ont des tailles et des proportions corporelles similaires une fois les effets du sexe et de la localité d'échantillonnage enlevés. Le sexe a été montré comme influençant significativement la taille et la conformation des animaux. Les mâles sont plus gros et ont une morphologie différente des femelles avec une forme de corps plus allongée et une oreille proportionnellement plus petite. Les animaux diffèrent en taille et conformation entre les localités et particulièrement entre celle de Krong Preah Sihanouk, la localité la plus au sud du Cambodge et les autres localités. Une variation clinale du nord au sud des conformations a pu être mise en évidence par la représentation en boîtes à moustaches en fonction de la longitude et de la latitude qui montre que les animaux du sud possèdent une queue relativement plus courte que les autres animaux. Cette tendance morphologique est à mettre en relation avec la variation géographique de la fréquence relative de présence des deux mitotypes (figure 2). Le mitotype *R. tanezumi* est majoritairement capturé au nord-est de la zone d'échantillonnage et les captures du mitotype *R3* décroissent du sud-est où elles sont maximales vers le nord-est. L'attendu est que si ces mitotypes possédaient une morphologie propre, la variation géographique des morphologies observées devrait être marquée d'un plateau lors du passage dans la zone d'hybridation. Or l'évolution morphologique est ici continue, ce qui suggère qu'il n'existe pas de zone hybride étroite entre les deux mitotypes dans la zone échantillonnée ou que l'ensemble de l'échantillonnage est réalisé de manière quasi tangentielle à la zone hybride, auquel cas un tel cline pourrait être également observé. Un échantillonnage géographique plus étendu des deux mitotypes au Vietnam, au Myanmar et dans la Province du Yunnan (Chine) permettrait de préciser l'existence ou non d'une zone hybride.



## Conclusion

L'analyse de la taille et des proportions corporelles des animaux a permis de mettre en évidence un effet du sexe et de la localité d'échantillonnage sur ces deux variables. Cependant elle n'a pas permis de mettre en évidence des différences de taille et de proportions significatives pour les deux mitotypes observés. Cependant, certains traits varient suivant un cline du nord au sud, mais l'analyse n'a pas permis de mettre en évidence une zone hybride claire entre les deux mitotypes. Lorsque l'on s'intéresse à leur morphologie externe, les deux mitotypes représentent donc une seule espèce ou bien ce sont des espèces cryptiques pour ce marqueur.

### *1.3. Variation de la conformation crânienne en vue ventrale entre les deux clades mitochondriaux R. tanezumi et R3 (analyse non publiée)*

La morphométrie géométrique estimant les variations de forme du crâne en vue ventrale a un pouvoir discriminant très supérieur à l'analyse en log shape-ratio sur les mesures externes (voir partie 1 chapitre C5). L'analyse de la morphologie externe a donc été complétée par l'analyse de la morphologie crânienne en vue ventrale en utilisant cette méthode.

#### *Matériel et méthode*

Un groupe d'individus identifiés moléculairement à l'aide de l'outil web service ([http://www.ceropath.org/barcoding\\_tool/rodentsea](http://www.ceropath.org/barcoding_tool/rodentsea)) et appartenant aux mitotypes *R. tanezumi* (n= 38 individus parmi lesquels 17 femelles et 21 mâles) et *R3* (n= 34 individus dont 20 femelles et 14 mâles) a été considéré pour réaliser cette analyse (voir tableau 1 Partie 1). Ces individus proviennent du Cambodge, du Laos et de la Thaïlande pour le mitotype *R3* et du Laos et de la Thaïlande pour *R. tanezumi*. La conformation de la vue ventrale du crâne des animaux a été estimée à partir de photographies des crânes (voir Partie 1 pour l'acquisition et le traitement des données) sur lesquelles a été digitalisée une configuration comprenant 40 points homologues (voir annexe III). Chaque crâne a été photographié et les points homologues ont été digitalisés deux fois pour pouvoir estimer l'erreur de mesure. Les conformations des individus et de leurs répliqués ont alors été superposées en utilisant la superposition généralisée Procrustes partielle (voir Partie 1 paragraphe A1.4.3). La taille centroïde a été utilisée comme estimateur synthétique de la taille des individus (voir Partie 1 paragraphe A1.4.3). Pour rendre compte de la conformation des crânes, les coordonnées Procrustes des individus ont été synthétisées à l'aide d'une analyse en composantes principales sur les résidus Procrustes. Afin d'éviter toute redondance, quatre variables de conformation ont été retirées de l'analyse (voir Partie 1 paragraphe A1.4.3). Le pourcentage d'erreur de mesure a été calculé pour la taille et la conformation des crânes, en suivant le protocole de Yezerinac (Yezerinac et al., 1992 ; voir partie 1 paragraphe A2.2.3). Par la suite, les configurations moyennes par individu ont été utilisées. Une analyse de variance de la taille a été réalisée en utilisant comme facteurs explicatifs l'appartenance à l'un des deux mitotypes, l'âge et le sexe des animaux. L'échantillonnage étant réduit à celui utilisé pour la mise en place de l'outil d'identification pour les deux mitotypes (soit 72 individus), l'effet de la localité qui comporte un nombre important de modalités n'a pas été étudié ici. La variance de la conformation des crânes sous l'effet des mêmes facteurs (appartenance à l'un des deux mitotypes, âge et sexe des animaux) a été également inférée par une analyse de variance multivariée réalisée sur les 60 premières composantes principales de l'ACP de conformation.

**Tableau 1:** analyse de variance de la taille des crânes en vue palatine des mitotypes *Rattus tanezumi* et R3. Les variables explicatives étudiées sont l'âge, le sexe, le mitotype des animaux et les interactions entre ces facteurs.

<i>facteur</i>	<i>Df</i>	<i>Somme des carrés</i>	<i>Moyenne Des carrés</i>	<i>Valeur de F</i>	<i>Pr(&gt;F)</i>
âge	2	38,09	19,046	86,556	<2x10-16***
sexe	1	1,91	1,910	8,674	0,00459**
mitotype	1	0,40	0,398	1,808	0,18381
sexe:âge	2	0,08	0,041	0,187	0,82984
sexe:mitotype	1	0,08	0,084	0,383	0,53844
âge:mitotype	2	0,38	0,189	0,859	0,45878
sexe:âge:mitotype	2	0,91	0,453	2,057	0,13674
Residuals	60	13,21	0,220		

**Tableau 2:** analyse de variance multivariée sur les composantes principales de conformation des crânes en vue palatine des mitotypes *Rattus tanezumi* et R3. Les variables explicatives étudiées sont l'âge, le sexe, le mitotype des animaux et les interactions entre ces facteurs.

<i>facteur</i>	<i>Df (facteur)</i>	<i>Pillai</i>	<i>valeur de F</i>	<i>Df (num)</i>	<i>Df (individus)</i>	<i>Pr(&gt;F)</i>
âge	2	1,97632	2,7818	120	4	0,1630
sexe	1	0,99678	5,1630	60	1	0,3386
mitotype	1	0,99876	13,4034	60	1	0,2143
sexe:âge	2	1,90076	0,6385	120	4	0,8123
sexe:mitotype	1	0,99242	2,1824	60	1	0,4989
âge:mitotype	2	1,96078	1,6664	120	4	0,3367
sexe:âge:mitotype	2	1,95712	1,5214	120	4	0,3771
Residuals	70					

Pour les mêmes raisons que pour la taille, l'effet localité n'a pas été pris en compte et une seconde analyse sans tenir compte du sexe des animaux a aussi été réalisée.

### *Résultats*

L'erreur de taille atteint 0,16% de la variance totale de taille et l'erreur de conformation 21,7% de la variance de ce trait. L'erreur est dans les deux cas très inférieure à la variance inter-individuelle (99,84% de la variance pour la taille 78,3% pour la variance de conformation) pour les deux traits.

La taille est expliquée logiquement par l'âge des animaux ( $F(2,60) = 86,556$ ,  $P < 2 \cdot 10^{-16}$ ) leur sexe ( $F(1,60) = 8,674$ ,  $P = 0,0046$ ) mais pas par l'appartenance à un mitotype ( $F(1,60) = 1,808$ ,  $P = 0,18$ ). D'autre part, aucune interaction entre l'âge et le sexe ou/et le mitotype et entre le sexe et le mitotype ne s'est avérée significative (tableau 1). La taille évolue donc de la même manière chez les deux mitotypes en fonction de l'âge et du genre des individus. Un test post-hoc de Tukey montre que chaque classe d'âge a une taille significativement différente des autres et que les mâles sont plus gros que les femelles.

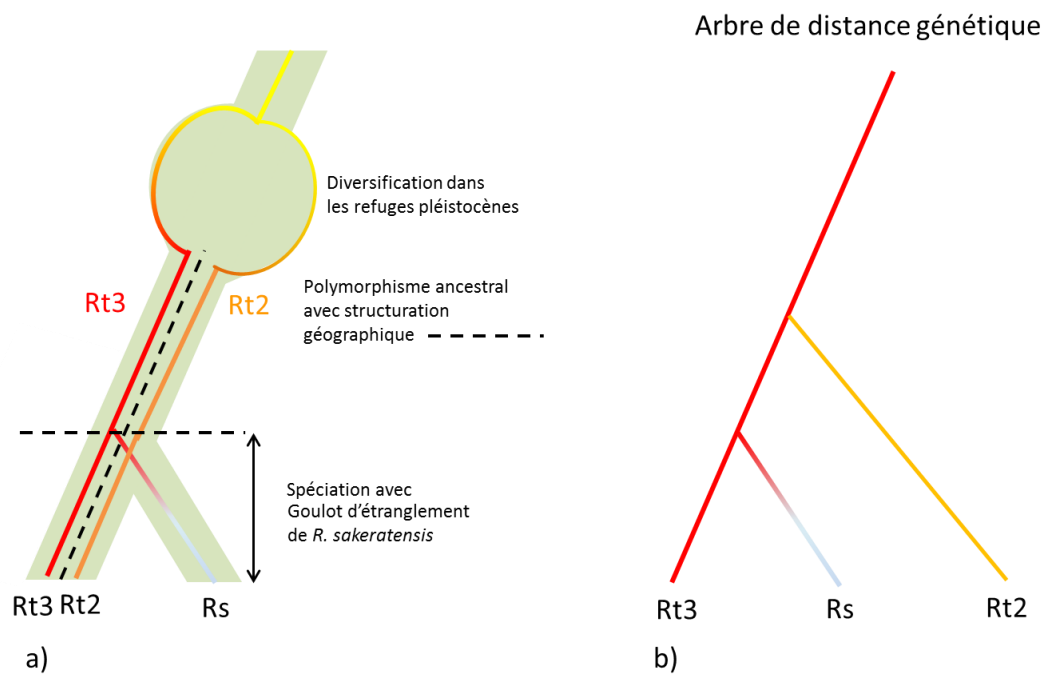
La conformation du crâne des animaux n'est pas expliquée par leur sexe (Pillai (1, 60) = 0,10,  $p=0,34$ ) ni par leur âge (Pillai (2, 60) = 1,98,  $p=0,16$ ) ni par leur appartenance à un mitotype (Pillai (1, 60) = 0,10,  $p=0,21$ ) ni par les interactions entre l'âge et le sexe ou/et le mitotype et entre le sexe et le mitotype (tableau 2).

### *Discussion*

L'erreur de mesure est beaucoup plus élevée que lorsque le jeu de données de référence de l'outil d'identification morphométrique est considéré dans sa globalité, que ce soit pour la taille (0,16% contre 0,04%) ou pour la conformation des crânes (21,7 contre 11,6%). Cet accroissement de l'erreur de mesure, par rapport à l'erreur présente dans le jeu de données de référence, provient en partie du fait que la variance entre les mitotypes étudiés ici est beaucoup plus faible que la variance inter-spécifique présente dans le jeu de données de référence.

La taille est expliquée par l'âge des animaux mais cela n'est pas le cas pour la conformation de leur crâne. Il n'y a donc pas de changement de morphologie en fonction de l'âge des individus. Cette isométrie de croissance est peu fréquente chez les rongeurs et lorsque l'on considère le jeu de données dans son ensemble (i.e. celui utilisé pour la mise en place de l'outil d'identification), l'évolution de la conformation est affectée par l'âge des animaux. Une hypothèse pour expliquer la différence observée ici est un manque de **puissance** de l'analyse de variance qui pourrait favoriser un **erreur de type II**. Les animaux sont divisés en deux sexes, deux mitotypes et trois classes d'âge et les effectifs par classe d'âge sont relativement faibles (28 individus adultes, 26 sub-adultes et 17 jeunes). Le sexe des rongeurs n'influence pas la conformation des crânes ce qui confirme les résultats trouvés pour le jeu de données de référence de l'outil d'identification (Partie 1, paragraphe C4.2). Lorsque le facteur sexe est retiré des facteurs explicatifs et que seuls les mitotypes et l'âge sont conservés alors ce dernier a un effet hautement significatif (Pillai (1,66) = 1,96  $p = 7,4 \cdot 10^{-5}$ ). Enfin le mitotype n'influence ni la taille, ni la conformation des animaux. Les deux mitotypes ne représentent donc qu'un seul morphotype. Cependant, il convient de relativiser les résultats de cette étude préliminaire qui devra être complétée en augmentant la taille de l'échantillonnage pour augmenter la puissance du test et pouvoir également inférer les interactions entre le phénotype et les localités d'échantillonnages comme cela était le cas pour la morphologie externe.

### Représentation du phénomène de tri imparfait des lignées ancestrales.



**Figure 3:** L'histoire évolutive des gènes ne reflète pas toujours l'histoire des espèces. Ici deux allèles d'un même gène (a) restent présents au sein d'une espèce et la font apparaître paraphylétique (b).

## Conclusion

Les résultats de morphométrie externe et les résultats préliminaires de morphométrie géométrique vont dans le sens des résultats moléculaires obtenus sur les marqueurs neutres nucléaires (Pagès et al., 2013) et plaident pour la conspécificité des deux mitotypes. Cette mise en synonymie avait été anticipée pour le développement de l'outil d'identification morphométrique (voir Partie 1, chapitre 5).

### *1.4. Scénarios évolutifs permettant la réconciliation des données mitochondriales et nucléaires*

La seconde voie d'investigation employée dans l'article de Pagès et al. (2013) est moléculaire et a permis, à partir de marqueurs neutres hypervariables, les microsatellites et d'un marqueur nucléaire codant (IRBP), d'identifier une seule lignée nucléaire pour les mitotypes *R. tanezumi* (ou R2) et R3. Ces résultats sont en accord avec ceux suggérés par la morphologie corporelle (paragraphe 1.2) et crânienne en face ventrale (paragraphe 1.3) pour les deux mitotypes. Les processus d'évolution permettant d'expliquer l'absence de congruence entre les signaux nucléaires codant et non codant d'une part et le signal mitochondrial d'autre part sont présentés ici.

#### *1.4.1. Tri imparfait des lignées ancestrales (ILS : Incomplete Lineage Sorting)*

### *Théorie*

Le premier scénario évolutif susceptible de réconcilier les données moléculaires nucléaires, morphométriques et mitochondriales est le tri imparfait des lignées ancestrales (figure 3). Dans l'hypothèse d'une population ancestrale importante, plusieurs haplotypes (copies différentes) d'un même gène mitochondrial peuvent se maintenir et être fixés car la dérive génétique est faible et les temps de **coalescence** sont donc plus importants. Ce polymorphisme ancestral est rare mais peut être observé chez des espèces dont certaines populations ont été isolées les unes des autres suffisamment longtemps pour permettre une diversification des gènes mitochondriaux sans qu'un isolement reproductif, signature d'un événement de spéciation, n'ait eu le temps de se mettre en place (figure 3a). Lors d'un événement de spéciation, le polymorphisme peut perdurer dans l'une des espèces filles alors que dans la seconde espèce, l'un des haplotypes est éliminé sous l'action de la dérive (figure 3a). Ce cas de figure est observé lorsque la seconde espèce se diversifie à partir de l'isolement d'une population aux effectifs restreints (effet fondateur). Une fois le phénomène de spéciation intervenu, l'haplotype mitochondrial commun aux deux espèces suivra sa propre évolution dans chacune d'elle. Cependant, si le temps de diversification de cet haplotype est court (figure 3a) alors les espèces distinctes qui le possèdent apparaîtront plus proches au regard de ce marqueur que les deux haplotypes mitochondriaux appartenant à la même espèce (figure 3b).

### *Application à notre modèle*

Dans notre étude, nous pouvons faire l'hypothèse que la grande population ancestrale à l'origine des lignées mitochondriales R2, R3, Rs (figure 3a) était forestière car ce biome, à la transition Plio- Pléistocène (2,6 Ma), couvrait alors la majeure partie de l'Asie du Sud-Est (Woodruff, 2010, et voir Annexe I pour l'histoire biogéographique de la zone). Les cycles successifs d'expansion - contraction de la forêt durant les soubresauts climatiques du





Pléistocène auraient confiné les individus dans des refuges forestiers distincts (figure 3a). L'existence de tels refuges est attestée chez d'autres animaux forestiers tels que les éléphants d'Asie (Vidya et al., 2009). L'haplotype ancestral aurait alors évolué indépendamment sous différentes formes dans chaque refuge sans toutefois qu'il y ait spéciation. Par la suite, la remise en contact des populations aurait donné une espèce ancestrale polymorphe structurée géographiquement avec des haplotypes R2 plus fréquents au nord et Rt3 au sud (par effet de diffusion à la sortie des refuges Pléistocènes). Un petit groupe d'individus se serait alors isolé de ce pool commun, s'adaptant à l'exploitation de milieux ouverts humides. Cette spéciation péripatrique (Pagès et al., 2013) serait ainsi à l'origine de l'apparition de *R. sakeratensis* il y a cinq cents mille ans (Aplin et al., 2011). Le goulot d'étranglement étant fort lors de l'évènement de spéciation, la dérive agit fortement et l'ADN mitochondrial est très vite trié pour cette nouvelle espèce: un des haplotypes mitochondriaux est fixé (en l'occurrence R3), l'autre éliminé (R2). Les pools génétiques mitochondriaux et nucléaires de *R. sakeratensis* évoluent distinctement depuis cet évènement de spéciation alors que les deux mitotypes R2 et R3 se maintiennent dans la large population dérivée de la population ancestrale. Ainsi sur la base de l'ADN mitochondrial, *R. sakeratensis* apparaît plus proche de R3 que de R2 (figure 3b) et se distingue de ces deux mitotypes sur la base des marqueurs nucléaires ou des données morphométriques et morphologiques (voir Partie 1). Puisque les mitotypes R2 et R3 sont maintenus au sein de la même population, les individus R2 et R3 possèdent le même génome nucléaire et ne peuvent être distingués quels que soient les marqueurs (IRBP, microsatellites) ou données morphométriques et morphologiques. Ils appartiennent donc à la même espèce qui apparaît paraphylétique lorsque son histoire évolutive est retracée à partir de gènes mitochondriaux (COI et Cytb).

#### 1.4.2. L'introgression mitochondriale

##### *Théorie*

Lors de l'hybridation entre deux espèces proches, le génome nucléaire des hybrides est fourni pour moitié par chacune des espèces. Les gènes mitochondriaux pour leur part ne sont transmis à la descendance que par les femelles. Si l'hybridation est directionnelle c'est-à-dire que l'accouplement se fait préférentiellement entre les mâles d'une espèce (A) et les femelles de l'autre espèce (B), alors un phénomène d'introgression mitochondriale peut être observé. Cette hybridation directionnelle peut être le fruit de plusieurs phénomènes agissant seuls ou de façon concomitante. Elle peut, par exemple, être induite par une différence dans la disponibilité des partenaires sexuels. Si l'espèce B a une taille de population plus importante que A, alors les mâles de l'espèce A rencontreront plus aisément des femelles de l'espèce B que des femelles de leur propre espèce. Les hybrides engendrés posséderont majoritairement la mitochondrie de l'espèce B. D'autres phénomènes peuvent expliquer ou venir renforcer le sens de l'hybridation comme un biais de dispersion efficace entre les deux sexes d'une espèce, phénomène couramment observé chez les rongeurs (Nunes, 2007) ou encore la préférence sexuelle des femelles de l'espèce B pour les mâles de l'espèce A. Si les hybrides peuvent garder la faculté de se reproduire avec l'espèce mère A. Si tel est le cas, alors la contribution nucléaire de l'espèce B va progressivement disparaître du génome nucléaire des hybrides. On observe alors des individus ayant le matériel nucléaire de l'espèce A et le type mitochondrial de l'espèce B.

**Tableau 3:** échantillons morphologiques utilisés pour étudier l'évolution des conformations en vue ventrale de la tribu des Rattini en Asie du Sud-Est. Tous les individus proviennent d'Asie du Sud-Est à l'exception de ceux appartenant aux mitotype *R. rattus* qui proviennent du Niger et ceux appartenant au mitotype *M. minutus* qui proviennent de France.

Espèces	nombre d'individus
<i>Bandicota indica</i> (Bi)	53
<i>Bandicota savilei</i> (Bs)	42
<i>Berylmys berdmorei</i> (Bbe)	41
<i>Berylmys bowersi</i> (Bbo)	38
<i>Leopoldamys edwardsi</i> (Le)	31
<i>Leopoldamys neilli</i> (Ln)	41
<i>Leopoldamys sabanus</i> (Ls)	4
<i>Maxomys surifer</i> (Ms)	39
<i>Niviventer huang</i> (Nh)	44
<i>Niviventer fulvescens</i> (Nf)	1
<i>Rattus andamanensis</i> (Ran)	5
<i>Rattus argentiventer</i> (Rar)	42
<i>Rattus exulans</i> (Re)	35
<i>Rattus nitidus</i> (Rni)	14
<i>Rattus norvegicus</i> (Rn)	25
<i>Rattus R3</i> (Rr3)	34
<i>Rattus tanezumi</i> (Rt)	38
<i>Rattus rattus</i> (Rr)	32
<i>Rattus tiomanicus</i> (Rti)	61
<i>Rattus sakeratensis</i> (Rs)	41
<i>Micromys minutus</i> (Mn)	3

## *Application à notre modèle*

Deux scénarios alternatifs sont possibles. Soit une espèce porteuse du génome mitochondrial R3 provenant du sud-est de la zone étudiée a introgressé partiellement l'espèce *R. tanezumi* installée sur le continent asiatique, soit l'ADN mitochondrial R2 de l'espèce *R. tanezumi* provenant du nord-est a introgressé l'espèce porteuse du mitotype R3 installée dans la péninsule indochinoise. Le premier cas (le second est symétrique du premier) impliquerait que les femelles de l'espèce porteuse du mitotype R3 se sont reproduites majoritairement avec les mâles de *R. tanezumi*. L'espèce porteuse du mitotype R3 n'a pas été échantillonnée et les hybrides en se reproduisant avec les *R. tanezumi* toujours présents ont éliminé la partie du génome nucléaire provenant de l'espèce porteuse du mitotype R3. Ce phénomène est relativement courant. Ainsi, les travaux de Lack et al. (2012) ont montré que le patron de distribution spatiale qu'ils observent entre lignées hybrides et espèces mères sur le continent américain et la répétabilité de ce patron pour différents loci étudiés était un signal fort en faveur d'une hybridation entre *R. rattus* et *R. tanezumi*. Dans cette étude, l'espèce *R. tanezumi* englobe les mitotypes R2 et R3. L'hybridation observée par ces auteurs entre *R. rattus* et *R. tanezumi* laisse supposer que l'hybridation des deux lignées mitochondriales R2 et R3 est probable. Lack et al. (2012) ont étudié plusieurs régions éloignées du génome et retrouvent un signal équivalent. Or, sous l'hypothèse de l'ILS, qui est un phénomène stochastique, les patrons observés pour des loci éloignés dans le génome ne devraient pas être les mêmes.

Les deux hypothèses restent cependant valides car l'étude de Pagès et al. (2013) n'a pas permis de trancher en faveur de l'une ou de l'autre.

## **2. Facteurs influençant la variabilité morphologique**

L'un des objectifs des programmes de recherche que nous poursuivons en Asie du Sud-Est est de comprendre les mécanismes qui permettent la structuration des communautés de rongeurs pour comprendre l'influence du changement global sur celles-ci. Les travaux présentés dans les parties précédentes ont permis de faciliter la caractérisation des communautés de rongeurs en fournissant une trame systématique modernisée et les outils d'identification nécessaires à cette tâche. Cependant la diversité taxonomique décrite dans ces parties n'a été interprétée ni d'un point de vue fonctionnel, ni à la lumière des relations phylogénétiques des espèces. Dans cette partie, le rôle du milieu de vie et des relations phylogénétiques sur la morphologie des crânes des animaux sont évalués à l'échelle inter-spécifique.

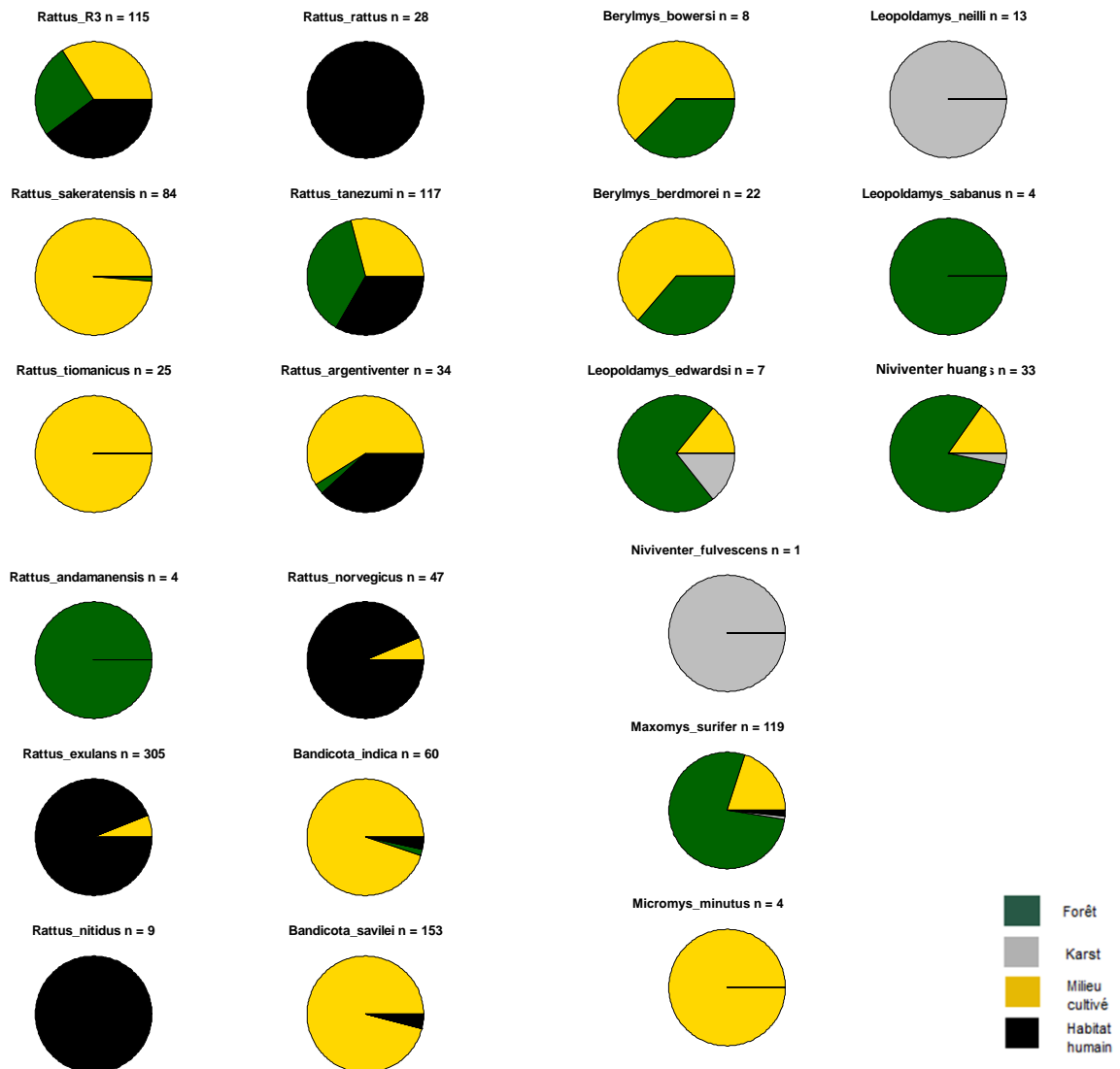
### *2.1. Données utilisées dans cette étude*

La morphologie d'une espèce donnée résulte à la fois de ses adaptations à l'environnement et aussi de l'histoire évolutive qu'elle a partagée avec d'autres espèces. Afin d'évaluer la relation entre ces deux facteurs et la morphologie des espèces, un jeu de données inter-spécifiques a été constitué.

#### *2.1.1. Données morphométriques*

Le jeu de données de morphométrie décrivant le phénotype des espèces repose sur la face ventrale du crâne qui est celui utilisé pour réaliser l'outil d'identification. Il a cependant été augmenté de certaines espèces (tableau 3) qui n'avaient pu être prises en compte lors de la réalisation de celui-ci (voir partie 1, tableau 1) de façon à couvrir les taxons étudiés dans les travaux de Pagès et al. (2010). Parmi ces taxons figurent des espèces échantillonnées dans des

# Habitats des espèces



**Figure 4:** Milieux fréquentés par chaque espèce d'après 1192 captures réalisées durant le programme CERoPath et dont la localisation est d'une précision inférieure ou égale à 100 mètres. Dans ce graphique, rizières et cultures sèches sont groupés au sein de la classe milieux cultivés.

zones géographiques différentes de la zone couverte par l'outil d'identification morphométrique : *Rattus rattus* (n= 32 individus), *R. tiomanicus* (n= 63) qui proviennent respectivement du Niger et de Sumatra (Indonésie). S'ajoutent également des espèces dont les effectifs étaient trop faibles pour appréhender la variation intra-spécifique dans le cadre de la mise en place de l'outil d'identification : *R. andamanensis* (n= 4), *Leopoldamys sabanus* (n= 4), *Niviventer fulvescens* (n= 1). Enfin des individus de *Micromys minutus* (n= 3) provenant de France ont également été ajoutés car cette espèce représente le taxon basal de la tribu des Rattini d'après les travaux de Lecompte et al. (2008) et est le groupe externe des travaux de Pagès et al. (2010).

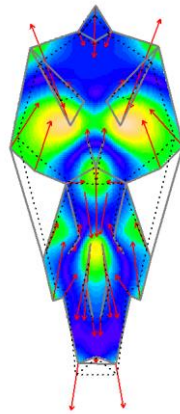
### 2.1.2. Données moléculaires

Dans la publication de Pagès et al. (2010), les relations phylogénétiques sont décrites pour trois divisions de la tribu des Rattini (*Dacnomys*, *Maxomys* et *Rattus*). La couverture taxonomique de cet échantillonnage moléculaire a été reprise à l'exception du clade N4 de cette publication (voir tableau 1 en annexe 1) pour lequel aucun crâne n'était disponible et qui a donc été ôté du jeu de données moléculaires. La nomenclature des clades est celle présentée en annexe 1 (tableau 1).

### 2.1.3. La variable environnementale : le milieu de vie des espèces

Les données écologiques de 2124 individus appartenant à l'ensemble des taxons récoltés durant le programme Ceropath ont été utilisées pour caractériser les habitats utilisés préférentiellement par les animaux. Durant le travail de terrain de ce programme, quatre grands types d'habitats ont été échantillonnés avec des pressions de piégeage comparables (4200 nuits pièges/habitat) et un milieu (karst) a été échantillonné de façon plus opportuniste et dans le cadre d'un travail parallèle (Latinne, 2012a). Les quatre grands habitats sont l'habitat humain (caractérisé par: les maisons isolées, les villages, les villes), les cultures irriguées (correspondant aux zones de rizières : riz, légumineuses, maïs, piments, cultures maraîchères), les cultures sèches (cultures de pente : maïs, manioc, friches herbacées ou arbustives, jeunes plantations, vergers), et les forêts (**sempervirentes** sèches, mixtes décidues, **diptérocarpes** sèches, plantations anciennes à canopée jointive). Le dispositif de piégeage est composé de lignes de dix pièges espacés chacun de dix mètres et appâtés avec du maïs ou du manioc auxquels ont été ajoutés des fruits pour le piégeage en forêt. Chaque ligne a été géo-référencée en son centre et le paysage a été décrit à cette position (quatre photos en direction des points cardinaux ont été prises pour chaque ligne). La précision de capture de chaque piège composant la ligne est alors estimée à 100 mètres. Des piégeages supplémentaires ont été réalisés avec des pièges isolés, avec, dans ces conditions, une précision de dix mètres (qui est celle du GPS). Enfin des animaux ont été récupérés auprès de chasseurs dont la maison était géolocalisée avec une précision de l'ordre de 1 à 10 km suivant le moyen de locomotion utilisé par le chasseur pour poser ses pièges. Afin d'estimer le milieu de vie des animaux, les données d'habitat ont été filtrées pour ne garder que celles qui affichaient une précision de capture inférieure ou égale à cent mètres du point de géo-référencement. L'échantillonnage est très inégal suivant les espèces, et l'information porte parfois sur très peu d'individus. C'est le cas pour les taxons *L. sabanus* ou *N. hinpoon* mais les données recueillies pour ces espèces sont concordantes avec des données mieux documentées de la littérature (voir Latinne, 2012a). Les données sont représentées sous la forme de diagrammes camembert pour chaque espèce (figure 4). Le milieu de vie des animaux a été décrit comme un pourcentage d'occupation de chaque habitat par les espèces et transformé en fonction logistique, de type  $\ln(x/(1-x))$ ,  $x$  étant le pourcentage représenté par l'habitat considéré dans le milieu de vie de l'espèce. Cette transformation permet l'analyse par régression d'une variable quantitative par

**Représentation de la transformation d'une conformation de crâne en face ventrale d'une espèce à une autre ou le long des composantes de conformation de l'ACP de conformation.**



**Figure 5:** La conformation source est représentée en trait gris continu et la conformation cible en pointillés noirs. Les vecteurs rouges sont orientés dans la direction de déplacement des points homologues d'une conformation à l'autre et leur longueur représente l'intensité de leur mouvement (amplifié 3 fois dans ce travail). Les couleurs représentent l'intensité relative des transformations : du bleu pour les transformations minimales au rouge pour les transformations les plus intenses en passant par des niveaux intermédiaires en vert et jaune.

rapport à des données de pourcentages (régression logistique). Pour éviter les valeurs infinies lors de la transformation logarithmique, les valeurs nulles (l'espèce ne fréquente jamais l'habitat) ont été transformées arbitrairement en 0,1% et les valeurs de 100 en 99,9% en considérant qu'il n'y a jamais eu 1000 individus capturés dans le même milieu.

## *2.2. Représentation de l'espace des conformations des espèces*

La représentation des données de morphométrie géométrique à l'aide de l'ACP de conformation a été décrite en détails Partie 1, paragraphe C3.3. Cependant l'objectif n'est plus ici de représenter les individus mais l'espèce à laquelle ils appartiennent. Seules les conformations moyennes des espèces ont donc été considérées. Celles-ci sont calculées pour chaque espèce comme la moyenne arithmétique des coordonnées de l'ensemble des individus qui lui sont assignés sur des bases moléculaires. La variation des conformations moyennes le long des deux composantes du plan principal informent sur l'évolution des morphologies crâniennes dans l'espace des conformations. Leur représentation peut se faire par le calcul respectif des formes théoriques correspondant à la plus faible et à la plus forte valeur des composantes de conformation 1 et 2. La plus grande valeur est alors ajoutée à la conformation moyenne et la plus petite lui est soustraite. Le déplacement des points homologues pour passer de la conformation minimale à la conformation maximale est ensuite matérialisé par des champs vectoriels (voir Partie 1 paragraphe C8.3). Pour rendre le graphique visuellement plus parlant, les changements ont été localement interpolés à l'aide d'un spectre de couleur qui varie du bleu au rouge en passant par le vert et le jaune (voir figure 5). Ces couleurs représentent les variations relatives des zones concernées par rapport à la variation totale de la structure. Une zone qui se transforme peu, proportionnellement au reste de la structure, sera colorée en bleu, une zone soumise à un changement proportionnellement plus important en vert, puis jaune et enfin rouge. Pour réaliser les interpolations du changement de conformation entre les points homologues, la fonction `spsample{sp}` a été utilisée. Elle permet d'échantillonner des points (ici 10000) à l'intérieur du polygone défini par les points homologues. La fonction `tps2d` (Claude, 2008) permet ensuite de générer une grille de déformation d'Arcy Thompson en deux dimensions sur la surface considérée. La grille de déformation d'Arcy Thompson traduite mathématiquement par les « thin plates-splines » (Bookstein, 1989) permet de matérialiser la déformation entre une valeur d'origine et une valeur cible. Ces déformations appliquées aux dix mille points régulièrement échantillonnés dans le polygone de la conformation d'origine permettent à l'aide de la fonction `image{graphics}` de générer les variations de couleur (figure 5).

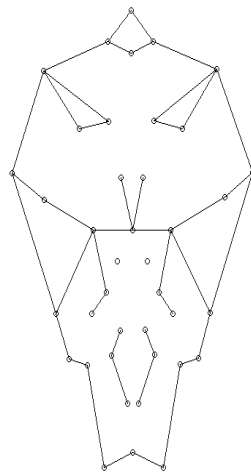
## *2.3. Représentations conjointes de l'évolution des conformations de crâne, du milieu de vie et de la phylogénie*

La représentation des données est la première étape de leur analyse. Une représentation de l'espace des conformations intégrant la représentation de la phylogénie (Rohlf, 2002 ; Klingenberg et Gidaszewski, 2010 ; Klingenberg, 2010 ; Klingenberg et al., 2011 ; Klingenberg et Mararugan-lobon, 2013 ; Monteiro, 2013) et le milieu de vie des espèces est une aide précieuse à l'interprétation de la mise en place de la diversité biologique observée.

### *Méthode*

Pour la représentation du milieu de vie des espèces, les diagrammes des milieux ont simplement été affichés sur les coordonnées des conformations moyennes des espèces dans le

**Reconstruction en face ventrale de l'ancêtre commun des Rattini.**



**Figure 6:** reconstruction de la conformation ancestrale (nommée R0) de l'ancêtre commun au Rattini étudiés grace à la méthode des contrastes indépendants (Felsenstein, 1985).



plan de l'ACP Procrustes. La représentation des relations phylogénétiques dans l'espace des conformations est plus complexe car elle exige l'estimation des conformations ancestrales.

### 2.3.1. La reconstruction des caractères ancestraux

Pour reconstruire les caractères ancestraux la méthode des moindres carrés généralisés (GLS) appliquée aux données phylogénétiquement structurées, élaborée par Martins et Hansen (1997) et décrite par Claude (2008) a été utilisée. Ce modèle utilise la variance attendue entre les taxons ancestraux (Cunningham et al., 1998). Cette variance est estimée par la matrice de variance/covariance entre les taxons actuels. Cette méthode possède l'avantage, par rapport à des méthodes basées sur le maximum de vraisemblance, de pouvoir être adaptée à une large gamme de scénarios macro-évolutifs dans la mesure où ceux-ci sont intégrés lors du calcul de la matrice de variance-covariance (Pagel, 2002).

Dans ce modèle, l'état des caractères ancestraux est déterminé comme étant une déviation de chaque nœud à la valeur de la racine.

$$\mathbf{Anc} = \mathbf{R0} + \boldsymbol{\mu} * (\mathbf{Y} - \mathbf{R0}) \quad (1)$$

Où:

**Anc** est la matrice des états de caractères des ancêtres (matrice des coordonnées Procrustes des ancêtres).

**R0** est la valeur de la racine (figure 6), à savoir une matrice des états de caractères de l'ancêtre commun (matrice des coordonnées Procrustes de l'ancêtre commun des Rattini).

**Y** est la matrice des états de caractères des espèces actuelles (matrice des coordonnées Procrustes des espèces actuelles).

**$\mu$** : est une variable de positionnement dans le chronogramme et représente la déviation à la racine.

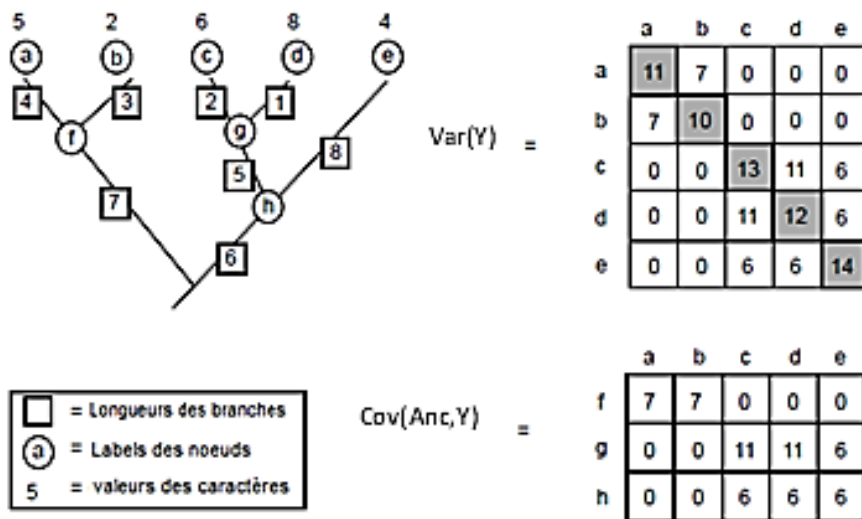
La première étape consiste à estimer la conformation ancestrale de l'ancêtre commun R0. Pour cela, il a été fait l'hypothèse que la conformation du crâne suit un mouvement d'évolution Brownien (la validité de cette hypothèse sera vérifiée au paragraphe 2.6.3). Sous cette hypothèse, il est possible d'utiliser une méthode d'ajustement aux moindres carrés implémentée dans la fonction `ace{ape}` en sélectionnant la méthode des contrastes indépendants PIC (Felsenstein, 1985).

La seconde étape est l'estimation de la proportion d'histoire commune entre les espèces actuelles et les nœuds ancestraux, représentée par la matrice  **$\mu$** .

$$\boldsymbol{\mu} = \frac{\mathbf{Cov}(\mathbf{Anc}, \mathbf{Y})}{\mathbf{Var}(\mathbf{Y})}$$

Où :

**Principe de reconstruction des caractères ancestraux à l'aide un modèle linéaire généralisé.**



Adapté de:

Cunningham, C. W., K. E. Omland, et al. (1998). "Reconstructing ancestral character states: a critical reappraisal." *Trends in Ecology and Evolution* 13(9): 361-366.

**Figure 7:** La zone grisée de la matrice (diagonale) en haut à droite représente la variance totale entre les feuilles et la racine sous l'hypothèse brownienne. La matrice de covariance en bas à droite représente l'histoire commune entre deux taxons de la racine à leur plus proche ancêtre commun.

**Cov(Anc, Y)** est la matrice de covariance qui contient les distances patristiques entre les taxons terminaux et les ancêtres. Elle correspond à l'histoire partagée entre les taxons actuels et les nœuds.

**Var(Y)** est la matrice de variance/covariance qui contient les distances patristiques entre les taxons terminaux. Elle représente l'histoire partagée entre les taxa sur l'ensemble de l'arbre.

La figure 7 présente la méthode utilisée pour le calcul de ces matrices (adapté de Cunningham et al., 1998).

La matrice de covariance entre les taxons terminaux et les ancêtres **Cov (Anc, Y)** est calculée à l'aide de la fonction `mrca{ape}` (most recent common ancestor) qui renvoie l'ancêtre commun de chaque paire de nœuds et de la fonction `dist.nodes{ape}` qui donne les distances entre chaque nœud de l'arbre. La matrice de variance/covariance attendue des espèces actuelles, **Var(Y)**, est donnée par la fonction `vcv.phylo{ape}` appliquée au chronogramme. Le chronogramme est estimé par une méthode non paramétrique de lissage des taux d'évolution (NPRS), proposée par Sanderson (1997). Cette méthode est basée sur la minimisation des changements des taux d'évolution locaux entre ancêtres et descendants et la vraisemblance que les taux d'évolution soient auto-corrélés dans le temps (Sanderson, 1997). L'application de la formule (1) permet alors le calcul des états des caractères ancestraux, ici les conformations ancestrales des crânes en face ventrale.

D'autres méthodes que celle des moindres carrés généralisés peuvent également être implémentées (quelques exemples sont donnés en annexe IX). C'est le cas de la méthode de « squared change parsimony » qui est la méthode la plus utilisée aujourd'hui pour la reconstruction de caractères ancestraux continus (Maddison, 1991, Klingenberg et Gidaszewski, 2010) et qui est un cas particulier d'estimation par maximum de vraisemblance dans lequel les longueurs de branche sont ignorées (voir Schluter et al., 1997).

### 2.3.2. Représentation

La reconstruction des états de caractère ancestraux des conformations des crânes en vue ventrale des rongeurs de la tribu des Rattini permet de visualiser l'espace des conformations occupé par la tribu au cours de son histoire évolutive. Cette représentation aide à estimer le poids de la phylogénie sur les ressemblances entre groupes taxonomiquement proches. La superposition des milieux de vie actuels permet d'apprécier leur rôle dans la structuration de l'espace des conformations. Le premier plan principal de l'ACP ne représente pas la totalité de la variance des conformations des crânes, cependant une grande partie de l'information morphologique est captée par les deux premières composantes de conformation (figure 8, ).

### 2.3.3. Rôle du milieu de vie sur la morphologie des crânes

Une première estimation de la structuration de l'espace des conformations moyennes en fonction du milieu a été inférée à l'aide d'analyses de variance en considérant la taille ou les composantes de conformation comme variables à expliquer et la fonction logistique du milieu de vie comme facteur explicatif.



## Résultats

La première composante principale de l'ACP capture 64,6% de la variance totale (figure 8) et la seconde composante en capture 26,4% (figure 8). L'espace des conformations semble structuré par les habitats et par le genre des espèces. Ainsi les espèces à dominante forestière et karstique ont des ordinations principalement positives sur la première composante et nulles à faiblement négatives sur la seconde. Les espèces qui vivent dans les milieux les plus anthropisés (habitats humains et agricoles) ont des ordinations proches de l'origine sur les deux composantes pour les espèces commensales et nulles à négatives sur la seconde composante principale pour les espèces occupant les milieux agricoles. *M. minutus* fait cependant exception car son score est minimal sur la première composante et maximale sur la seconde. La morphologie crânienne des espèces vivant dans les milieux agricoles varie principalement le long de la seconde composante de conformation qui sépare les morphologies opposées de *B. indica* et de *M. minutus*. Le milieu de vie n'explique pas significativement la taille des animaux ( $F(5, 15) = 1,295$ ,  $P = 0,317$ , tableau 4). Cependant cette même variable explique la distribution des conformations observées lorsque l'on considère les quatre premières composantes de conformation de l'ACP qui représentent 98,72% de la variance de conformation pour le crâne en vue ventrale ( $Pillai(5, 15) = 1,543$ ,  $p = 0,03101$ , tableau 5).

D'autre part, les espèces sont groupées par genre au sein de l'espace des conformations. Ainsi les genres *Leopoldamys*, *Niviventer* et *Maxomys* occupent l'espace des espèces forestières et karstiques et les genres *Berylmys* et *Bandicota*, possèdent les ordinations les plus négatives sur la seconde composante. Le genre *Rattus*, très représenté dans l'échantillonnage avec 9 espèces, se trouve en position centrale, proche de la forme moyenne qu'il influence donc fortement. Finalement la morphologie crânienne de *M. minutus* s'isole fortement des autres conformations dans le premier plan principal car ses coordonnées correspondent au maximum des valeurs positives pour la seconde composante et au maximum des valeurs négatives pour la première composante.

Les transformations de la morphologie crânienne en vue ventrale se traduisent, le long de la première composante de conformation dans le sens croissant, par un amincissement général du crâne, un raccourcissement du maxillaire et un allongement de la partie prémaxillaire. Les bulles tympaniques se réduisent considérablement le long de cet axe ainsi que la périphérie de la boîte crânienne. La limite postérieure du palais est ramenée vers l'avant. Les foramens incisifs accompagnent l'allongement de l'os prémaxillaire. Ces foramens sont cependant considérablement raccourcis par le mouvement vers l'avant proportionnellement plus important de leur partie postérieure. Les rangées dentaires sont peu affectées par ces changements de conformation mais l'on remarque un rapprochement des troisièmes molaires supérieures ce qui se traduit par un rétrécissement de la partie postérieure du palais à ce niveau. Les variations de conformation le long de la seconde composante principale, de son minimum vers son maximum, se caractérisent par un changement des proportions entre la boîte crânienne et la partie frontale du crâne formée par le palais et le maxillaire. En effet la jonction entre l'alisphénoïde et l'os squamosal ainsi que l'apophyse zygomatique de l'os squamosal migrent vers l'avant et vers l'extérieur du crâne ce qui augmente le volume de la boîte crânienne. La jonction de la bulle tympanique et de l'os basioccipital participent également, mais dans une moindre mesure, à cette augmentation du volume crânien. La partie postérieure du foramen magnum est tirée vers l'arrière ce qui a pour effet d'agrandir la partie visible de cette structure. Les bulles tympaniques s'élargissent

**Tableau 4:** Influence du milieu de vie des espèces sur la variance de la taille des crânes en vue ventrale. Le pourcentage d'occupation des milieux pour chaque espèce (x) a subi une transformation logistique de type  $\ln(x/(1-x))$ .

<i>facteur</i>	<i>Df</i>	<i>Somme des carrés</i>	<i>Moyenne Des carrés</i>	<i>Valeur de F</i>	<i>Pr(&gt;F)</i>
$\ln(x/(1-x))$	5	16,41	3,283	1,295	0,317
Residuals	15	38,02	2,534		

**Tableau 5:** Influence du milieu de vie des espèces sur la variance des 4 premières composantes de conformation des crânes (98,7% de la variance du trait) en vue ventrale. Le pourcentage d'occupation des milieux pour chaque espèce (x) a subi une transformation logistique de type  $\ln(x/(1-x))$ .

<i>facteur</i>	<i>Df (facteur)</i>	<i>Pillai</i>	<i>valeur de F</i>	<i>Df (num)</i>	<i>Df (individus)</i>	<i>Pr(&gt;F)</i>
$\ln(x/(1-x))$	5	1,543	1,884	20	60	0,03101
Residuals	15					

légèrement. Au niveau de la partie maxillaire, on note un élargissement modéré de la structure. L'os prémaxillaire est raccourci et les foramens palatins avancent en s'élargissant dans leur partie médiane. Les rangées dentaires formées par les molaires sont raccourcies et le palais s'élargit légèrement au niveau de la M1. Le changement le plus remarquable se traduit par un recul important des incisives.

### *Discussion*

Le premier plan principal de l'espace des conformations des mitotypes étudiés est fortement structuré le long des deux premières composantes de conformation qui représentent à elles seules plus de 90% de la variance de conformation pour la vue ventrale des crânes. Trois morphotypes principaux se dégagent. Le premier est représenté par une seule espèce, *Micromys minutus*, espèce paléarctique qui occupe une vaste aire de répartition, de l'Atlantique au Pacifique et semble d'une grande homogénéité morphologique et génétique (Quéré et Le Louarn, 2011). Cependant les populations du sud-est asiatique, pour lesquelles certaines différences de coloration du pelage et de mensurations externes avaient été mises en évidence par Corbet et Hill (1992), ont récemment été élevées au rang d'espèce valide à partir d'analyses morphologiques et génétiques solides (Abramov et al., 2009). L'origine paléarctique des crânes étudiés pour cette espèce pourrait donc être une cause partielle de la distance morphologique observée entre cette morphologie et les autres membres de la tribu des Rattini. Cette espèce de petite taille, possède une boîte crânienne proportionnellement plus développée que les autres espèces, des os prémaxillaires et maxillaires plus courts et des incisives supérieures beaucoup plus orientées vers l'intérieur de la mâchoire. La variation de ce dernier trait est le changement de conformation le plus important caractérisant *M. minutus* par rapport aux autres espèces présentes ici. Cette particularité anatomique ainsi que le raccourcissement des os prémaxillaires et maxillaires et le changement de proportion et de position des rangées dentaires pourraient être liés à une différence de régime alimentaire de l'espèce, trait confondu ici avec le milieu de vie. Chez les rongeurs, le régime alimentaire peut influencer fortement la face inférieure des crânes (Caumul et Polly, 2005). *M. minutus* vit principalement dans les champs de blé et leur pourtour et son régime alimentaire est composé principalement de différentes graines auxquelles s'ajoutent des arthropodes (Quéré et Le Louarn, 2011). La culture du blé est absente de la zone étudiée et ce mode de vie différent pourrait expliquer l'éloignement de ce morphotype par rapport aux autres morphologies crâniennes des espèces anthropophiles et forestières.

Le second morphotype est constitué par celui des espèces anthropophiles, utilisant soit directement l'habitat humain soit les milieux agricoles. Ces espèces occupent un espace de conformation relativement important dans le premier plan principal des conformations et peut être séparé en deux sous-groupes : 1/ les espèces principalement commensales, du genre *Rattus* qui possèdent des bulles tympaniques et une boîte crânienne plus développées ainsi que des incisives d'orientation médiane par rapport à *M. minutus* et 2/ un sous-groupe composé des genres *Bandicota* et *Berylmys* et de certaines espèces du genre *Rattus* caractérisées par des bulles tympaniques volumineuses. Ces bulles tympaniques tendent à se réduire pour les espèces de ce groupe à tendance plus forestière, comme *R. andamanensis* et plus encore pour *B. bowersi*.

Le dernier groupe morphologique est celui des espèces clairement forestières et karstiques. Elles possèdent une boîte crânienne proportionnellement plus réduite, un maxillaire plus fin, un prémaxillaire plus allongé et surtout des bulles tympaniques beaucoup moins développées. L'augmentation de volume des bulles tympaniques a été mise en relation avec la densité des animaux (Petter, 1961 ; Dobigny et al., 2002) et la transmission à longue distance des signaux sonores en basses fréquences (Petter, 1961 ; Schleich et Busch, 2004 ;





Begall et Burda, 2006). Or, si les basses fréquences se propagent mieux en milieu ouvert, un sol forestier de type humus les atténue fortement (Barrière et al., 2000; Defrance et Premat, 2001). En conséquence la diminution du volume des bulles tympaniques des espèces forestières pourrait être une adaptation à la vie en forêt et au développement d'un système de communication basé sur les hautes fréquences beaucoup moins absorbées. Les espèces karstiques pour leur part possèdent des bulles tympaniques plus développées que leurs espèces sœurs vivant en forêt ce qui peut être vu comme une tendance adaptative du trait influencée par ce milieu sans humus forestier.

Cette hypothèse adaptative ne prend pas en compte la structuration phylogénétique de l'espace des conformations. Toutes les espèces appartenant à un même genre restent groupées dans le plan principal de l'ACP de conformation suggérant que la morphologie n'est pas indépendante des relations phylogénétiques entre les espèces. De même les espèces phylogénétiquement proches semblent occuper des milieux de vie similaires. Cette organisation de la distribution laisse à penser que les variations morphologiques et environnementales pourraient ne pas être indépendantes des corrélations phylogénétiques entre espèces. Le premier groupe morphologique est constitué de la seule espèce *M. minutus*. Cette espèce est le taxon basal de la tribu des Rattini (Lecompte et al., 2008) et a donc été utilisée pour enracer l'arbre de la phylogénie moléculaire de Pagès et al. (2010). Cette espèce est la première à diverger dans l'arbre phylogénétique des Rattini et possède donc la plus longue histoire évolutive indépendante des autres espèces ce qui pourrait par exemple expliquer sa divergence morphologique. D'autre part, le groupe morphologique des espèces forestières et karstiques est représenté par les espèces des divisions *Maxomys* et *Dacnomys* (Musser et Carleton, 2005 ; Pagès et al., 2010) qui sont des divisions qui s'enracinent très tôt dans la phylogénie moléculaire de Pagès et al. (2010). On trouve ensuite les sous-groupes *Berylmys* et *Bandicota* qui divergent plus tardivement dans l'arbre des Rattini et enfin la radiation récente des *Rattus* qui forme un groupe morphologiquement relativement homogène. Pour tester l'hypothèse de l'évolution concomitante des changements de milieu de vie et de morphologie, il semble donc nécessaire de tenir compte du fait que l'histoire évolutive des taxons n'est pas indépendante (voir Felsenstein, 1985). Cependant, la nécessité de replacer l'évolution des traits étudiés dans leur contexte historique n'est pas toujours avérée (Westoby et al., 1995). Ainsi si les espèces sont le fruit d'une radiation unique et rapide, alors l'évolution de leur morphologie n'est pas influencée par la phylogénie (Polly et al., 2013). D'autre part Westoby et al. (1995) montrent que la portion phylogénétique de la variance totale de la variable d'intérêt, ici la morphologie, peut contenir une composante non indépendante de l'écologie (voir également Desdevises et al, 2003). Ainsi, pour ces auteurs, la variation d'un trait se décompose en: une part de variation qui lui est propre; une part de variation corrélée avec la phylogénie, une part corrélée aux facteurs environnementaux ou écologiques (adaptations), et enfin une part corrélée à l'interaction entre l'écologie et la phylogénie. Ce dernier compartiment de la variation est ce que Harvey et Pagel (1991) nomment le conservatisme de niche écologique, concept depuis développé et affiné dans différents travaux (Westoby et al., 1995 ; Desdevises et al, 2003 ; Cooper et al., 2010 ; Wiens et al., 2010). Ce concept peut être décrit de la manière suivante : deux espèces sœurs peuvent développer des adaptations phénotypiques à leur environnement propre, cependant, parce qu'elles ont une longue histoire évolutive commune, elles vont tendre à occuper des niches écologiques proches et vont développer des adaptations phénotypiques voisines (Desdevises et al, 2003). Il est donc important, pour comprendre le rôle du milieu de vie des espèces sur leur morphologie, de tenir compte des relations phylogénétiques entre les espèces et des interactions de celles-ci avec chacun des compartiments de la variation. L'étude des patrons macro-évolutifs qui gouvernent l'évolution du phénotype renseignera sur la méthode qui doit être mise en œuvre dans l'analyse comparative.



## *Conclusion*

Le premier plan principal de l'ACP de conformation offre une représentation de la structuration de l'espace des conformations par le milieu de vie des animaux. L'analyse de variance confirme cette structuration pour près de 99% de la variance de conformation. Cependant cette structuration peut également être causée par le facteur confondant que représentent les relations phylogénétiques entretenues entre les différentes espèces au cours de leur histoire évolutive. Afin d'estimer le signal phylogénétique et de mieux comprendre le rôle de l'environnement dans l'évolution des conformations crâniennes, il va être nécessaire de : (1) tester la dépendance des matrices de variance-covariance phylogénétiques et morphologiques ; (2) si ces matrices ne sont pas corrélées alors l'évolution des morphologies pourra être considérée comme principalement adaptative. Si au contraire ses matrices sont corrélées alors il faudra rechercher quel scénario macro-évolutif correspond le mieux à la variance morphologique observée et évaluer ainsi la part d'adaptation qu'elle contient ; (3) en fonction de ce résultat les matrices morphologiques et environnementales seront corrigées par la variance phylogénétique et l'effet de l'environnement pourra alors être testé à nouveau sur les matrices corrigées.

## *2.5. Corrélation des matrices morphologiques et moléculaires*

### *Méthode*

Le test de Mantel est communément employé en biologie comparative pour tester l'équivalence de deux matrices de distances. Cependant, jusqu'à récemment, les propriétés statistiques de celui-ci pour évaluer le signal phylogénétique contenu dans un trait continu n'avaient jamais été étudiées. Une étude récente basée sur des simulations (Harmon et Glor, 2010a) montre que ce test peut, sous certaines conditions, augmenter les erreurs de type I (rejet de  $H_0$  alors que celle-ci est vrai). Les erreurs de type I sont induites par des auto-corrélations entre les éléments de la matrice génétique qui sont liées à la structure phylogénétique sous-jacente (Harmon et Glor, 2010a). L'augmentation de ce type d'erreur peut être maîtrisée par des permutations phylogénétiques (Lapointe et Garland, 2001). Les auteurs de cette étude proposent un test de Mantel phylogénétique où la matrice moléculaire n'est plus permutée en ligne et en colonne au hasard, mais où les permutations entre espèces tiennent compte de la longueur des branches qui les séparent dans la phylogénie. Les espèces proches sont plus susceptibles d'être permutées que les espèces plus distantes. L'augmentation de l'erreur de type I est constatée principalement dans l'usage d'un test de Mantel partiel (Smouse et al., 1986) où le signal phylogénétique est estimé pour plusieurs matrices de traits en même temps. D'autre part, Harmon et Glor (2010a) montre que le test de Mantel possède un faible pouvoir statistique qui est lié à la nature même du test qui transforme les matrices de traits en matrices de distances deux à deux. L'utilisation du test de Mantel n'est donc, dans le cadre des études comparatives, conseillée que pour les données ne pouvant être représentées qu'en matrice de distances deux à deux. Le test de Mantel phylogénétique, s'il résolve le problème d'inflation de l'erreur de type I n'augmente pas le pouvoir statistique du test. Le test K (Blomberg et al., 2003), conseillé par Harmon et Glor (2010a) comme une alternative au test de Mantel a donc été utilisé pour tester si deux espèces proches phylogénétiquement ont une taille plus proche et/ou se ressemblent plus morphologiquement que ce qui est attendu par hasard. Ce test n'est pas basé sur des matrices de distances deux à deux et a ainsi une meilleure puissance statistique. Il permet en outre d'extraire une statistique descriptive informative, le paramètre K, qui permet d'inférer l'histoire évolutive du trait étudié (voir paragraphe suivant). Ce test a été réalisé sur les



composantes de conformation de l'ACP Procrustes et sur les tailles centroïdes. Cependant ce test est implémenté sous une forme univariée par la fonction `phylosignal{picante}`. Les variables de conformation des crânes ont donc été testées une par une. Ce test permet d'estimer si l'arbre d'origine (la topologie et les longueurs de branches) correspond mieux, pour le trait étudié, à un agencement des espèces par rapport à un arbre où les espèces ont été permutées au hasard pour produire l'hypothèse nulle (absence de tout signal phylogénétique ; Blomberg et al., 2003).

### Résultats

Le test K de Blomberg réalisé sur la première composante de conformation (résumant 62,2% de la variance totale de conformation) est très hautement significatif ( $p=0,001$ ). Il le reste pour les 6 premières composantes de conformation qui résument, lorsque leur information est cumulée, 99,84% de la variance totale. Il est également hautement significatif pour la taille centroïde ( $p=0,001$ ). L'évolution de la morphologie des animaux est ici très corrélée à la phylogénie des espèces comme cela est régulièrement mis en évidence chez les rongeurs (Cardini, 2003 ; Caumul et Polly, 2005 ; Hautier et al, 2012 ; Casanovas-Vilar et van Dam, 2013). Ainsi les espèces proches phylogénétiquement ont une conformation et une taille de crâne plus proche que ce qui est attendu dans le cas d'une ressemblance fortuite.

### Discussion

L'espace des conformations est fortement structuré par la phylogénie, aussi il va être nécessaire de tenir compte de cette corrélation lors de l'évaluation de l'effet du milieu de vie sur la morphologie des espèces. De plus, la corrélation qui est observée entre la morphologie et la phylogénie laisse penser que la morphologie pourrait être un bon candidat pour reconstruire la phylogénie des espèces. Cependant les traits continus, de par leur nature même, ont longtemps été considérés comme de mauvais candidats pour la reconstruction phylogénétique (Pimentel and Riggins 1987; Crisp and Weston, 1987 ; Felsenstein 1988) qui est traditionnellement basée sur des **synapomorphies** de caractères discrets. Cependant les données morphométriques remplissent le seul critère d'intégration obligatoire dans les analyses phylogénétiques, à savoir la présence d'états de caractères homologues et leur nature ne permet donc pas de les exclure de ce genre d'analyses (Rae, 1998). Il est aujourd'hui reconnu que les interrogations à ce sujet sont plus en relation avec les méthodes d'analyse qu'avec la valeur intrinsèque des données (Felsenstein, 2002). La présence d'**homoplasies** des caractères par exemple rend délicate l'utilisation des données de morphométrie pour l'établissement des **synapomorphies** qui sont le socle de l'approche cladistique de la reconstruction phylogénétique. La plupart des espèces de la tribu des Rattini exhibent des caractères morphologiques reflétant une mosaïque d'états de caractères convergents et de caractères **plésiomorphes**, ce qui rend incertaines les reconstructions phylogénétiques basées sur la morphologie (Musser et Newcomb, 1983; Musser et Holden, 1991; Fabre et al., 2013). Ainsi dans l'espace des conformations de la tribu des Rattini (figure 8), les espèces ne sont pas systématiquement groupées suivant leurs relations phylogénétiques. Les espèces des genres *Maxomys* et *Leopoldamys* par exemple, ont une ordination proche dans le plan principal de l'espace des conformations alors qu'ils appartiennent à deux divisions distinctes. Cette convergence évolutive des traits morphologiques se retrouve également au sein du genre *Rattus* chez lequel les espèces *R. rattus* et *R. tiomanicus* sont proches morphologiquement alors qu'elles sont phylogénétiquement distantes. L'évolution des connaissances et des techniques de reconstruction phylogénétique, comme par exemple la reconstruction d'arbres parcimonieux qui permettent de détecter les **homoplasies** (Lecointre et Le Guyader, 2009) ont permis l'utilisation des caractères continus en phylogénie. Cependant les algorithmes de



reconstruction ayant été historiquement conçus pour traiter des caractères discrets (Felsenstein, 2002; Smith et Hendricks, 2013), ces données ont le plus souvent été rendues discrètes par des choix arbitraires qui peuvent induire un bruit important masquant le signal phylogénétique (Gaubert et al., 2005). La possibilité d'utiliser les caractères continus sans étape de discrétisation (Goloboff et al., 2006) ouvre aujourd'hui la voie à leur utilisation pour inférer l'histoire évolutive des espèces y compris lorsque ces traits sont estimés à partir de données de morphométrie géométrique (Macleod, 2002; González-José et al., 2008; Aguilar-Medrano et al., 2011; Smith and Hendricks, 2013).

## Conclusion

La morphologie des espèces actuelles n'est pas indépendante de leur histoire évolutive commune. Il y a donc un signal phylogénétique dans la morphologie des espèces actuelles et la possibilité de l'utilisation de la morphologie pour la reconstruction des relations phylogénétiques entre espèces est aujourd'hui avérée. La part des adaptations propres aux espèces ne peut donc pas se faire sans intégrer la phylogénie dans l'analyse comparative des morphologies.

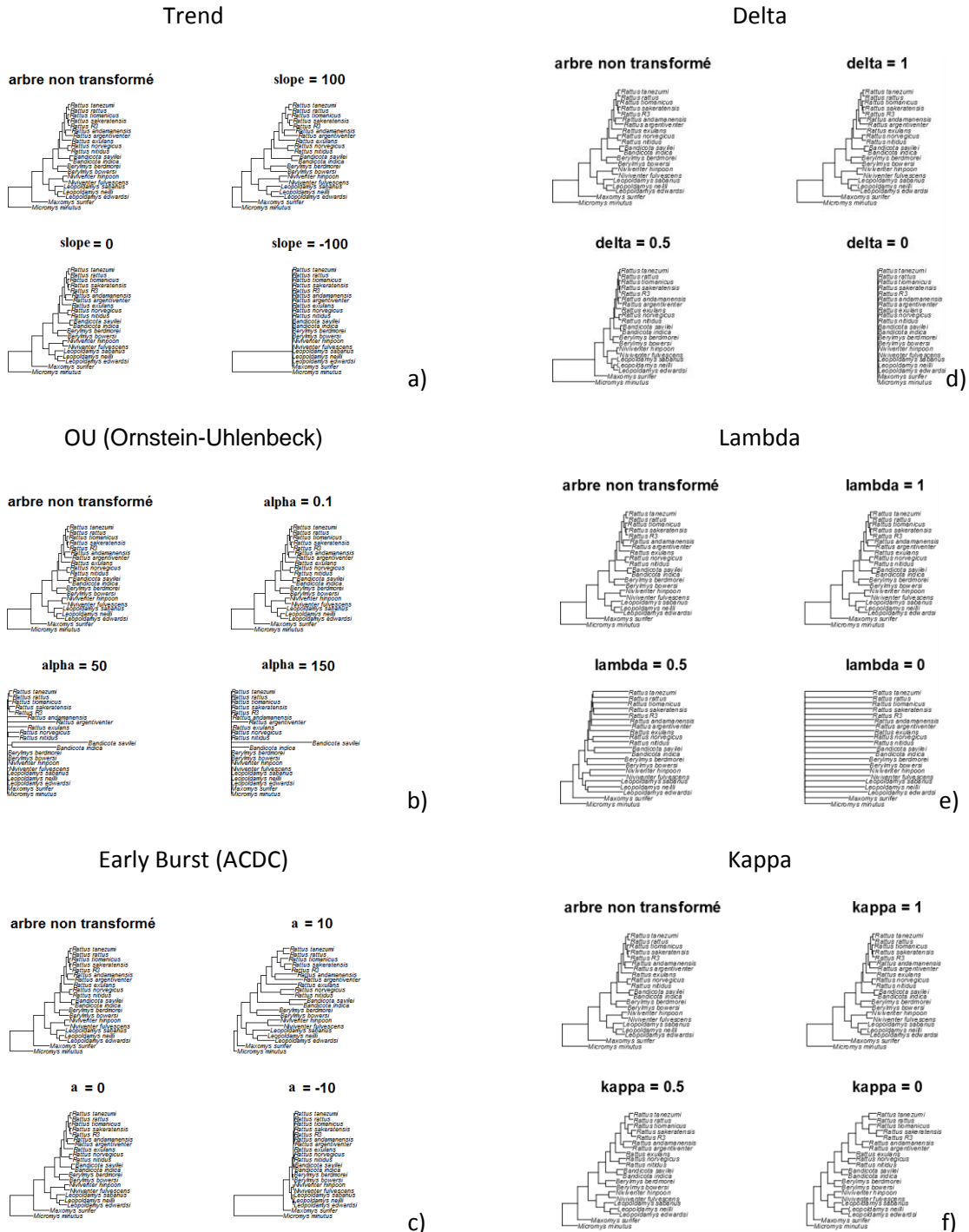
### *2.6. Scénarios macro-évolutifs et évolution des conformations et de la taille*

Le test K de Blomberg (Blomberg et al., 2003) montre que la matrice morphologique est corrélée avec la matrice de divergence et d'évolution partagée entre les espèces. Les variations phénotypiques de taille et de conformation doivent donc être évaluées en considérant que plus les espèces sont distantes phylogénétiquement, plus leur histoire évolutive est indépendante. Si la dimension phylogénétique est ignorée, le rôle de l'environnement sur la morphologie sera mal appréhendé et le risque de conclure que le lien entre morphologie et environnement est fort alors que cette hypothèse est fautive augmentera (inflation de l'erreur de **type I**). En effet, plus les espèces sont proches, plus elles se comportent comme des **pseudo-répliquations**. Cependant afin d'apporter une correction adéquate, il est nécessaire au préalable d'évaluer s'il est possible ou non d'établir une relation linéaire entre la variabilité des traits et le temps de divergence entre l'ancêtre commun et les espèces actuelles. Si une relation linéaire existe, alors l'évolution des traits peut être assimilée à un mouvement Brownien et les matrices de variance-covariance morphologique et celles décrivant le milieu de vie pourront être corrigées par la méthode traditionnelle des contrastes indépendants (Felsenstein, 1985). Si ces relations ne sont pas linéaires, alors ces matrices devront être corrigées en fonction du modèle évolutif qui décrit le mieux l'évolution des traits.

#### *Méthode*

Une première tendance de l'évolution des traits est donnée par le test K de Blomberg (Blomberg et al., 2003). En effet ce test permet, en plus du test de la corrélation des matrices, l'extraction d'une statistique descriptive, le paramètre K qui autorise la comparaison du taux de signal phylogénétique dans le trait étudié entre les feuilles de l'arbre (Blomberg et al., 2003). Ce paramètre exprime le signal phylogénétique comme le rapport entre l'erreur quadratique moyenne des données des feuilles de l'arbre et la matrice de variance-covariance dérivée d'un arbre phylogénétique donné sous l'hypothèse d'un mouvement Brownien (Münkemüller et al., 2012) et donc le rapport entre l'erreur quadratique moyenne observée sur l'erreur quadratique moyenne attendue (Blomberg et al., 2003; Krasnov et al., 2011). Si ce paramètre est inférieur à 1 alors les espèces se ressemblent moins qu'attendu sous l'hypothèse d'un mouvement Brownien. Au contraire si ce paramètre est supérieur à 1, elles se

# Déformation de l'arbre phylogénétique en fonction de différents paramètres de modèles macro-évolutifs



**Figure 9:** déformation de l'arbre phylogénétique en fonction des paramètres a) de la pente du modèle Brownien dirigé (trend), b) du paramètre alpha du modèle d'Ornstein-Uhlenbeck c) du paramètre a du modèle Early Burst ; des paramètres du modèle de Pagel d) Delta e) Lambda f) Kappa.



ressemblent plus qu'attendu sous ce même modèle macro-évolutif. Si  $K$  est égal à 1 alors le trait enregistre exactement la quantité de signal phylogénétique attendu sous l'hypothèse d'un mouvement Brownien. Le mouvement Brownien sous-entend que l'évolution du trait se fait de façon stochastique unimodale autour de la valeur ancestrale (Butler et King, 2004; Pagel, 1994, 1999). L'évolution n'est influencée que par la dérive sans être affectée par des forces sélectives (Butler et King, 2004; Harmon et al., 2010b) ou des forces sélectives très variables. La variance du trait est donc proportionnelle au temps de divergence. C'est un modèle à un paramètre, la variance évolutive instantanée (Pagel, 2002).

Une démarche complémentaire au calcul du paramètre  $K$  est de tester la vraisemblance de différents modèles macro-évolutifs pour identifier ceux qui sont le plus à même de rendre compte de l'évolution des conformations des crânes d'une part et de la taille des espèces d'autre part en fonction de la phylogénie. La fonction `fitContinuous {geiger}` permet de tester l'adéquation de différents modèles macro évolutifs aux données observées, parmi lesquels les modèles Brownien (Felsenstein, 1973), Ornstein-Uhlenbeck (OU ou alpha, Butler and King, 2004), Early Burst (EB ou ACDC, Harmon et al., 2010b), et le mouvement Brownien dirigé (voir Pagel, 1999) ainsi que la méthode de Pagel basée sur les paramètres Lambda, Delta, Kappa (Pagel, 1997, 1999, 2002; voir à titre d'exemple Piras et al., 2012).

### *2.6.1. Description des modèles testés*

Le modèle Brownien dirigé (« trend » ou « directional random-walk » ou SSP « single stationary peak » suivant les auteurs) est un modèle de diffusion où le trait étudié varie selon une tendance linéaire au cours du temps. Il est donc soumis à une force sélective de faible intensité et n'est pas influencé par la seule dérive comme dans le cas d'un modèle Brownien pur (Harmon et al., 2010b; Butler et King, 2004; Pagel, 1999). Il représente une évolution à tendance générale (les espèces sont de plus en plus grosses de plus en plus petites etc...). Ce modèle possède un paramètre supplémentaire par rapport au modèle Brownien qui est le paramètre de changement de direction. Ce paramètre est une mesure des valeurs de la régression du trait entre les espèces par la longueur évolutive (somme des longueurs de branches) de la racine aux espèces actuelles (Pagel, 2002). Il est comparable à un modèle OU qui n'aurait qu'une modalité (voir Hansen, 1997). La figure 9a donne une représentation de l'arbre phylogénétique de Pagès et al. (2010) en fonction de la valeur des paramètres du modèle.

Le modèle d'Ornstein-Uhlenbeck (OU) est un modèle d'évolution aléatoire à tendance centrale. Avant un événement de spéciation, les traits évoluent sous un régime de sélection unique (modèle Brownien dirigé). Après l'événement de spéciation, l'évolution des traits se fait de façon bimodale sous l'effet de deux régimes de sélection distincts (un mode pour chaque espèce) avec des pics proches des optimums de chaque espèce sous la contrainte d'une sélection stabilisatrice (Butler and King, 2004). Si la sélection stabilisante est forte le modèle fait tendre le trait vers un état de stase (Wiens et al., 2010). Ce modèle est utilisé pour détecter une évolution du trait plus lente que le taux de changement d'un facteur externe comme les changements environnementaux. Le trait évolue alors sous la contrainte de l'inertie phylogénétique qui est une forme de conservatisme de niche écologique (Cooper et al., 2010). La figure 9b donne une représentation de l'arbre phylogénétique de Pagès et al. (2010) en fonction de la valeur du paramètre du modèle.

Le modèle Early Burst (Harmon et al., 2010b), également appelé ACDC (ACcelerating-DeCelerating), est un modèle d'évolution où le trait évolue de façon



exponentielle en fonction du temps et d'un taux d'évolution variable :  $r(t) = r_0 * \exp(a * t)$ . avec  $r_0$  qui représente le taux d'évolution initial et  $a$  qui est le paramètre de changement de taux d'évolution qui varie entre -1 et 1. Si  $a = 1$  l'évolution s'accélère avec le temps (adaptation spécifique), si  $a = -1$  elle diminue avec le temps (radiation adaptative). La figure 9c donne une représentation de l'arbre phylogénétique de Pagès et al. (2010) en fonction de la valeur du paramètre  $a$  du modèle.

### 2.6.2. La méthode de Pagel, signification et évolution des paramètres $\lambda$ , $\delta$ , $\kappa$

Pagel (1994, 1997, 1999, 2002) a développé une méthode pour estimer le modèle évolutif du trait étudié, basée sur les moindres carrés généralisés (GLS) dont les paramètres sont estimés par maximum de vraisemblance. Le principe est de tester la vraisemblance des modèles Brownien (Pagel, 1994) ou Brownien dirigé (Pagel, 1999) aux données observées (Pagel, 2002). Il permet également en utilisant trois paramètres correctifs (par défaut chacun de valeur 1) de mise à l'échelle et de transformation de la phylogénie d'origine de tester l'adéquation de celle-ci à ces deux modèles (i.e Brownien et Brownien dirigé) pour savoir si le trait étudié doit être corrigé ou non et pour tester des hypothèses évolutives sur le trait lui-même (évolution graduelle, ponctuelle, radiation adaptative, constance du taux d'évolution). Ces trois paramètres, nommés kappa ( $\kappa$ ), lambda ( $\lambda$ ) et delta ( $\delta$ ) permettent de tester le mode ( $\kappa$ ), le tempo ( $\delta$ ) et l'association avec la phylogénie ( $\lambda$ ) de l'évolution du trait. Si, pour un trait donné, les trois paramètres ont la valeur égale à 1, cela signifie que la phylogénie et ses longueurs de branches représentent un mouvement évolutif Brownien ou Brownien dirigé (suivant le modèle sous-jacent choisi comme référence). Dans ce cas un contraste indépendant corrigera les données du trait (Felsenstein, 1985). Cependant si l'évolution du trait ne suit pas la topologie ou la longueur des branches de l'arbre phylogénétique alors ces paramètres vont varier (voir Pagel, 2002). La fonction `fitContinuous{geiger}` estime la valeur de ces paramètres de façon empirique en faisant varier les différents paramètres du modèle de Pagel et en estimant pour chacune des valeurs prises la vraisemblance de l'arbre obtenu pour décrire les données observées du trait étudié.

Le paramètre  $\lambda$  de Pagel (figure 9d) quantifie la contribution phylogénétique pour le trait étudié. Il est donc redondant avec le test K utilisé au chapitre précédent. Cependant il semble qu'il soit un estimateur plus sensible que le K de Blomberg pour de faibles écarts au modèle Brownien (Münkemüller et al., 2012) lorsque celui-ci est plus apte à détecter des évolutions de traits dirigés par une sélection stabilisatrice. Il peut varier de :

- zéro (indépendant de la phylogénie, une seule radiation, phylogénie en étoile);
- à 1 (le trait évolue selon la topologie de l'arbre et en suivant un modèle Brownien).

Une valeur intermédiaire de ce paramètre signifie que la topologie de la phylogénie surestime la variance inter-spécifique.

Le paramètre  $\delta$  (figure 9e) est une mise à l'échelle par l'histoire évolutive (i.e. de tous les chemins évolutifs de la racine aux espèces actuelles). Il permet de détecter si le taux d'évolution du trait étudié a augmenté ou au contraire diminué au cours du temps de la racine aux feuilles de l'arbre.



Une valeur estimée inférieure à 1 signifie que les chemins évolutifs plus courts contribuent proportionnellement beaucoup plus à l'évolution du trait. Les changements évolutifs du trait sont anciens, ce qui est la signature d'une radiation adaptative ancienne avec un ralentissement au cours du temps du taux d'évolution pour le trait étudié ;

- une valeur de 1 signifie les changements sont répartis tout au long de l'histoire évolutive du trait (mouvement Brownien);

- une valeur supérieure à 1 signifie que l'évolution du trait est récente et liée à une adaptation spécifique ;

Le paramètre delta ( $\delta$ ) est un paramètre qui détecte les changements de taux d'évolution et met à l'échelle la phylogénie pour qu'elle corresponde à un taux d'évolution constant (mouvement Brownien). Faire varier ce paramètre revient à tester un modèle Early Burst (Harmon et al., 2013).

Le paramètre Kappa  $\kappa$  (figure 9f) est une mise à l'échelle de la longueur des branches, il étire et compresse l'histoire évolutive entre deux nœuds de l'arbre, il permet de détecter si l'évolution du trait est ponctuelle ou graduelle :

- une valeur de 0 indique une évolution instantanée du trait suite à l'événement de spéciation. L'évolution du trait est complètement indépendante de la longueur des branches de l'arbre phylogénétique ;

- une valeur du paramètre inférieure à 1 va compresser les longues branches plus que les courtes et indiquer un ralentissement de l'évolution du trait sur les branches longues,

- une valeur égale à 1 représente une évolution graduelle (mouvement Brownien) ;

- une valeur du paramètre supérieure à 1 va étirer les longues branches de manière plus importante que les courtes. Cela indique que les longues branches contribuent plus à l'évolution du trait que les branches courtes, comme si l'évolution accélérât le long de ces branches longues. L'évolution est donc positivement corrélée à la longueur des branches.

### *2.6.3. Sélection du modèle qui correspond le mieux aux données observées*

Le modèle qui décrit au mieux l'histoire évolutive de chaque trait étudié est sélectionné comme étant celui dont le critère d'information d'Akaike corrigé (AICc) est le plus faible. Ce critère de qualité du modèle est dérivé de l'AIC (Akaike, 1987), très souvent utilisé, qui se base sur la déviance qui est le rapport entre la variance expliquée par le modèle et la variance totale suivant la relation :

$$\text{AIC} = \text{déviance} + 2*n$$

Où n représente le nombre de facteurs du modèle.

La déviance est le rapport entre la variance expliquée par le modèle et la variance totale.

L'AIC est un compromis entre l'ajustement du modèle aux données et le nombre de paramètres utilisés pour réaliser cet ajustement. En effet plus un modèle utilise de paramètres



pour décrire une relation, moins il est parcimonieux et plus sa qualité baisse. L'AIC rend compte de ce compromis entre nombre de paramètres utilisés et proximité aux données.

L'AIC peut être corrigé pour tenir compte du faible nombre d'espèces que nous avons. La pénalité générée par le nombre de facteurs (k) du modèle dans l'AIC est amplifiée par la faible taille (n) de l'échantillon dans l'AICc (Hurvich et Tsai, 1989). Plus le nombre d'individus croît plus l'AICc converge avec l'AIC.

$$\text{AICc} = \text{AIC} + \frac{2k*(k+1)}{n-k-1}$$

Une fois testée la correspondance entre des modèles d'évolution et les données observées, la distribution des différences de vraisemblance de ces modèles et du modèle Brownien est testée pour estimer si elle est significativement différente de la loi de distribution de Chi deux avec un degré de liberté.

$$\Delta (\text{trait}) = 2*(\text{MLmodèle}-\text{MLBrownien})$$

Avec :

MLm : maximum de vraisemblance du modèle testé.

MLb : maximum de vraisemblance du mouvement Brownien.

### *Résultats*

Le paramètre K est égal à 1,39 pour la taille et à 0,97 pour la conformation. Les modèles Brownien, Brownien dirigé, EB (ACDC), OU, et le modèle de Pagel pour les différentes valeurs de ses paramètres kappa ( $\kappa$ ), lambda ( $\lambda$ ), delta ( $\delta$ ) ont été testés sur les 6 premières composantes principales qui représentent 99,84% de la variance totale et sont significativement corrélées à la phylogénie d'après les résultats du test K (voir paragraphe 2.5). Les résultats sont présentés table 6. Pour toutes les composantes de conformation étudiées ainsi que pour la taille, c'est le modèle Brownien qui représente le mieux l'évolution des données observées. Seule la cinquième composante de conformation fait exception. L'évolution de la cinquième composante de conformation est mieux modélisée par la méthode de Pagel qu'elle ne le serait par un mouvement Brownien avec une valeur de delta ( $\delta$ ) égale à 2,99 (borne haute du paramètre), ce qui est la signature d'une adaptation spécifique du trait. Cependant cette tendance n'est pas plus vraisemblable que l'hypothèse d'un mouvement Brownien p (delta)= 0,0659 et la composante concernée ne capture que 0,8% de la variance totale.

### *Discussion*

Le paramètre K de Blomberg est supérieur à 1 pour la taille ce qui est le signe que le signal phylogénétique est très fort et qu'il y a un conservatisme de la taille. Les espèces se ressemblent plus que sous l'hypothèse d'un mouvement Brownien (Cooper et al., 2010) ce qui veut dire que la taille évolue plus lentement qu'attendu sous l'effet d'une évolution stochastique. Ce résultat peut être la signature d'un conservatisme de niche phylogénétique se traduisant par une rétention de niche écologique (Cooper et al., 2010) où l'évolution du trait est soumise à une sélection stabilisante ou à des contraintes évolutives qui peuvent être d'origines développementales (Cooper et al., 2010; Wiens et al., 2010). Pour ce même trait, la





valeur du paramètre delta ( $\delta$ ) pour la meilleure correspondance par la méthode de Pagel (qui est celle qui correspond le mieux aux données observées après le modèle Brownien) prend la valeur 0,79 (données non montrées) ce qui suggère un léger signal de radiation adaptative ancienne et donc une diversification de ce trait avec un ralentissement de l'évolution au cours du temps comme le supposait le paramètre K précédemment. Cependant cette tendance n'est pas significative. En effet, l'écart entre la distribution de la différence des maximums de vraisemblance du delta de Pagel et du modèle Brownien n'est pas significativement différent [ $p(\delta) = 0,69$ ], de la distribution de Chi deux avec un degré de liberté. Il en va de même des autres modèles macro-évolutifs et, malgré les valeurs prises par les paramètres K et  $\delta$ , aucun scénario d'évolution de la taille du crâne des animaux ne décrit mieux celle-ci que le mouvement Brownien. La valeur du paramètre K pour la conformation du crâne est proche de 1 ce qui signifie que ce trait évolue conformément à un mouvement Brownien (Blomberg et al., 2003; Wiens et al., 2010; Cooper et al., 2010). Les tests de correspondance à des modèles macro-évolutifs corroborent majoritairement le scénario d'évolution Brownien pour l'évolution de la conformation du crâne. Cependant il convient de garder à l'esprit que si le mouvement Brownien a été montré ici comme le modèle représentant le mieux les données observées, des modèles faisant intervenir de faibles pressions de sélection stabilisante (trend) ou même des adaptations comme le delta ( $\delta$ ) de Pagel ont des valeurs d'AICc proches du modèle Brownien (respectivement delta AICc égal à 0,13 et 0,40 voir tableau 6) pour la première composante de conformation qui résume la part de variance la plus importante du jeu de données. Ceci signifie que d'autres modèles d'évolutions donnant plus de place à l'écologie pourraient être sélectionnés sans que ce choix soit déraisonnable (Westoby et al., 1995).

### *Conclusions*

L'évolution de la taille et de la conformation des crânes en vue ventrale (i.e de leur forme) ne présente pas de tendance adaptative directionnelle forte mais évolue à partir des traits ancestraux sous la seule action stochastique de la dérive ou sous l'action de pressions de sélection variables au cours du temps. La valeur du K de Blomberg est très proche de 1 et celle du lambda de Pagel également pour les six premières composantes de conformation ( $0,92 \leq \lambda_{(Cc1 \text{ à } Cc6)} \leq 1$ ) ce qui suggère une évolution purement Brownienne pour ce trait et donc un conservatisme de niche écologique faible voire totalement absent. La valeur du paramètre  $\lambda$  (égal à 1) pour la taille suggère une évolution qui suit un mouvement Brownien alors que le K de Blomberg  $>1$  suggère que les espèces proches phylogénétiquement ont une taille plus proche que ne le suggérerait une évolution purement Brownienne. La taille du corps est un trait qui entre, parmi d'autres traits, dans la définition de la niche écologique des espèces (Wiens et al., 2010) car elle conditionne l'accès aux ressources alimentaires. Or la taille corporelle et la taille centroïde du crâne sont fortement corrélées (voir partie 1 paragraphe C2.3). Aussi la taille du crâne peut-elle être considérée comme un marqueur de la niche écologique des espèces. Une évolution lente de ces traits est souvent observée (Wiens et Graham, 2005; Wiens et al., 2010). Cependant cette évolution peut également être rapide (Schluter, 2000). Wiens et al. (2010) critiquent l'usage des paramètres K et  $\lambda$  utilisés seuls pour la mise en évidence du conservatisme de niche écologique car le calcul de ces estimateurs est fondé sur les écarts à l'hypothèse nulle (mouvement Brownien) ce qui sous-entend des changements phénotypiques substantiels du trait au cours du temps et de la phylogénie (Revell et al., 2008). Cependant le signal attendu pour mettre en évidence un conservatisme de niche écologique est l'absence de changement au court du temps. Aussi

**Tableau 7 :** Influence du milieu de vie des espèces sur la variance de la taille des crânes en vue ventrale. La taille et le milieu ont été corrigés par la méthode des contrastes indépendants (Felsenstein, 1985).

<i>facteur</i>	<i>Df</i>	<i>Somme des carrés</i>	<i>Moyenne Des carrés</i>	<i>Valeur de F</i>	<i>Pr(&gt;F)</i>
PIC milieu	5	63,88	12,78	0,687	0,641
Residuals	14	260,20	18,59		

**Tableau 8:** Influence du milieu de vie des espèces sur la variance des 4 premières composantes de conformation des crânes en vue ventrale. La taille et le milieu ont été corrigés par la méthode des contrastes indépendants (Felsenstein, 1985).

<i>facteur</i>	<i>Df (facteur)</i>	<i>Pillai</i>	<i>valeur de F</i>	<i>Df (num)</i>	<i>Df (individus)</i>	<i>Pr(&gt;F)</i>
PIC milieu	5	1,2926	1,3367	20	56	0,1956
Residuals	14					

Wiens et al. (2010) et Cooper et al. (2010) conseillent une approche fondée sur le test de différents modèles macro-évolutifs (OU, BM et « white noise »). Un modèle de stase phénotypique comme celui d'Ornstein-Uhlenbeck (OU) qui comprend une force de sélection stabilisatrice de nature à traduire le conservatisme de niche devrait donc correspondre mieux à l'évolution observée que le modèle Brownien. Cependant dans notre étude, ces modèles n'expliquent pas significativement mieux l'évolution observée que le modèle Brownien. En conséquence les données morphologiques et environnementales peuvent être corrigées à l'aide de la méthode de contrastes indépendants (Felsenstein, 1985).

## *2.7. Effet du facteur environnemental sur la morphologie en tenant compte de l'information phylogénétique*

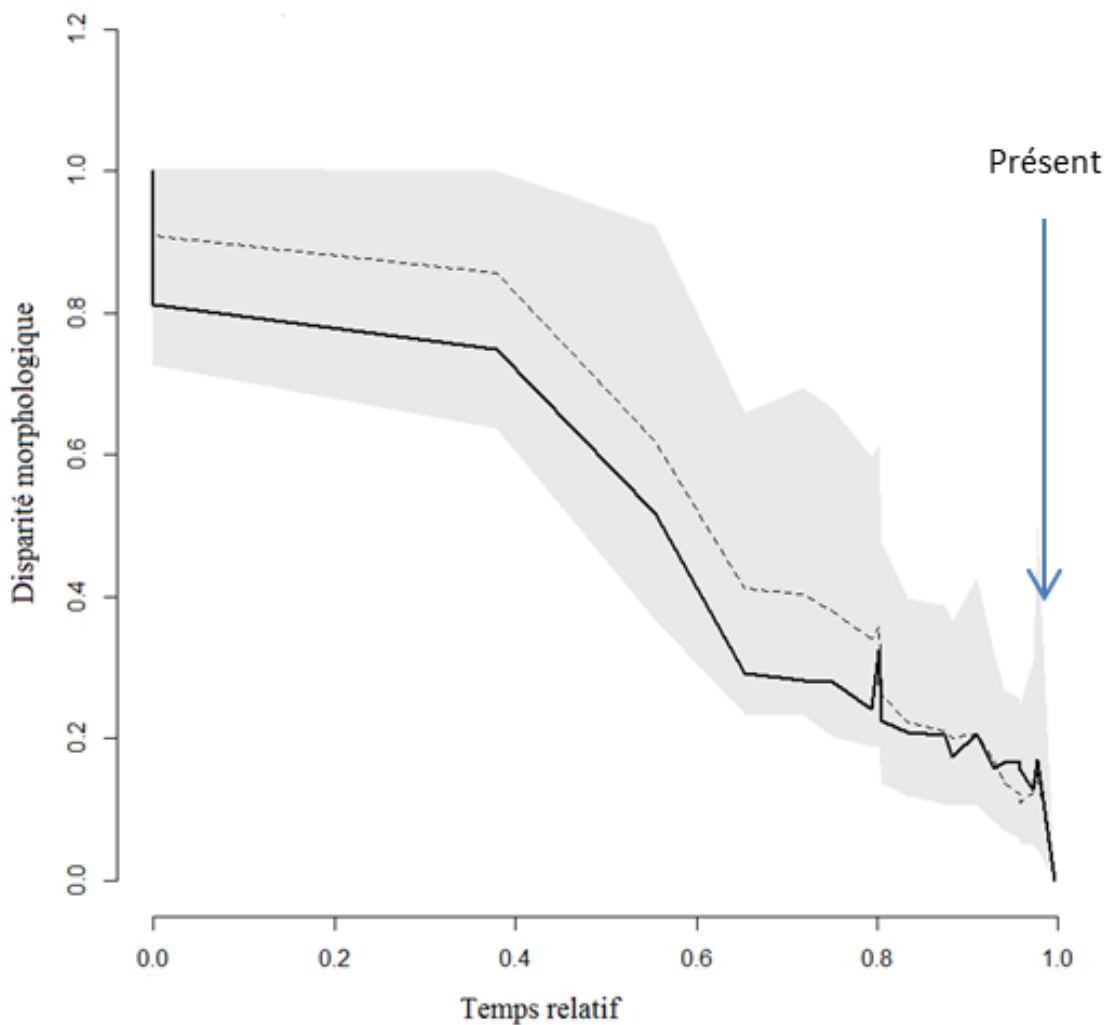
### *Méthode*

La taille, la transformation logistique traduisant les données environnementales, la matrice des scores de l'ACP Procrustes et la matrice de variance-covariance phylogénétique ont été utilisées pour calculer les contrastes phylogénétiquement indépendants entre espèces (Felsenstein, 1985). Pour cela la fonction `pic{ape}` a été utilisée. Les contrastes ont ensuite été utilisés pour tester à nouveau l'effet de l'environnement sur la taille et la morphologie rendues indépendantes de la phylogénie. De plus la disparité morphologique au cours du temps a été évaluée à l'aide de deux indices, l'indice DTT et l'indice MDI (Harmon, 2003; Rowe et al., 2011). Afin de comparer nos données à celles de Rowe et al. (2011) portant sur le genre *Rattus* des îles de la sous-région biogéographique du Sahul, les données de conformation crâniennes ont été complétées par la longueur tête/corps, la longueur de la queue et la longueur de la patte arrière ainsi que la taille moyenne des espèces (voir partie 1 paragraphe B.3.1). L'indice DTT (Disparity Thought Time) est une mesure de la disparité qui est calculée à partir de la moyenne des distances euclidiennes par paire entre les espèces. Il représente le rapport entre la dispersion intra et inter-clades. Une mesure proche de 0 signifie que la dispersion inter-spécifique est très supérieure à la dispersion intra-spécifique. Au contraire un DTT proche de 1 signifie que la dispersion intra-spécifique représente une grosse partie de la variance totale étudiée pour l'ensemble des espèces. Cette valeur (DTT proche de 1) indique que la morphologie des espèces est susceptible d'un recouvrement dans l'espace des conformations, ce qui veut dire que les espèces ont évolué indépendamment pour occuper des portions équivalentes de l'espace morphologique (Harmon, 2003). L'écart de cet indice à l'hypothèse nulle au cours du temps a été représenté en construisant une représentation à partir de 1000 simulations permettant de définir la distribution de la disparité morphologique sous l'hypothèse nulle (évolution Brownienne) au cours du temps (Harmon, 2003). L'indice MDI (Morphological Disparity Index) représente la différence globale de disparité relative d'un clade comparée à celle attendue sous l'hypothèse nulle. Il permet de tester si la disparité au sein des espèces est inférieure ou supérieure à la médiane de celle attendue sous l'hypothèse d'une évolution sous mouvement Brownien. C'est un indice très corrélé au delta de Pagel (Harmon, 2003). Des valeurs négatives de cet indice indiquent que la disparité morphologique intra-spécifique évolue moins qu'attendu sous l'hypothèse d'un mouvement Brownien et une valeur positive indique que la disparité intra-spécifique évolue plus qu'attendu sous l'hypothèse nulle.

### *Résultats*

L'environnement est un facteur qui n'explique significativement ni la taille (tableau 7) des animaux ( $F(5,14) = 0,69$ ,  $p = 0,64$ ) ni la distribution des conformations (tableau 8)

## Evolution de la disparité morphologique en fonction du temps.



**Figure 10:** disparité morphologique observée (courbe en traits pleins) et disparité morphologique attendue sous l'hypothèse d'une évolution suivant un mouvement Brownien (courbe en pointillés). La zone grisée représente l'intervalle de confiance à 95% (Harmon, 2008).

observées lorsque l'on considère les quatre premières composantes de conformation (Pillai (5, 14)= 1,29,  $p = 0,20$ ). L'indice de disparité morphologique relatif observé pour les espèces actuelles (disparité intra-spécifique/disparité globale) est compris entre 0,02 pour les espèces du genre *Rattus* et 0,2 pour *Micromys*. La valeur de l'indice MDI est négative dans notre étude comme dans celle de Rowe et al. (2011) qui obtiennent -0,08 pour les Rattini de la zone indochinoise et -0,07 pour les *Rattus* du Sahul. La disparité intra-spécifique observée est donc inférieure à celle attendue sous l'hypothèse d'un mouvement Brownien. Cette tendance se retrouve lorsque l'on représente la disparité morphologique des taxons au cours du temps (figure 10), cependant cette tendance se trouve comprise dans l'intervalle de confiance d'une distribution de la disparité morphologique sous mouvement Brownien estimé à partir de 1000 simulations (Harmon, 2003).

### *Discussion*

Ces résultats montrent que les différences d'environnement n'expliquent pas les différences de morphologie, une fois que l'histoire évolutive est prise en compte. Ils illustrent ainsi l'importance de la prise en compte de la phylogénie lors de l'analyse comparative de traits inter-spécifiques sans laquelle on aurait facilement pu conclure que les différences morphologiques inter-spécifiques résultaient de différences d'environnement (paragraphe 2.3.3). Ces résultats ont été obtenus avec la méthode des contrastes indépendants (Felsenstein, 1985) qui est une méthode de correction statistiquement conservative (Westoby et al., 1995). Si le trait évolue en partie sous la contrainte d'un conservatisme phylogénétique de niche, à la fois corrélé à la phylogénie et à l'écologie, celui-ci est considéré par cette méthode comme étant uniquement lié à la phylogénie. Cependant cette correction s'applique dans le cadre d'un mouvement d'évolution Brownien et aucun autre modèle ne s'est montré plus vraisemblable pour expliquer l'évolution des traits observés. D'autres choix auraient pu être faits comme celui de considérer la variation liée à l'écologie comme l'hypothèse nulle à la place des relations phylogénétiques (Westoby et al., 1995; Martins, 2000), donnant ainsi plus de puissance statistique pour la détection de variations adaptatives.

Une autre étude portant sur les représentants du genre *Rattus* présents dans les îles qui forment le Sahul (Rowe et al., 2011) montre que le patron de diversification observé pour ces espèces partage certaines caractéristiques liées aux radiations adaptatives, comme une partition environnementale des lignées et une diversification importantes des lignées au moment de la radiation suivi d'une décroissance du taux de spéciation au cours du temps. Notre échantillonnage d'espèces indochinoises enregistre également cette partition environnementale mais l'étude de la disparité morphologique de notre jeu de données ne permet pas de mettre en évidence une radiation adaptative pour ces espèces. En effet chaque espèce possède un indice de disparité morphologique relatif proche de 0 ce qui signifie que la part d'évolution des conformations inter-spécifique est très supérieure à la part d'évolution intra spécifique. Cette tendance est confirmée par la valeur négative de l'indice de disparité morphologique (MDI). Rowe et al. (2011) s'appuient sur la valeur négative de MDI pour étayer l'hypothèse d'une proportion d'évolution morphologique totale inférieure à celle attendue sous l'hypothèse d'une évolution à mouvement Brownien. Le tracé de la disparité morphologique au cours du temps pour les espèces indochinoises (figure 10), à savoir la moyenne inter-spécifique des distances euclidiennes deux à deux (Harmon et al., 2003), montre que la courbe observée est contenue dans l'intervalle de confiance à 95% de la courbe décrivant l'évolution sous mouvement Brownien. Elle montre que rapidement après la diversification des Rattini de la zone indochinoise, la disparité morphologique intra-spécifique est inférieure à celle attendue sous l'hypothèse d'un mouvement Brownien ce qui est souvent



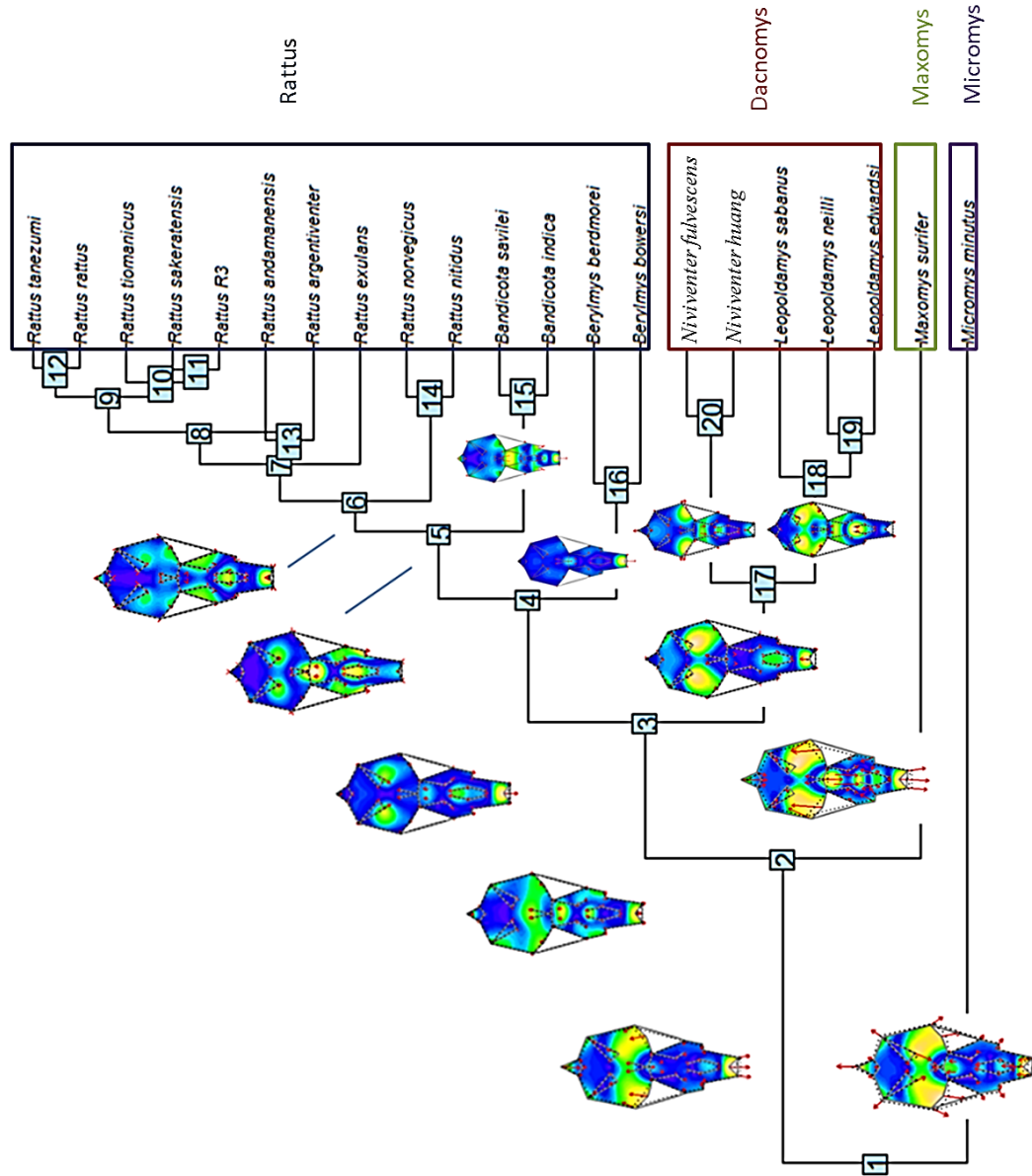
interprété comme un indice en faveur d'une divergence éco-morphologique (Rowe et al., 2011). Cependant cette tendance pour les Rattini indochinois n'est pas significative et la disparité observée reste comprise dans l'intervalle de confiance d'une évolution Brownienne. Ce résultat confirme celui obtenu lors du test de différents scénarios macro-évolutifs pour décrire l'évolution des conformations observées. La différence entre les résultats obtenus dans notre étude et celle de Rowe et al. (2011) peut être attribuée à plusieurs facteurs. Le premier est la présence d'un biais d'échantillonnage de l'un ou l'autre des jeux de données. Notre jeu de données couvre la tribu des Rattini de son taxon le plus basal à son taxon le plus sommital. Cependant il est très incomplet puisque seulement 21 sur les 167 espèces de la tribu sont représentées (Musser et Carleton, 2005). L'échantillonnage de Rowe et al. (2011), est plus exhaustif en ce que qui concerne les rats Sahuliens mais très concentré sur une petite partie de la diversité des Rattini, le genre *Rattus*. Une étude moléculaire récente (Schenk et al., 2013) et qui intègre les espèces dont la disparité morphologique a été traitée par Rowe et al. (2011), montre que le signal significatif de changement du taux de diversification obtenu par un test de cladogénèse relative (Purvis et al., 1995) est lié ici à un biais d'échantillonnage des *Rattus* Sahuliens. De plus, les zones biogéographiques étudiées n'ont pas la même histoire géologique et climatique (voir annexe 1). Ainsi l'Asie du Sud-Est est considérée comme le foyer de diversification du genre *Rattus* où l'on trouve les plus vieux fossiles du genre (Chaimanee, 1998). Des études moléculaires récentes date l'apparition du genre *Rattus* à 3 millions d'année (Verneau et al., 1998; Robins et al., 2008). Ce n'est que récemment, entre 1 et 2 millions d'années, que les *Rattus* sont apparus dans le Sahul (Rowe et al., 2011). Cette date de colonisation est incertaine car les fossiles sont rares. Cependant il est possible d'imaginer des scénarios évolutifs différents pour les Rattini des sous-régions Indochinoises et Sundaïques d'une part et pour ceux du Sahul d'autre part. Pour les premiers, la radiation au sein du genre *Rattus* aurait été lente et influencée par des évènements de vicariance comme ceux mis en évidence par l'étude de Gorog et al. (2004) pour les genres *Maxomys* et *Leopoldamys* (voir également Lohman et al., 2011 ; Latinne, 2012b). A l'inverse, la radiation des *Rattus* Sahuliens pourrait avoir été plus rapide car elle serait liée à des phénomènes de dispersion impulsés par le franchissement des multiples bras de mers séparant le Sahul du continent asiatique à cette époque (Rowe et al., 2011). Les espèces se seraient adaptées à des niches écologiques vacantes dans un environnement nouveau conduisant à une radiation adaptive rapide suivie d'un ralentissement de l'évolution du trait.

### Conclusion

L'histoire évolutive des traits morphologiques en relation avec l'écologie est parfois complexe. Toutefois, quelle que soit la nature des facteurs étudiés, la représentation d'un facteur pouvant influencer la morphologie des animaux dans l'espace des conformations devrait toujours être accompagnée de la représentation de la phylogénie dans ce même espace pour en faciliter l'interprétation (Rohlf, 2002; Klingenberg et Gidaszewski, 2010; Klingenberg, 2010; Klingenberg et al., 2011; Klingenberg et Mararugan-lobon, 2013; Monteiro, 2013). Le test de divers scénarios macro-évolutifs devrait être la règle pour expliquer l'histoire évolutive des espèces et comprendre comment la matrice de variance-covariance de traits morphologiques est influencée par l'histoire commune des espèces (Wiens et al., 2010; Cooper et al., 2010).

## Evolution des conformations en vue ventrale des crânes de rongeurs de la tribu des Rattini

### Divisions



**Figure 11:** La conformation de l'ancêtre commun est représentée en gris clair et la conformation cible en pointillés. Les flèches rouges indiquent le déplacement, amplifié 3 fois, de chaque point homologue pour passer de la conformation source à la cible. Les couleurs, du bleu foncé au rouge clair, représentent la quantité relative de déformation interpolée par 10000 points placés régulièrement sur la surface.



## 2.8. Evolution des morphologies crâniennes

### Méthode

L'évolution des morphologies crâniennes (par la méthode décrite paragraphe 2.2 et illustrée figure 5) pour chaque branche de l'arbre phylogénétique de Pagès et al. (2010) a été inférée en prenant comme conformation de référence celle de l'ancêtre commun et comme cible celle de la conformation des espèces filles. L'évolution est ainsi décrite de l'ancêtre commun des Rattini à l'ancêtre commun de chaque genre. Cette description a été complétée par la description des espèces actuelles sœurs ou des espèces susceptibles de poser des problèmes d'identification.

### Résultats

Le premier évènement de diversification rencontré lorsque l'on parcourt l'arbre de l'ancêtre commun des Rattini aux espèces actuelles est l'isolement de la lignée qui a évolué pour donner la division *Micromys* et d'une seconde lignée dont la diversification engendre la radiation de la tribu des Rattini (figure 11). L'évolution de la conformation des crânes en vue ventrale à partir de l'ancêtre commun des Rattini et jusqu'à la conformation de *Micromys minutus* se caractérise par une augmentation du volume de la boîte crânienne, un élargissement du maxillaire et un raccourcissement de l'os prémaxillaire accompagné par une inclinaison des incisives vers le palais. La structure palatine reste inchangée, et les foramens mésoptéridoïdes se raccourcissent. Les foramens incisifs s'allongent légèrement et les bulles tympaniques se développent. L'évolution de la conformation de la lignée soeur (entre le nœud 1 et 2 figure 11) de la division *Micromys* qui évolue vers l'ancêtre commun des divisions *Maxomys*, *Dacnomys* et *Rattus* se traduit par un allongement de l'os prémaxillaire, un redressement des incisives et une diminution du périmètre crânien au niveau de la jonction de l'os squamosal et de l'os alisphénoïde ainsi qu'au niveau de l'apophyse zygomatique du squamosal. Les foramens mésoptéridoïdes s'étirent également en arrière alors que les bulles tympaniques diminuent de volume. La seconde lignée qui diverge donne naissance à la division *Maxomys* dont les membres représentés ici par *Maxomys surifer* ont un périmètre crânien qui se réduit encore de par le mouvement de la jonction de l'os squamosal et de l'alisphénoïde ainsi que par celui de l'apophyse zygomatique du squamosal. La jonction de la bulle tympanique et de l'os basioccipital ainsi que la partie postérieure de la boîte crânienne participent à cette réduction du périmètre crânien en se déplaçant vers l'intérieur de la boîte crânienne. Les bulles tympaniques se réduisent de façon importante. La partie frontale du crâne est également affectée par ces transformations. Le palais se raccourcit et les foramens palatins se déplacent vers les foramens incisifs qui se raccourcissent dans leur partie postérieure tout en s'élargissant pour adopter une forme caractéristique du genre *Maxomys*. Le maxillaire se rétrécit et l'arcade zygomatique recule dans le même temps. Enfin l'os prémaxillaire s'allonge et les incisives se redressent. Le groupe frère de la division *Maxomys*, évolue vers l'ancêtre commun des divisions *Rattus* et *Dacnomys* (entre les nœuds 2 et 3). Cette évolution de la conformation du crâne se traduit par un changement très modéré et caractérisé par un très léger redressement des incisives. La séparation des divisions *Dacnomys* et *Rattus* conduit à la naissance de l'ancêtre commun des genres *Leopoldamys* et *Niviventer* (entre les nœuds 3 et 17). L'évolution de la conformation crânienne le long de cette branche est marquée par un mouvement des incisives vers les foramens palatins et une réduction du volume des bulles tympaniques. Dans le même temps, les changements morphologiques qui conduisent à l'ancêtre commun de la division *Rattus* (entre les nœuds 3 et 4) subissent des transformations opposées, c'est-à-dire que le volume des bulles tympaniques augmente alors

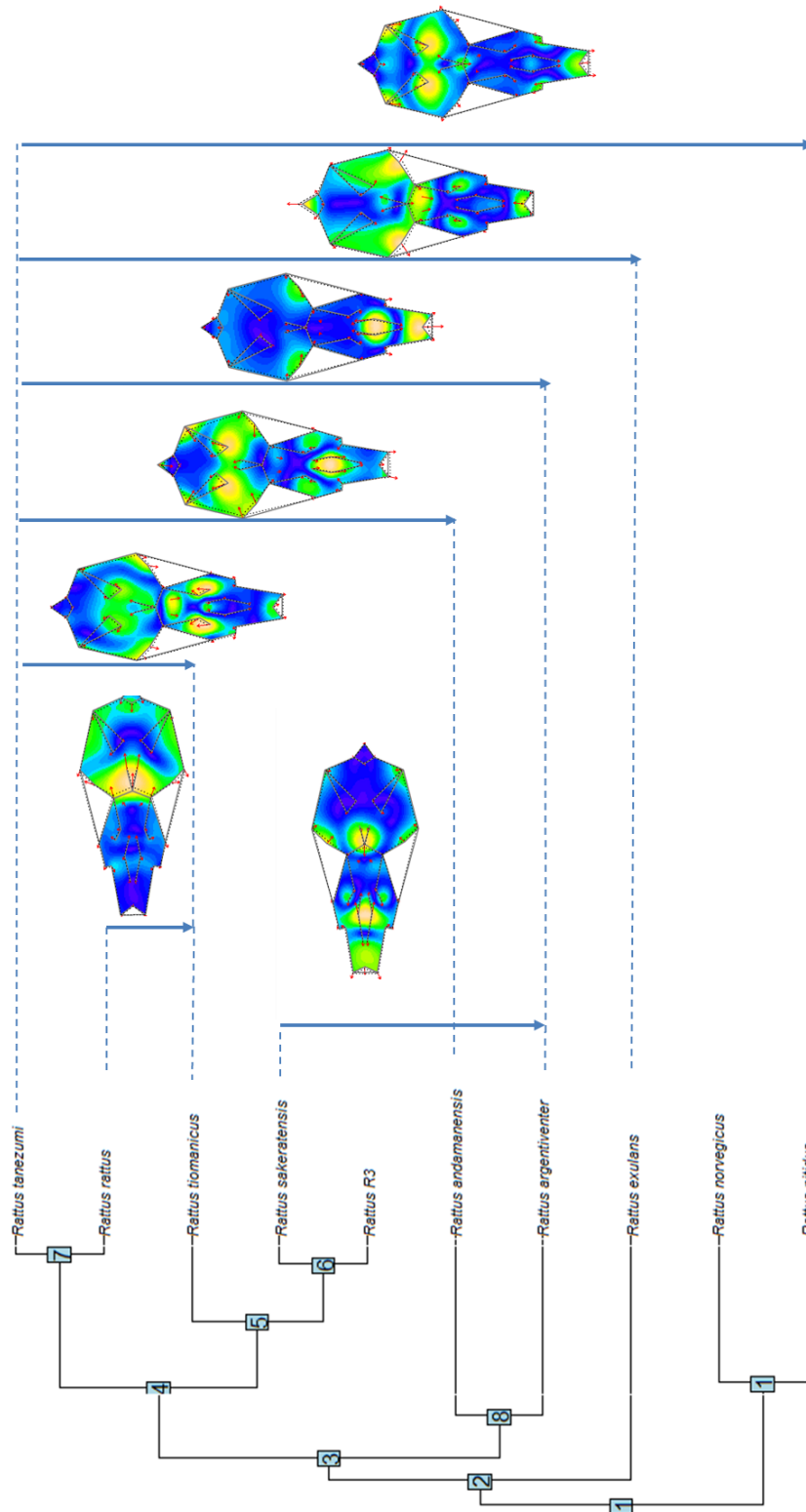


que les incisives évoluent en se redressant. Entre les nœuds 17 et 20, la conformation du crâne évolue au niveau de la jonction de l'os squamosal et de l'apophyse zygomatique du squamosal qui s'élargit un peu. Les foramens palatins avancent légèrement et les foramens incisifs s'étirent en direction du palais. Les incisives continuent à s'incliner vers l'intérieur. Ces transformations conduisent à la conformation de l'ancêtre commun au genre *Niviventer*. L'évolution de la conformation du crâne le long de la branche sœur entre les nœuds 17 et 18 conduit à l'ancêtre commun au genre *Leopoldamys* et se caractérise par une réduction des bulles tympaniques, une légère réduction du périmètre crânien, une migration de foramens palatins vers l'arrière du palais, un raccourcissement des foramens incisifs dans leur partie postérieure et un très faible allongement de l'os prémaxillaire. La séparation d'une lignée évoluant vers l'ancêtre commun au genre *Berylmys* (entre les nœuds 4 et 16) à partir de la conformation ancestrale de la division *Rattus* est marquée par une évolution qui se concentre au niveau des incisives qui se redressent fortement. La conformation portée par la branche comprise entre les nœuds 4 et 5 du chronogramme évolue peu. Cette évolution est caractérisée par le recul des foramens palatins, l'accroissement des bulles tympaniques et l'avancée de l'arche zygomatique. L'évolution de la conformation du crâne attribuable à la divergence du taxon qui évolue vers l'ancêtre commun du genre *Bandicota* est concentrée principalement dans les parties prémaxillaires et maxillaires du crâne à l'exception d'un rétrécissement du crâne au niveau de la jonction entre l'os squamosal et de l'alisphénoïde. Les foramens palatins reculent vers l'arrière du palais qui se rétrécit, l'arcade zygomatique avance, les foramens incisifs s'allongent dans leur partie frontale, l'os prémaxillaire s'élargit et les incisives se redressent. Entre les nœuds 5 et 6 la conformation évolue vers l'ancêtre commun du genre *Rattus*. Cette évolution se concentre également sur les parties prémaxillaires et maxillaires et plus particulièrement les incisives qui s'incurvent vers les foramens palatins antérieurs.

### *Discussion*

Le chemin évolutif qui mène à l'ancêtre commun des *Rattus* est marqué par une évolution globale qui n'est pas directionnelle pour l'ensemble du crâne. A titre d'exemple, la position relative des incisives par rapport à l'os prémaxillaire évolue dans un mouvement vers l'avant dans les nœuds profonds (entre les nœuds 1 à 4) avant de se stabiliser (entre les nœuds 4 et 5) et de s'incurver vers les foramens incisifs (entre 5 et 6). Lors de ce changement évolutif, c'est tout d'abord la jonction entre l'os squamosal et l'alisphénoïde qui se transforme, accompagné par un rétrécissement des bulles tympaniques et un allongement du prémaxillaire (entre 1 et 3). Ces structures se stabilisent ensuite et les transformations successives se concentrent sur le développement, qui reste cependant modéré, des bulles tympaniques (entre 3 et 6). Dans le même temps, certains traits évoluent dans le même sens au cours du temps. Ainsi les foramens incisifs ne cessent de s'allonger et les foramens palatins migrent vers le fond du palais tout au long de l'histoire évolutive des taxons. Ce chemin évolutif est donc caractérisé par des comportements différenciés des traits qui dessinent la morphologie crânienne. Ainsi on observe une évolution directionnelle de certains traits, alors que d'autres évoluent dans une direction puis dans une autre opposée. Certaines structures évoluent de façon concomitantes ou au contraire l'une après l'autre. Ce patron d'évolution est en accord avec une évolution décrivant un mouvement Brownien ou aucune force de sélection constante n'est à l'œuvre sur l'ensemble de la structure. La sélection peut exister au cours de l'évolution de la morphologie sous mouvement Brownien mais elle n'est pas une force constante et c'est donc la dérive, phénomène stochastique, qui est le mécanisme principal de l'évolution de la morphologie. Une étude plus approfondie portant sur des modules évolutifs (parties identifiées de la structure étudiée ayant la co-variation la plus faible possible) pourrait permettre d'affiner l'histoire évolutive de la face inférieure du crâne et éventuellement de mettre en évidence le caractère adaptatif de l'évolution de certains traits.

Evolution des conformations inter-spécifiques des crânes en vue ventrale pou le genre *Rattus*.



**Figure 12:** Evolution des conformations en vue ventrale des crânes pour les espèces du genre *Rattus* susceptibles de poser des problèmes d'identification sur le terrain.

Cependant cette démarche n'a pas été entreprise ici car si l'évolution des conformations estimées entre l'ancêtre commun à la tribu des Rattini et les ancêtres communs de chaque genre actuel donne une idée de l'évolution de la face inférieure du crâne, une interprétation très approfondie de l'évolution des conformations est rendue difficile par le fait que la phylogénie des Rattini utilisée est très loin d'être exhaustive. La conformation de l'ancêtre commun est estimée comme étant une moyenne pondérée par la variance de l'ensemble des espèces actuelles présentes dans la phylogénie. Ainsi l'ancêtre commun des Rattini estimé dans ce travail est probablement très éloigné de l'ancêtre réel dont la morphologie ne pourrait être rigoureusement estimée qu'à l'aide d'un échantillonnage moléculaire et morphologique beaucoup plus large couvrant l'ensemble de la tribu des Rattini qui compte 35 genres et 167 espèces (Musser et Carleton, 2005). Les espèces du genre *Rattus* ont ici quasiment le même contour que la conformation en vue inférieure du crâne de l'ancêtre commun. A l'opposé, la conformation du crâne de *Micromys minutus* en est très éloignée. Le clade des *Rattus* est représenté par 10 espèces proches morphologiquement et qui influencent donc fortement la conformation ancestrale. *Micromys minutus* est le seul représentant de la division *Micromys* et l'influence beaucoup moins. Il faut donc prendre ces inférences avec beaucoup de précautions mais garder en mémoire cette capacité de la morphométrie géométrique à décrire et à quantifier non seulement la biodiversité actuelle et passée qu'elle soit réelle (fossiles) ou estimée (reconstruction) et également à pouvoir matérialiser les transformations des objets biologiques à l'échelle de l'évolution des espèces.

## 2.9. Comparaison des conformations actuelles et implications pour l'identification des espèces

La même méthode de visualisation des changements de conformation permet de compléter la description des espèces sœurs pour lesquelles les erreurs d'identification sur le terrain ou en collection sont fréquentes. D'autres espèces plus éloignées phylogénétiquement mais aux morphologies externes proches, au moins pour certaines classes d'âge, ont également fait l'objet de comparaisons afin de compléter la description des espèces réalisées dans la première partie.

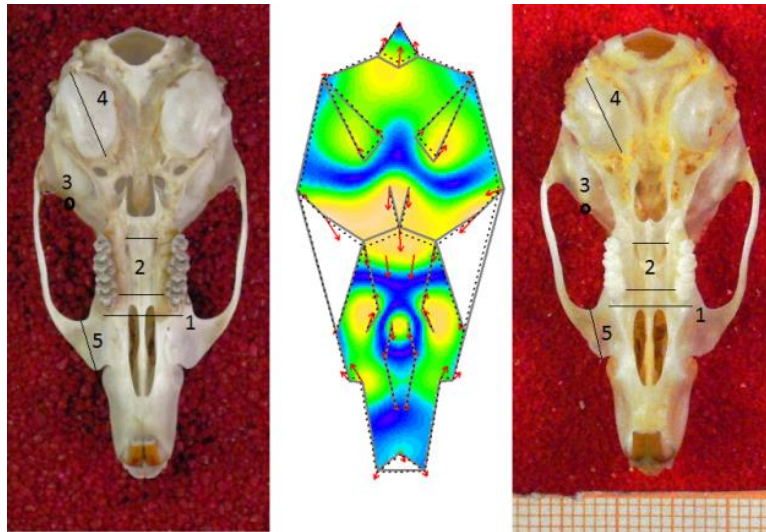
### Résultats

La figure 12 représente l'évolution des conformations entre les espèces de *Rattus* qui sont couramment confondues sur le terrain.

#### 2.9.1. Différences entre *Rattus tanezumi* et *R. rattus* (Figure 13)

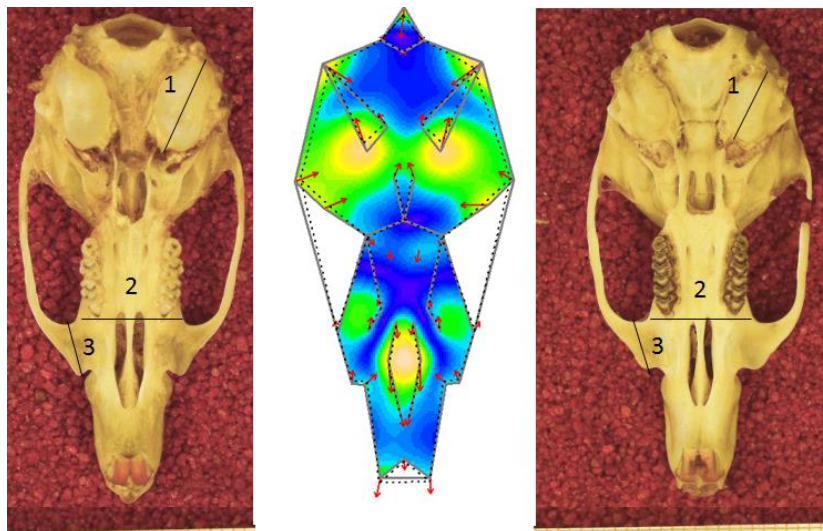
La distinction morphologique entre ces deux espèces est très intéressante car elles sont régulièrement confondues dans la littérature et sont connues pour vivre en sympatrie et s'hybrider dans les zones d'introduction (Lack et al., 2012). L'analyse montre que la morphologie crânienne est proche chez les deux espèces mais que certaines différences peuvent être observées. La partie prémaxillaire est identique chez les deux espèces mais chez *R. rattus*, les foramens incisifs se raccourcissent dans leur partie postérieure (voir la figure 13 critère n°1), de même que les rangées dentaires supérieures. La structure du palais est très différente chez les deux espèces. Ainsi chez *R. rattus* celui-ci est raccourci dans sa partie postérieure ce qui s'accompagne d'une migration des foramens palatins vers l'avant du palais. Cette migration vers l'avant des foramens palatins et le raccourcissement des rangées de molaires fait diminuer la distance entre la partie la moins large du palais et les foramens palatins chez *R. rattus* par rapport à *R. tanezumi* (voir figure 13 critère n°2). Le palais est d'autre part légèrement plus large chez *R. rattus*. La jonction entre l'os squamosal et l'alisphénoïde est plus avancée chez *R. rattus* et elle est plus proche de la troisième molaire

## Différences entre *R. tanezumi* et *R. rattus*.



**Figure 13:** Critères de différenciation entre *R. tanezumi* (à gauche) et *R. rattus* (à droite) à la même échelle. 1) les foramens palatins antérieurs ne s'allongent pas au-delà de la M1.chez *R. rattus* 2). La distance des foramens palatins à la ligne matérialisant la plus petite largeur du palais est plus grande chez *R. tanezumi* que chez *R. rattus*. 3). La jonction entre l'os squamosal et l'alisphénoïde est plus proche de la base de l'apophyse zygomatique de l'os squamosal chez *R.tanezumi*. 4) les bulles tympaniques de *R. tanezumi* sont plus développées. 5) l'arcade zygomatique est plus fine chez *R.rattus*.

## Différences entre *R. tanezumi* et *R. andamanensis*.



**Figure 14:** Critères de différenciation entre *R. tanezumi* (à gauche) et *R. andamanensis* (à droite) à la même échelle. 1) les bulles tympaniques de *R. tanezumi* sont nettement plus développées. 2) les foramens palatins antérieurs ne s'allongent pas au-delà de la M1.chez *R. andamanensis*. 3) la partie frontale de l'arcade zygomatique fait un angle obtus chez *R. andamanensis* et aigu chez *Rattus tanezumi* et l'arc est donc plus large chez cette dernière espèce.

(figure 13 critère n°3). Les foramens mésoptéridés sont plus courts. Le crâne de *R. rattus* est légèrement moins large au niveau des arcades zygomatiques et légèrement plus allongé vers l'arrière. Les bulles tympaniques sont plus petites (figure 13 critère n°4). L'arcade zygomatique est cependant légèrement plus fine chez *R. rattus* (voir figure 5 critère n°5). Les différences majeures résident au niveau du palais et des rangées dentaires, notamment au niveau de la première molaire. Le raccourcissement de la rangée dentaire s'accompagne en effet d'un mouvement vers l'intérieur du palais de la partie frontale de la première cuspide labiale chez *R. rattus*. Il y a donc dans l'évolution de cette structure une allométrie interspécifique. Une étude approfondie de la M1 supérieure de ces deux espèces par des transformés elliptiques de Fourier permettraient certainement de dissocier les deux espèces.

#### 2.9.2. Différences entre *Rattus tanezumi* et *R. tiomanicus*

Les différences observées entre ces deux espèces sont comparables à celles observées entre *R. tanezumi* et *R. rattus*. Ce résultat était attendu car l'homoplasie des morphologies pour cette vue et pour *R. tiomanicus* et *R. rattus* avait déjà été notée au chapitre 2.5 (voir figure 8). Cependant quelques différences sont à noter. Ainsi la structure du palais n'évolue pas dans sa partie postérieure entre *R. tanezumi* et *R. tiomanicus*. Les rangées dentaires se raccourcissent pour passer d'une morphologie à l'autre mais cette fois-ci il n'y a pas d'avancée de la M3. Les foramens ptéridés s'allongent chez *R. tiomanicus* par rapport à *R. tanezumi*. Cependant ces différences ne permettent pas d'établir un caractère diagnostique sur cette face du crâne. Là encore une étude approfondie de la M1 supérieure de ces deux espèces par des transformés elliptiques de Fourier permettraient certainement de dissocier les deux espèces.

#### 2.9.3. Différences entre *Rattus rattus* et *R. tiomanicus*

*R. tiomanicus* et *R. rattus* sont très proches morphologiquement pour cette vue. Cependant chez *R. tiomanicus*, le palais et les rangées dentaires s'allongent vers l'arrière de même que les foramens ptéridés. La boîte crânienne est très légèrement élargie au niveau de la jonction de l'os squamosal, de l'alisphénoïde et de l'apophyse zygomatique du squamosal. Cependant ces différences sont trop subtiles pour être utilisées à des fins diagnostiques.

#### 2.9.4. Différences entre *Rattus tanezumi* et *Rattus andamanensis* (figure 14)

*R. andamanensis* est une espèce qui appartenait anciennement au complexe *Rattus* (Aplin et al., 2011). Ce complexe d'espèces est un ensemble de taxons dont les limites ne sont pas clairement définies sur des bases morphologiques. Cette espèce principalement forestière peut parfois être rencontrée dans l'habitat humain (observation personnelle basée sur les piégeages réalisés lors du programme CERoPath). L'évolution de la morphologie du crâne en vue inférieure entre les deux espèces est marquée par des transformations importantes au niveau des bulles tympaniques qui se réduisent chez *R. andamanensis* (voir figure 14 critère n°1). Cette transformation est accompagnée par une réduction de l'ensemble du périmètre de la boîte crânienne. Les rangées dentaires et les foramens incisifs se raccourcissent légèrement de sorte ceux-ci en s'allonge pas au-delà de la M1 (figure 14 critère n°2). La partie prémaxillaire du crâne s'allonge par un mouvement concomitant de recul de l'arcade zygomatique et d'allongement de l'os pré-maxillaire. L'arcade zygomatique est enfin moins saillante chez *R. andamanensis* (figure 14 critère n°3).





#### 2.9.5. Différences entre *Rattus tanezumi* et *R. argentiventer*

*Rattus argentiventer* est une espèce typique des zones rizicoles alors que *R. tanezumi* fréquente peu ce milieu. Cependant ces espèces peuvent se trouver en sympatrie à la saison des pluies où l'humidité permet à *R. argentiventer* de fréquenter des milieux de friches arborées fréquentées également par *R. tanezumi*. *R. argentiventer* présente des incisives beaucoup plus en avant que *R. tanezumi* et les foramens incisifs sont élargis dans leur partie postérieure.

#### 2.9.6. Différences entre *Rattus tanezumi* et *R. exulans*

Les erreurs d'identification sont rares entre *R. tanezumi* et *R. exulans* sur le terrain car la morphologie corporelle de ces deux espèces est bien différenciée. Cependant certaines erreurs peuvent subvenir lorsque l'on est en présence de jeunes *R. tanezumi* qui peuvent alors être confondus avec des *R. exulans* adultes. La morphologie crânienne des deux espèces en face ventrale se différencie par une boîte crânienne plus développée chez *R. exulans* notamment au niveau de la jonction entre l'os squamosal et l'alisphénoïde. Cette différence donne un aspect beaucoup plus arrondi aux crânes de *R. exulans*. De plus les foramens palatins sont plus en avant chez cette espèce et les foramens ptéridoïdes plus en arrière. D'autre part, la rangée dentaire est plus courte chez *R. exulans* (Claude, 2013), le palais légèrement plus court et les incisives plus recourbées vers les foramens palatins antérieurs.

#### 2.9.7. Différences entre *Rattus tanezumi* et *R. nitidus*

*R. nitidus* est une espèce commensale qui remplace *R. exulans* dans le nord de la zone d'étude. Chez cette espèce, les bulles tympaniques sont moins développées que chez *R. tanezumi*, le cortex est moins arrondi au niveau de la jonction de la bulle tympanique et de l'os basisoccipital. L'arcade zygomatique à une position plus postérieure et l'os prémaxillaire s'allonge légèrement dans le même temps chez *R. nitidus*.

#### 2.9.8. Différences entre *Rattus sakeratensis* et *R. argentiventer*

*R. sakeratensis* et *R. argentiventer* sont deux espèces typiques des zones de cultures irriguées. *R. argentiventer* possède des foramens incisifs très légèrement plus allongés en avant et en arrière et qui ont une forme plus élargie sur l'arrière des foramens. La partie postérieure du palais et les foramens ptéridoïdes sont plus courts. La M1 à une position légèrement plus antérieure chez *R. argentiventer* que chez *R. sakeratensis*. Les bulles tympaniques sont identiques chez les deux espèces mais le périmètre crânien est légèrement plus petit chez *R. argentiventer* et l'os prémaxillaire est légèrement plus avancé que chez *R. sakeratensis*.

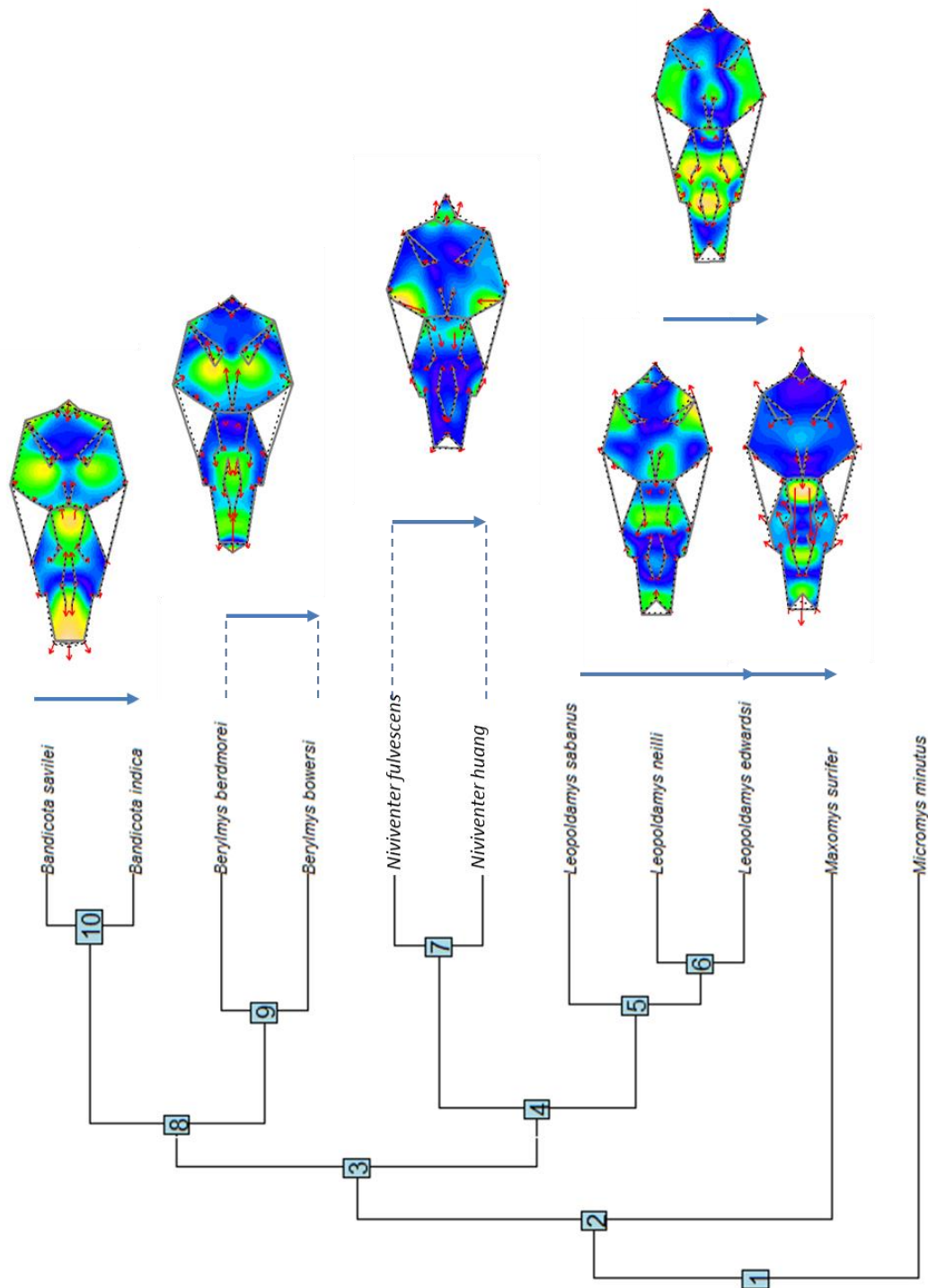
#### 2.9.9. Différences entre *Bandicota savilei* et *B. indica*.

*B. savilei* et *B. indica* sont deux espèces qui peuvent être sympatriques. Le cortex et les bulles tympaniques de *B. savilei* sont légèrement plus développés que chez *B. indica*. Le palais possède la même conformation chez les deux espèces mais les foramens palatins sont placés plus en arrière chez *B. indica*. Les foramens incisifs et l'os prémaxillaire sont plus allongés chez *B. indica*.

#### 2.9.10. Différences entre *Berylmys berdmorei* et *B. bowersi*.

*B. berdmorei* présente une boîte crânienne et des bulles tympaniques plus développées que *B. bowersi*. C'est au niveau de l'apophyse zygomatique de l'os squamosal que la

**Evolution des conformations inter-spécifiques des crânes en vue ventrale de la tribu des Rattini.**



**Figure 15:** Evolution des conformations en vue ventrale des crânes entre les espèces sœurs ou des espèces pouvant être confondues sur le terrain de la tribu des Rattini à l'exception de celles du genre *Rattus*.

réduction du périmètre crânien est la plus marquée. Les foramens ptéridoïdes s'allongent plus en arrière. Les rangées dentaires sont légèrement plus allongées et les foramens incisifs migrent vers l'avant chez *B. bowersi*.

#### 2.9.11. Différences entre *Niviventer fulvescens* et *N. huang*.

Les différences entre *N. fulvescens* et *N. huang* sont données à titre indicatif car *N. fulvescens* n'est représenté que par un individu. Il semble cependant que l'arcade zygomatique soit placée plus en arrière chez *N. huang* et que la jonction entre l'os squamosal et l'alisphénoïde soit plus étroite chez cette espèce.

#### 2.9.12. Différences entre *Leopoldamys sabanus* et *L. edwardsi*.

L'échantillonnage de *L. sabanus* est également limité (n = 4) et les différences observées ne sont donc données qu'à titre indicatif. Il semble que *L. edwardsi* présente des incisives plus incurvées vers l'arrière que *L. sabanus*, ses rangées dentaires sont légèrement allongées vers l'avant et la jonction des bulles tympaniques et de l'os basioccipital est plus en arrière.

#### 2.9.13. Différences entre *L. neilli* et *L. edwardsi*.

Les différences entre *L. neilli* et *L. edwardsi* se concentrent principalement dans les parties maxillaire et prémaxillaire du crâne. Les rangées dentaires de *L. edwardsi* sont légèrement plus allongées, l'arcade zygomatique est plus courte, le maxillaire et le prémaxillaires sont plus fins. Le prémaxillaire est légèrement plus court chez *L. edwardsi*. L'écartement maximum des foramens incisifs est plus en avant chez *L. edwardsi*.

#### 2.9.14. Différences entre *L. edwardsi* et *M. surifer*.

La morphologie externe de *M. surifer* est celle d'un *L. edwardsi* miniature et les patrons de coloration des fourrures sont identiques. La confusion est possible aussi est-il intéressant de discriminer ces espèces sur des critères de morphologie crânienne. Chez *M. surifer*, les foramens palatins sont placés beaucoup plus en avant du palais que chez *L. edwardsi*. Les foramens palatins antérieurs foramens incisifs ont une forme caractéristique en goutte d'eau, très arrondis dans leur partie postérieure et effilés vers l'avant. Les incisives de *M. surifer* sont plus redressées que celles de *L. edwardsi* et la partie maxillaire est plus large. Les rangées dentaires sont raccourcies et le palais est légèrement plus large et plus court. Les bulles tympaniques sont plus développées et la boîte crânienne plus allongé vers l'arrière.

#### 2.9.15. Différences entre les espèces du genre *Mus*

Les trois espèces de la tribu des Murini capturées en sympatrie durant le travail de terrain du projet CERoPath sont décrites car leur identification est souvent problématique sur le terrain.

#### 2.9.17. Différences entre *Mus cookii* et *M. cervicolor*

*M. cervicolor* possède des foramens incisifs qui s'allongent beaucoup plus en arrière que *M. cookii*. Les incisives sont plus redressées et l'arcade zygomatique s'arrête plus en avant chez *M. cervicolor* que chez *M. cookii*.



#### 2.9.18. Différences entre *M. cookii* et *M. caroli*

*M. caroli* possède un os prémaxillaire plus court et des incisives plus redressées que *M. cookii*. L'arcade zygomatique est plus en avant et le périmètre crânien plus important comparativement au reste du crâne chez *M. caroli*.

#### 2.9.19. Différences entre *M. cervicolor* et *M. caroli*

*M. caroli* possède un os prémaxillaire plus court que *M. cervicolor*. Elle possède également un périmètre crânien est plus grand et son augmentation est plus importante dans sa partie postérieure. Les rangées dentaires sont légèrement moins longues. Les foramens incisifs sont nettement plus courts dans leur partie postérieure.

### Discussion

La distinction entre *R. rattus* et *R. tanezumi* n'est pas aisée (Mostert, 2009) et ces espèces sont régulièrement confondues (Musser et Carleton, 2005). Cependant un critère discret, la longueur des foramens palatins antérieurs, semble être en mesure de discriminer au moins partiellement les deux espèces. Pour confirmer cela une étude morphologique devra être menée sur des populations des deux espèces vivant en sympatrie ce qui n'est pas le cas ici où les échantillons de *R. tanezumi* proviennent d'Asie du Sud-Est et les *R. rattus* du Niger. De plus ce critère ne permet pas de dissocier l'ensemble des individus car la variabilité individuelle peut conduire à une mauvaise identification dans certains cas. Avec la configuration de points homologues utilisée pour la vue ventrale, *R. tanezumi* et *R. rattus* sont trop semblables à *R. tiomanicus* pour que l'on puisse définir des critères d'identification utilisables visuellement. Leur discrimination porte sur la forme du crâne (taille et conformation associée). Ces espèces très semblables sont groupées par les taxonomistes dans le complexe *Rattus* (Musser et Carleton, 2005 ; Aplin et al., 2011) au sein duquel les espèces ont des écologies proches. *R. tanezumi* est par ailleurs connu pour entrer en concurrence et remplacer *R. tiomanicus* dans les plantations de palmiers à huile (Wood, 2001) ce qui montre que ces espèces ont des niches écologiques suffisamment semblables pour entrer en compétition. *R. andamanensis* ressemble fortement à *R. tanezumi* sur des critères de morphologie externe. Cependant *R. andamanensis* semble se différencier par ses foramens incisifs qui ne s'étendent pas vers l'arrière au-delà de la M1. Cependant un échantillonnage plus large de cette espèce sera nécessaire pour confirmer la stabilité de ce critère au regard de la variabilité intra-spécifique de l'espèce. Le crâne de *R. tanezumi* ne se différencie pas de celui de *R. argentiventer* par des critères discrets mais par des différences quantitatives qui rendent ces espèces difficiles à distinguer. Au sein du genre *Rattus*, c'est entre ces deux espèces que l'assignation morphologique pour les individus identifiés moléculairement donne les résultats les plus divergents lorsque l'on réalise une analyse discriminante (tableau 9, voir partie 1). La différenciation de *R. tanezumi* avec *R. exulans* et *R. nitidus* est plus aisée sur le terrain et les erreurs restent marginales. Là encore aucun caractère discret ne permet de différencier *R. tanezumi* des deux autres espèces et ce sont seulement des différences de proportions, entre le cortex et la partie frontale du crâne pour la première ou les bulles tympaniques pour la seconde qui, associée à des différences de taille, permettent de les discriminer. Pour le genre *Bandicota* la différenciation des espèces pose problème mais des critères de morphologie externe (Aplin et al., 2003a,b ; Chaval, 2011) ainsi que des critères moléculaires et de morphologie mandibulaire ont été avancés pour permettre de différencier efficacement ces taxons (voir partie 1 et Chaval et al., en prép. en annexe 7). Le genre *Berylmys* est caractérisé par des incisives à l'émail blanc et *B. bowersi* à des incisives plus orientées vers l'intérieur que *B. berdmorei*. Les individus des espèces *N. hinpoon* et *L. sabanus* sont trop faiblement représentés dans l'échantillonnage étudié pour que les



différences observées puissent être avancées comme critères de discrimination. Les différences entre *L. edwardsi* et *M. surifer* sont évidentes lorsque l'on considère la forme des foramens incisifs et la position des foramens palatins. Finalement, les différences entre les espèces de la tribu des Murini sont elles aussi subtiles puisqu'elles reposent également sur des différences de proportions.

### *Conclusion*

Le croisement des données moléculaires et morphométriques est riche d'enseignement pour la connaissance des espèces et pour la compréhension de l'assemblage de la biodiversité. Cette approche intégrative permet d'une part de préciser les limites d'espèces et en cela elle représente un allié précieux du systématicien. D'autre part grâce aux méthodes comparatives brièvement abordées ici, l'approche intégrative permet aussi une visualisation et une quantification de l'effet de divers facteurs, qu'ils soient historiques ou écologiques, façonnant le phénotype. Finalement, l'approche intégrative donne aux données moléculaires et morphométriques une puissance d'investigation qu'aucune d'entre elles ne pourrait avoir seule.





## *Conclusion*

Ce travail a été réalisé dans le cadre de deux programmes de recherche, roboviroses à hantavirus (ANR, dir. Jean-Pierre Hugot) et CERoPath (ANR, dir. Serge Morand). L'objectif de ces programmes est de comprendre la structuration des communautés de rongeurs et de leurs communautés de parasites dans le contexte des changements globaux. Cette tâche nécessitait la mise à jour de la taxonomie des rongeurs de la zone étudiée ainsi que la production d'outils d'identification adaptés à l'identification de terrain, au laboratoire et à l'actualisation des collections ostéologiques de référence. Un guide de synthèse des connaissances disponibles a été réalisé pour faciliter l'identification de terrain ([http://www.ceropath.org/references/rodent\\_field\\_guide](http://www.ceropath.org/references/rodent_field_guide)). L'identification au laboratoire a été rendue accessible à une large communauté d'utilisateurs par la création d'un outil, RodentSEA ([http://www.ceropath.org/barcoding\\_tool/rodentsea](http://www.ceropath.org/barcoding_tool/rodentsea)), permettant l'identification rapide des animaux capturés sur le terrain à partir de la séquence d'un gène mitochondrial (COI). Cet outil est destiné aux utilisateurs capables de produire des données moléculaires et possède une base systématique qui repose sur le croisement : de l'identification de terrain ; des données d'habitat; des travaux de délimitation d'espèces et de phylogénie moléculaires et des travaux de cytogénétique entrepris durant les programmes de recherche. L'outil est couplé à un Museum virtuel qui permet la présentation des espèces de rongeurs hôtes ([http://www.ceropath.org/virtual\\_museum/rodent\\_information\\_center](http://www.ceropath.org/virtual_museum/rodent_information_center)) et de leurs parasites ([http://www.ceropath.org/virtual\\_museum/parasite\\_information\\_center](http://www.ceropath.org/virtual_museum/parasite_information_center)).

Le travail présenté dans ce mémoire avait pour objectif principal d'étudier la faisabilité de la mise au point d'un nouvel outil web-service destiné à la diffusion d'une expertise morphométrique indépendante de l'identification moléculaire mais réalisée à partir d'un échantillonnage validé par celle-ci. Un outil de ce type était attendu pour permettre l'actualisation des collections ostéologiques de référence et assurer une identification fiable pour des utilisateurs dans l'incapacité de produire des données moléculaires. L'analyse de la faisabilité d'un tel outil s'est articulée autour de deux approches, l'étude de la variabilité morphologique des espèces réalisée à l'aide de techniques de morphométrie traditionnelle et de morphométrie géométrique, et le questionnement relatif à l'encadrement de l'erreur de mesure lors de la production de données morphométriques par des opérateurs multiples.

L'étude de la variabilité morphologique avait pour objectif d'appréhender la variation intra-spécifique (âge, sexe) et inter-spécifique. En effet un certain nombre de facteurs confondant ont une influence sur notre capacité à distinguer les espèces et leur étude permet de savoir lesquels sont à prendre en compte, pour chaque structure étudiée, lors de la mise en place de l'outil d'identification. Cette étude a montré que chez les rongeurs étudiés, les tailles corporelle et crânienne en vue ventrale des individus adultes sont influencées par l'espèce à laquelle ils appartiennent et par leur sexe. La conformation et la forme du corps sont influencées par l'espèce, l'âge et le sexe. La conformation et la forme du crâne en vue ventrale sont influencées par l'âge et l'espèce mais ne présente pas de dimorphisme sexuel significatif. Les tailles corporelle et crânienne sont fortement corrélées, cependant certaines espèces expriment des différences de proportions plus ou moins marquées. L'espace des formes et celui des conformations corporelles sont relativement similaires, par ailleurs des allométries intra-spécifiques sont différentes entre groupes d'espèces au niveau de la face ventrale du crâne. En particulier, la tribu des Murini et celle des Rattini, diffèrent dans leur patron de variation lié à la taille. La conformation des crânes ne suit donc pas les mêmes patrons de variation au cours de la croissance chez les espèces de ces deux tribus.



L'outil d'identification morphologique proposé repose sur l'analyse discriminante linéaire prédictive. Celle-ci utilise l'appartenance mitotypique des individus comme à priori. Cette analyse permet de discriminer les espèces avec plus ou moins de robustesse suivant le jeu de données considéré (points homologues saisis sur le crâne ou la mandibule, mesures corporelles). Ainsi, l'information collectée sur la face inférieure du crâne et dans une moindre mesure celle collectée sur la face dorsale sont plus robustes pour discriminer les taxons étudiés que l'information apportée par les héli-mandibules qui elle-même l'est plus que l'information apportée par les mesures corporelles traditionnelles. L'étude fait néanmoins ressortir que, pour certains genres, l'information portée par les héli-mandibules est parfois aussi pertinente que l'information crânienne pour l'identification. L'assignation au genre est toujours réalisée par l'analyse discriminante lorsque l'on considère la conformation de la face ventrale du crâne. Son pouvoir discriminant s'améliore lorsque l'on considère la forme de cette structure. Au Final, sur 17 espèces différentes, les informations collectées sur la forme de cette structure permettent d'identifier correctement 98% (i.e. de manière congruente avec leur identification mitotypique) des 660 individus de l'échantillon, quel que soit leur âge et leur sexe.

Ces résultats prometteurs permettent d'envisager la mise en place d'un outil web-service pour l'identification morphologique des rongeurs en Asie du sud-est. Cet outil permet de gérer la variance induite par la présence d'opérateurs multiples en limitant l'erreur de mesure systématique et stochastique. En particulier, la standardisation de la prise de données permet de réduire ces erreurs et des protocoles prévus à cet effet sont fournis. De plus, un module d'apprentissage permet à l'utilisateur d'apprendre à placer ses points homologues de manière analogue à l'opérateur de référence. L'utilisateur est invité à digitaliser un sous-échantillonnage des vues de référence, correspondant à environ 5% du jeu de données de référence. L'outil est alors capable d'estimer et de visualiser l'erreur produite et d'indiquer à l'utilisateur quels sont les points homologues dont le placement s'éloigne sensiblement de la référence et dans quelle direction ils s'en éloignent. Une fois l'utilisateur entraîné, les points homologues sont placés de manière analogue à la référence et l'outil est en mesure d'assigner un individu inconnu à l'échantillonnage de référence. Par ailleurs, l'outil a été paramétré pour pouvoir gérer des individus qui n'appartiennent pas aux espèces présentes dans le jeu de données de référence. Pour cela nous avons défini des restrictions quant aux possibilités d'assignation d'un individu à un groupe. La première restriction empêche l'outil d'assigner un individu au groupe auquel l'assignerai l'analyse discriminante si celui-ci est trop éloigné morphologiquement de la valeur centrale groupe. La seconde restriction est probabiliste et empêche l'outil d'assigner un individu inconnu à un groupe lorsque la probabilité postérieure d'assignation est trop faible. Ces restrictions limitent fortement la tolérance de l'outil à la variance induite par l'effet 'opérateurs'. Cependant elle autorise l'outil à fonctionner dans un mode « dégradé » lorsque l'utilisateur présage qu'il est suffisamment expert dans la prise de données morphométriques. Dans ce mode, l'outil pourra fournir des indications sur la proximité de la morphologie de l'individu aux morphologies des genres de référence sans toutefois assigner les individus soumis par l'utilisateur à un groupe.

Ce travail, en plus de permettre l'étude de faisabilité de l'outil d'identification, a également permis de compléter l'approche moléculaire pour la délimitation d'espèces et de mieux définir certains taxons. Entre autre, il propose une taxonomie intégrative du genre *Bandicota* et il permet de mieux de préciser la limite d'espèce de *Rattus tanezumi*. Cette dernière espèce possède la particularité d'être polyphylétique sur la base du signal mitochondrial. La morphologie externe et la morphologie crânienne de l'espèce ne sont pas influencées par l'appartenance à l'un ou l'autre de ces mitotypes. Les résultats de la morphométrie confirment ceux obtenus par l'étude de marqueurs nucléaires hypervariables et



permettent de fournir des arguments supplémentaires en faveur de la conspécificité des deux mitotypes. Finalement la possibilité offerte par la morphométrie géométrique pour visualiser les transformations crâniennes inter-spécifiques a été mise à profit pour préciser les variations observées entre espèces proches. Lorsque cela était possible, des caractères diagnostiques pour l'identification ont été proposés.

L'évolution des espèces définit le cadre conceptuel de la biologie contemporaine, ce que Dobzhansky résumait par le célèbre aphorisme: "Rien n'a de sens en biologie, si ce n'est à la lumière de l'évolution". Le programme de recherche CERoPath a permis la réalisation d'un travail moléculaire robuste décrivant les relations phylogénétiques au sein de la tribu des Rattini pour la sous-région indochinoise. Ces connaissances nous ont permis d'interpréter la diversité morphologique présente au regard de l'histoire évolutive commune entre les taxons et de tester la corrélation entre l'évolution des phénotypes et l'environnement. A l'échelle macro-évolutive, la phylogénie est fortement corrélée aux différences de morphologie crânienne actuelle tandis que le milieu de vie semble avoir une influence moindre pour comprendre ces différences. L'histoire évolutive des morphologies crâniennes pour les espèces de la tribu des Rattini de la sous-région indochinoise a été comparée à celle d'espèces appartenant à la même tribu étudiées dans la sous-région du Sahul. Les différences de patrons évolutifs observées ont été discutées dans le contexte de l'histoire distincte des deux zones biogéographiques considérées qui sont séparées par la célèbre ligne de Wallace.

Dans le contexte évolutif des différences intra-spécifiques, des variations morphologiques peuvent être liées à l'ontogénèse des espèces, à l'adaptation de populations à des conditions locales ou à la plasticité phénotypique. Ce travail ouvre ainsi des questions sur l'impact des changements de paysage sur les espèces et leur variabilité morphologique.



## Bibliographie.

- ABRAMOV A.V., MESCHERSKY I.G., ROZHNOV V.V. (2009). On the taxonomic status of the harvest mouse *Micromys minutus* (Rodentia: Muridae) from Vietnam. *Zootaxa*, 2199, 58-68.
- ADAMS D.C., ROHLF F.J., SLICE D.E. (2004). Geometric morphometrics: Ten years of progress following the 'revolution'. *Italian Journal of Zoology*, 71, 5-16.
- AGUILAR-MEDRANO R, FRÉDÉRICH B, DELUNA E, EF B. (2011). Patterns of morphological evolution of the cephalic region in damselfishes (Perciformes: Pomacentridae) of the Eastern Pacific. *Biological Journal of the Linnean Society*, 102, 593–613.
- AKAIKE H. (1987). Factor analysis and AIC. *Psychometrika*, 52, 317-332.
- ALTSCHUL S.F., MADDEN T.L., SCHAFFER A.A., ZHANG J., ZHANG Z., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, 25, 3389-3402.
- APLIN K., BROWN P., JACOB J., KREBS C., SINGLETON G. (2003a). Field methods for rodent studies in Asia and the Indo-Pacific. *ACIAR Monograph*. Canberra 132 p.
- APLIN K.P., FROST A., PHU TUAN N., LAN L.P., MANH HUNG N. (2003b). Identification of rodents of the genus *Bandicota* in Vietnam and Cambodia. In: "Rats, mice and people: rodent biology and management". *Australian Centre for International Agricultural Research (ACIAR)*. Canberra: SINGLETON, G. R., HINDS, L. A., KREBS, C. J. and SPRATT, D. M. (eds). 531-535.
- APLIN K., SUZUKI H., CHINEN A.A., CHESSER R.T., HAVE J.T., et al. (2011). Multiple Geographic Origins of Commensalism and Complex Dispersal History of Black Rats. *PLoS One*, 6.
- ARNQVIST G., MARTENSSON T. (1998). Measurement error in geometric morphometrics: empirical strategies to assess and reduce its impact on measure of shape. *Acta Zoologica Academiae Scientarum Hungaricae*, 44, 73-96.
- AUFFRAY J.-C., ORTH A., CATALAN J., GONZALEZ J.-P., DESMARAIS E., et al. (2003). Phylogenetic position and description of a new species of subgenus *Mus* (Rodentia, Mammalia) from Thailand. *Zoologica Scripta*, 32, 119-127.
- AUFFRAY J.C., BRITTON-DAVIDIAN J. (2012). The house mouse and its relatives: systematics and taxonomy. In: BAIRD S.J., MACHOLAN M., editors. *Evolution of the house mouse*. *Cambridge University Press*. Cambridge. 1-34.
- BADENHORST D., HERBRETEAU V., CHAVAL Y., PAGES M., ROBINSON T.J., et al. (2009). New karyotypic data for Asian rodents (Rodentia, Muridae) with the first report of B-chromosomes in the genus *Mus*. *Journal of Zoology*, London, 79, 44-56.
- BADENHORST D., DOBIGNY G., ADEGA F., CHAVES R., O'BRIEN P.C.M., et al. (2011). Chromosomal evolution in Rattini (Muridae, Rodentia). *Chromosome Research*, 19, 709–727.
- BADENHORST D., DOBIGNY G., ROBINSON T.J., CARDOSO A. (2012). Karyotypic Evolution of *Hapalomys* Inferred from Chromosome Painting: A Detailed Characterization Contributing New Insights into the Ancestral Murinae Karyotype. *Cytogenetic and Genome Research*, 136, 83-88.
- BALAKIREV A.E., ROZHNOV V.V. (2010). Phylogenetic relationships and species composition in the genus *Niviventer* (Rodentia, Muridae) based on studies of the cytochrome b gene of mtDNA. *Moscow University Biological Sciences Bulletin*, 65, 170-173.
- BALAKIREV A.E., ABRAMOV A.V., ROZHNOV V.V. (2011). Taxonomic revision of *Niviventer* (Rodentia, Muridae) from Vietnam: a morphological and molecular approach. *Russian Journal of Theriology*, 1, 1-26.
- BALAKIREV A.E., ROZHNOV V.V. (2012). Contribution to the species composition and taxonomic status of some *Rattus* inhabiting Southern Vietnam and Sundaland. *Russian Journal of Theriology*, 11, 33-45.
- BALAKIREV A.E., ABRAMOV A.V., ROZHNOV V.V. (2013a). Revision of the genus *Leopoldamys* (Rodentia, Muridae) as inferred from morphological and molecular data, with a special emphasis on the species composition in continental Indochina. *Zootaxa*, 3640, 521-549.
- BALAKIREV A.E., ANISKIN V.V., TIEN T.Q., ROZHNOV V.V. (2013b). The taxonomic position of *Tonkinomys daovantieni* (Rodentia: Muridae) based on karyological and molecular data. *Zootaxa*, 3734, 536-544.
- BARRIÈRE N., GABILLET Y., DEFRANCE J. (2000). La forêt : un écran antibruit météorologique. *Acoustique & Technique, CIDB*, 23, 41-48.
- BEGALL S., BURDA H. (2006). Acoustic communication and burrow acoustics are reflected in the ear morphology of the coruro (*Spalacopus cyanus*, Octodontidae), a social fossorial rodent. *Journal of Morphology*, 267, 382-390.
- BLOMBERG S.P., GARLAND T., IVES A.R. (2003). Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution*, 57, 717-745.

- BOOKSTEIN F.L. (1981). Looking at mandibular growth: some new geometrical methods. *Craniofacial Biology*, University of Michigan, Press Ann Arbor 74-79.
- BOOKSTEIN F.L. (1982). Foundations of morphometrics. *Annual Review of Ecology and Systematics*, 13, 451-470.
- BOOKSTEIN F.L., CHERNOFF B., ELDER R.L., HUMPHRIES J.M.J., SMITH G.R., et al. (1985). *Morphometrics in evolutionary biology*. Academy of Natural Sciences Press, Philadelphia.
- BOOKSTEIN F.L. (1986). Size and shape spaces for landmark data in two dimensions. *Stat Sci*, 1, 181-222.
- BOOKSTEIN, F.L. (1989). Principal warps: thin-plate splines and the decomposition of deformations. *Pattern Analysis and Machine Intelligence*, IEEE Transactions on 11(6), 567-585.
- BOOKSTEIN F.L. (1991). *Morphometric tools for landmark data: Geometry and Biology*. Cambridge: Cambridge University Press. 435 p.
- BOOKSTEIN F.L. (1998). A hundred years of morphometrics. *Acta Zool Acad Sci Hung*, 44, 7-59.
- BRADLEY R., BAKER R. (2001). A test of the genetic species concepts: cytochrome-b sequences and mammals. *Journal of Mammalogy*, 82, 960-973.
- BURNABY T.P. (1966). Growth-invariant discriminant functions and generalized distances. *biometrics*, 22, 96-110.
- BUTLER M.A., KING A.A. (2004). Phylogenetic Comparative Analysis: A Modeling Approach for Adaptive Evolution. *The american naturalist*, 164, 683-695.
- CARDINI A. (2003). The geometry of marmot (Rodentia: Sciuridae) mandible: phylogeny and patterns of morphological evolution. *Systematic Biology*, 52, 186-205.
- CARLETON M.D., MUSSER G.G. (2005). Order Rodentia. In: "Mammal species of the world: A taxonomic and geographic reference" WILSON DE, REEDER, DM. (eds).3rd edition. Johns Hopkins University. Baltimore. 894-1531.
- CASANOVAS-VILAR I., VAN DAM J. (2013). Conservatism and Adaptability during Squirrel Radiation: What Is Mandible Shape Telling Us? doi:10.1371/journal.pone.0061298. *PLoS One*. pp. e61298.
- CAUMUL R., POLLY P.D. (2005). Phylogenetic and environmental components of morphological variation: skull, mandible, and molar shape in marmots (*Marmota*, Rodentia). *Evolution*, 59, 2460-2472.
- CHAIMANEE Y. (1998). Plio-Pleistocene rodents of Thailand. *Thai Studies in Biodiversity*, 3, 1 - 303.
- CHAVAL Y., DOBIGNY G., MICHAUX J., PAGÈS M., CORBISIER C., et al. (2010). A multi-approach survey as the most reliable tool to accurately assess biodiversity: the example of Thai murine rodents. *Kasetart journal (Nat Sci)*, 44, 590-603.
- CHAVAL Y. (2011). South East Asian rodent field guide. [http://www.wceropath.org/references/rodent\\_field\\_guide](http://www.wceropath.org/references/rodent_field_guide): ANR BiodivhealthSEA.
- CHEVRET P., JENKINS P., CATZEFLIS F. (2003). Evolutionary systematics of the Indian mouse *Mus famulus* Bonhote, 1898: molecular (DNA/DNA hybridization and 12S rRNA sequences) and morphological evidences. *Zool J Linn Soc*, 137, 385 - 401.
- CHIARI Y., CLAUDE J. (2012). Morphometric identification of individuals when there are more shape variables than reference specimens: A case study in Galapagos tortoises. *Comptes Rendus Biologies*, 335, 62-68.
- CIA (2012). The World Factbook. *Central Intelligence Agency*, Washington, DC. <https://www.cia.gov/library/publications/the-world-factbook/index.html>
- CLAUDE J., PRITCHARD P.C., TONG H., PARADIS E., AUFFRAY J.C. (2004). Ecological correlates and evolutionary divergence in the skull of turtles: a geometric morphometric assessment. *Systematic Biology*, 53, 933-948.
- CLAUDE J., (2008). *Morphometrics with R*. Springer. 318 p.
- CLAUDE J. (2013). Log-Shape Ratios, Procrustes Superimposition, Elliptic Fourier Analysis: Three Worked Examples in *R*. *Hystrix*. 24(1), 94-102.
- COKER R.J., HUNTER B.M., RUDGE J.W., LIVERANI M., HANVORAVONGCHAI P. (2011). Emerging infectious diseases in southeast Asia: regional challenges to control. *Lancet*, 377, 599-609.
- COOPER N., JETZ W., FRECKLETON R.P. (2010). Phylogenetic comparative approaches for studying niche conservatism. *Journal of Evolutionary Biology*, 23, 2529-2539.
- CORBET G.B. (1990). The relevance of metrical chromosomal and allozyme variation to the systematics of the genus *Mus*. *Biological Journal of the Linnean Society*, 41, 5-12.
- CORBET G.B., HILL J.E. (1992). Mammals of the Indomalayan region. A systematic review. *Oxford University Press*. 488 p.
- CRISP M., WESTON P. (1987). part 3: Cladistics and legume systematics, with an analysis of the Bossiaceae, Brongniartieae and Mirbelieae', in 'Advances in legume systematics', part 3, Stirton, C. (eds). *Royal Botanical Gardens*. Kew: 65-130.



- CUNNINGHAM C.W., OMLAND K.E., OAKLEY T.H. (1998). Reconstructing ancestral character states: a critical reappraisal. *Trends in Ecology and Evolution*, 13, 361-366.
- DASMAHAPATRA K.K., MALLETT J. (2006). DNA barcodes: recent successes and future prospects, *Heredity*, 97, 254–255.
- DASZAK P., CUNNINGHAM A.A., HYATT A.D. (2000). Emerging Infectious Diseases of Wildlife. Threats to Biodiversity and Human Health. *Science*, 287, 443-449.
- DAVIS S. (1983). Morphometric variation of populations of house mice *Mus domesticus* in Britain and Faroe. *Journal of Zoology*, 199, 521–534.
- DEFRANCE J., PREMAT E. (2001). La forêt: un écran antibruit météorologique. *Scientific american*, dossier hors-série, juil./oct., 122-128.
- DENYS C., LECOMPTE E., GRANJON L., BAYLAC M., CORDEIRO P., et al. (2003). Integrative systematics: the importance of combining techniques for increasing knowledge of African Murinae. In: "Rats, mice and people: rodent biology and management". In: "Rats, mice and people: rodent biology and management". *Australian Centre for International Agricultural Research (ACIAR)*. Canberra: SINGLETON, G. R., HINDS, L. A., KREBS, C. J. and SPRATT, D. M. (eds). 499-506.
- DESDEVEISES Y., LEGENDRE P., AZOUZI L., MORAND S. (2003). Quantifying phylogenetically structured environmental variation. *Evolution*, 57, 2647-2652.
- DOBIGNY G., BEYLAC M., DENYS C. (2002). Geometric morphometrics, neural networks and diagnosis of sibling *Taterillus species* (Rodentia, Gerbillinae). *Biological Journal of the Linnean Society*, 77, 319–327.
- DRYDEN I.E., MARDIA K.V., editors (1998). *Statistical Shape Analysis*. London.
- DUJARDIN J.-P.A., KABA D., HENRY A.B. (2010). The exchangeability of shape. *BMC Research Notes*, 3, 2-66.
- DUPUY S., HERBRETEAU V., FEYFANT T., MORAND S., TRAN A. (2012). Land-cover dynamics in Southeast Asia: Contribution of object-oriented techniques for change detection. 4th International Conference on Geographic Object-Based Image Analysis (*GEOBIA*), Rio de Janeiro - Brazil, May 7-9 2012. p. 217.
- EDGAR R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32, 1792-1797.
- FABRE P.H., HAUTIER L., DIMITROV D., DOUZERY E.J. (2012). A glimpse on the pattern of rodent diversification: a phylogenetic approach. *BMC Evolutionary Biology*, 12, 88.
- FABRE P.H., PAGÈS M., MUSSER G.G., FITRIANA Y.S., FJELDSÅ J., et al. (2013). A new genus of rodent from Wallacea (Rodentia: Muridae: Murinae: Rattini), and its implication for biogeography and Indo-Pacific Rattini systematics. *Zoological Journal of the Linnean Society*, 169, 408-447.
- FAHN J.D. (2003). *A Land on Fire: The Environmental Consequences of the Southeast Asian Boom*; *Basic Books*, 384 p.
- FELSENSTEIN J. (1973). Maximum Likelihood and Minimum-Steps Methods for Estimating Evolutionary Trees from Data on Discrete Characters. *Systematic Biology*, 22.
- FELSENSTEIN J. (1985). Phylogenies and the Comparative Method. *American Naturalist*, 125, 1-15.
- FELSENSTEIN J. (1988). Phylogenies and quantitative characters *Annual Review of Ecology and Systematics*, 19, 445-471.
- FELSENSTEIN J. (2002). Quantitative characters, phylogenies, and morphometrics. In: "Morphology, Shape and Phylogeny" NORMAN MACLEOD P.L.F. (eds) *Systematics Association Special Volumes*. London: 27–44.
- FELSENSTEIN J. (2007). {PHYMLIP}: phylogenetic inference package, version 3.67.
- FINK W.L. (1990). Chapter 1: Data Acquisition for Morphometric Analysis in Systematic Biology. In: F.J. R., BOOKSTEIN F.L., editors, University of Michigan, 1990. pp. 9-19.
- FISHER R.A. (1934). Two New Properties of Mathematical Likelihood. *Proc R Soc Lond A*, 144.
- FRANCIS C.M., editor (2008). *a field guide to the mammals of South-East Asia*. *New Holand Publishers (UK) Ltd*. 392 p.
- GARDNER S., SIDISUNTHORN P., ANUSARNSUNTHORN V., editors (2000). *A field guide to forest trees of northern Thailand: Kobfai Publishing Project*. 545 p.
- GASCUEL, O. (1997). BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. *Molecular biology and evolution*, 14(7), 685-695.
- GASTON K.J., O'NEILL M.A. (2004). Automated species identification: why not? *Phil Trans R Soc Lond B*, 259.
- GAUBERT P., WOZENCRAFT W.C., CORDEIRO-ESTRELA P., VERON G. (2005). Mosaics of convergences and noise in morphological phylogenies: What's in a viverrid-like carnivoran ? *Systematic Biology*, 54, 865-894.
- GAUTHIER-CLERC M., THOMAS F. (2010). *Ecologie de la santé et Biodiversité: DeBoeck*. 522 p.

- GOLOBOFF P.A., MATTONI C.I., QUINTEROS A.S. (2006). Continuous characters analyzed as such. *Cladistics*, 22, 589-601.
- GONZÁLEZ-JOSÉ R., ESCAPA I., NEVES W.A., CÚNEO R., PUCCIARELLI H.M. (2008). Cladistic analysis of continuous modularized traits provides phylogenetic signals in Homo evolution. *Nature*, 453, 775-778.
- GOROG A.J., SINAGA M.H., ENGSTROM M.D. (2004). Vicariance or dispersal? Historical biogeography of three Sunda shelf murine rodents (*Maxomys surifer*, *Leopoldamys sabanus* and *Maxomys whiteheadi*). *Biological Journal of the Linnean Society*, 81, 91-109.
- GOUY R. (2013). Relations génétiques au sein du complexe d'espèces de Rats Noirs *Rattus rattus* (Mammifères, Rongeurs) en Asie du Sud-Est. *Université de Liège*. Liège. 82 p.
- GOWER J.C. (1975). Generalized Procrustes analysis. *Psychometrika*, 40, 33-51.
- HANSEN T.F. (1997). Stabilizing selection and the comparative analysis of adaptation. *Evolution*, 51, 1341-1351.
- HARVEY P.H., PAGEL M.D. (1991). The comparative method in evolutionary biology. *Oxford series in ecology and evolution*, 1.
- HARMON L.J. (2003). Tempo and Mode of Evolutionary Radiation in Iguanian Lizards. *Science*, 301, 961-964.
- HARMON L.J., GLOR R.E. (2010a). Poor statistical performance of the Mantel test in phylogenetic comparative analyses. *Evolution*, 64, pp. 2173-2178.
- HARMON L.J., LOSOS J.B., DAVIES J., GILLESPIE R.G., GITTLEMAN J.L., et al. (2010b). Early bursts of body size and shape evolution are rare in comparative data. *Evolution*, 64, 2385-2396.
- HARMON L.J., WEIR J., BROCK C., GLOR R., CHALLENGER W., et al. (2013). Documentation for package 'geiger' version 1.99-3. [www.cran.r-project.org/web/packages/geiger/geiger.pdf](http://www.cran.r-project.org/web/packages/geiger/geiger.pdf)
- HAUTIER L., LEBRUN R., COX P.G. (2012). Patterns of covariation in the masticatory apparatus of hystricognathous rodents: implications for evolution and diversification. *Journal of Morphology*, 273, pp. 1319-1337.
- HE K., JIANG X.L. (2013). Mitochondrial phylogeny reveals cryptic genetic diversity in the genus *Niviventer* (Rodentia, Muroidea). *Mitochondrial DNA*, 0, 1-8.
- HEBERT P.D.N., CYWINSKA A., BALL S., L., DEWAARD J., R. (2003a). Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences*, 270, 313-321.
- HEBERT P.D.N., RATNASINGHAM S., DEWAARD J.R. (2003b). Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society B: Biological Sciences*, 270, S96-S99.
- HEBERT P.D.N., GREGORY T.R. (2005). The Promise of DNA Barcoding for Taxonomy. *Systematic Biology*, 54, pp. 852 - 859.
- HERBRETEAU V. (2007). Géographie de zoonoses en Thaïlande: de la distribution des rongeurs, vecteurs et hôtes, au risque de transmission. Université de Nanterre - Paris X.
- HERBRETEAU V., RERKAMNUAYCHOKE W., JITTAPALAPONG S., CHAVAL Y., COSSON J.-F., MORAND S. (2011). Protocols for field and laboratory rodent studies. Bangkok: Kasetart University Press. 46 p.
- HOTELLING H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24, 498-520.
- HURVICH C.M., TSAI C.-L. (1989). Regression and time series model selection in small samples. *Biometrika*, 76, 297-307.
- JING M., HON-TSEN Y., SHENG-HAI W., WANG W., ZHENG X. (2007). Phylogenetic relationships in genus *Niviventer* (Rodentia: Muridae) in China inferred from complete mitochondrial cytochrome b gene. *Molecular Phylogenetics and Evolution*, vol. 44, 521-529.
- JOLICOEUR P., MOSIMANN J.E. (1960). Size and shape variation in the painted turtle. A principal component analysis. *Growth*, 24, 339-354.
- JONES K.E., PATEL N.G., LEVY M.A., STOREYGARD A., BALK D., et al. (2008). Global trends in emerging infectious diseases. *Nature*, 451, 990-993.
- KENDALL D.G. (1984). Shape manifolds, Procrustean métrics and complex projective spaces. *Bulletin of London Mathematical Society*, 16, 81-121.
- KENDALL D.G. (1989). A Survey of the Statistical Theory of Shape. *Statistical Science*, 4, 87-120.
- KIMURA M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, 16, 111-120.
- KLINGENBERG C.P., MC INTYRE G.S. (1998). Geometric morphometrics of developmental instability: analyzing patterns of fluctuating asymmetry with procrustes methods. *Evolution*, 54, 1363-1375.
- KLINGENBERG C.P. (2010). Evolution and development of shape: integrating quantitative approaches. *Nature Reviews Genetics*, 11. 623-635.

- KLINGENBERG C.P., GIDASZEWSKI N.A. (2010). Testing and quantifying phylogenetic signals and homoplasy in morphometric data. *Systematic Biology*, 59, 245-261.
- KLINGENBERG C.P., DUTTKE S., WHELAN S., KIM M. (2011). Developmental plasticity, morphological variation and evolvability: a multilevel analysis of morphometric integration in the shape of compound leaves. *Journal of Evolutionary Biology*, 25, 115-129.
- KLINGENBERG C.P., MARUGÁN-LOBÓN J. (2013). Evolutionary covariation in geometric morphometric data: analyzing integration, modularity and allometry in a phylogenetic context. *Systematic Biology*, 62, 591-610.
- KRASNOV B.R., POULIN R., MOUILLOT D. (2011). Scale-dependence of phylogenetic signal in ecological traits of ectoparasites. *Ecography*, 34.
- KRESS W.J., WURDACK K.J., ZIMMER E.A., WEIGT L.A., JANZEN D.H. (2005). Use of DNA barcodes to identify flowering plants. *Proceedings of the National Academy of Sciences of the United States of America*, 102, 8369-8374.
- KRESS W.J., ERICKSON D.L. (2008). DNA barcodes: Genes, genomics, and bioinformatics. *Proceedings of the National Academy of Sciences of the United States of America*, 105.
- LACK J.B., GREENE D.U., CONROY C.J., HAMILTON M.J., BRAUN J.K., et al. (2012). Invasion facilitates hybridization with introgression in the *Rattus rattus* species complex. *Molecular Ecology*, 21(14), 3545-3561.
- LAPOINTE F.-J., GARLAND T.J. (2001). A generalized permutation model for the analysis of cross-species data. *Journal of Classification*, 18, 109-127.
- LATINNE A., WAENGSOOTHORN S., HERBRETEAU V., MICHAUX J.R. (2011a). Evidence of complex phylogeographic structure for the threatened rodent *Leopoldamys neilli*, in Southeast Asia. *Conservation Genetics*, 12, 1495-1511.
- LATINNE A., WAENGSOOTHORN S., HERBRETEAU V., MICHAUX J.R. (2011b). Thai limestone karsts: an impending biodiversity crisis, In 'The 1st Environment Asia International Conference on Environmental Supporting in Food and Energy Security: Crisis and Opportunity' pp. 176-187.
- LATINNE A., WAENGSOOTHORN S., RISTERUCCI A.M., MICHAUX J.R. (2011c). Isolation, characterization and PCR multiplexing of polymorphic microsatellite markers in the threatened murine rodent, *Leopoldamys neilli*. *Conservation Genetics Resources*, 3, 511-513.
- LATINNE A. (2012a). Diversity and endemism of Murinae rodents in Thai limestone karsts: a genetic approach. *Université de Liège. Liège*. 256 p.
- LATINNE A., WAENGSOOTHORN S., ROJANADILOK P., EIAMAMPAI K., SRIBUAROD K., et al. (2012b). Combined Mitochondrial and Nuclear Markers Revealed a Deep Vicariant History for *Leopoldamys neilli*, a Cave-Dwelling Rodent of Thailand. *PLoS One*, 7, e47670.
- LATINNE A., CHAVAL Y., WAENGSOOTHORN S., ROJANADILOK P., EIAMAMPAI K., et al. (2013a). Is *Leopoldamys neilli* (Rodentia, Muridae) a synonym of *Leopoldamys herberti*? A reply to Balakirev et al. (2013). *Zootaxa*, 3731, 589-598.
- LATINNE A., WAENGSOOTHORN S., ROJANADILOK P., EIAMAMPAI K., SRIBUAROD K., et al. (2013b). Diversity and endemism of Murinae rodents in Thai limestone karsts. *Systematics and Biodiversity*, 11, 323-344.
- LECOINTRE G., LE GUYADER H. (2009). Classification phylogénétique du vivant. 3 ème edition, illustrations VISSET D., BELIN, 559 p.
- LECOMPTE E., DENYS C., GRANJON L., VOLOBOUEV V. (2003). Integrative systematics: contributions to *Mastomys* phylogeny and evolution. In: "Rats, mice and people: rodent biology and management". *Australian Centre for International Agricultural Research (ACIAR)*. Canberra: SINGLETON, G. R., HINDS, L. A., KREBS, C. J. and SPRATT, D. M. (eds). 536-540.
- LECOMPTE E., DENYS C., GRANJON L. (2005a). Confrontation of morphological and molecular data: The *Praomys* group (Rodentia, Murinae) as a case of adaptive convergences and morphological stasis. *Mol Phylogenet Evol*, 37, 899 - 919.
- LECOMPTE E., BROUAT C., DUPLANTIER J.-M., MAXIME G., GRANJON L., et al. (2005b). Molecular identification of four cryptic species of *Mastomys* (Rodentia, Murinae). *Biochemical Systematics and Ecology*, 33, 681-689.
- LECOMPTE E., APLIN K., DENYS C., CATZEFLIS F., CHADES M., et al. (2008). Phylogeny and biogeography of African Murinae based on mitochondrial and nuclear gene sequences, with a new tribal classification of the subfamily. *BMC Evolutionary Biology*, 8, 199.
- LEKAGUL B., MCNEELY J.A., (1977). Mammals of Thailand. *Kurusapha Ladprao: the Association for the Conservation of Wildlife*. 758 p.
- LOHMAN D.J., DE BRUYN M., PAGE T., VON RINTELEN K., HALL R.N., P. K., et al. (2011). Biogeography of the Indo-Australian archipelago. *Annual Review of Ecology, Evolution, and Systematics*, 42, 205-226

- LORIUS C., CARPENTIER L. (2011). Voyage dans l'Anthropocène: Cette nouvelle ère dont nous sommes les héros: *Actes Sud Sciences Humaines*. 200 p.
- LU L., CHESTERS D., ZHANG W., LI G., MA Y., et al. (2012). Small Mammal Investigation in Spotted Fever Focus with DNA-Barcoding and Taxonomic Implications on Rodents Species from Hainan of China. *PLoS One*, 7, e43479.
- LUNDE D.P., SON N.T. (2001). An Identification Guide to the Rodents of Vietnam. *Center for Biodiversity and Conservation at the American Museum of Natural History*. New York. 80 p.
- LUNDRIGAN B.L., JANSA S.A., TUCKER P.K. (2002). Phylogenetic Relationships in the Genus *Mus*, Based on Paternally, Maternally, and Biparentally Inherited Characters. *Systematic Biology*, 51, 410 - 431.
- MACLEOD N. (2002). phylogenetic signals in morphometric data. In: "Morphology, Shape and Phylogeny" NORMAN MACLEOD P.L.F. (eds) *Systematics Association Special Volumes*. London.100-138.
- MACLEOD N., O'NEILL M., WALSH A.S. (2007). Automated tools for the identification of taxa from morphological data: face recognition in wasps. In: MACLEOD N., editor. Automated taxon identification in systematics: theory, approaches, and applications. London: Taylor and Francis. pp. 153-188.
- MADDISON W.P. (1991). Square-change parsimony reconstructions of ancestral states fro continuous-valued characters on a phylogenetic tree. *Systematic Zoology*, 40, 304-314.
- MARCUS L.F. (1990). Chapter 4: Traditional Morphometrics. In: F.J. R., BOOKSTEIN F.L., editors, University of Michigan, 1990. pp. 77-122.
- MARSHALL J.T. (1977a). Family Muridae: Rats and mice In: "Mammals of Thailand" *Kurusapha Ladprao: the Association for the Conservation of Wildlife*, LEKAGUL B., MCNEELY J.A. (eds). Bangkok. pp. 396-487.
- MARSHALL J.T. (1977b). A synopsis of Asian species of *Mus* (Rodentia, Muridae). Bulletin of the AMNH, 158.
- MARTINS E.P., HANSEN T.F. (1997). Phylogenies and the comparative method: A general approach to incorporating phylogenetic information into the anlysis of interspecific. *American Naturalist*, 149, 646-667 ERRATUM *Am. Nat.* 153:448.
- MARTINS E.P. (2000). Adaptation and the comparative method. *Trends in Ecology & Evolution*, 15, 296-299.
- MAYR E. (1942). Systematics and the Origin of Species, from the Viewpoint of a Zoologist.; Cambridge: *Harvard University Press* n°13.
- MAYR E., editor (1963). Animal Species and Evolution. Cambridge.
- MAYR E., ASHLOCK P.D., (1969). Principles of systematic zoology. second edition ed. New York.
- MEERBURG B., SINGLETON G.R., KIJLSTRA A. (2009). Rodent-borne diseases and their risks for public health. *Crit Rev Microbiol*, 35.
- MEYER C.P., PAULAY G. (2005). DNA Barcoding: Error Rates Based on Comprehensive Sampling. *PLoS Biology*, 3, e422.
- MICHAUX J., CHEVRET P., RENAUD S. (2007). Morphological diversity of Old World rats and mice (Rodentia, Muridae) mandible in relation with phylogeny and adaptation. *J Zool Syst Evol Research*, 45, 263 - 279.
- MITTEROECKER P., GUNZ P., BERNHARD M., SCHAEFER K., BOOKSTEIN F.L. (2004). Comparison of cranial ontogenetic trajectories among great apes and humans. *Journal of Human Evolution*, 46, 679–698.
- MITTEROECKER P., GUNZ P., WINDHAGER S., SCHAEFER K. (2013). A brief review of shape, form, and allometry in geometric morphometrics, with applications to human facial morphology. *Hystrix, the Italian Journal of Mammalogy*, 24, 59-66.
- MONTEIRO L.R., DUARTE L.C., DOS REIS S.E.F. (2003). Environmental correlates of geographical variation in skull and mandible shape of the punar ´e rat *Thrichomys apereoides* (Rodentia: Echimyidae). *Journal of The Zoological Society of London*, 261, 47-57.
- MONTEIRO L.R. (2013). Morphometrics and the comparative method: studying the evolution of biological shape. *Hystrix, the Italian Journal of Mammalogy*, 24, 8.
- MORITZ C., DOWLING T.E., BROWN W.M. (1987). Evolution of Animal Mitochondrial DNA: Relevance for Population Biology and Systematics. *Annual Review of Ecology and Systematics*, 18, 269-292.
- MORITZ C., CICERO C. (2005). Promises and pitfalls on DNA barcoding. *Genes & Genetic Systems*, 80, 443-443.
- MOSIMANN J.E. (1970). Size allometry: size and shape variables with characterizations of the lognormal and generalized gamma distributions. *Journal of the American Statistical Association*, 65, 930–948.
- MOSTERT M.E. (2009). Molecular and morphological assessment of invasive, inland *Rattus* (Rodentia: Muridae) congeners in South Africa and their reservoir host potential with respect to *Helicobacter* and *Bartonella*. *University of Pretoria*. Pretoria.160 p.

- MÜNKEMÜLLER T., LAVERGNE S., BZEZNIK B., DRAY S., JOMBART T., et al. (2012). How to measure and test phylogenetic signal. *Methods in Ecology and Evolution*, 3, 743-756.
- MUSSER G.G. (1970). Species-limits of *Rattus brahma*, a murid rodent of northeastern India and northern Burma. *American Museum Novitates*, 2406, 1-27.
- MUSSER G.G. (1973). Species-limits of *Rattus cremoriventer* and *Rattus langbianis*, murid rodents of Southeast Asia and the Greater Sunda Islands. *American Museum novitates. American Museum Novitates*, 2525.
- MUSSER G.G., MARSHALL J.T., BOEADI J. (1979). Definition and Contents of the Sundaic Genus *Maxomys* (Rodentia, Muridae). *Journal of Mammalogy*, 60, 592-606.
- MUSSER G.G. (1981). Results of the Archbold expeditions, no.105. Notes on systematics of Indo-Malayan murid rodents, and descriptions of new genera and species from Ceylon, Sulawesi, and the Philippines. *Bulletin of the American Museum of Natural History*, 168, 225-334.
- MUSSER G.G., NEWCOMB C. (1983). Malaysian murids and the giant rat of Sumatra. *Bulletin of the American Museum of Natural History*, 174, 327-598.
- MUSSER G.G., HOLDEN M.E. (1991). Sulawesi rodents (muridae, murinae)-morphological and geographical boundaries of species in the *rattus-hoffmanni* group and a new species from pulau-peleng. *Bulletin of the American Museum of Natural History*, 206, 322-413.
- MUSSER G.G., CARLETON M.D. (1993). Family Muridae. . In: "Mammal Species of the World: a taxonomic and geographic reference. WILSON DE, REEDER, DM. (eds).2nd ed". Washington, D.C. *Smithsonian Institution Press*. 501-755.
- MUSSER G.G., BROTHERS E.M. (1994). Identification of bandicoot rats from Thailand (Bandicota, Muridae, Rodentia). *American Museum Novitates*, 3110.
- MUSSER G.G., CARLETON M.D. (2005). Superfamily Muroidea. In: "Mammal species of the world: A taxonomic and geographic reference" WILSON DE, REEDER, DM. (eds).3rd edition. Johns Hopkins University. Baltimore. 894-1531.
- MUSSER G.G., SMITH A.L., ROBINSON M.F., LUNDE D.P. (2005b). Description of a New Genus and Species of Rodent (Murinae, Muridae, Rodentia) from the Khammouan Limestone National Biodiversity Conservation Area in Lao PDR. *American Museum Novitates*, 3497, pp. 1-31.
- MUSSER G.G., LUNDE D.P., SON N.T. (2006). Description of a New Genus and Species of Rodent (Murinae, Muridae, Rodentia) from the Tower Karst Region of Northeastern Vietnam. *American Museum Novitates*, 3517, pp. 1-41.
- MYERS N., MITTERMEIER R.A., MITTERMEIER C.G., DA FONSECA G.A.B., KENT J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, 403, 853-858.
- NAVARRO N., ZATARAIN X., MONTUIRE S. (2004). Effects of morphometric descriptor changes on statistical classification and morphospaces. *Biological Journal of the Linnean Society*, 83, 243-260.
- NUNES S. (2007). Dispersal and Philopatry. In: "Rodent Societies: An Ecological and Evolutionary Perspective" JERRY, O. WOLFF, P.W.S (eds). *University of Chicago*. Chicago. 150-162.
- PAGEL M. (1994). Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. *Proceedings of the Royal Society (B)*, 255 37-45.
- PAGEL M. (1997). Inferring evolutionary processes from phylogenies. *Zoologica Scripta*, 26, 331-348
- PAGEL M. (1999). Inferring the historical patterns of biological evolution. *Nature*, 401.
- PAGEL M. (2002). User's Manual for Continuous. *School of Animal and Microbial Sciences*.
- PAGÈS M., CHAVAL Y., HERBRETEAU V., WAENGSTHORN S., COSSON J.-F., et al. (2010). Revisiting the taxonomy of the Rattini tribe: a phylogeny-based delimitation of species boundaries. *BMC Evolutionary Biology*, 10, 184.
- PAGÈS M., LATINNE A., JOHAN M. (2011). Inter- and Intraspecific Genetic Biodiversity in South East Asian Rodents: New Insights for Their Conservation. In: "Biodiversity Hotspots" ZACHOS F.E., HABEL, J. C. (eds). *Springer*. Heidelberg: 363-382.
- PAGÈS M., BAZIN E., GALAN M., CHAVAL Y., CLAUDE J., et al. (2013). Cytonuclear discordance among Southeast Asian black rats (*Rattus rattus* complex). *Mol Ecol*, 22, 1019-1034.
- PALANISWAMY S., THACKER N.A., KLINGENBERG C.P. (2007). Automatic Identification of Morphometric Landmarks in Digital Images.: Imaging Science and Biomedical Engineering Division, Medical School, University of Manchester, Stopford Building, Oxford Road, Manchester, M13 9PT.
- PALMER A.R. (1994). Fluctuating asymmetry analyses: A primer. In: "Developmental Instability: Its Origins and Evolutionary Implications" MARKOW T.A., editor. Dordrecht, Netherlands.: Kluwer. pp. 335-364.
- PETTER F. (1961). Répartition géographique et écologie des rongeurs désertiques (du Sahara occidental à l'Iran oriental). *MAMMIFÈRES. S.P.L.É.E.L.P.D.* Paris. 167 p.
- PEARSON (1901). On Lines & Planes of Closest Fit to Systems of Points in Space. *Philosophical magazine*, 2, 557-572.

- PIMENTEL R.A. (1979). Morphometrics, the multivariate analysis of biological data. Dubuque, Iowa, USA: Kendall/Hunt Publishing Company. p. 276 p.
- PIMENTEL R.A., RIGGINS R. (1987). The nature of cladistic data. *Cladistics*, 201-209.
- POLLY P.D., LAWING A.M., FABRE A.C., GOSWAMI A. (2013). Phylogenetic principal components analysis and geometric morphometrics. *Hystrix, the Italian Journal of Mammalogy*, 24, 9. 33-41
- PURVIS A., NEE S., HARVEY P.H. (1995). Macroevolutionary inferences from primate phylogeny. *Proceedings of the Royal Society of London Series B: Biological Sciences*, 260, 329-333.
- QUÉRÉ J.P., LE LOUARN H. (2011). Les rongeurs de France Faunistique et biologie. *QUAE*, Versailles. 310 p.
- R CORE TEAM (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, [URL http://www.R-project.org/](http://www.R-project.org/).
- RAE T.C. (1998). The logical basis for the use of continuous characters in phylogenetic systematics. *Cladistics*, 14, 221-228.
- RABINOVICH S.G. (2005). Measurement errors and Uncertainties: Theory and Practice. New York: Springer. 308 p.
- RASMUSSEN P., WHEELER W., MOSER T., VINE L., SULLIVAN B., et al. (2001). Measurements of Canada goose morphology - Sources of error and effects on classification of subspecies. *Journal of Wildlife Management*, 65, 716-725.
- RENAUD S., MICHAUX J., MEIN P., AGUILAR J.P., AUFFRAY J.C. (1999). Patterns of size and shape differentiation during the evolutionary radiation of the European Miocene Murine rodents. *Lethaia*, 32, 2.p. 100.
- RENAUD S. (2005). First upper molar and mandible shape of wood mice (*Apodemus sylvaticus*) from northern Germany: ageing, habitat, and insularity. *Mammalian Biology* 70, 157-170.
- REVELL L.J., HARMON L.J., COLLAR D.C. (2008). Phylogenetic signal, evolutionary process, and rate. *Systematic Biology*, 57, 591-601.
- RICKART E.A., MUSSER G.G. (1993). Philippine rodents: chromosomal characteristics and their significance for phylogenetic inference among 13 species (Rodentia, Muridae, Murinae). *American Museum Novitates*, 3064, 34 p.
- ROBINS J.H., HINGSTON M., MATISOO-SMITH E., ROSS H.A. (2007). Identifying *Rattus* species using mitochondrial DNA. *Molecular Ecology Notes*, 7, 717-729.
- ROBINS J.H., MCLENACHAN P.A., PHILLIPS M.J., CRAIG L., ROSS H.A., et al. (2008). Dating of divergences within the *Rattus* genus phylogeny using whole mitochondrial genomes. *Molecular Phylogenetics and Evolution*, 49, 460-466.
- ROHLF F.J. (1990). Chapter 3: An overview of image Preprocessing and Analysis Techniques for Morphometrics. In: F.J. R., BOOKSTEIN F.L., editors, University of Michigan, 1990. pp. 37-60.
- ROHLF F.J., SLICE D. (1990). Extensions of the Procrustes method for the optimal superimposition of landmarks. *Systematic Zoology*, 39, 40-59.
- ROHLF F.J. (2002). Geometric morphometrics and phylogeny. In: "Morphology, shape, and phylogeny" MACLEOD N F.P. (eds). London: *Taylor & Francis*. pp. 175-193.
- ROHLF F.J. (2013). TPSdig2. Version 2.17.
- ROWE K.C., APLIN K.P., BAVERSTOCK P.R., MORITZ C. (2011). Recent and rapid speciation with limited morphological disparity in the genus *Rattus*. *Systematic Biology*, 60, 188-203.
- SANDERSON M.J. (1997). A nonparametric approach to estimating divergence times in the absence of rate constancy. *Molecular biology and evolution*, 14, 1218-1231.
- SCHENK J.J., ROWE K.C., STEPPAN S.J. (2013). Ecological Opportunity and Incumbency in the Diversification of Repeated Continental Colonizations by Murine Rodents. *Systematic Biology*, 62, 837-864.
- SCHLEICH C.E., BUSCH C. (2004). Functional morphology of the middle ear of *Ctenomys talarum* (RODENTIA: OCTODONTIDAE). *Journal of Mammalogy*, 85, 290-295.
- SCHLUTER D., PRICE T., MOOERS A.O., LUDWIG D. (1997). Likelihood of ancestral states in adaptive radiation. *Evolution*, 51, 1699-1711.
- SCHLUTER D., (2000). The Ecology of Adaptive Radiations. *Oxford University Press*. 288 p.
- SCHWARZ E. (1960). Classification, Origin and Distribution of Commensal Rats. *Bulletin of the World Health Organization*, 23, 411-416.
- SEBERG O., PETERSEN G. (2009). How Many Loci Does it Take to DNA Barcode a Crocus? *PLoS ONE*, 4.
- SIAHSARVIE R., AUFFRAY J.-C., DARVISH J., RAJABI-MAHAM H., YU H.-T., et al. (2012). Patterns of morphological evolution in the mandible of the house mouse *Mus musculus* (Rodentia: Muridae). *Biological Journal of the Linnean Society*, 105, 635-647.
- SMITH A.T., XIE Y. (2008). A Guide to the Mammals of China. *Princeton University Press*. 576 p.
- SMITH U.E., HENDRICKS J.R. (2013). Geometric Morphometric Character Suites as Phylogenetic Data: Extracting Phylogenetic Signal from Gastropod Shells. *Systematic biology*, 62, 366-385.

- SMOUSE P., LONG J., SOKAL R. (1986). Multiple Regression and Correlation Extensions of the Mantel Test of Matrix Correspondence. *Systematic Zoology*, 35, 625-632
- SODHI N.S., KOHA L.P., BROOK B.W., NG P.K.L. (2004). Southeast Asian biodiversity: an impending disaster. *Trends in Ecology and Evolution*, 19, 654-660.
- SUZUKI H., SHIMADA T., TERASHIMA M., TSUCHIYA K., APLIN K. (2004). Temporal, spatial, and ecological modes of evolution of Eurasian *Mus* based on mitochondrial and nuclear gene sequences. *Molecular Phylogenetics and Evolution*, 33, 626-646.
- SUZUKI H., APLIN K.P. (2012). Phylogeny and biogeography of the genus *Mus* in Eurasia. In: "Evolution of the house mouse". BAIRD S.J., MACHOLAN M. (eds). *Cambridge University Press*. Cambridge. 35-64.
- SUZUKI H., NUNOME M., KINOSHITA G., APLIN K.P., VOGEL P., KRYUKOV A.P., JIN M.L., HAN S.H., MARYANTO I., TSUCHIYA K., IKEDA H., SHIROISHI T., YONEKAWA H., MORIWAKI K. (2013). Evolutionary and dispersal history of Eurasian house mice *Mus musculus* clarified by more extensive geographic sampling of mitochondrial DNA. *Heredity*, 111, pp. 375-390.
- THOMAS F., GUÉGUAN J.F., RENAUD F. (2009). Introduction à l'épidémiologie intégrative des maladies infectieuses et parasitaires. 2ed edition, *De Boeck*, IRD, 522 p.
- THOMPSON A.W. (1917). On Growth and Form.
- TOFILSKI A. (2007). Automatic Measurement of Honeybee Wings. In: "Automated taxon identification in systematics: theory, approaches, and applications" LEOD. N.M., editor. London: Taylor and Francis. pp. 289-296.
- USDIN K., CHEVRET P., CATZEFLIS F.M., VERONA R., FURANO A.V. (1995). L1 (LINE-1) retrotransposable elements provide a "fossil" record of the phylogenetic history of murid rodents. *Molecular biology and evolution*, 12, 73-82.
- USHER A.D. (2009). Thai Forestry: A Critical History; Silkworm Books., 248 p.
- VERNEAU O., CATZEFLIS F., FURANO A.V. (1997). Determination of the Evolutionary Relationships in *Rattus sensu lato* (Rodentia : Muridae) Using L1 (LINE-1) Amplification Events. *Journal of Molecular Evolution*, 45, 424-436.
- VERNEAU O., CATZEFLIS F., FURANO A.V. (1998). Determining and dating recent rodent speciation events by using L1 (LINE-1) retrotransposons. *PNAS*, 95, 11284-11289.
- VIDYA T.N.C., SUKUMAR R., MELNICK D.J. (2009). Range-wide mtDNA phylogeography yields insights into the origins of Asian elephants. *Proceedings of the Royal Society B: Biological Sciences*, 276, 893-902.
- WESTOBY M., LEISHMAN M.R., LORD J.M. (1995). On misinterpreting the 'phylogenetic correction'. *Journal of Ecology*, 83., 531-534.
- WIENS J.J., GRAHAM C.H. (2005). Niche conservatism: integrating evolution, ecology, and conservation biology. *Annual Review of Ecology, Evolution, and Systematics* 519-539.
- WIENS J.J., ACKERLY D.D., ALLEN A.P., ANACKER B.L., BUCKLEY L.B., et al. (2010). Niche conservatism as an emerging principle in ecology and conservation biology. *Ecology letters*, 13, 1310-1324.
- WILL K.W., MISHLER B.D., WHEELER Q.D. (2005). The perils of DNA barcoding and the need for integrative taxonomy. *Syst Biol*, 54, 844-851.
- WILSON D.E., REEDER D.M. (2005). Mammal Species of the World. A Taxonomic and Geographic Reference (3rd ed). *Johns Hopkins University Press*. 2142 p.
- WOOD B. (2001). Rat control in oil palms and rice fields. *pesticide outlook*.12(1). 71-74.
- WOODRUFF D.S. (2010). Biogeography and conservation in Southeast Asia: how 2.7 million years of repeated environmental fluctuations affect today's patterns and the future of the remaining refugial-phase biodiversity. *Biodiversity and Conservation*, 19, 919-941.
- WOOLHOUSE M.E. (2002). Population biology of emerging and re-emerging pathogens. *Trends Microbiology*, 10, S3-7.
- YEZERINAC S.M., LOOGHEED S.C., HANDFORD P. (1992). Measurement error and morphometric studies: statistical power and observer experience. *Systematic Biology*, 41, 471-482.
- YONG H.-S. (1969). Karyotypes of Malayan rats (Rodentia-Muridae, genus *Rattus* Fischer). *Chromosoma*, 27, 245-267.
- YU H.T., FANG Y.P., CHOU C.W., HUANG S.W., YEW F.H. (1996). Chromosomal evolution in three species of murid rodents of Taiwan. *Zoological Studies*, 35, 195-199.



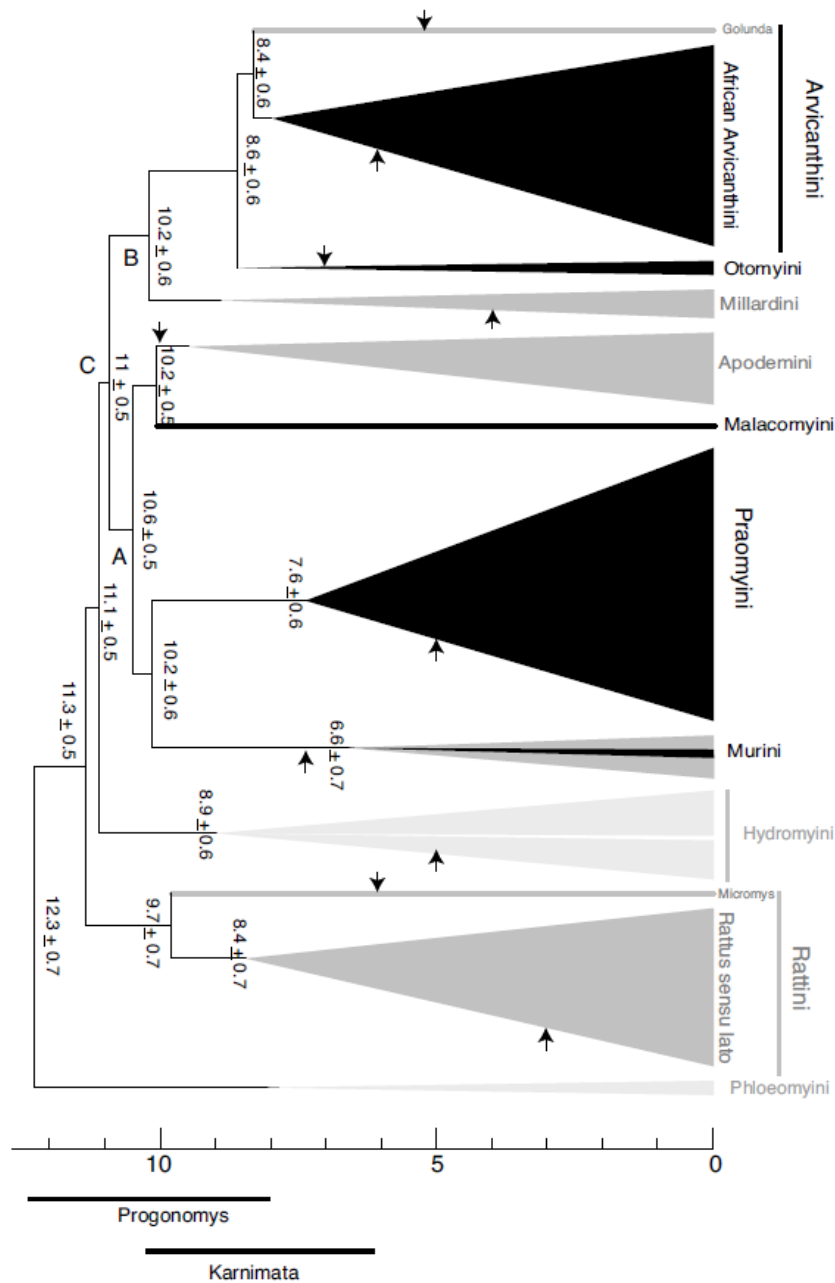


## ***Annexe I – Systématique utilisée dans ce mémoire et biogéographie des Rattini et des Murini en Asie du Sud-Est.***

### ***1- Etat des lieux des connaissances sur la systématique des rongeurs étudiés***

Les rongeurs étudiés dans le cadre des projets Roboviroses à Hantavirus et CERoPath et par voie de conséquence dans le présent travail sont regroupés au sein de deux tribus de la sous-famille des murinés, la tribu des Rattini et celle des Murini telles qu'elles ont été définies lors de l'étude phylogénétique moléculaire réalisée par Lecompte et al. (2008). Les relations phylogénétiques de ces deux tribus au sein de la sous-famille des murinés sont présentées à la figure 1 (au verso). Ces deux tribus hébergent chacune au moins un modèle biologique important en écologie et pour la recherche médicale (*Rattus norvegicus* Berkenhout, 1769 pour les Rattini et *Mus musculus* Linnaeus, 1758 pour les Murini). Toutefois, les relations phylogénétiques et la systématique au sein de ces deux tribus demeurent très parcellaires. Ainsi, la récente étude de l'histoire évolutive du génome des Rattini et la reconstruction du génome ancestral (Badenhorst et al., 2011) montrent que les Rattini ont un taux d'évolution de leur génome globalement plus lent que les Murini et que ce taux alterne des périodes d'accélération suivies de périodes de stases. Sa structure est par ailleurs plus proche de celle de l'humain que celui des Murini, ce qui souligne l'intérêt du modèle *Rattus* pour des études de génomique comparative (Badenhorst et al., 2011). Plusieurs ouvrages généraux constituent de précieuses sources de connaissances pour l'identification des rongeurs en Asie du Sud-Est. Concernant la zone étudiée, l'un des ouvrages majeurs a été publié par Boonsong et McNeely (1977). Ce travail comprend une partie traitant des rongeurs, très détaillée et réalisée par J.T. Marshall (1977a). Elle porte sur 39 espèces de Muroides dont 23 espèces sont rattachées au genre *Rattus*. La morphologie externe et crânienne ainsi que l'étude cytogénétique des espèces sont explorés dans cet ouvrage pour la définition des espèces. D'autres travaux généraux sont pertinents pour l'identification des rongeurs de la zone étudiée. Corbet et Hill (1992) offrent une revue systématique approfondie des mammifères de la zone indo-malayenne ; Lunde et Nguyen Truong Son (2001) un guide d'identification des rongeurs du Vietnam ; Aplin (2003a) un guide pour l'étude des rongeurs en Asie et indo-pacifique ; Francis (2008) un guide de terrain des mammifères d'Asie du Sud-Est ; Smith et Xie (2008) un guide de terrain des mammifères de Chine. Finalement Wilson et Reeder (2005) est l'ouvrage de référence pour la systématique des mammifères. Toutefois, de nouvelles espèces et de nouveaux genres au sein de ces deux tribus sont régulièrement décrites en Asie du Sud-Est, tant et si bien que les ouvrages de référence sont rapidement incomplets voir obsolètes. Ainsi, dans la zone étudiée, la tribu des Murini s'est enrichie récemment d'un taxon, *Mus fagilicauda* (Auffray et al., 2003) et *Mus pahari* Thomas, 1916 est aujourd'hui regardé comme un complexe d'espèces (Suzuki and Aplin, 2012). Au sein de la tribu des Rattini, les populations du Sud-Est asiatique de *Micromys minutus* Pallas, 1771 ont été élevées au rang spécifique et rattachées à *M. erythrotis* (Blyth, 1856) par Abramov et al. (2009). Une partie des populations de *Rattus losea* Swinhoe, 1870 ont été également élevées au rang d'espèce et nommées *Rattus sakeratensis* Gyldenstolpe, 1917 par Aplin et al. (2011). D'autre part, la sous-structuration géographique de clades génétiques décrivant certaines espèces comme *Leopoldamys neilli* J. T. Marshall, Jr., 1976, *Maxomys surifer* Miller, 1900 et *Berylmys bowersi* Anderson, 1879 invitent à des études taxonomiques plus approfondies pour juger du rang taxonomiques à attribuer à certaines lignées (Pagès et al., 2010 ; 2011 ; Latinne, 2012a).

De plus, de nouveaux genres sont encore régulièrement décrits, comme *Saxatilomys* par Musser et al. (2005) ou



**Figure 1 :** Chronogramme simplifié des principales tribus au sein de la sous-famille des murinés réalisé à partir des gènes *Cytb*, *IRBP* exon I et *growth hormone receptor* (*GHR*). Pour chacune d'elle le plus vieux fossile est indiqué par une flèche noire. Les taxons africains sont représentés en noir, les australiens en gris clair et les eurasiens en gris foncé. (Extrait de Lecompte et al., 2008).

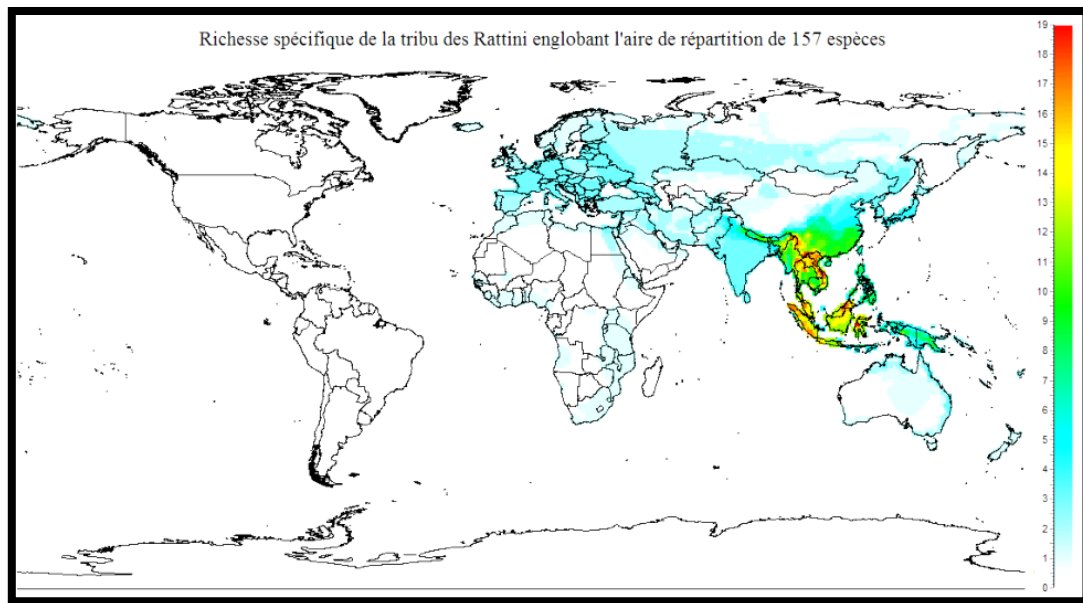
*Tonkinomys* (Musser et al., 2006) tous deux de la Province de Khammouane au Laos et plus récemment *Halmaheramys* (Fabre et al., 2013) sur l'île indonésienne d'Halmahera dans l'archipel des Moluques. Enfin, nombre de taxons n'ont pas de positionnement phylogénétique solide et sont placés dans la catégorie taxonomique *incertae sedis* (Lecompte et al., 2008) ou font l'objet de groupements artificiels (Musser et Newcomb, 1983 ; Musser et Carleton, 2005). Les paragraphes ci-après font donc la synthèse de l'état des lieux de l'ensemble des connaissances disponibles à ce jour sur la systématique et les données d'identification des représentants de ces deux tribus.

### 1.1- Systématique de la tribu des Rattini pour la zone étudiée

La tribu des Rattini comprend 35 genres correspondant à 167 espèces murines (Musser et Carleton, 2005). La diversité spécifique de la tribu est maximale dans le Sud-Est asiatique (voir Figure 2 au verso) Les relations phylogénétiques des taxons indo-chinois de la tribu sont présentées figure 3. En 1977, Marshall (1977a) ne distingue pour la Thaïlande que les genres *Rattus* Fischer, 1803 et *Bandicota* Gray, 1873 au sein de la tribu des Rattini considérant les genres *Leopoldamys* Ellerman, 1947-1948, *Maxomys* Sody, 1936, *Berylmys* Ellerman, 1947, *Sundamys* Musser and Newcomb, 1983 et *Niviventer* J. T. Marshall, Jr., 1976 comme des sous genres du genre *Rattus*. Les travaux de Musser et collaborateurs (voir les paragraphes 1.1.1 à 1.1.6) ont fait évoluer cette systématique en définissant les limites d'espèces et de genres sur des bases principalement morphométriques. Les **Divisions** *Maxomys*, *Dacnomys*, *Rattus* ont été formalisées (Musser et Carleton, 2005) d'après leurs affinités phylogénétiques estimées d'après la présence ou l'absence d'éléments transposables de type LINE-1 (Verneau et al., 1997, 1998). De nombreux travaux et notamment ceux réalisés durant nos programmes de recherche en cytogénétiques (Badenhorst et al., 2009, 2011, 2012) ou alliant données moléculaires et morphologiques (Aplin et al. 2011 ; Pagès et al, 2010, 2011, 2013 ; Latinne et al., 2011c ; Latinne, 2012a ; Latinne et al., 2012b ; 2013a ; 2013b ; Fabre et al., 2012 ; 2013 ; Balakirev et al., 2010 ; 2011 ; 2012 ; 2013) ont considérablement fait évoluer la systématique de la tribu, et soulevé de nombreuses questions toujours en suspens. Le tableau 1 permet de faire la correspondance entre les clades de la figure 3 et les noms utilisés dans chaque partie de ce mémoire.

#### 1.1.1- Le genre *Maxomys* Sody, 1936, Division *Maxomys* (Musser et Carleton, 2005)

Musser et collaborateurs (1979) distinguent le sous-genre *Maxomys* (Marshall, 1977a) du genre *Rattus* et le caractérisent sur la base de critères cytogénétiques et morphologiques. Des études ultérieures viennent confirmer la validité de ce genre et renforcer sa distinction des autres groupes d'espèces définis par Musser et Carleton (2005). La Division *Maxomys* définie à partir de l'étude des éléments transposables de type LINE-1 (Verneau et al., 1997, 1998) repris par Musser et Carleton (2005) est confirmée par les travaux moléculaires ultérieurs (Lecompte et al. 2008, Pagès et al., 2010). Dix-sept espèces dont cinq sont présentes en Asie du Sud-Est continentale sont aujourd'hui considérées comme valides (Musser et Carleton, 2005). Le présent travail s'intéresse principalement à *M. surifer* et, avec des effectifs beaucoup plus faibles, à *M. rajah* Thomas, 1894 et *M. whiteheadi* Thomas, 1894. La variabilité génétique et la structuration géographique des populations de *M. surifer* révélées lors de nos travaux suggèrent que dans le futur plusieurs espèces pourraient être amenées à être décrites et que des études complémentaires sont donc nécessaires (Musser et al., 1979 ; Musser, 1981; Gorog et al., 2004 ; Pagès et al., 2010; 2011; Latinne, 2012a ; Latinne et al., 2013a). La présence d'un chromosome X acrocentrique chez l'ensemble des individus de notre échantillonnage pour cette espèce (Badenhorst et al., 2009, 2011) porte le nombre fondamental N<sub>Fa</sub> à 64 contre 62 dans les études précédentes (Marshall, 1977a ; Rickart et



**Figure 2:** richesse spécifique mondiale de la tribu des Rattini réalisée à partir des aires de répartition de 157 espèces décrites par l’UICN (<http://www.iucnredlist.org/technical-documents/spatial-data>, site visité le 14/09/2013). La présence absence des espèces est estimée par l’intersection des aires de répartition des espèces et d’une grille carrée de 1 degré de côté. La richesse spécifique est représentée par un code couleur selon l’échelle située à droite de la carte (Chaval, données non publiées).

**Tableau1 :** correspondance entre la dénomination des clades moléculaires de la figure 3 et les noms utilisés pour les nommer dans l’article de Pagès et al. (2010) et dans les parties de ce mémoire.

fig.3 Annexe 1 (Pagès et al., 2010)	Pagès et al., 2010	Ce travail partie 1	Ce travail partie 2
R1	<i>Rattus rattus</i>		<i>Rattus rattus</i>
R2	<i>Rattus tanezumi</i>	<i>Rattus tanezumi</i>	<i>Rattus tanezumi</i>
R3	<i>Rattus sp.</i>	<i>Rattus tanezumi</i>	<i>Rattus tanezumi</i>
R4	<i>Rattus losea</i> ou <i>Rattus "losea-like"</i>	<i>Rattus sakeratensis</i>	<i>Rattus sakeratensis</i>
R5	<i>Rattus tiomanicus</i>		<i>Rattus tiomanicus</i>
R6	<i>Rattus argentiventer</i>	<i>Rattus argentiventer</i>	<i>Rattus argentiventer</i>
R7	<i>Rattus andamanensis</i>		<i>Rattus andamanensis</i>
R8	<i>Rattus exulans</i>	<i>Rattus exulans</i>	<i>Rattus exulans</i>
R9	<i>Rattus norvegicus</i>	<i>Rattus norvegicus</i>	<i>Rattus norvegicus</i>
R10	<i>Rattus nitidus</i>	<i>Rattus nitidus</i>	<i>Rattus nitidus</i>
B1	<i>Bandicota indica</i>	<i>Bandicota indica</i>	<i>Bandicota indica</i>
B2	<i>Bandicota savilei</i>	<i>Bandicota savilei</i>	<i>Bandicota savilei</i>
Be1	<i>Berylmys berdmorei</i>	<i>Berylmys berdmorei</i>	<i>Berylmys berdmorei</i>
Be2a	<i>Berylmys bowersi</i>	<i>Berylmys bowersi</i>	<i>Berylmys bowersi</i>
Be2b	<i>Berylmys sp.</i>		
L1	<i>Leopoldamys edwardsi</i>	<i>Leopoldamys edwardsi</i>	<i>Leopoldamys edwardsi</i>
L2	<i>Leopoldamys neilli</i>	<i>Leopoldamys neilli</i>	<i>Leopoldamys neilli</i>
L3	<i>Leopoldamys sabanus</i>		<i>Leopoldamys sabanus</i>
N1	<i>Niviventer fulvescens</i>	<i>Niviventer huang</i>	<i>Niviventer huang</i>
N2	<i>Niviventer sp.1</i>		<i>Niviventer fulvescens</i>
N3	<i>Niviventer sp. 2</i>		
N4	<i>Niviventer langbianis</i> ou <i>Chiromyscus chiropus</i>		
M1	<i>Maxomys sp.</i>		
M2	<i>Maxomys surifer</i>	<i>Maxomys surifer</i>	<i>Maxomys surifer</i>
racine	<i>Micromys minutus</i>		<i>Micromys minutus</i>
		<i>Mus cervicolor</i>	
		<i>Mus cookii</i>	
		<i>Mus caroli</i>	

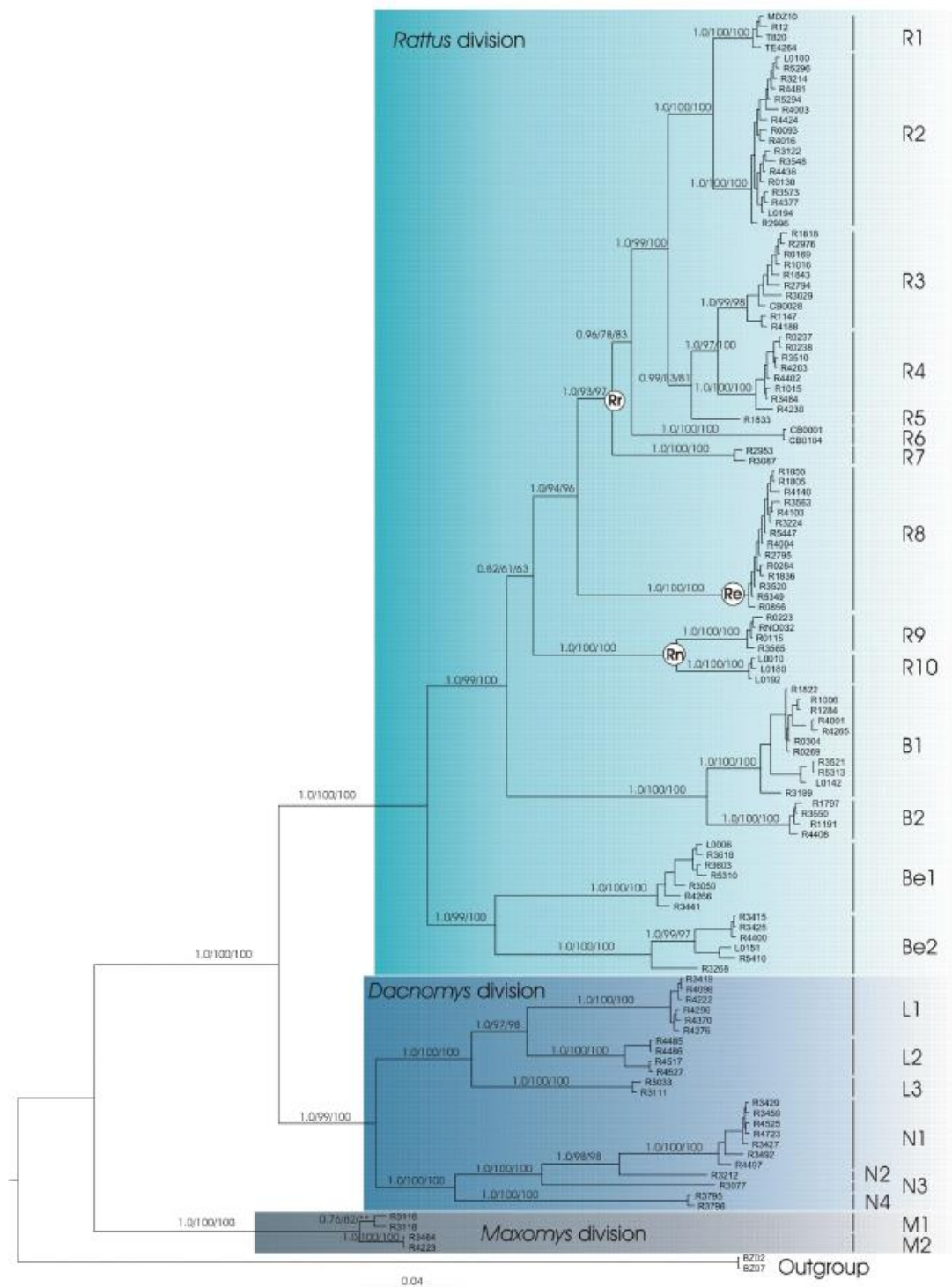
Musser, 1993) et supporte l'hypothèse de l'existence d'espèces cryptiques au sein du taxon. De même, la phylogénie moléculaire réalisée par Pagès et al. (2010) sur la délimitation moléculaire des espèces de Rattini sur la base de deux gènes mitochondriaux cytochrome b (Cytb) et Cytochrome C Oxydase I (COI) et d'un gène nucléaire Interphotoreceptor retinoid-binding protein (IRBP) met en évidence deux clades moléculaires M1 et M2 pour *M. surifer* (l'échantillonnage couvre l'ensemble de nos sites d'étude), qui pourraient correspondre à deux espèces distinctes. Pour le moment, aucune étude morphométrique exhaustive ne permet de trancher clairement sur le nombre de taxons en présence.

1.1.2- Le genre *Niviventer* J. T. Marshall, Jr., 1976, Division *Dacnomys* (Musser et Carleton, 2005)

Musser (1981) exclut le sous-genre *Niviventer* (Marshall, 1977a) du genre *Rattus* pour l'élever au rang de genre. La différenciation des genres *Rattus* et *Niviventer* est attestée par des caryotypes distincts (Yong, 1969 ; Badenhorst et al., 2009 ; 2011). D'après la systématique de Musser et Carleton (2005), ce genre compte 17 espèces valides dont 11 peuvent être rencontrées en Asie du Sud-Est continentale. Les espèces sont classées en deux groupes d'espèces par Musser et Carleton (2005), « andersoni » (*N. andersoni* Thomas, 1911, *N. excelsior* Thomas, 1911) et « niviventer » (toutes les autres espèces). Ce dernier groupe d'espèces est lui-même subdivisé en deux. D'une part les espèces dont les limites morphologiques et les aires de distribution sont bien définies : *N. brahma* Thomas, 1914 et *N. eha* Wroughton, 1916 (Musser, 1970); *N. cremoriventer* Miller, 1900 et *N. langbianis* Robinson and Kloss, 1922 (Musser, 1973); *N. hinpoon* J. T. Marshall, Jr., 1976 (Marshall, 1977a ; Pagès et al., 2010 ; Latinne, 2012a). D'autre part celles pour lesquelles les limites d'espèces restent floues et qui sont groupées par localités géographiques (*N. niviventer* Hodgson, 1836, *N. fulvescens* Gray, 1847, *N. confucianus* Milne-Edwards, 1871, *N. tenaster* Thomas, 1916, *N. cameroni* Chasen, 1940, *N. coninga* Swinhoe, 1864, *N. rapit* Bonhote, 1903, *N. lepturus* Jentink, 1879, *N. bukit* Bonhote, 1903, *N. fraternus* Robinson and Kloss, 1916). La taxonomie du genre reste donc très confuse et nécessite une révision systématique (Musser et Carleton, 2005). He et Jiang (2013) montrent à l'aide d'une analyse bayésienne partitionnée réalisée sur le gène Cytb qu'il y a certainement bien plus de taxons que d'espèces valides décrites aujourd'hui (21 phylo-espèces d'après cette étude qui n'inclut pas *N. hinpoon*) au sein du genre *Niviventer* au sens de Musser et Carleton (2005).

Balakirev et al. (2011) ont entrepris une révision taxonomique du genre basée sur la morphométrie crânienne et l'analyse de séquences de gènes mitochondriaux (Cytb et COI). Cette étude est dans les grandes lignes congruente avec la taxonomie de Musser et Carleton (1993). Selon Balakirev et al. (2011), *N. huang* ne serait pas un synonyme de *N. fulvescens* comme dans la taxonomie proposée par Musser et Carleton (2005) mais serait un nom d'espèce valide pour désigner les populations échantillonnées lors de nos programmes et désignées sous l'appellation N1 (Pagès et al., 2010). Les critères crâniens utilisés par Balakirev et al. (2011) pour différencier *N. huang* Bonhote, 1905 et *N. fulvescens* sont :

- la différence de taille des bulles tympaniques chez les deux espèces. Cependant les mesures prises par Balakirev et al. (2011 voir Appendix 3) ne permettent pas de distinguer les deux taxons de façon non ambiguës car la variance de ce trait pour les deux espèces reste très chevauchante ;
- la taille des foramens palatins dont la partie postérieure ne dépasse pas la limite antérieure de la première molaire supérieure (M1) de plus de 0,1 mm chez *N. huang* alors qu'elle la dépasse de plus de 0,4 mm chez *N. fulvescens*. En accord avec Balakirev et collaborateurs, notre échantillonnage morphométrique du clade N1 (sensu Pagès et al., 2010) s'approche en moyenne de la valeur observée pour *N. huang* avec un dépassement de la limite antérieure de la M1 par les foramens palatins de 0,01 mm.



**Figure 3** : Arbre des relations phylogénétiques de la tribu des Rattini indochinois basé sur une analyse combinée des gènes (Ctyb, COI, IRBP) et reconstruit par des méthodes Bayésiennes. Les nombres au-dessus des branches donnent le support par trois méthodes de reconstruction différentes (Bayésien, Maximum de vraisemblance non-partitionné, Maximum de vraisemblance partitionné). « \*\* » indiquent que les relations phylogénétiques ne sont pas soutenues pour les analyses de maximum de vraisemblance partitionnées. **Rr** indique le groupe *Rattus rattus*, **Re** le groupe *Rattus exulans*, **Rn** le groupe *Rattus norvegicus* (Extrait de Pagès et al, 2010).

- la jonction de la crête supra-orbitale et de l'os supra-occipital chez *N. huang* mis en évidence par Balakirev et al. (2011) semble se vérifier sur les crânes du clade N1 de notre échantillonnage.

L'un des échantillons que nous avons étudié (R3212) a été collecté dans la province de Kanchanaburi lors du travail de doctorat de V. Herbreteau (2007). Il correspond à un autre clade moléculaire, N2, suivant la nomenclature de Pagès et al. (2010). Pour Balakirev et al. (2011), N2 appartiendrait à *N. fulvescens*.

Cependant les critères de morphologie crânienne proposés par ces auteurs pour discriminer *N. huang* et *N. fulvescens*, et donc les clades N1 et N2 sensu Pagès et al. (2010), ne permettent pas de différencier R3212, notre seul représentant du clade N2 du reste de notre jeu de données (N1). De même l'assignation du clade N2 à l'aide de la forme du crâne en notre possession (voir Partie 1 chapitre D1) le place dans la variabilité du clade N1 décrit dans notre échantillonnage de référence. Concernant la morphologie externe, N1 est en tous points semblables au *N. huang* décrit par Balakirev et al. (2011) et N2 ne diffère de *N. fulvescens* que par la coloration de la fourrure dorsale. La morphologie externe et les distances moléculaires (distances Kimura à deux paramètres ou K2P) estimées à partir des séquences de Cytb produites lors de différentes études (Pagès et al., 2010 ; Latinne 2012a ; Balakirev et al., 2011) sont résumées dans le tableau 2 . Les distances génétiques mettent en évidence des proximités génétiques entre clades mais dont les valeurs sont toujours comprises dans un intervalle de distance de 2 à 11% que Bradley et Baker (2001) définissent comme une zone d'incertitude à l'intérieure de laquelle les taxons étudiées peuvent avoir un statut conspécifique ou interspécifique. Ainsi on note une proximité des clades *N. hinpoon* (Latinne, 2012) et N2 (Pagès et al., 2010) dont la distance est de 4,3 %. Ces deux mitotypes sont également proches du clade *N. fulvescens* (Balakirev et al., 2012) avec des distances respectives à celui-ci de 6,6 % et 5,1 %. Le clade N1 est pour sa part proche de *N. huang* (Balakirev et al., 2012) avec une distance de 3 %. Concernant la morphologie externe, N1 est en tous points semblables au *N. huang* décrit par Balakirev et collaborateurs (2012) et N2 ne diffère de *N. fulvescens* que par la coloration de la fourrure dorsale. Cependant ces deux derniers taxons génétiquement proches de *N. hinpoon* en sont phénotypiquement bien distincts. En conséquence dans ce travail, le taxon nommé N1 ou *N. fulvescens* dans le travail de Pagès et al. (2010) sera renommé *N. huang* et le taxon N2 sera dénommé *N. fulvescens*.

Le clade N3 de Pagès et al. (2010) n'a pu être nommé sur des bases phylogénétiques car sa distance avec les taxons publiés (Jing et al., 2007, Pagès et al., 2010 ; Balakirev et al., 2011, Latinne, 2012a) est comprise entre 12,6 et 17,5 % (données non montrées) ce qui est dans tous les cas de l'ordre des distances inter-spécifiques constatées pour ce gène (Bradley et Baker, 2001).

Le clade N4, nommé *N. langbianis* (Pagès et al., 2010) est proche (5,2 % de divergence sur le gène Cytb) du taxon nommé *N. langbianis* par Balakirev et al. (2012). Cependant les analyses de He et Jiang (2013) placent ce clade à part du clade *N. langbianis* (Balakirev, 2012). Cependant la description de la morphologie externe des animaux examinés par Balakirev et al. (2011) en tout point semblable à celle des animaux formant le clade N4 nous invite à conserver ce nom.

### 1.2.3- Le genre *Leopoldamys* Ellerman, 1947-1948, Division *Dacnomys* (Musser et Carleton, 2005)

Musser (1981) exclut le sous-genre *Leopoldamys* (Ellerman, 1947) du genre *Rattus* et l'élève au rang de genre à partir de critères morphologiques, morphométriques et cytogénétiques. *Leopoldamys* est classé au sein de la division *Dacnomys* sur la base de ses affinités moléculaires (Verneau et al., 1997,1998) et morphologiques (Musser et Carleton, 2005) avec le genre *Niviventer* ainsi que par la présence chez ces taxons de synapomorphies

**Tableau 2 :** Distance de Kimura à deux paramètres par paires (délétion par paires) de différentes lignées phylogénétiques (Chaval, non publié) définies à partir de l'analyse du gène mitochondrial Cyt B. *N. fulvescens* et *N. huang*, *N. langbianis* (Balakirev et al, 2011) ; N1 et N2, N4 (Pagès et al., 2010); *N. hinpoon* qui inclut l'holotype de l'espèce (Latinne, 2012a).

<b>clades</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>morphologie corporelle</b>
<i>N.fulvescens</i>							dos brun rougeâtre, queue longue, 148 % de la longueur tête corps, généralement faiblement contrastée et complètement bicolore, ventre blanc pur.
<i>N.huang</i>	9,2						dos fauve, queue 123% de la longueur tête corps, queue complètement bicolore très contrastée, ventre blanc pur.
<i>N.langbianis</i>	13,3	13,8					Dos gris brun séparé du ventre par une ligne nette, queue 140% de la taille du corps, queue complètement sombre, ventre blanc pur
N1	8,8	3	13,8				dos brun rougeâtre, queue 115 % de la longueur tête corps, queue complètement bicolore très contrastée, ventre blanc pur.
N2	5,1	9,2	14,1	9,2			dos fauve, queue longue, 144 % de la longueur tête corps, faiblement contrastée mais complètement bicolore. Ventre blanc pur.
N4	14,6	15,1	5,2	14,7	15,4		Dos gris brun séparé du ventre par une ligne plus claire, taille de la queue inconnue, queue complètement sombre, ventre blanc pur
<i>N.hinpoon</i>	6,6	10,1	3,5	10,4	4,3	13,5	dos gris marron, longueur de la queue équivalente à la longueur tête corps (Marshall, 1977a), queue complètement bicolore. Ventre crème à base grise.



crâniennes et dentaires (Musser et Newcomb, 1983). Le genre contient 6 espèces d'après la systématique de Musser et Carleton (2005) parmi lesquelles trois (*L. neilli*, *L. edwardsi* Thomas, 1882 et *L. sabanus* Thomas, 1887) sont traitées dans le présent travail. Les relations phylogénétiques entre ces trois espèces au sein de la tribu des Rattini ont fait l'objet de travaux récents (Pagès et al., 2010 ; Latinne, 2012a ; Latinne et al. 2013b). La composition en espèces du genre et la dénomination de celles-ci est revue par Balakirev et al. (2013) mais partiellement remise en question par Latinne et al. (2013a). En l'absence d'une taxonomie suffisamment étayée dans l'étude de Balakirev et al. (2013), la systématique suivie ici est celle de Musser et Carleton (2005). La diversité intra-spécifique et la structuration des populations de *L. neilli* a été abondamment étudiée (Latinne et al., 2011a, 2011b ; Latinne 2012a ; Latinne et al. 2012b, 2013b ).

#### 1.1.4- Le genre *Bandicota* Gray, 1873, Division *Rattus* (Musser et Carleton, 2005)

Une synthèse des travaux antérieurs basés sur la morphologie et la cytogénétique permet à Musser et Brothers (1994) de confirmer les affinités de ce genre avec le genre *Rattus*. Ces travaux sont renforcés par des travaux de génétique (Verneau et al., 1997, 1998, Lecompte, 2008, Pagès, 2010) et cytogénétiques (Badenhorst et al., 2011). Ce genre comprend trois espèces en Asie du Sud-Est : *B. indica* Bechstein, 1800, *B. savilei* Thomas, 1916, *B. bengalensis* Gray, 1835 dont les deux premières sont présentes dans la zone étudiée. Les deux espèces s'identifient partiellement sur le terrain à l'aide du rapport de la longueur tête corps sur la longueur de la patte arrière mais elles restent cryptiques morphologiquement pour une partie des classes d'âge (Aplin et al. 2003; Chaval et al., 2010, en prep.).

#### 1.1.5- Le genre *Berylmys* Ellerman, 1947, Division *Rattus* (Musser and Carleton, 2005)

Musser et Newcomb (1983) excluent le sous-genre *Berylmys* (Ellerman, 1947) du genre *Rattus* en l'élevant au rang de genre à partir de critères morphologiques, morphométriques et cytogénétiques. Des travaux moléculaires postérieurs concluent à l'affinité phylogénétique du genre *Berylmys* avec la Division *Rattus* (Usdin et al., 1995 ; Verneau et al., 1997, 1998 ; Lecompte et al., 2008 ; Pagès et al., 2010). Des travaux cytogénétiques récents confirment cette affinité (Badenhorst et al., 2011). Ce genre comprend quatre espèces en Asie du Sud-Est : *B. bowersi*, *B. berdmorei* Blyth, 1851, *B. mackenziei* Thomas, 1916, *B. manipulus* Thomas, 1916 (Musser et Carleton, 2005 ; Chaval, 2011) dont les deux premières sont présentes dans la zone étudiée, les deux autres espèces se rencontrent au centre et dans la partie ouest du Myanmar. Le genre s'identifie aisément sur le terrain car il a la particularité de posséder une couche d'émail blanc sur les incisives, alors qu'elle est orange chez les autres genres (Marshall, 1977a ; Aplin et al., 2003a). *B. berdmorei* est plus petit que *B. bowersi* et possède une queue plus courte que la longueur tête corps alors que celle-ci est plus longue chez *B. bowersi* (Marshall, 1977a ; Aplin et al., 2003a).

#### 1.1.6- Le genre *Rattus*, Fischer, 1803, Division *Rattus* (Musser et Carleton, 2005)

Bien que nombre d'espèces ait été réaffecté au sein de 25 genres différents depuis la création du genre *Rattus* par Fischer, ce genre demeure, possède la diversité spécifique la plus importante au sein de la Classe des mammifères (Pagès et al., 2013) avec 66 espèces valides recensées (Musser et Carleton, 2005). Cette richesse spécifique rend délicate la définition des limites d'espèces et nombre de taxons appartenant au genre *Rattus* possèdent un très grand nombre de synonymes (Pagès et al., 2013). La délimitation des espèces du genre *Rattus* demeure elle-même incertaine (voir Fabre et al., 2012 ; Schenk et al., 2013). Par voie de

conséquence, l'identification des espèces est elle-même rendue délicate. Le genre est subdivisé en sept groupes d'espèces, le dernier groupe (*xanthurus*) correspondant à un groupe où sont placées toutes les espèces dont la phylogénie n'est pas résolue. Seuls trois groupes d'espèces seront étudiés dans le présente rapport: les groupes *Rattus norvegicus*, *Rattus rattus* et *Rattus exulans*. Des études génétiques récentes confirment la monophylie de ces trois groupes (Robins et al., 2007 ; Lecompte et al., 2008 ; Pagès et al., 2011, Aplin et al., 2011, Fabre et al., 2012). La taxonomie des espèces du groupe *Rattus rattus* a récemment évolué.

Ainsi, au sein de ce groupe, le taxon R4 selon la nomenclature de Pagès et al. (2010) nommé « *R. losea-like* », a été renommé *R. sakeratensis* à la suite d'une révision taxonomique réalisée par Aplin et al. (2011). L'ancien taxon *R. losea* (Musser et Carleton, 2005) renfermait en effet deux entités taxonomiques qui présentent des traits morphologiques et génétiques caractéristiques et distincts bien que les deux espèces se ressemblent. Les *R. losea* vrais du Nord Vietnam s'avèrent être plus massifs que les *R. sakeratensis*. *R. losea* se rencontre au Cambodge (Province de Pursat), Nord du Vietnam, de la Thaïlande et du Laos, Sud et Est de la Chine ainsi que dans les îles du Hainan et de Taiwan (Aplin et al., 2011, et obs pers.). *R. sakeratensis* est une version miniature du précédent taxon que l'on rencontre au sud Vietnam, dans l'ouest du Cambodge et en Thaïlande et dans les zones rizicoles du sud et du centre du Laos (Aplin et al., 2011 et obs pers.). Les deux espèces sont possiblement sympatriques au centre du Cambodge et dans le nord de la Thaïlande mais aucune observation confirmant cette hypothèse n'a encore été publiée. Une comparaison entre les caryotypes des deux taxons (Yu et al., 1996) montre que le chromosome X est plus petit chez les *R. losea* (échantillon de Taiwan) que chez les *R. sakeratensis* (échantillons de Thaïlande). Les auteurs émettent l'hypothèse d'un isolement reproducteur entre ces deux taxons considérés alors comme des sous-espèces.

De plus, l'étude de Pagès et al. 2010 puis des études plus récentes (Aplin et al., 2011; Lack et al., 2012) ont confirmé l'existence d'un clade mitochondrial supplémentaire au sein du genre *Rattus* nommée R3 dans le travail de Pagès et al. (2010). Ce clade avait déjà été mis en évidence par Robins et al. (2007). R3 est morphologiquement identique aux *Rattus tanezumi* Temminck, 1844, nommé R2 dans l'étude Pagès et al. (2010), mais proche phylogénétiquement de R4, autrement dit, *R. sakeratensis*. Cette étude utilise des gènes mitochondriaux (Cytb, COI) mais également un gène nucléaire (IRBP). Cependant se sont surtout les gènes mitochondriaux qui soutiennent les nœuds terminaux des arbres phylogénétiques et donc la proximité de R3 et *R. sakeratensis*. Récemment Pagès et al. (2013) ont confronté les signaux mitochondriaux et nucléaires (microsatellites) sur un plus large échantillonnage et révélé qu'il existait une discordance cyto-nucléaire. Les données mitochondriales discriminent *R. tanezumi* et R3 alors que les données nucléaires ne le permettent pas et suggèrent la conspécificité de ces deux mitotypes. Deux hypothèses permettent de réconcilier ces signaux (voir partie 2 chapitre 1). Une étude préliminaire, qui intègre un échantillonnage élargi incluant le Vietnam (Gouy, 2013) et basée sur les mêmes marqueurs mitochondriaux et nucléaires que Pagès et al. (2013) confirme le caractère conspécificique des deux lignées mitochondriales en Thaïlande au Laos et à l'Ouest du Cambodge. Cependant cette étude met en évidence des populations de la lignée R3 du Sud Vietnam qui sont génétiquement très différenciées sans qu'il soit possible pour le moment de dire si cette différence est due à un isolement géographique de deux populations de la même espèce ou s'il est la signature d'un isolement reproductif et donc de deux espèces au sens de Mayr (1942).

## 2- Biogéographie de la tribu des *Rattini*

Les archives fossiles (Musser and Carleton, 2005a) et les analyses phylogénétiques (Schenck, 2013) s'accordent sur une origine sud-est asiatique de la sous-famille des Murinés vers le milieu du Miocène (23 à 5,3 Ma). Les membres de la tribu murine des *Rattini* sont répartis, en majorité, sur deux régions biogéographiques, la région indomalayenne, foyer de diversification de ce clade (Musser and Carleton, 2005a ; Rowe et al, 2011 ; Fabre et al., 2013) et la région australienne (introduction Figure 2 et Annexe 8 Figure 4) qui n'a été colonisée que plus tardivement. Ces deux régions biogéographiques sont portées par deux plaques tectoniques distinctes: la plaque eurasiennne et la plaque australienne. La zone de contact de ces deux plaques, qui se situe à la jonction de la sous-région du Wallacea et de la sous-région du Sahul est une zone de subduction dont l'activité se caractérise d'une part par la surrection de morceaux détachés du plancher océanique, les terranes, et d'autre part par une activité volcanique intense (Lohman, 2011) depuis l'Eocène moyen (43 Ma). A l'intérieur de la région indo malayenne (portée par la plaque eurasiennne), les sous-régions Sundaïque (regroupant les îles de Bornéo, Java, Sumatra et la partie continentale de la Malaisie) et indochinoise (s'étendant au nord de la ligne Kangar-Pattani qui barre l'isthme de Kra) ont une histoire commune longue et des contacts terrestres répétés (Woodruff, 2010, Lohman, 2011) depuis le paléocène ancien (60 Ma) et jusqu'à la fin du dernier maximum glaciaire (19-23 Ka). A l'opposé, la sous-région du Wallacea à l'extrême Est de la région indomalayenne, marque la transition avec la région australienne et est constituée par des îles océaniques d'origine volcanique sur lesquelles se sont accrétés des terranes. Ces îles sont séparées par de multiples bras de mer qui forment des barrières plus ou moins perméables au rythme des oscillations climatiques et marines du Quaternaire.

Le climat de la région indomalayenne durant le Miocène (23 à 5,3 Ma) et le Pliocène (5,3 à 2,6 Ma) était plus humide qu'aujourd'hui et la température supérieure en moyenne aux températures actuelles d'environ 3°C. Dans ces conditions, la végétation était dominée par la forêt tropicale humide (Woodruff, 2010). Le clade formant la tribu des *Rattini* se diversifie probablement vers 9 Ma (Lecompte et al, 2008, Fabre et al., 2013) dans ce paysage où la forêt prédomine en Asie du sud-est continentale. A partir de 3,2 Ma, la température commence à décroître et à partir de 2,7 Ma, la calotte glaciaire se forme dans l'hémisphère nord ce qui s'accompagne d'une baisse des niveaux marins. Le climat est alors progressivement plus froid et sec et c'est à cette époque qu'apparaît le genre *Rattus* (Verneau et al, 1998, Robins et al., 2008) probablement dans le sud-est asiatique (Rowe et al. , 2011). C'est dans cette zone que sont localisés les premiers fossiles du genre (Chaimanee, 1998). La transition climatique du milieu du Pléistocène (il y a 800 mille ans environs) est marquée par une remontée des eaux froides océaniques profondes qui accentue la croissance et l'épaississement des banquises en surface des océans dans l'hémisphère nord et engendre la baisse du niveau marin (Woodruff, 2010). Dans la sous-région sundaïque, cette baisse importante du niveau marin fait émerger le plancher océanique formant le « sunda shelf » qui relie par voie terrestre le continent asiatique et les îles les plus importantes de la sonde ainsi que Bornéo et les Philippines (Annexe 8 Figure 5). Ce pont entre les sous-régions est maintenu jusqu'à la fin du dernier maximum glaciaire il y a 19 mille ans. Cependant la période Quaternaire est marquée par de nombreuses oscillations climatiques. Plus de 50 cycles glaciaires séparés ont été enregistrés durant cette période (Woodruff, 2010). Ces changements de climat s'accompagnent de cycles de contraction puis d'expansion des forêts et des savanes (Wurster, 2010) ce qui a pour effet de séparer ou de réunir les populations animales avec des pas de temps variables suivant la longueur de chaque cycle. La création de refuges pléistocènes est une explication à la radiation de certains taxons par vicariance. Ces événements ont été mis en évidence chez les

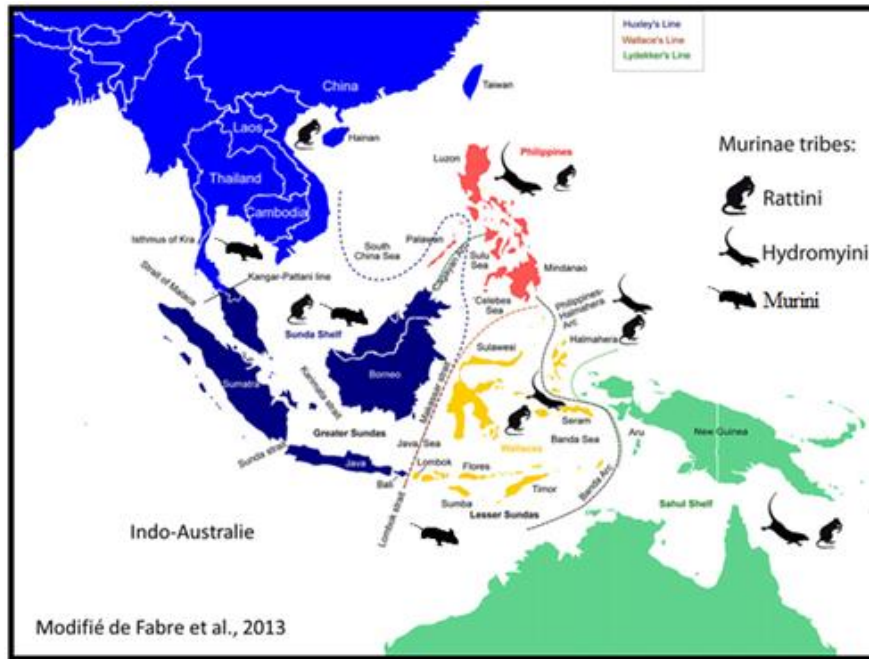


Figure 4: Distribution actuelle de la tribu des Rattini, Hydromyini et Murini dans les différentes sous-régions de la zone indo-australienne (Fabre et al., 2013).

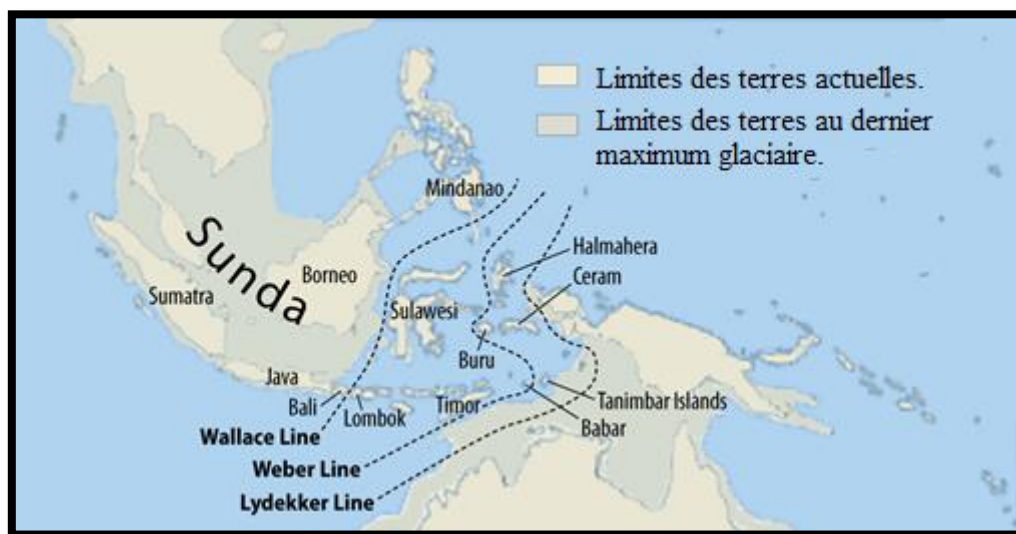


Image: Maximilian Dörrbecker

Source: Used under CC BY-SA 3.0 from [Wikimedia Commons](https://commons.wikimedia.org/wiki/File:Map_of_Sunda_and_Sahul.png).

Figure 5: Limite des terres au dernier maximum glaciaire (19-23Ka). Les sous-régions sundaique et indochinoise sont largement connectées entre elles mais restent isolées de la sous-région du Sahul par les multiples bras de mer parcourant la sous-région du Wallacea ([http://en.wikipedia.org/wiki/File:Map\\_of\\_Sunda\\_and\\_Sahul.png](http://en.wikipedia.org/wiki/File:Map_of_Sunda_and_Sahul.png))

rongeurs de la tribu des Rattini des zones sundaïque (Gorog et al., 2004) et indochinoise (Latinne, 2012). L'histoire biogéographique des Rattini des sous-régions du Wallacea et du Sahul est différente. La ligne de Wallace qui sépare ces deux sous-régions du Sunda et qui suit à cet endroit le détroit de Makassar séparant l'île de Bornéo de la Sulawesi est une frontière faiblement perméable (Annexe 8, Figure 4 et Figure 5). Une connexion brève (quelques Millions d'années) entre le sud de la Sulawesi et Bornéo a été mise en évidence au début de l'Eocène il y a 50 Ma (Wilson et Moss, 1999) mais les deux îles sont restées séparées depuis l'apparition des rongeurs en Asie du Sud-est (il y a 34 Ma). Le fort taux d'endémisme enregistré dans la zone (90 % des espèces en Sulawesi) est le signe d'une radiation adaptative qui a suivi la colonisation par un petit nombre d'espèces migrantes à l'Eocène tardif (Lohman, 2011). Les événements de vicariance à partir des terranes australiens sont rares et n'ont été démontrés que pour les escargots pachychilides en Sulawesi (Rintelen & Glaubrecht, 2005). D'autre part, les espèces de la sous famille des murinés sont les seules mammifères non volant asiatiques à avoir réussi à traverser les bras de mer multiples partageant ces deux sous régions (Rowe et al., 2011). La sous-région biogéographique du Wallacea à laquelle appartient la Sulawesi a connu des événements de colonisation multiples au cours du Miocène tardif et du plio-pléistocène notamment par des taxons de la tribu des Rattini (Fabre, 2013). Le Sahul de son côté a fait l'objet de deux colonisations distinctes par deux tribus de la sous-famille des murinés (Rowe et al., 2011). Une première colonisation est attribuée à un groupe très diversifié morphologiquement et écologiquement et nommé « old endemics » (Simpson 1961 ; Jansa, 2006 ; Rowe et al., 2008 ; Rowe et al., 2011), dont le foyer de diversification est situé aux Philippines et dont la monophylie est attestée (Rowe, 2008 ; Fabre et al., 2013). Ces taxons forment la tribu des Hydromyini (Musser & Carleton, 2005 ; Rowe, 2008 ; Fabre et al., 2013). La seconde colonisation est attribuable à des taxons de la tribu des Rattini. Elle est plus récente et daterait de la transition plio-pléistocène (Rowe et al., 2011 ; Fabre et al., 2013). La recolonisation pléistocène du Wallacea par un taxon de la tribu des Rattini à partir du Sahul est attestée par Fabre et collaborateur (2013). Les Rattini semblent avoir supplanté les premiers colonisateurs de la tribu des Hydromyini dont la répartition géographique était certainement plus étendue qu'elle ne l'est aujourd'hui (Fabre et al., 2013). Cependant, parmi les Rattini colonisateurs du Sahul, les espèces du genre *Rattus* sont régulièrement rencontrées vivant en sympatrie avec les premiers colonisateurs (Rowe et al., 2011). Cependant les *Rattus* diffèrent des autres espèces en ce qui concerne leur taux de reproduction (Geffen et al., 2011). Ces différences reproductives pourraient leur permettre d'utiliser des composantes différentes de niche écologique (Schenke et al., 2013) et au travers de l'histoire biogéographique complexe de la zone avoir favorisé la mise en place d'une faune incroyablement riche et diversifiée.

### ***3- Systématique de la tribu des Murini pour la zone étudiée***

La tribu des Murini se limite au seul genre *Mus* (Linnaeus 1758) qui comprend 41 espèces valides (Auffray et Britton-Davidian, 2012 ; Suzuki et Aplin, 2012). Les espèces qui constituent ce genre possèdent une morphologie (Auffray et Britton-Davidian, 2012) et une écologie (Suzuki et Aplin, 2012) relativement conservées. A titre d'exemple, le poids des différentes espèces est compris dans une petite gamme de variation (de 5 à 35 grammes) et la plupart ont adopté un mode de vie terrestre. Ce genre a été subdivisé en quatre sous-genres (*Coelomys*, *Mus*, *Pyromys*, *Nannomys*) par Marshall (1977b) d'après des critères crânio-dentaires. Cette classification a ultérieurement été validée par des analyses phylogénétiques (Lundrigan et al., 2002 ; Suzuki et al., 2004 ; Suzuki et Aplin, 2012). Une seule espèce, *Mus famulus*, initialement placée par Marshall dans le sous-genre *Coelomys* a par la suite été reclassée dans le sous-genre *Mus* (Chevret et al., 2003). Pour le sous-genre *Mus*, Suzuki et al. (2004) reconnaissent quatre groupes d'espèces phylogénétiques (les groupes: *musculus* ,

*cervicolor*, *lepidoides* et *booduga*). Les espèces formant les groupes phylogénétiques *cervicolor*, *lepidoides*, *booduga* sont, à quelques exceptions près, distribuées dans des zones géographiques proches en Asie du Sud et du Sud-Est. Il n'en est pas de même pour le groupe *musculus*, dont les espèces occupent une zone s'étendant sur l'ensemble de la zone paléarctique. Parmi ce groupe d'espèces, *M. musculus* est une espèce à très large répartition géographique (Suzuky et Aplin, 2012 ; Suzuky et al. 2013) chez laquelle cinq sous-espèces ont été décrites (Corbet, 1990; Auffray et Britton-Davidian, 2012). Au cours de nos programmes de recherche, seules quatre espèces du sous-genre *Mus* ont été échantillonnées. Trois appartiennent au groupe *cervicolor* (Suzuky et Aplin, 2012) : *M. cookii* Ryley, 1914, *M. cervicolor* Hodgson, 1845, *M. caroli* Bonhote, 1902. *M. fragilicauda* appartient quant à elle au groupe d'espèces de *M. booduga* et elle est la seule espèce de ce groupe qui est majoritairement indien à posséder une répartition qui soit à ce point orientale. Les quatre espèces sont difficilement distinguables sur le terrain à partir de leur morphologie externe. Les relations phylogénétiques des taxons étudiés sont représentées à la figure 6.

#### **4- Biogéographie de la tribu des Murini**

La tribu des Murini se caractérise par une répartition géographique étendue et une présence naturelle dans la plupart des biomes de la forêt tropicale à la steppe et aux milieux semi-désertiques (Suzuki et Aplin, 2012). L'une des espèces du genre *Mus*, *Mus musculus* au mode de vie commensal a été largement introduit sur tous les continents par l'action de l'homme. La plus ancienne cladogénèse enregistrée au sein du genre *Mus* est celle du sous-genre *Coelomys* vers 7,5 Ma (Veyrunes et al., 2006). Cette époque du Miocène tardif, est marquée par la surrection de la chaîne Tibéto-Himalayenne (Suzuki and Aplin, 2012) qui a permis la mise en place d'un climat de type saisonnier marqué par l'apparition de la mousson et une alternance de périodes humides et sèches. Les espèces constituant le sous-genre *Coelomys* fréquentent les milieux forestiers les plus humides allant de la forêt à feuillage permanent, classe typologique à laquelle sont incluses les forêts de montagnes, à la forêt mixte semi-décidue. Il se rencontre principalement en Asie du Sud-Est dans les sous-régions indochinoise et sondaïque (Figure 4) mais certaines espèces dont l'une du complexe *M. pahari* (voir Suzuki et Aplin, 2012) se rencontre en Inde et *M. mayori* au Sri Lanka. Le second sous-genre à se séparer est sous-genre *Nannomys* vers 7 Ma (Veyrunes et al., 2006) qui se diversifie ensuite sur le continent africain exclusivement. Les espèces de ce sous-genre occupent des milieux plus secs à l'exception de *M. setulosus*, espèce basale dont l'habitat est la forêt tropicale toujours verte. Les sous-genres *Mus* et *Pyromys* sont les derniers à diverger vers 6,5 Ma. Le sous-genre *Pyromys* présente une diversifié spécifique maximale dans le centre-sud du continent Eurasien (Inde, Pakistan) mais l'un de ses représentants, *M. shorridgei*, peut être rencontré en Asie du Sud-Est. Il occupe des habitats plus secs que le sous-genre *Coelomys* et l'habitat de *M. shorridgei* est la forêt sèche à diphérocarpes. Enfin, le sous-genre *Mus* dont l'origine est probablement située en Asie centrale a, par la suite, colonisé une grande partie de l'Eurasie (Suzuki et Aplin, 2012). Des quatre groupes d'espèces reconnus au sein sous-genre *Mus* par Suzuki et collaborateurs (2004), *M. musculus* est incontestablement le groupe le plus connu et le plus étudié car il comprend la souris domestique (*M. musculus*) espèce commensale au comportement invasif, domestiquée en tant qu'animal de compagnie et comme animale de laboratoire au même titre que *R. norvegicus* chez les Rattini. Cette espèce est également un modèle privilégié d'étude de la spéciation et des mécanismes d'évolution (Macholan and Baird, 2012). Le groupe d'espèces *M. musculus* provient probablement de l'adaptation de taxons du groupe d'espèces *M. Booduga* au climat tempéré au Cambrien vers 1,7 Ma. Il se diversifie alors en colonisant l'Europe de l'Ouest et le nord de l'Afrique (Suzuki et Aplin, 2012). L'espèce *M. musculus*, qui apparaît probablement en Asie centrale sur le plateau Iranien (Siahsarvie et al., 2012). Elle colonise à son tour la

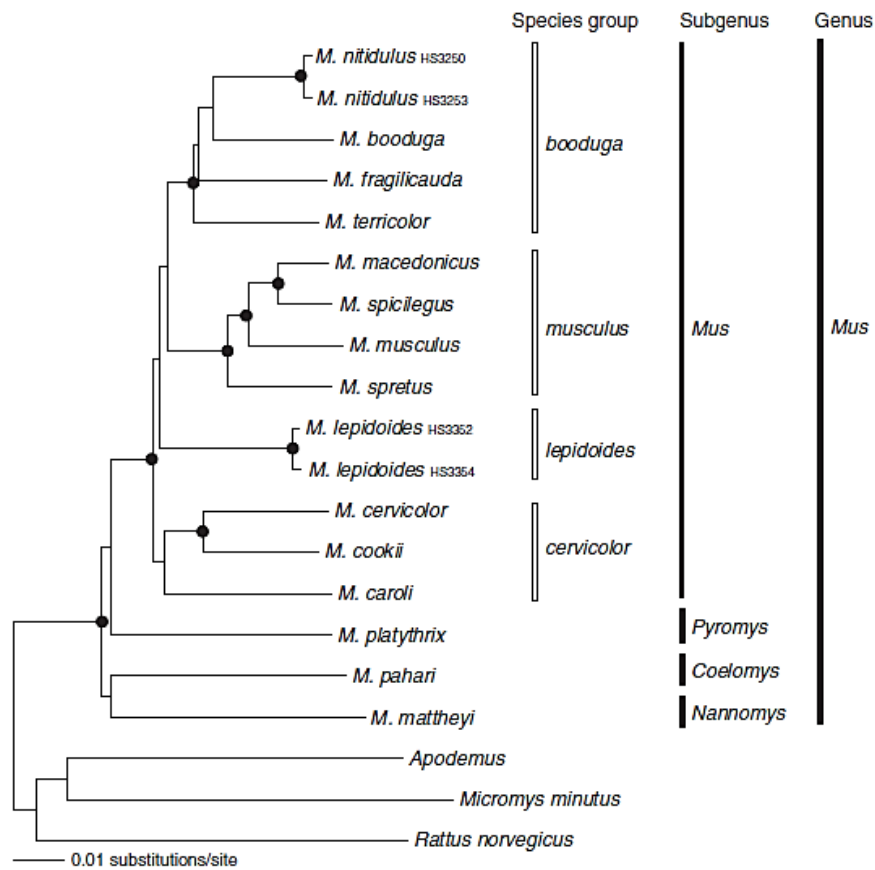


Figure 6 : Arbre phylogénétique de consensus strict construit par la méthode du maximum de vraisemblance sur un jeu de données non-partitionné des gènes Cytb, IRBP et recombinaison-activating (RagI). Les nœuds matérialisés par des cercles noirs possèdent une valeur de bootstrap supérieure à 80 % pour les analyses en neighbour-joining (1000 pseudo-réplicats), maximum de parcimonie (1000 pseudo-réplicats) et maximum de vraisemblance (100 pseudo-réplicats). D'après Simadha et al. (2010) et extrait de Suzuki et Aplin (2012).

partie sud de l'Eurasie et se diversifie pour donner naissance à 5 sous-espèces (Auffray and Britton-Davidian, 2012) dont une seule, *M. musculus castaneus* est présente en Asie du Sud-Est et est la seule espèce qui se rencontre dans la sous-région du Wallacea (Figure 4 ; Suzuki et al., 2013).



## Annexe II

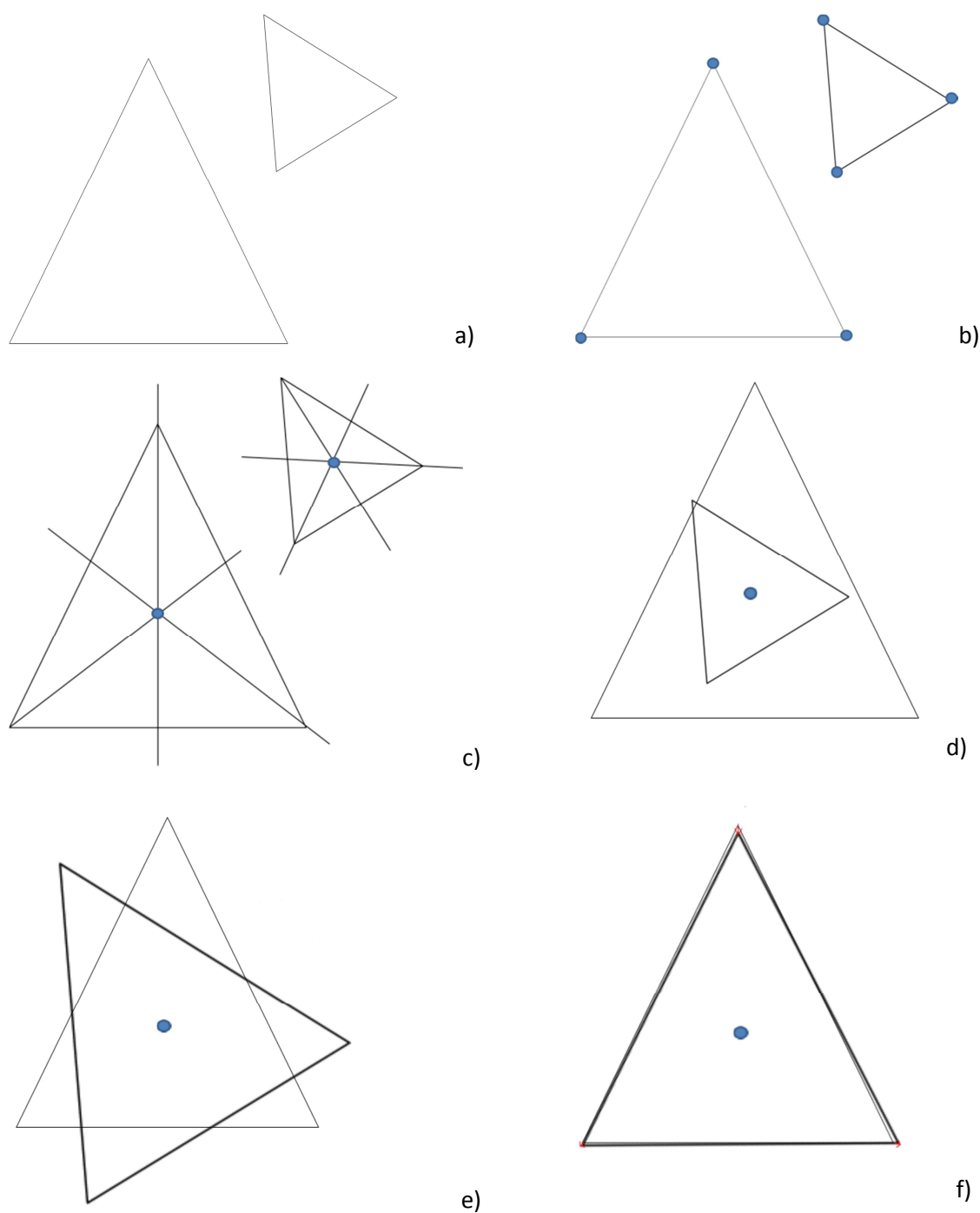


Figure 1: Les étapes de la superposition Procrustes. (a) L'objectif est de comparer deux objets qui diffèrent par leur position, leur orientation et leur taille. (b) L'information est résumée par un ensemble de points homologues \*. (c) Le centre de gravité des figures est calculé (d) une **translation** permet de faire correspondre les centres de gravité des formes (e) une **mise à l'échelle** permet de séparer la variable de taille et celle de conformation (f) Enfin une **rotation** permet de faire correspondre les conformations par un ajustement au moindre carré. Les flèches rouges représentent les différences entre conformations aussi appelés les résidus Procrustes.

\* Ici trois points placés aux sommets résument l'ensemble de l'information de forme. Ici les triangles ont été conservés pour visualiser de façon plus parlante les étapes de la superposition mais en réalité l'analyse ne porte que sur les configurations de points homologues faisant abstraction de l'objet proprement dit.



## Annexe III

Lentilles qu'il sera possible de corriger dans l'outil web service (base de données libre PTLensDB\_06-02-08 : <http://sourceforge.net/projects/hugin/files/PTLens%20Database/>)

### Canon EF/TS:

Canon EF-S 10-22mm f/3.5-4.5 USM  
Canon EF 16-35mm f/2.8L USM  
Canon EF 17-35mm f/2.8L USM  
Canon EF 17-40mm f/4L USM  
Canon EF-S 17-85mm f/4-5.6 IS USM  
Canon EF-S 18-55mm f/3.5-5.6  
Canon EF 24-70mm f/2.8L USM  
Canon EF 24-85mm f/3.5-4.5 USM  
Canon EF 28-70mm f/2.8L  
Canon EF 24-105mm f/4L IS USM  
Canon EF 24-105mm f/4L IS USM  
Canon EF 28-105mm f/3.5-4.5 II USM  
Canon EF 28-300mm f/3.5-5.6L IS USM  
Canon EF 28-135mm f/3.5-5.6 IS USM  
Canon EF 35-70mm f/3.5-4.5  
Canon EF 35-80mm f/4-5.6 III  
Canon EF 35-105mm f/3.5-4.5  
Canon EF 35-135mm f/4-5.6 USM  
Canon EF 55-200mm f/4.5-5.6  
Canon EF 70-200mm f/2.8L IS USM  
Canon EF 70-200mm f/2.8L IS USM + EF 1.4x ext  
Canon EF 70-200mm f/4L USM  
Canon EF 70-200mm f/4L USM + EF 1.4x ext  
Canon EF 70-300mm f/4.5-5.6 DO IS USM  
Canon EF 75-300mm f/4-5.6 IS USM  
Canon EF 80-200mm f/2.8L  
Canon EF 100-300mm f/5.6L  
Canon EF 100-400mm f/4.5-5.6L IS USM  
Canon EF 20mm f/2.8 USM  
Canon EF 24mm f/1.4L USM  
Canon EF 24mm f/2.8  
Canon EF 28mm f/1.8  
Canon EF 28mm f/2.8  
Canon EF 35mm f/1.4L USM  
Canon EF 35mm f/2.0  
Canon EF 50mm f/1.4  
Canon EF 50mm f/1.4

Canon EF 50mm f/1.8 I and II  
Canon EF-S 60mm f/2.8 Macro  
Canon EF 85mm f/1.2L USM  
Canon EF 85mm f/1.8 USM  
Canon EF 100mm f/2.8 USM Macro  
Canon EF 100mm f/2.8 Macro  
Canon EF 135mm f/2.8 Soft Focus  
Canon EF 200mm f/2.8L II USM  
Canon TS-E 24mm f/3.5L  
Canon TS-E 45mm f/2.8  
Canon TS-E 90mm f/2.8

### Canon:

WC-DC58N  
Geobart 0.5x Converter  
Tiffen MegaPlus 0.56 Converter  
Canon TC-DC 10 2x Converter  
Tiffen .56x converter  
Tiffen 2x converter  
TC-DC58B

### 6x6:

Schneider 28mm Digitar f/2.8  
Mamiya 35mm f/3.5  
Mamiya 80mm f/2.8  
Mamiya 55-110mm f/4.5  
Schneider 80mm Xenotar f/2.8

### NikonSLR:

Nikkor 12-24mm f/4G ED-IF AF-S DX  
Nikkor 17-35mm f/2.8 ED-AFS  
Nikkor 17-55mm f/2.8G ED-IF AF-S DX  
Nikkor 18-35mm f/3.5-4.5D ED-IF AF  
Nikkor 18-55mm f/3.5-5.6G ED AF-S DX  
Nikkor 18-70mm f/3.5-4.5G ED-IF AF-S DX  
Nikkor 18-200mm f/3.5-5.6 G ED-IF AF-S VR DX  
Nikkor 24-50mm f/3.3-4.5D Ai AF  
Nikkor 24-85mm f/2.8-4.0D AF  
Nikkor 24-85mm f/3.5-4.5G ED-IF AF-S  
Nikkor 24-120mm f/3.5-5.6G ED-IF AF-S VR  
Nikkor 28-200mm f/3.5-5.6G ED-IF AF  
Nikkor 35-70mm f/2.8D AF  
Nikkor 50-135mm f/3.5 AI-S

Nikkor 55-200mm f/4-5.6 G ED AF-S DX  
Nikkor 70-180mm f/4.5-5.6 ED-AF-D Micro  
Nikkor 70-200mm f/2.8G ED-IF AF-S VR  
Nikkor 70-210mm f/4.5-5.6  
Nikkor 70-210mm f/4 AF  
Nikkor 70-300mm f/4-5.6D ED AF  
Nikkor 70-300mm f/4-5.6G AF  
Nikkor 80-200mm f/2.8D ED AF  
Nikkor 80-400mm f/4.5-5.6 ED-AF-D-VR  
Nikkor 14mm f/2.8D ED AF  
Nikkor 15mm f/3.5 AI  
Nikkor 16mm f/2.8D AF Fisheye  
Nikkor 20mm f/2.8D AF  
Nikkor 24mm f/2.8D AF  
Nikkor 24mm f/2.0 AI-S  
Nikkor 28mm f/2.8D AF  
Nikkor 28mm f/2.0 AI-S  
Nikkor 28mm f/3.5 AI-S PC (unshifted)  
Nikkor 35mm f/1.4 AI-S  
Nikkor 35mm f/2.0D AF  
Nikkor 45mm f/2.8 AI GN  
Nikkor 50mm f/1.4D AF  
Nikkor 50mm f/1.4 AI-S  
Nikkor 50mm f/1.8 AI-S  
Nikkor 50mm f/2.0 non-AI  
Nikkor 55mm f/1.2 AI  
Nikkor 55mm f/2.8 AI-S Micro  
Nikkor 58mm f/1.2 Noct  
Nikkor 85mm f/1.8D AF  
Nikkor 105mm f/2.8D AF  
Nikkor 135mm f/2.0 AI-S  
Nikkor 180mm f/2.8 ED-AIS  
Nikkor 500mm f/8.0 Reflex

### Nikon:

WM-E80  
WC-E75  
TC-E15ED (full tele)  
WC-E80 (full wide)  
WC-E80  
WC-E68 (full wide)  
TC-E2  
TC-3ED  
WC-E63  
WC-E63  
WC-E24  
TC-E2

WC-E63

### **Sony:**

VCL-DH0758 0.7x Converter  
VCL-DH1758 1.7x Converter  
std macro: 1 inch lens-to-subject  
std macro: 2 inches lens-to-subject  
std macro: 4 inches lens-to-subject  
std macro: 6 inches lens-to-subject  
std macro: 8 inches lens-to-subject  
std macro: 12 inches lens-to-subject  
std macro: 16 inches lens-to-subject  
std macro: 20 inches lens-to-subject  
std macro: 24 inches lens-to-subject  
std macro: 32 inches lens-to-subject  
VCL-HGD0758 Wide Angle Sakar 1858W  
VCL-DEH07V Converter  
VCL-DEH17V Converter  
VCL-MHG07a

### **Casio:**

WC-DC58A  
Raynox 0.66x

### **ContaxSLR:**

Zeiss 21mm f/2.8 Distagon  
Zeiss 28mm f/2.8 Distagon

### **Fujifilm:**

WL-FXE01 (full wide)

### **GenericSLR:**

Schneider 28mm f/2.8 PC  
Sigma 10-20mm f/4-5.6 EX DC  
Sigma 12-24mm f/4.5-5.6 EX DG  
Sigma 15-30mm f/3.5-4.5 EX DG  
Sigma 17-35mm f/2.8-4 EX DG  
Sigma 17-70mm f/2.8-4.5 DC Macro  
Sigma 18-50mm f/2.8 EX DC  
Sigma 18-50mm f/2.8 EX DC  
Sigma 18-50mm f/3.5-5.6 DC  
Sigma 18-125mm f/3.3-5.6 (3.5-5.6) DC  
Sigma 18-125mm f/3.3-5.6 (3.5-5.6) DC  
Sigma 18-125mm f/3.3-5.6 (3.5-5.6) DC  
Sigma 18-200mm f/3.5-6.3 DC  
Sigma 18-200mm f/3.5-6.3 DC  
Sigma 24-70mm f/2.8 EX DG Macro  
Sigma 28-70mm f/2.8 EX DG  
Sigma 28-300mm f/3.5-6.3 Macro ASP IF

Sigma 50-300mm f/4-6.3 EX DG/HSM  
Sigma 55-200mm f/4-5.6 DC  
Sigma 70-300mm f/4-5.6 APO Macro Super II  
Sigma 14mm f/2.8 EX  
Sigma 14mm f/3.5 EX  
Sigma 15mm f/2.8 EX Diagonal Fisheye  
Sigma 15mm f/2.8 EX Diagonal Fisheye  
Sigma 20mm f/1.8 EX DG  
Sigma 28mm f/1.8 EX DG  
Sigma 30mm f/1.4 EX DC HSM  
Tamron 17-35mm f/2.8-4 Di LD  
Tamron 18-200mm f/3.5-6.3 XR Di II LD  
Tamron 18-200mm f/3.5-6.3 XR Di II LD  
Tamron 19-35mm f/3.5-4.5  
Tamron 24-135mm f/3.5-5.6 SP AD  
Tamron 28-75mm f/2.8 XR Di  
Tamron 28-105mm f/2.8 SP LD IF  
Tamron 28-300mm f/3.5-6.3 XR Di LD  
Tokina 12-24mm f/4 AT-X 124 AF PRO DX  
Tokina 19-35mm f/3.5-4.5 AF 193  
Tokina 28-70mm f/2.8 AT-X 287 PRO SV  
Tokina 17mm f/3.5 AT-X AF PRO  
Tokina 17mm f/3.5 AT-X AF PRO  
Zenitar 16mm f/2.8 MC Fisheye

### **KonicaMinolta:**

ZCW-100  
ZCW-100  
ZCW-200  
ACW-100 Converter  
KM 24-105mm 3.5-4.5 AF D  
KM 28-100mm 3.5-5.6 AF D  
KM 80-200mm 2.8  
KM 20mm 2.8

### **Olympus:**

A-28 iS/L Converter  
WCON-08B

### **OlympusSLR:**

Zuiko Digital 7-14mm f/4.0  
Zuiko Digital 11-22mm f/2.8-3.5  
Zuiko Digital 14-45mm f/3.5-5.6  
Zuiko Digital 14-54mm f/2.8-3.5  
Zuiko Digital 40-150mm f/3.5-4.5  
Zuiko Digital 50-200mm f/2.8-3.5  
Zuiko Digital 50mm f/2.0 Macro WCON-08B  
B-300 (full tele)

macro (full tele): 8 inches lens-to-subject  
macro (full tele): 12 inches lens-to-subject  
macro (full tele): 16 inches lens-to-subject  
macro (full tele): 20 inches lens-to-subject

### **PentaxSLR:**

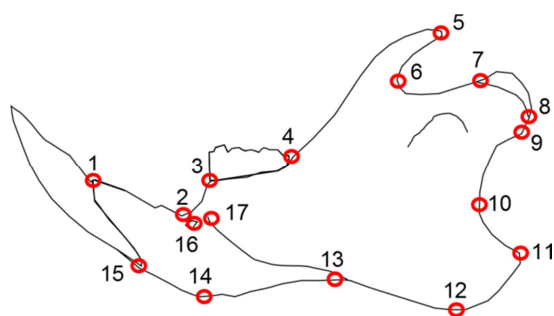
Pentax 16-45mm f/4.0 P-DA J ED/AL  
Pentax 18-55mm f/3.5-5.6  
Pentax 50-200mm f/4-5.6 P-DA ED

### **Ricoh:**

DW-4

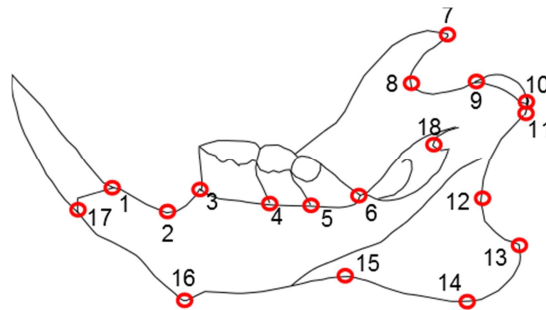
## Annexe IV : Dénomination des points homologues

vue externe de la mâchoire



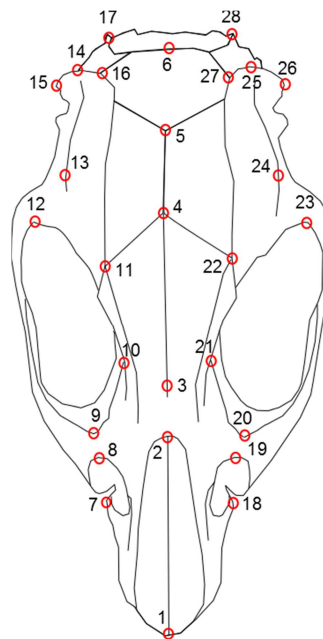
- 1-jonction entre la partie interne de l'incisive et de l'os mandibulaire ;
- 2-point le plus bas de l'os mandibulaire ;
- 3-jonction de la première molaire et de l'os mandibulaire ;
- 4-Point de superposition du Ramus et de la troisième molaire ;
- 5-extrémité de l'apophyse coronoïde ;
- 6-base postérieure de l'apophyse coronoïde, sur l'incisure mandibulaire;
- 7-extrémité antérieure du cartilage de l'apophyse condyalaire ;
- 8-extrémité postérieure du cartilage de l'apophyse condyalaire ;
- 9-élargissement de l'apophyse condyalaire ;
- 10-point le plus antérieur du Ramus postérieur ;
- 11-apex de l'apophyse angulaire ;
- 12-point le plus bas du Ramus ;
- 13-jonction du Ramus, de l'os mandibulaire et de la marge mastoïde ;
- 14-point le plus bas de l'os mandibulaire ;
- 15- jonction entre la partie externe de l'incisive et de l'os mandibulaire.

## Vue interne de la mâchoire



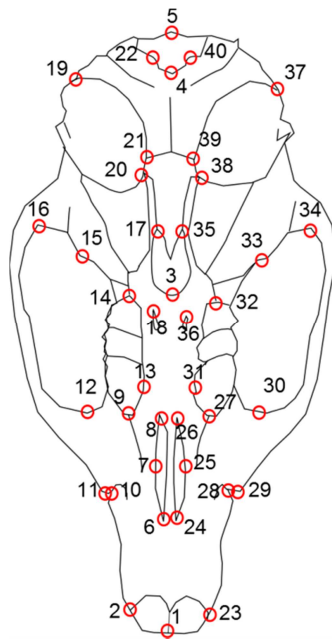
- 1-jonction entre la partie interne de l'incisive et de l'os mandibulaire ;
- 2-point le plus bas de l'os mandibulaire ;
- 3-jonction de la première molaire et de l'os mandibulaire ;
- 4- jonction de la première et de la seconde molaire ;
- 5- jonction de la seconde et de la troisième molaire ;
- 6-base postérieure de la troisième molaire ;
- 7- extrémité de l'apophyse coronoïde ;
- 8- base postérieure de l'apophyse coronoïde sur l'incisure mandibulaire ;
- 9-extrémité antérieure du cartilage de l'apophyse condyloïde ;
- 10-extrémité postérieure du cartilage de l'apophyse condyloïde ;
- 11-élargissement de l'apophyse condyloïde ;
- 12-point le plus antérieur du Ramus postérieur ;
- 13-apex de l'apophyse angulaire ;
- 14-point le plus bas du Ramus ;
- 15-point le plus haut de la limite inférieure du Ramus;
- 16-point le plus bas de l'os mandibulaire ;
- 17- jonction entre la partie externe de l'incisive et de l'os mandibulaire ;

### Vue dorsale du crâne



- 1-apex du rostre ;
- 2-limite postérieure du rostre ;
- 3-limite antérieure du bulbe rachidien ;
- 4-jonction des sutures frontales et pariétales ;
- 5-jonction des sutures pariétales et interpariétales ;
- 6-point le plus postérieur de l'interpariétal ;
- 7-point le plus externe du renflement frontal du maxillaire ;
- 8-point le plus postérieur du foramen infra-orbital ;
- 9-base de l'apophyse lacrymale ;
- 10-forme avec 21 la plus petite largeur inter-orbitale ;
- 11-point de jonction entre la limite externe frontale et suture antérieure pariétale ;
- 12-point le plus postérieur de l'orbite, base de l'apophyse squamosal ;
- 13-forme avec 24 la plus grande largeur de la boîte crânienne ;
- 14-point le plus postérieur de la crête supra-occipitale ;
- 15-point le plus externe de la crête supra-occipitale ;
- 16-jonction de la crête supra-occipitale et de l'os interpariétal ;
- 17- condyle exoccipital
- 18-28- symétriques des points précédents par l'axe de symétrie du crâne.

Vue ventrale du crâne

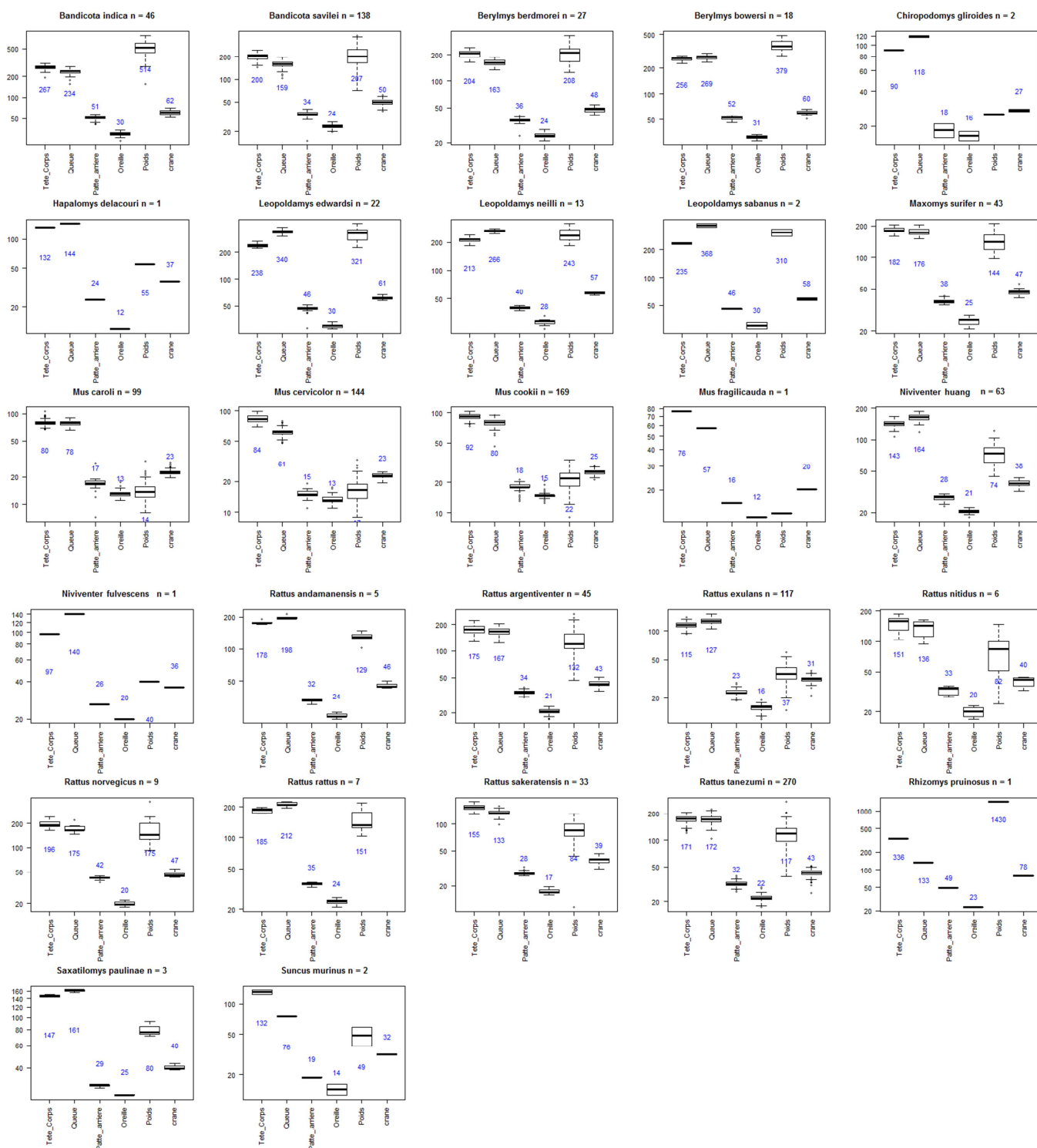


- |   |  |
|---|--|
| 1-jonction des incisives ;  | 14-limite postérieure de la troisième molaire ;  |
| 2-jonction émail, dentine os pré-maxillaire ;                                       | 15-jonction de l'os squamosal et de l'alisphénoïde;  |
| 3-limite postérieure du palais ;  | 16- Apophyse zygomatique de l'os squamosal, point le plus postérieur de l'orbite;                            |
| 4-partie la plus antérieure du foramen magnum ;                                     | 17- foramen mésoptéridoïde ;   |
| 5-partie la plus postérieure du foramen magnum ;                                    | 18- foramen palatin ;  |
| 6-partie antérieure des foramens incisifs;  | 19-jonction de la bulle tympanique et de l'os basioccipital;   |
| 7-forme avec 25 la plus grande largeur des foramens incisifs ;                      | 20-base externe de l'os du tube eustachien, il forme avec 19 la plus grande longueur de la bulle tympanique; |
| 8- partie postérieure des foramens incisifs;  | 21-jonction entre la bulle tympanique, le basisphénoïde et le basisoccipital;                                |
| 9-limite antérieure de la première molaire ;  | 22-renflement du condyle exoccipital ;   |
| 10-limite la plus interne du renflement maxillaire;                                 | 23-40- symétriques des points précédents par l'axe de symétrie du crâne.                                     |
| 11-point le plus antérieur de l'arche zygomatique ;                                 |  |
| 12-point le plus antérieur de l'orbite ;  |  |
| 13-forme avec 31 la plus petite largeur du palais au niveau de la première molaire; |  |



## Annexe V

Représentation des mesures et du poids des animaux adultes échantillonnés durant le programme CERoPath. La valeur médiane est donnée en bleu pour chaque trait mesuré. Les mensurations de *Suncus murinus* sont données à titre indicatif car l'espèce fait partie de l'échantillonnage du programme.



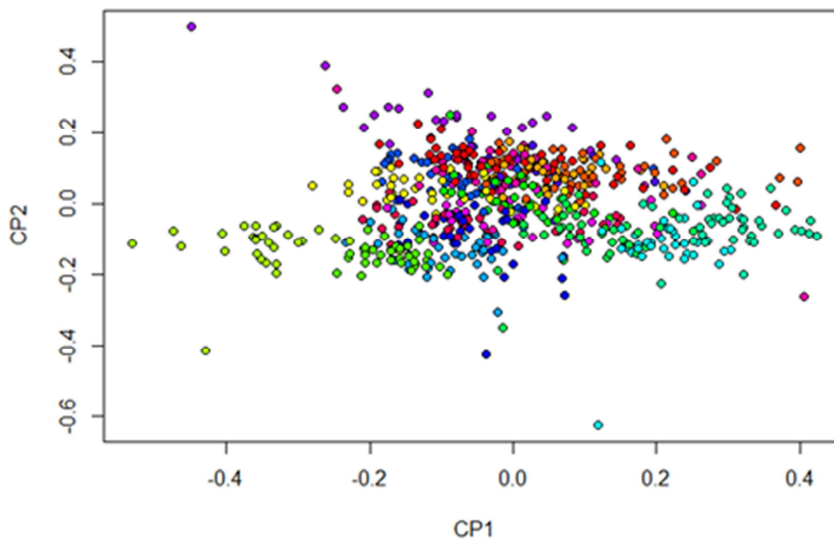


## Annexe VI

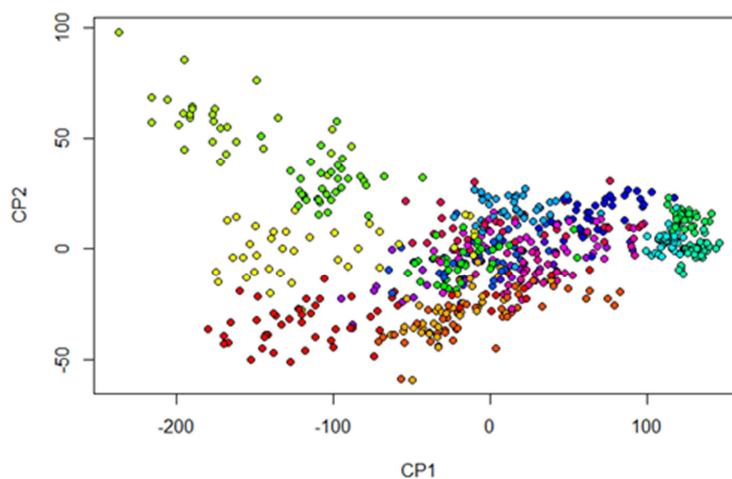
Représentation du plan principal de l'ACP pour les données corporelles brutes (a), transformées en logshape ratio (b), des prises de vues des conformations en face ventrale (c), en face dorsale (d), en vue internes pour les hémimandibules(e) en vue externe pour les hémimandibules (f). Les codes couleurs des espèces sont les même que pour la représentation des ellipses et donnés dans la légende.

- *Bandicota indica*
- *Bandicota savilei*
- *Berylmys berdmorei*
- *Berylmys bowersi*
- *Leopoldamys edwardsi*
- *Leopoldamys neilli*
- *Maxomys surifer*
- *Mus caroli*
- *Mus cervicolor*
- *Mus cookii*
- *Niviventer fulvescens*
- *Rattus argentiventer*
- *Rattus exulans*
- *Rattus nitidus*
- *Rattus norvegicus*
- *Rattus sakeratensis*
- *Rattus tanezumi*

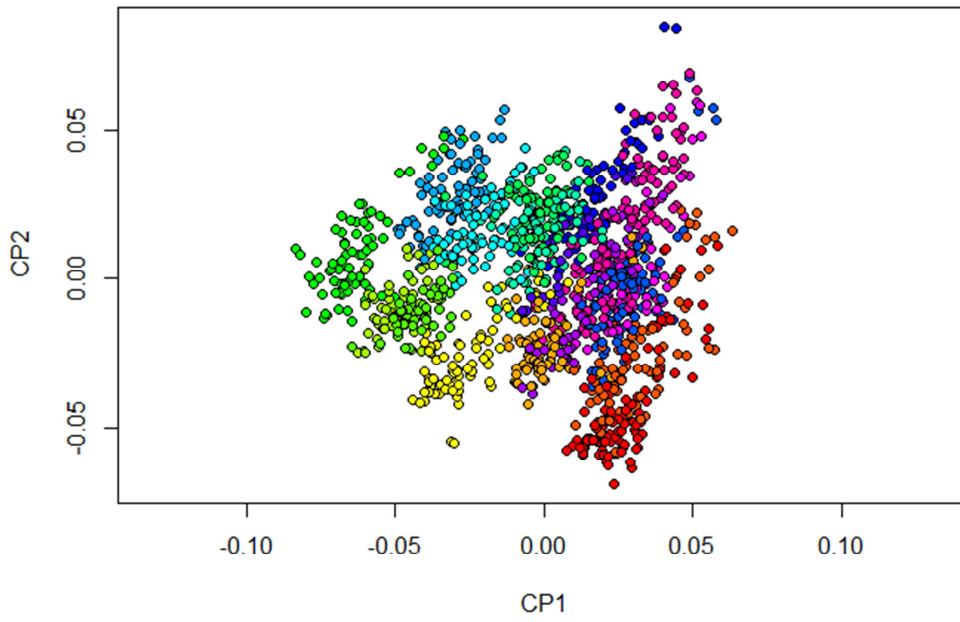
ACP des données corporelles brutes



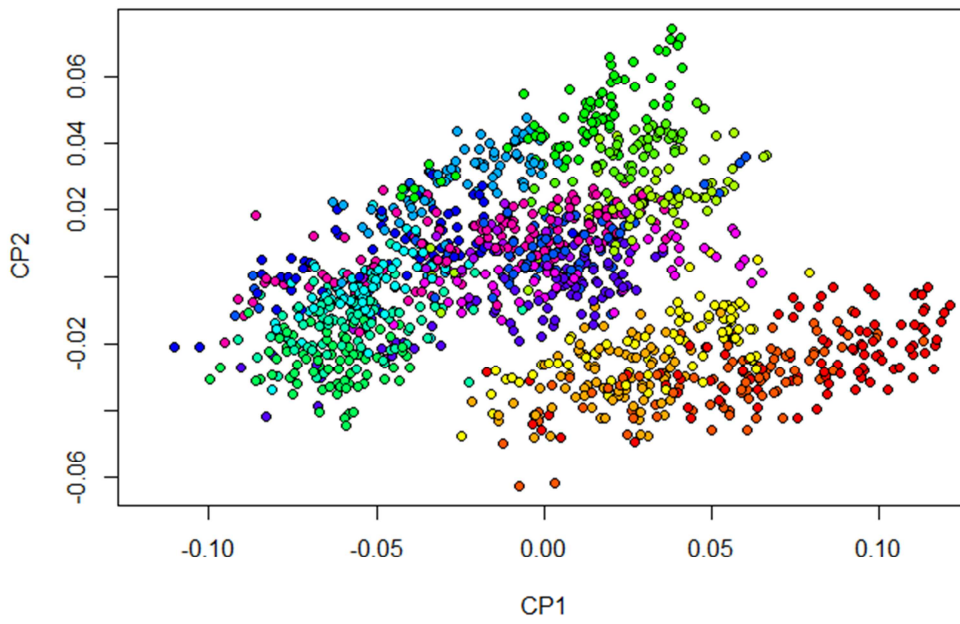
ACP des logshape-ratio



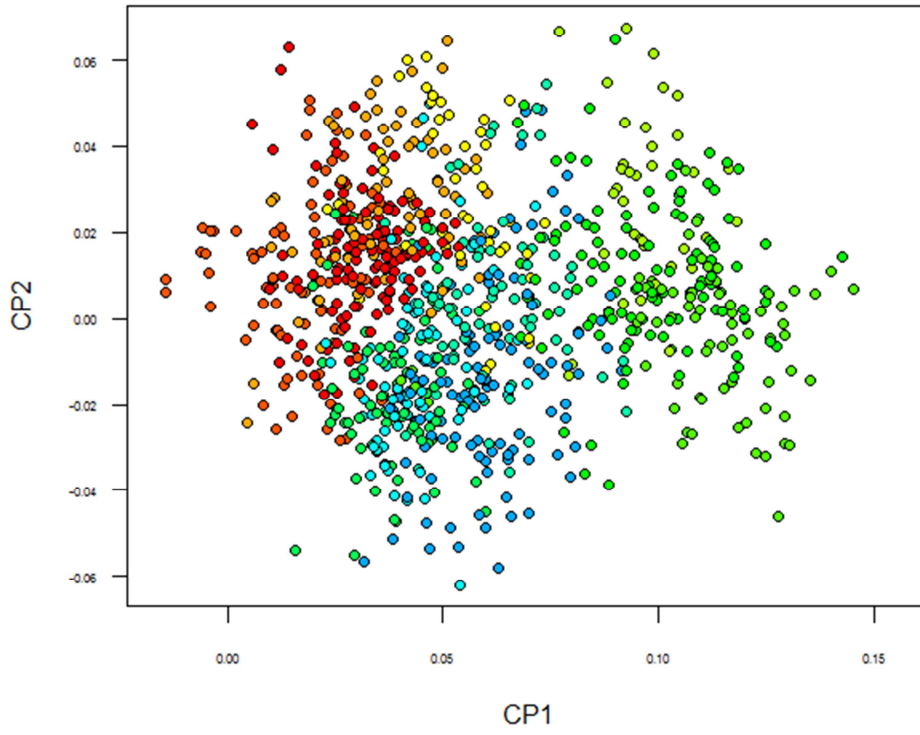
**ACP sur les prises de vues des conformations en vue ventrale**



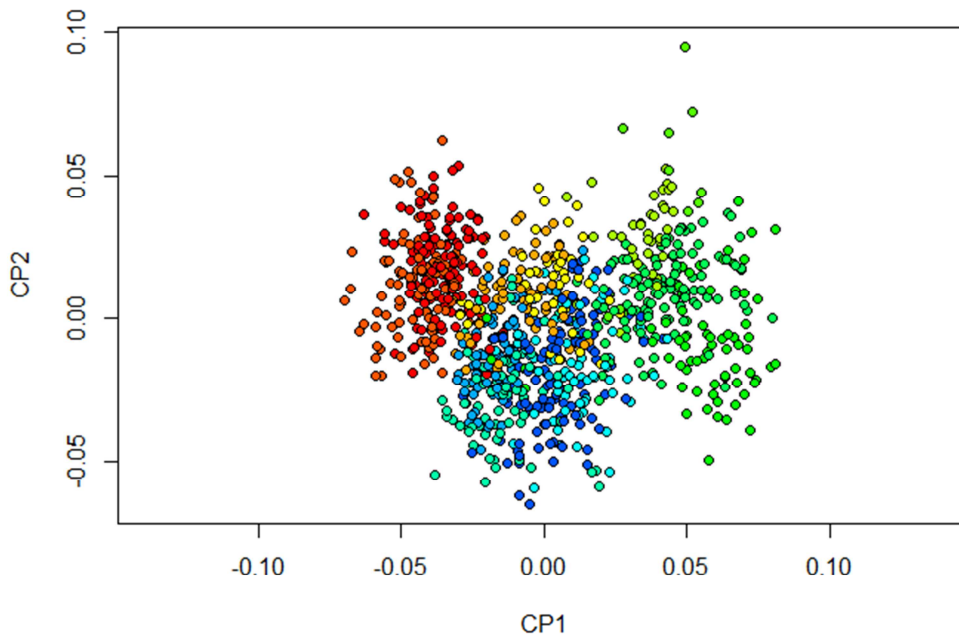
**ACP sur les prises de vues des conformations en vue dorsale**



ACP sur les prises de vues des conformations des mâchoires en vue interne



ACP sur les prises de vues des conformations des mâchoires en vue externe





## Annexe VII

Article soumis.

### **From the field to the laboratory and collection: a taxonomic toolkit to accurately identify *Bandicota indica* and *Bandicota savilei*.**

**(Rodentia-Muridae, genus *Bandicota* Gray, 1873)**

Chaval Yannick<sup>1</sup>, Waengsothorn Surachit<sup>5</sup>, Morand Serge<sup>235</sup>, Cosson Jean François<sup>1</sup>, Claude Julien<sup>2</sup>.

<sup>1</sup>INRA, UMR CBGP (INRA/IRD/Cirad/Montpellier SupAgro), Campus International de Baillarguet, F-34988 Montferrier-sur-Lez, France.

<sup>2</sup>CNRS - Université Montpellier 2 - IRD, UMR ISEM (Institut des Sciences de l'Evolution), F-34095 Montpellier, France.

<sup>3</sup>CIRAD, UR AGIRs, F-34398 Montpellier, France.

<sup>4</sup>Thailand Institute of Scientific and Technological Research (TISTR), Pathum Thani 12120, Thailand.

<sup>5</sup>Department of Parasitology, Faculty of Veterinary Sciences Kasetsart University, Bangkok, Thailand.

Keywords: species identification, primers specific, field criteria, morphometrics, *B. indica*, *B. savilei*.

Corresponding author: Chaval Yannick, [chaval@supagro.inra.fr](mailto:chaval@supagro.inra.fr)

Abstract (277 words) Bandicoot rats (Rodentia, Muridae) are widespread species in Asia where they live in close contact with human populations and can occasionally produce significant crop damage. All over their distribution bandicoot rats are of high sanitary importance as reservoirs of at least ten human diseases. In Thailand, Cambodia, Laos and Vietnam, two species coexists, the great Bandicoot, *Bandicota indica* (Bechstein, 1800), and the lesser Bandicoot, *Bandicota savilei* (Thomas, 1916). Species distinction based on external characters is often difficult and may lead to a certain proportion of species misidentification. Thus, a reliable, fast and inexpensive identification at the specific level of specimens collected in the field is in need to develop effective strategies for population management or to map sanitary risks during epidemiologic surveys. The purpose of this paper is to evaluate the efficiency of classical methods for identifying these bandicoot species and provide new, simple and cheap tools for discriminating both species based on molecular, morphological ratio and morphometrics variation. Two PCR-test are developed in this study: a first (PCR1) for discriminating between *Bandicota* and *Rattus* genera; and a second (PCR2) to discriminate between *B. indica* and *B. savilei* species. In addition, we evaluated two morphometric data sets for discrimination between species. a) one based on the ratio between head and body and hind foot length proposed by Aplin and collaborators (2003a); and b) a second one based on landmark analysis of mandible forms. Accuracy of both morphological tests was fairly high and better for the landmark analysis than the traditional one. Altogether the three identification tools would probably be useful for a large scientific community, including museologists, agronomists, ecologists and epidemiologists willing to ensure their species identifications.





## Annexe VIII

Liste des individus possédant une identification morphologique différente de l'assignation mitotypique pour l'une des structures ou l'une des vues étudiées. L'individu surligné en rouge est celui pour lequel toutes les identifications morphologiques sont fausses, en orange les individus pour lesquels toutes les identifications crâniennes sont fausses. Les cases vertes signifient que l'identification est congruente avec l'identification moléculaire pour la vue considérée.

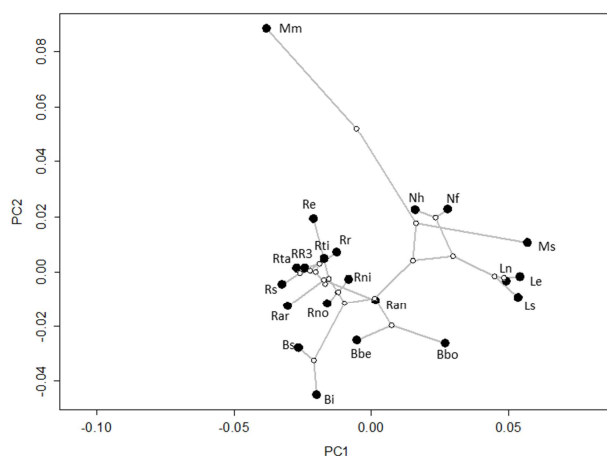
code individu	mitotype	Mâchoire		Crâne	
		interne	externe	dorsale	ventrale
vue	identification vue	identification vue	identification vue	identification vue	identification vue
C0044	<i>Rattus argentiventer</i>	NA	NA		
C0322	<i>Niviventer fulvescens</i>			<i>Rattus tanezumi</i>	
C0448	<i>Niviventer fulvescens</i>			<i>Rattus tanezumi</i>	
C0478	<i>Maxomys surifer</i>	NA	NA	<i>Niviventer fulvescens</i>	
C0488	<i>Rattus argentiventer</i>	<i>Rattus tanezumi</i>	<i>Rattus tanezumi</i>	<i>Rattus losea</i>	
C0492	<i>Rattus argentiventer</i>	NA	NA	<i>Rattus tanezumi</i>	<i>Rattus tanezumi</i>
C0504	<i>Rattus phylogenetic R3</i>	<i>Rattus argentiventer</i>		<i>Rattus argentiventer</i>	<i>Rattus argentiventer</i>
C0506	<i>Rattus argentiventer</i>			<i>Rattus tanezumi</i>	<i>Rattus tanezumi</i>
C0514	<i>Rattus argentiventer</i>			<i>Rattus tanezumi</i>	<i>Rattus tanezumi</i>
C0532	<i>Rattus argentiventer</i>				<i>Rattus tanezumi</i>
C0547	<i>Rattus argentiventer</i>			<i>Rattus tanezumi</i>	<i>Rattus tanezumi</i>
C0569	<i>Rattus phylogenetic R3</i>	NA	NA	<i>Rattus argentiventer</i>	<i>Rattus argentiventer</i>
C0581	<i>Rattus argentiventer</i>	<i>Rattus tanezumi</i>	<i>Rattus tanezumi</i>	<i>Rattus tanezumi</i>	<i>Rattus tanezumi</i>
C0582	<i>Maxomys surifer</i>	<i>Rattus argentiventer</i>	<i>Rattus argentiventer</i>		
C0693	<i>Rattus phylogenetic R3</i>	<i>Rattus argentiventer</i>			
L0048	<i>Rattus tanezumi</i>	NA	NA	<i>Niviventer fulvescens</i>	
L0158	<i>Rattus tanezumi</i>			<i>Rattus argentiventer</i>	
L0181	<i>Rattus nitidus</i>	NA	NA	<i>Rattus tanezumi</i>	
L0265	<i>Rattus losea</i>		<i>Rattus argentiventer</i>		
L0266	<i>Rattus losea</i>	NA			<i>Rattus tanezumi</i>
R4182	<i>Rattus tanezumi</i>	<i>Rattus losea</i>	<i>Rattus losea</i>		
R4230	<i>Rattus losea</i>	<i>Rattus tanezumi</i>			
R4245	<i>Rattus losea</i>				<i>Rattus tanezumi</i>
R4250	<i>Rattus losea</i>		<i>Rattus tanezumi</i>		
R4265	<i>Bandicota indica</i>	NA		<i>Bandicota savilei</i>	
R4324	<i>Bandicota indica</i>	NA	<i>Bandicota savilei (Chaval et al., in prep)</i>	<i>Bandicota savilei</i>	<i>Bandicota savilei</i>
R4436	<i>Rattus tanezumi</i>				<i>Rattus argentiventer</i>
R4447	<i>Rattus tanezumi</i>	<i>Rattus losea</i>	<i>Rattus losea</i>		
R4471	<i>Leopoldamys neilli</i>		<i>Leopoldamys edwardsi</i>		
R4542	<i>Rattus tanezumi</i>	<i>Rattus argentiventer</i>			
R4764	<i>Mus cookii</i>	NA	NA	<i>Mus cervicolor</i>	
R4831	<i>Bandicota indica</i>	NA	<i>Bandicota savilei (Chaval et al., in prep)</i>	<i>Bandicota savilei</i>	<i>Bandicota savilei</i>
R4895	<i>Berylmys bowersi</i>	NA	NA	<i>Berylmys berdmorei</i>	
R4983	<i>Berylmys berdmorei</i>	<i>Berylmys bowersi</i>		<i>Berylmys bowersi</i>	<i>Berylmys bowersi</i>
R4995	<i>Leopoldamys edwardsi</i>			<i>Leopoldamys neilli</i>	<i>Leopoldamys neilli</i>
R4996	<i>Rattus tanezumi</i>	NA	NA	<i>Rattus argentiventer</i>	
R5046	<i>Rattus losea</i>	<i>Rattus argentiventer</i>			
R5071	<i>Rattus losea</i>	NA	NA	<i>Rattus tanezumi</i>	<i>Rattus tanezumi</i>
R5090	<i>Rattus losea</i>	<i>Rattus argentiventer</i>	<i>Rattus argentiventer</i>		
R5098	<i>Leopoldamys neilli</i>				<i>Leopoldamys edwardsi</i>
R5184	<i>Rattus losea</i>	NA	NA		<i>Rattus tanezumi</i>
R5196	<i>Rattus losea</i>		<i>Rattus argentiventer</i>		
R5202	<i>Niviventer fulvescens</i>	<i>Rattus exulans</i>	<i>Rattus exulans</i>		
R5211	<i>Rattus phylogenetic R3</i>	NA	NA	<i>Rattus nitidus</i>	
R5330	<i>Rattus tanezumi</i>				<i>Rattus argentiventer</i>
R5348	<i>Berylmys berdmorei</i>	<i>Berylmys berdmorei</i>	<i>Berylmys berdmorei</i>		
R5537	<i>Rattus tanezumi</i>	NA	NA	<i>Rattus nitidus</i>	
R5592	<i>Mus cervicolor</i>	NA	NA	<i>Mus caroli</i>	
R5555	<i>Rattus losea</i>				<i>Rattus tanezumi</i>
R5600	<i>Rattus phylogenetic R3</i>		<i>Rattus argentiventer</i>		
R5607	<i>Mus cervicolor</i>	NA	NA		<i>Mus cookii</i>
R5621	<i>Mus caroli</i>	NA	NA	<i>Mus cervicolor</i>	<i>Mus cervicolor</i>
R5623	<i>Mus caroli</i>	NA	NA	<i>Mus cervicolor</i>	
R5669	<i>Mus caroli</i>	NA	NA	<i>Mus cervicolor</i>	
R5679	<i>Rattus argentiventer</i>	<i>Rattus tanezumi</i>	<i>Rattus tanezumi</i>		<i>Rattus tanezumi</i>
R5848	<i>Mus caroli</i>	NA	NA	<i>Mus cervicolor</i>	
R5893	<i>Mus cervicolor</i>	NA	NA	<i>Mus cookii</i>	
R5908	<i>Mus cervicolor</i>	NA	NA	<i>Mus cookii</i>	
R5953	<i>Mus caroli</i>	NA	NA	<i>Mus cervicolor</i>	
R5995	<i>Rattus losea</i>	NA	NA	<i>Rattus tanezumi</i>	
R6005	<i>Niviventer fulvescens</i>			<i>Rattus tanezumi</i>	
R6019	<i>Rattus phylogenetic R3</i>	<i>Rattus argentiventer</i>			
R6068	<i>Mus cookii</i>	NA	NA		<i>Mus caroli</i>
R6176	<i>Rattus argentiventer</i>		<i>Rattus tanezumi</i>		
R6177	<i>Rattus argentiventer</i>	<i>Rattus tanezumi</i>			
R6208	<i>Rattus argentiventer</i>		<i>Rattus losea</i>		
R6220	<i>Rattus losea</i>	<i>Rattus tanezumi</i>			<i>Rattus tanezumi</i>
R6234	<i>Rattus phylogenetic R3</i>	<i>Rattus argentiventer</i>			<i>Rattus argentiventer</i>
R6235	<i>Rattus phylogenetic R3</i>	<i>Rattus argentiventer</i>	<i>Rattus argentiventer</i>		
R6237	<i>Bandicota indica</i>	NA		<i>Bandicota savilei</i>	
R6259	<i>Rattus tanezumi</i>		<i>Rattus losea</i>		
R6273	<i>Bandicota indica</i>	<i>Bandicota savilei</i>			
R6350	<i>Mus caroli</i>	NA	NA	<i>Mus cervicolor</i>	
R6426	<i>Mus cervicolor</i>	NA	NA	<i>Mus cookii</i>	
R6541	<i>Bandicota savilei</i>	NA		<i>Bandicota indica</i>	
R6548	<i>Mus cervicolor</i>	NA	NA	<i>Mus cookii</i>	
R6556	<i>Mus caroli</i>	NA	NA		<i>Mus cookii</i>
P00526	<i>Leopoldamys neilli</i>	NA	NA		<i>Leopoldamys edwardsi</i>
P00532	<i>Leopoldamys edwardsi</i>	NA	NA		<i>Leopoldamys neilli</i>
P00533	<i>Leopoldamys edwardsi</i>	NA	NA	<i>Leopoldamys neilli</i>	<i>Leopoldamys neilli</i>
P00543	<i>Leopoldamys neilli</i>	NA	NA	<i>Leopoldamys edwardsi</i>	<i>Leopoldamys edwardsi</i>



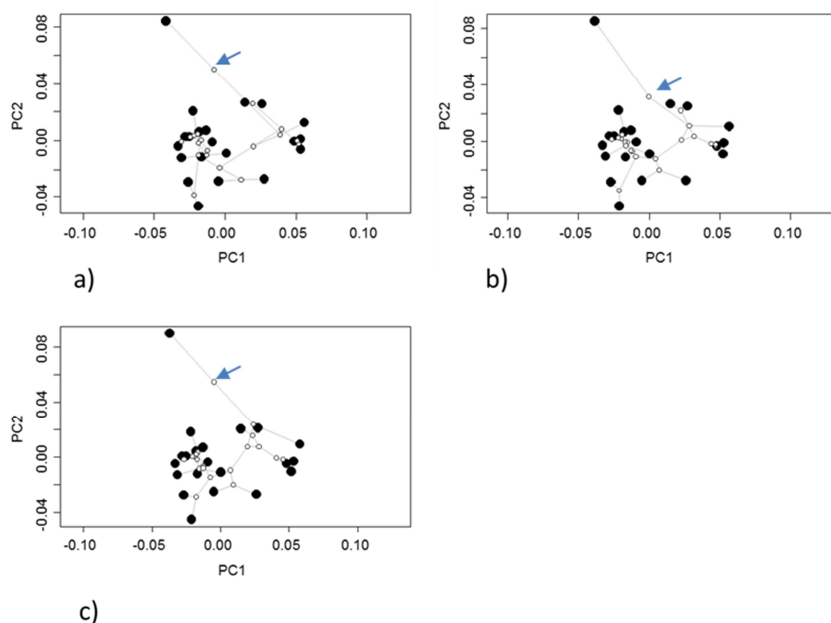
## Annexe IX

### Exemples de phylomorfo espaces reconstruits à partir de plusieurs méthodes d'estimation des caractères ancestraux

1- Représentation du phylomorfo espace de la vue palatine du crâne des rongeurs de la tribu des Rattini. Les ancêtres communs sont reconstruits avec la méthode présentée par Claude (2008) et sont représentés par de petits points blancs cerclés de noir. Les branches de la phylogénie sont matérialisées par des traits gris. Les espèces actuelles sont représentées par des points noirs et leur nom est codé avec la première lettre du genre en majuscule suivie d'une ou deux lettres du nom d'espèce en minuscule (voir le tableau 3 partie 2 pour une correspondance avec les noms d'espèces et leurs abréviations).



2- Représentation du phylomorfo espace de la vue palatine du crâne des rongeurs de la tribu des Rattini. Les ancêtres communs sont reconstruits avec la méthode a) Phylogenetically Independent Contrasts, b) Generalized least squared c) squared change parcimony. Les espèces sont positionnées au même endroit que sur la figure précédente. La flèche bleue indique l'ancêtre commun.





## Glossaire

**ACP:** L'analyse en composante principale est une méthode d'analyse de données qui permet la description des données contenues dans un tableau individus/variables avec des variables sous forme numérique. L'analyse crée des variables composites, les composantes principales, à partir des variables originales dans le but de maximiser la variance expliquée par chacune d'elle.

**ALD:** L'Analyse linéaire discriminante est une méthode d'analyse de données qui permet la discrimination de groupes définis à priori à partir de variables sous forme numérique. L'analyse décompose la variance totale en variance intra-groupe et en variance inter-groupe et définit des fonctions discriminantes en minimisant la première tout en maximisant la seconde.

**Allométrie :** désigne une corrélation entre la conformation d'un objet et sa taille. Un exemple classique est le changement de proportions du corps humain lors de la croissance. Lorsque l'homme grandit, son crâne devient proportionnellement plus petit que chez le nourrisson. On dit alors que la croissance du corps de l'homme est allométrique. À l'opposé, certains organismes possèdent une croissance dite isométrique durant laquelle aucun changement de proportion corporelle n'est observable. L'allométrie peut être également vue comme la variation d'un trait au sein des individus d'une même cohorte, on parle alors d'allométrie statique ou comme variation liée aux relations phylogénétiques entre les espèces, on parle alors d'allométrie inter-spécifique ou d'allométrie d'évolution.

**ANOVA:** ANalysis Of Variance est une analyse qui permet de vérifier si plusieurs échantillons (ici des mitotypes ou des espèces) sont extraits d'une même population (au sens statistique).

**Cladistique :** méthode de classification phylogénétique du vivant où les taxons doivent être monophylétiques.

**Correction de Bonferroni:** correction du seuil de significativité d'un test statistique tenant compte du nombre de comparaisons effectuées. Celles-ci sont susceptibles d'augmenter les erreurs de type I.

**Cryptique :** des espèces cryptiques ne sont pas distinguables morphologiquement.

**Dérive :** évolution d'une population d'individus causée par les phénomènes aléatoires.

**Diptérocarpes :** genre d'arbre de la famille des Dipterocarpaceae caractéristique des forêts sèches caducifoliées du sud-est asiatique.

**Division:** rang taxonomique non officiel défini comme un taxon intermédiaire entre la tribu et le genre. La division regroupe des genres possédant de fortes affinités phylogénétiques.

**Ecart-type:** racine carrée de la variance, l'écart-type mesure la dispersion par rapport à la valeur centrale d'un ensemble de données.

**Erreur de type I :** L'erreur de type I consiste, à rejeter l'hypothèse  $H_0$  alors que celle-ci est vraie.

**Erreur de type II :** L'erreur de type II consiste, à accepter l'hypothèse  $H_0$  alors que celle-ci est fautive.

**Espèce biologique:** concept le plus utilisé pour définir l'espèce comme des groupes de populations naturelles, effectivement ou potentiellement interfécondes, donnant naissance à une progéniture féconde et qui sont génétiquement isolées d'autres groupes similaires.

**estimate.missing{geomorph}:** estimate.missing est la fonction utilisée et geomorph est le package dans lequel se trouve la fonction.

**Homologie :** deux structures sont dites homologues (au sens de l'homologie primaire) lorsque, prises chez des êtres différents, elles conservent le même plan d'organisations fondamentales et les mêmes connexions avec les structures avoisinantes. L'homologie secondaire ou homologie de filiation est la similarité chez une ou plusieurs espèces de caractères lorsque l'on peut présumer que cette correspondance provient de l'héritage d'un ancêtre commun.

**Homoplasie :** similarité d'état de caractères chez plusieurs espèces lorsque l'on peut présumer que cette similarité ne provient pas d'un ancêtre commun.

**IC99 % :** Intervalle de Confiance à 99 %. 99% représente le degré de confiance de l'intervalle. Si on répète un grand nombre de fois la mesure en considérant que la valeur se trouve dans l'intervalle ainsi défini, on aura en moyenne 1 % de chance de se tromper.

**IDH :** l'Indice de Développement Humain est un indice composite mesurant le niveau moyen atteint dans trois dimensions essentielles du développement humain : santé et longévité, accès à l'éducation et niveau de vie décent (voir <http://hdr.undp.org/fr/content/l'indice-du-développement-humain-idh>).

**MANOVA :** Multivariate ANalysis Of Variance. Analyse de variance prenant en compte la variance simultanée de plusieurs traits.

**Ontogénique :** qui est lié à l'ontogénèse c'est-à-dire au développement des organismes.

**Procédure de Burnaby :** méthode statistique de « correction par la taille » qui permet la comparaison entre groupes libérée des relations d'allométrie. La procédure proposée par Burnaby élimine l'effet de la croissance d'un jeu de données multivarié en projetant les points dans un sous-espace orthogonal au vecteur de croissance. Ce vecteur est généralement la première composante principale de la matrice de covariance intra-groupe.

**Plasticité phénotypique :** la plasticité phénotypique est la capacité d'un génome donné à générer des phénotypes diverses en fonction des conditions environnementales.

**Plésiomorphes :** similarité d'état de caractères chez plusieurs espèces hérité d'un ancêtre commun.

**Paraphylétique :** groupe phylogénétique qui comprend une espèce ancestrale et seulement une partie de ses descendants.

**Pseudo-réplication :** effet lié à l'application de statistiques inférentielles sur des données non répliquées ou des réplicats non-indépendantes (cas des espèces phylogénétiquement proches).

**Puissance statistique :** La puissance statistique est la probabilité de rejeter l'hypothèse H0 (les groupes testés sont équivalents pour les variables explicatives testées) alors que cette hypothèse est fausse.

**Répétabilité :** la répétabilité d'une mesure est l'erreur de mesure en l'absence de variation des conditions de mesure. Exemple : le placement de points homologues par deux opérateurs différents à partir du même jeu de photographies.

**Reproductibilité :** la reproductibilité d'une mesure est l'erreur de mesure qui inclue la variation des conditions de mesure. Exemple : le placement de points homologues par deux opérateurs différents à partir de leur jeu de photographies propre réalisé sur les mêmes objets biologiques.

**Roboviroses:** « Rodent Borne Viruses ». Terme qui regroupe l'ensemble des zoonoses virales transmises à l'homme par les rongeurs. Les virus concernés appartiennent aux familles des Hantavirus et des Arenavirus.

**Synapomorphie :** en phylogénie la synapomorphie est un état de caractère dérivé hérité de l'ancêtre commun (une apomorphie) et présent chez des espèces sœurs.

**Temps de coalescence :** temps écoulé (en nombre de générations) entre les différents haplotypes d'un gène et leur ancêtre commun. Plus le nombre d'haplotypes est grand dans une population et plus le temps de coalescence sera important.

**TPS :** Thin Plate Spline représentation des grilles de déformation d'Arcy Thompson (1917) qui sont une quantification des déformations locales produites lors d'un changement dans la conformation d'un objet. Leur expression mathématique a été donnée par Bookstein (1998).

**Trace de Pillai :** la trace de Pillai est une approche paramétrique permettant de tester si plusieurs variables continues distinctes sont liées à une variable qualitative à K modalités lorsqu'elles sont considérées avec leurs différentes interactions multivariées, que les matrices de covariances respectives aux K groupes sont égales au sens statistique du terme et que les données suivent une loi multinormale.

La statistique de test de la trace de Pillai est:  $V = t(SCF((SCF+SCR)^{-1}))$

Avec SCF matrice de dispersion du facteur ou inter-groupes et SCR matrice de dispersion résiduelle ou intra-groupe.

**Variance :** mesure de l'étalement équivalant au carré moyen de l'écart de chaque nombre par rapport à la moyenne d'un ensemble de données.

**Zoonoses :** Maladies animales transmissibles à l'homme.







**ÉCOLE PRATIQUE DES HAUTES ÉTUDES**  
**SCIENCES DE LA VIE ET DE LA TERRE**

**Taxonomie intégrative de la tribu des Rattini  
(Rodentia, Muridae) en Asie du Sud-Est**

***Yannick Chaval***

**Soutenu le 25/04/2014**

**RÉSUMÉ**

L'Asie du Sud-Est est un point chaud de la biodiversité et, dans le même temps, cette région est un foyer actif d'émergence ou de réémergence de zoonoses (SRAS, grippe aviaire, leptospirose, roboviroses etc.). Le développement économique sans précédent et la croissance démographique humaine que connaît la zone biogéographique indochinoise sont un moteur de changements considérables des écosystèmes. Dans le contexte du changement global, il est primordial de comprendre comment ces modifications environnementales modifient les interactions entre les communautés d'hôtes et leurs communautés de parasites/pathogènes afin d'appréhender les transformations du paysage épidémiologique. Certaines espèces, qui sont des réservoirs de zoonoses constituent, de par leur mode de vie et leur statut de gibier des modèles particulièrement intéressants à étudier dans ce cadre. C'est le cas, pour la zone étudiée, des rongeurs et plus particulièrement des représentants de deux tribus de la sous-famille des Murinae : les Rattini et les Murini. La première partie de ce mémoire porte sur ces deux tribus, particulièrement diversifiés en Asie du Sud-Est. L'objectif a été d'étudier la faisabilité d'un outil d'identification web-service innovant basé sur la morphométrie. Cet outil viendra compléter des réalisations qui lui sont antérieures à savoir : un guide d'identification de terrain et un premier outil web-service basé sur l'identification moléculaire. Ces productions formeront un ensemble cohérent d'outils d'identification permettant d'offrir à la communauté des personnes désireuses d'étudier les rongeurs de la zone des accès multiples et adaptés à une identification de qualité. Pour la réalisation de cet outil d'identification morphométrique, nous avons tout d'abord étudié quels sont les facteurs intra et inter-spécifiques qui influencent la variation morphologique des espèces de rongeurs. Cette étude, au travers d'une approche de systématique intégrative a ainsi contribué à la stabilisation de la taxonomie des Rattini. L'étude morphométrique est entièrement basée sur un jeu de données de six cent soixante individus qui appartiennent à dix-sept espèces. L'ensemble des individus a fait au préalable l'objet d'une identification moléculaire réalisée à l'aide d'un gène mitochondrial. Dans le cadre de la création du web-service morphométrique, les sources d'erreurs relatives à la comparaison de jeux de données produits par des opérateurs différents sont identifiées et encadrées car elles pourraient compromettre l'identification des individus. Des seuils d'assignation sont également définis pour permettre à l'outil de gérer la confrontation avec des espèces qui ne sont pas représentées dans le jeu de données de référence. Dans un second temps nous avons souhaité replacer la disparité morphologique observée chez les espèces actuelles dans un contexte évolutif. L'objectif était de comprendre les patrons macro-évolutifs qui ont potentiellement gouverné la mise en place de la variation morphologique et ainsi d'évaluer le rôle de l'environnement sur celle-ci.

**MOTS-CLÉS : *Rattini, Murini, taxonomie intégrative, code barre moléculaire, analyse Procrustes, identification, web-service.***