L'Amélogénine, protéine majeure de l'émail dentaire.

Sidney Delgado

▶ **To cite this version:**

Sidney Delgado. L'Amélogénine, protéine majeure de l'émail dentaire.. Sciences du Vivant [q-bio]. Université paris 7 - paris diderot, 2002. Français. NNT : . tel-04157479

HAL Id: tel-04157479

https://hal.sorbonne-universite.fr/tel-04157479

Submitted on 10 Jul 2023

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

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

Thèse de doctorat de l'Université PARIS VII
Université Denis Diderot

Présentée par **Christophe Sidney DELGADO**

pour obtenir le grade de

DOCTEUR ES-SCIENCES de l'Université Paris VII

L'Amélogénine, protéine majeure de l'émail dentaire.

Origine, analyses évolutive et phylogénétique chez les Amniotes et recherche de son expression lors de la formation des dents de *Chalcides viridanus* (Squamate, Scincidé)

Soutenue le 19 décembre 2002
devant le jury composé de:

Dr Ariane Berdal	Examinateur
Prof. Ann Huysseune	Rapporteur
Prof. Marc Girondot	Examinateur
Prof. Michel Goldberg	Examinateur
Prof. Henry Magloire	Rapporteur
Dr. Jean-Yves Sire	Directeur de thèse

Équipe "Évolution et Développement du Squelette Dermique", UMR CNRS 8570,
2 Place Jussieu, 75251 Paris Cedex 05

“Nothing in biology makes sense except in the light of evolution“.

(«Rien n'a de sens en biologie qu'à la lumière de l'évolution»).

Theodosius Dobzhansky (1900-1975), 1973.

RÉSUMÉ

Cette thèse s'inscrit dans le cadre d'un vaste programme visant à étudier l'origine évolutive des tissus dentaires et les tissus apparentés. L'amélogénine représentant 90% de la matrice organique de l'émail dentaire en formation chez les mammifères, elle est apparue être une cible intéressante. Son rôle est mal connu, mais on suppose que cette protéine contrôle l'organisation des cristaux d'hydroxyapatite lors de la minéralisation de l'émail. Le but de ce travail a été de mieux connaître le gène qui code pour cette protéine et d'apporter des informations sur son évolution dans deux lignées de vertébrés, les mammifères et les reptiles. L'étude a été complétée par une analyse du signal phylogénétique que contient le gène de l'amélogénine dans ces deux lignées et par l'étude de son expression lors de l'amélogenèse chez un lézard.

Grâce à l'utilisation d'outils mathématiques, il est tout d'abord montré que le gène de l'amélogénine est âgé de plus de 630 millions d'années. Cette datation a été rendue possible parce que l'exon 2 de l'amélogénine a son origine dans une duplication de l'exon 2 de l'ostéonectine (SPARC). Ce résultat suggère que l'amélogénine était présente à une époque bien antérieure aux gisements fossiles qui renseignent sur les premiers tissus minéralisés recouverts d'émail chez les ancêtres des vertébrés. L'amélogénine devrait donc être présente dans toutes les lignées de vertébrés qui possèdent des dents, mais l'existence du gène n'est connue à ce jour que chez les tétrapodes. Il semble ainsi peu probable de trouver de l'amélogénine chez des espèces qui ne possèdent pas de dents ou d'émail (myxine, lamproies). Il est aussi montré qu'une séquence supposée d'amélogénine de myxine, et publiée dans la littérature, est le résultat d'une erreur.

L'évolution du gène de l'amélogénine chez les mammifères et les reptiles (crocodiliens et squamates) a été étudiée à partir de séquences extraites des bases de données et complétées par le séquençage chez 14 mammifères et 16 reptiles pour obtenir des données bien représentatives. Des séquences ancestrales ont été déduites des comparaisons de séquences. Certaines régions de l'amélogénine ont été bien conservées au cours de l'évolution. Elles contiennent des sites importants pour la fonction de la protéine : peptide signal, régions de protéolyse ou d'épissage alternatif, acides aminés connus, lorsqu'ils sont substitués, pour conduire à une maladie génétique, l'amélogenèse imparfaite. Une grande partie de l'exon 6 est beaucoup plus variable mais caractérisée par de nombreuses prolines et glutamines dont le pourcentage est très conservé chez les mammifères. Cette région possède également des répétitions de trois acides aminés. Certaines de ces répétitions sont très anciennes et se trouvent dans les deux lignées; d'autres, plus récentes, ont été générées indépendamment dans plusieurs groupes de mammifères. Alors que les séquences d'amélogénine de crocodiliens sont très semblables, celles des squamates le sont beaucoup moins en raison, probablement, de l'évolution plus rapide de leur genôme.

L'utilisation de l'exon 6 du gène de l'amélogénine comme outil dans l'étude de la phylogénie des mammifères et des squamates apporte des résultats, pour la plupart, en accord avec les récentes phylogénies moléculaires. Les deux groupes de mammifères, Laurasiathériens et Afrothériens, sont confirmés. L'emplacement des rongeurs est résolu. En revanche, la position des xénarthres est nouvelle. La phylogénie des squamates est aussi largement confirmée, mais elle est moins fiable en raison de l'évolution plus rapide du gène et d'un manque d'échantillonnage dans certains groupes.

Le dernier chapitre de résultat est une étude du développement dentaire chez un lézard, Chalcides viridanus. Il est tout d'abord démontré, par l'étude de la mise en place de la dentition et celle du remplacement dentaire, qu'il est possible de prévoir le stade de développement d'un bourgeon dentaire à partir de l'observation du stade de développement de la dent qui le précède. Ce travail est suivi d'une description, en microscopie photonique et électronique, à balayage et à transmission, des différents stades de développement d'une dent, depuis son initiation jusqu'à son remplacement. Enfin, une sonde d'ADNc d'amélogénine a été préparée pour détecter les ARNm présents dans les améloblastes à différents stades de l'amélogenèse par hybridation in situ. Les résultats montrent que le gène de l'amélogénine est exprimé au cours de la production de la matrice organique de l'émail et ne l'est plus lorsque la minéralisation de l'émail commence.

Mots clés. Mammifères - Reptiles - Dents - Email - Amélogénine - Evolution - Phylogénie moléculaire - Développement dentaire - Hybridation in situ - MET - MEB.

Sommaire

CHAPITRE 1 - INTRODUCTION..... 5

I. Histoire d'une protéine de l'émail dentaire : l'AMÉLOGÉNINE..... 6
II. Biologie moléculaire et Biochimie de l'amélogénine..... 15
III. Autres protéines de la matrice extracellulaire..... 31
IV. Amélogénine comme marqueur dans les analyses d'ADN..... 36
V. Choix de l'amélogénine comme sujet de recherche..... 39

CHAPITRE 2 - L'AMÉLOGÉNINE: origine évolutive, évolution moléculaire et son utilisation en phylogénèse..... 47

Première partie

ÉTUDE D'UNE SÉQUENCE D'AMÉLOGÉNINE DE MYXINE..... 48

Deuxième partie

L'ORIGINE DE L'AMÉLOGÉNINE..... 55
Résumé..... 57
Introduction..... 58
Matériel et méthodes..... 59
Résultats..... 63
Discussion..... 66

Troisième partie

I- ÉVOLUTION MOLÉCULAIRE DE L'AMÉLOGÉNINE CHEZ LES MAMMIFÈRES..... 69
Résumé..... 70
Introduction..... 71
Matériel et méthodes..... 72
Résultats..... 77
Discussion..... 88
Conclusion..... 95

II- ÉVOLUTION MOLÉCULAIRE DE L'AMÉLOGÉNINE CHEZ LES REPTILES..... 96
Résumé..... 97
Introduction..... 98
Matériel et méthodes..... 100
Résultats..... 104
Discussion..... 114
Conclusion..... 118

Quatrième partie..... 119

UN NOUVEL OUTIL MOLÉCULAIRE POUR L'ÉTUDE DE LA PHYLOGÉNÉSE DES MAMMIFÈRES ET DES SQUAMATES : L'AMÉLOGÉNINE

CHAPITRE 3 - AMÉLOGÉNINE ET DÉVELOPPEMENT DENTAIRE..... 138

Première partie

DENTITION ET PATRON DE REMPLACEMENT DENTAIRE CHEZ *CHALCIDES* (SQUAMATA ; SCINCIDAE) 139
Résumé..... 141
Introduction..... 142
Matériel et méthodes..... 143
Résultats..... 145
Discussion Conclusion..... 157

Deuxième partie

EXPRESSION DE L'AMÉLOGÉNINE LORS DU DÉVELOPPEMENT DES DENTS DE *CHALCIDES* 164
Résumé..... 166
Introduction..... 167
Matériels et méthodes..... 168
Résultats..... 171
Discussion-Conclusion..... 188

CHAPITRE 4 - DISCUSSION GENERALE..... 193

CHAPITRE 5 - BIBLIOGRAPHIE..... 204

CHAPITRE 6 - ANNEXES..... 227

ANNEXE 1 - Matériel et méthode détaillés..... 228
1/ Matériel biologique..... 228
2/ Biologie moléculaire..... 231
3/ Analyse de séquences et méthodes phylogénétiques..... 255
ANNEXE 2 - Analyse de quelques problèmes rencontrés..... 260
ANNEXE 3 - Tampons et solutions..... 273
ANNEXE 4 - Présentation de la séquence de *Chalcides sexlineatus*..... 275
ANNEXE 5 - Séquences obtenues dans Genbank..... 276
ANNEXE 6 - Liste des séquences clonées durant la thèse..... 277
ANNEXE 7 - Amorces de PCR..... 279
ANNEXE 8 - Diagramme des logiciels d'analyse moléculaire..... 280

Liste des publications personnelles..... 281

Remerciements

Je remercie chaleureusement mon directeur de thèse, Jean-Yves Sire, Directeur de recherche au CNRS, de m'avoir accueilli dans son équipe et de m'avoir donné un sujet de recherche aussi passionnant. Je le remercie également pour sa patience, ses nombreuses lectures et corrections du manuscrit, ainsi que pour sa disponibilité tout au long de mes années de thèse. Ses idées scientifiques ont toujours été très enrichissantes et ont apporté énormément à l'avancée de mes travaux. Je lui dois également toutes mes connaissances dans le domaine de l'histologie des tissus dentaires.

Cette thèse a été réalisée dans l'équipe “Formations Squelettiques“ dirigée par Armand de Ricqlès, Professeur au Collège de France. Je le remercie sincèrement de sa gentillesse et sa disponibilité et de m'avoir permis de travailler pour le Collège de France comme préparateur puis comme ATER, durant ma thèse. Sans son soutien, je n'aurais peut-être pas été jusqu'au bout.

J'associe à ces remerciements Marc Girondot, Professeur à l'Université de Paris XI, qui à dirigé mon stage de D.E.A et m'a apporté ses compétences en évolution moléculaire qui ont fortement enrichi cette thèse. Je lui suis également reconnaissant pour m'avoir appris à utiliser de nombreux logiciels de biologie moléculaire et de phylogénie et pour m'avoir initié aux techniques de biologie moléculaire, me mettant ainsi le “pied à l'étrier“. Je le remercie enfin pour sa présence dans mon jury de thèse.

Je remercie Henry Magloire, Professeur à l'Université de Lyon I, de m'avoir reçu à plusieurs reprises dans son laboratoire pour y réaliser la technique d'hybridation sur coupe. Sa bonne humeur, sa générosité et son sens de l'hospitalité m'ont permis de passer à chaque fois un très bon séjour à Lyon. Et je n'oublierai pas non plus sa façon de faire la cuisine ! Je le remercie également d'avoir accepté d'être membre de mon jury.

J'exprime toute ma gratitude à Ann Huysseune, Professeur à l'Université de Gand, Belgique. Elle a toujours apprécié mes recherches et m'a souvent encouragé à continuer quand les résultats tardaient à venir. Je la remercie également d'avoir accepté de participer à mon jury de thèse. J'associe à ces remerciements les autres membres de son laboratoire et en particulier Christine Van der Heyden qui est venue souvent travailler chez nous et avec qui j'ai apprécié la collaboration (bravo pour ta thèse et pour ta petite fille).

Je remercie Jacques Castanet, Professeur à l'Université Pierre et Marie Curie, qui a mis à ma disposition son élevage de lézards et dont les connaissances en herpétologie ont été très profitables. Je n'oublierai pas à cette occasion Olivier Babiar, animalier responsable des élevages, avec qui j'ai collaboré pour obtenir de nombreux individus.

Je voudrais exprimer toute ma gratitude à Marie-Lise Couble de l'équipe du professeur Henry Magloire à Lyon, qui m'a appris avec patience et gentillesse, la technique d'hybridation sur coupe. Ses compétences et son expérience dans ce domaine m'ont permis d'obtenir de superbes images d'expression. Je n'oublierais pas dans ces remerciements les autres membres du laboratoire qui m'ont toujours accueilli chaleureusement : Jean-Christophe, Françoise, ... Il est difficile de tous les citer, mais qu'ils sachent que je pense à eux.

Sans Françoise Allizard et sa maîtrise parfaite des techniques d'histologie, je n'aurais jamais eu à ma disposition de si magnifiques coupes de dents de lézard. Qu'elle en soit remerciée ici, ainsi que pour sa gentillesse lorsqu'elle m'a enseigné une partie de son savoir.

Un grand merci aux différents stagiaires de BTS qui sont venus travailler sous ma direction et à qui je dois une partie des séquences d'ADN mammalien présentées ici. Grâce à leur aide, mes travaux ont progressé plus rapidement dans une ambiance très sympathique. Merci donc à Carole et Johanna.

Tous les membres de l'UMR 8570 sont associés à ces remerciements. Merci à Michel Laurin qui m'a apporté ses grandes compétences en cladistique et a souvent corrigé mon anglais. Je n'oublie pas Alexandra Quilhac, ainsi que George Cubo avec qui j'ai eu de passionnantes discussions sur l'évolution. Merci également à Marie Madeleine Loth qui a toujours été de bon conseil et à Hélène Vieillot toujours prête à rendre service.

Je voudrais exprimer toute ma gratitude à Patrick Laurenti pour son aide technique et ses conseils qui m'auront aidé à comprendre certains “avatars“ de la microbiologie et de la biologie moléculaire. Il a toujours été disponible pour venir à mon secours quand les manips refusaient de fonctionner et m'a toujours apporté de bonnes solutions.

Je pense aussi aux autres membres du laboratoire de Biologie du développement de l'Université Paris 7: Christelle et Véronique bien sûr, pour leurs rapports amicaux et Fabien qui, en dehors de son amitié, m'a appris tellement de choses en informatique. Il nous manque beaucoup depuis qu'il est parti au Canada et l'ambiance n'est plus tout à fait la même ! L'humour de Nicolas aura été salutaire pour aider à oublier les problèmes de manips. Qu'ils en soient tous remerciés ici.

Merci à Tiphaine dont la présence et le sourire ont rendu l'ambiance dans la pièce 213 particulièrement agréable. J'espère que les amphibiens te livreront bientôt leurs secrets comme l'ont fait pour moi les reptiles.

J'ai une pensée particulière pour mes amis et compagnons de laboratoire Emmanuel “le canard“ et Grégoire “le microbrachis“. Sans leur amitié et leur soutien, les longues heures passées devant l'ordinateur auraient été bien pénibles ! Souviens toi des 5 synapomorphies de notre clade préféré Grégoire ! Évite les indigestions de crabe Emmanuel ! !

Bien sûr je pense à ma famille et en particulier à ma mère qui a soutenu, mais aussi supporté, mes longues années d'études tels les parents de “Tanguy“.

Enfin, comment ne pas associer à ces remerciements ma merveilleuse épouse qui a toujours été à mes côtés pour me soutenir, pour m'aider dans la traduction de l'anglais et dont les petites mains ont grandement contribué à la rédaction de ce manuscrit. Je t'embrasse très fort pour ta patience particulièrement mise à l'épreuve à la fin de ces travaux, même si tu as toujours compris que la recherche nécessite parfois des sacrifices. Ca y'est, c'est fini !

CHAPITRE 1

–

INTRODUCTION

Le sujet principal de ma thèse est le gène de l'amélogénine. L'amélogénine est une protéine impliquée exclusivement dans la formation de l'émail dentaire. De plus, elle est essentielle à la formation de ce tissu hautement minéralisé et certaines mutations dans le gène de l'amélogénine induisent une maladie génétique : l'amélogenèse imparfaite. A ce titre, l'amélogénine constitue donc un sujet de recherche majeur en médecine dentaire comme on le verra dans cette partie.

J'aborderais dans ce chapitre introductif tout d'abord quelques éléments sur l'histoire de la découverte de l'amélogénine ainsi que celle d'autres protéines de la matrice extracellulaire de l'émail. J'ai rassemblé ensuite les connaissances sur le gène de l'amélogénine et la protéine qu'il code. Enfin, j'aborderai le choix de l'amélogénine comme sujet de recherche dans la dernière partie de cette introduction.

L'originalité de ce travail a été d'analyser l'amélogénine d'un point de vue évolutif et phylogénétique. Le premier chapitre de résultats de cette thèse (Chapitre 2) est d'ailleurs consacré à l'étude comparative puis à l'étude phylogénétique de l'amélogénine au sein deux lignées : la lignée mammalienne, la seule dans laquelle il existait des données quand mes travaux ont commencé et la lignée reptilienne, le groupe le plus proche des mammifères. Avant d'aborder ces sujets, je traiterai le problème de l'origine évolutive du gène de l'amélogénine.

Le second chapitre des résultats (Chapitre 3) est consacré à l'expression du gène de l'amélogénine lors de l'amélogénèse dentaire chez un squamate (*Chalcides*), ce qui constitue le premier travail de ce genre dans une lignée non-mammalienne. En préambule à cette étude et afin de mieux connaître ce modèle, je présenterai les résultats obtenus sur la dentition, le remplacement et le développement dentaire de ce lézard.

I. HISTOIRE D'UNE PROTÉINE DE L'ÉMAIL DENTAIRE : L'AMÉLOGÉNINE

L'arracheur de dents 1523
Gravure sur cuivre originale de ***Lucas de Leyde*** *(1489-1533)*

I.1. Préhistoire…

Le terme "amélogénine" fut utilisé pour la première fois en 1965 par Eastoe pour décrire tous les produits sécrétés dans la matrice extracellulaire par les améloblastes. Mais en réalité, on peut considérer que l'histoire des protéines de l'émail dentaire, dont l'amélogénine est le composant dominant (Eastoe, 1963), avait commencé deux siècles plus tôt, en 1771, lorsque John Hunter, chirurgien anglais et anatomiste, suggère que l'émail dentaire contiendrait « une portion de substance animale ». Il faudra cependant attendre Deakins en 1942 pour pouvoir mesurer, par incinération, la quantité de matière organique présente dans la dent, qu'il estime a 19% du poids total. Il parle alors de « matrice biominéralisante ».

En 1960, les techniques biochimiques permettant alors d'analyser la composition « globale » en acides aminés des protéines, Piez réalise la première étude de ce genre sur la matrice organique de l'émail, à partir d'un extrait d'une troisième molaire humaine qui n'avait pas encore percé. Il conclut que la matrice de l'émail ne contient pas d'hydroxyproline ni d'hydroxylysine et que la protéine qui la compose ne ressemble à aucune autre protéine connue (cf. Tableau 1).

	Émail en formation d'un fœtus humain	Résidus acides insolubles de l'émail mature humain
	Piez (1960)	Hess *et al.* (1953)
Alanine	20,3	111,0
Glycine	65,0	304,0
Valine	39,6	33,8
Leucine	91,3	41,9
Isoleucine	32,7	16,3
Proline	251,0	47,0
Phénylalanine	23,4	45,8
Tyrosine	53,4	6,9
Serine	62,5	75,5
Thréonine	38,1	44,8
demi-cystine	<4,0	6,1
Méthionine	42,3	8,3
Arginine	23,3	47,0
Histidine	64,5	9,0
Lysine	17,7	31,2
Acide Aspartique	30,3	45,8
Acide Glutamique	142	78,0
Hydroxy-proline	0	47,3
Hydroxy-lysine	0	-
Amide	146,0	-

Tableau 1. Composition en acides aminés de l'émail en formation et des protéines insolubles de l'émail mature (les valeurs représentent le nombre de résidus d'acides aminés par milliers de résidus totaux). D'après Piez (1960).

L'étude suivante (Eastoe, 1963) a permis de mieux caractériser la composition en acides aminés du contenu protéique de l'émail humain, montrant ainsi une forte proportion de proline, glutamine et leucine.

Deux ans plus tard, Eastoe (1965) baptise cette protéine « amélogénine » étant donné sa composition unique qui la rend différente du collagène, de l'épiderme, des kératines et des protéines de l'émail mature. Ainsi, il est considéré comme le premier à avoir utilisé ce nom, *amélogénine* signifiant *ce qui génère l'émail*. Il considère « l'amélogénine » comme le composant protéique dominant des stades précoces de la matrice de l'émail dentaire.

	Émail dentaire de fœtus humains	Émail dentaire de fœtus humains
	Grammes d'acide aminé / 100 g	Moles d'acides aminés /10^5g
Alanine	1,62	18,2
Glycine	4,52	60,2
Valine	4,05	34,6
Leucine	10,7	81,6
Isoleucine	3,85	29,4
Proline	25,0	218
Phénylalanine	3,79	23,0
Tyrosine	9,6	52,8
Serine	6,2	59,3
Thréonine	4,04	33,9
demi-cystine	0,13	1,1
Méthionine	6,0	40,1
Arginine	3,64	20,9
Histidine	8,8	56,6
Lysine	2,37	16,2
Acide Aspartique	3,66	27,5
Acide Glutamique	18,2	124
Hydroxy-proline	0	0
Hydroxy-lysine	trace	0
Total	116,17	897,4
Amide	1,79	128

Tableau 2. Composition en acides aminés de l'émail en développement (d'après Eastoe, 1963).

I.2. Purification des protéines

A partir du milieu des années 60 et jusqu'à la fin des années 70, l'étude de l'amélogénine va se développer grâce à l'utilisation de techniques de biochimie comme l'exclusion de taille, la chromatographie échangeuse d'ion et l'électrophorèse. La possibilité de séparer les protéines en fonction de leur taille a alors montré que la matrice protéique de l'émail contenait un complexe de protéines dont les tailles variaient de 5 à 70 kDa.

De nombreuses équipes (Glimcher *et al.*, 1964 ; Burgess et MacLaren, 1965 ; Nikiforuk et Simmons, 1965 ; Mechanic, 1971 ; Eggert *et al.*, 1973 ; Papas *et al.*, 1977 ; Seyer et Glimcher, 1977 ; Fincham, 1979 ; Fukae *et al.*, 1979) ont tenté de séparer et de caractériser les différentes fractions de protéines de l'émail en utilisant divers modèles animaux expérimentaux comme la vache ou le porc, modèles plus accessibles que l'homme pour extraire des quantités suffisantes de protéines dentaires et pour travailler sur l'embryon. En effet, l'amélogénine n'est présente qu'au début de la formation de l'émail dentaire, puis elle est protéolysée et disparaît progressivement de l'émail mature dans la dent fonctionnelle (à l'exception des rongeurs).

Ces études ont apporté peu d'élément capital sur le fonctionnement de la protéine car elles présentaient souvent des analyses globales en acides aminés à partir de fragments d'amélogénine de taille diverse. Il y a donc peu de choses à en dire excepté qu'elles ont permis de mettre en lumière la présence de deux fragments d'amélogénine de petite taille (environ 5 KDa) ayant des compositions différentes en acides aminés et qui semblaient pourtant provenir d'un même fragment d'amélogénine de plus grande taille. Ces deux fragments ont été appelés ensuite « LRAP » pour "Leucine Rich Amelogenin Peptide" et « TRAP » pour "Tyrosine Rich Amelogenin" (Fincham *et al.*, 1980), résultent en réalité de la

protéolyse de l'amélogénine au cours de la maturation de l'émail (TRAP) et d'un épissage alternatif (LRAP). Bien que ces phénomènes n'etaient pas connus à cette période pour l'amélogénine et qu'on ne savait pas encore s'il existait plusieurs gènes codant pour une famille d'amélogénines ou une protéine unique, les premiers éléments permettant de comprendre son fonctionnement général venaient d'apparaître. De plus, bien que l'on ait eu aucune idée précise sur le rôle que jouait l'amélogénine dans la minéralisation de l'émail dentaire, la connaissance des acides aminés permettait de proposer quelques hypothèses sur son interaction possible avec le calcium ou sur les modifications chimiques post-traductionnelles, comme la phosphorylation de certains résidus (Papas *et al.*, 1977).

Malgré cela, les séparations des protéines étaient encore très incomplètes et les données confuses, ce qui conduisit à des difficultés d'interprétation. En effet, le mélange des différentes protéines de la matrice extracellulaire de l'émail faussait les analyses de composition globale en acides aminés et les agrégats de protéines conduisaient à surestimer le poids de l'amélogénine. Très tôt, Katz *et al.* (1965) suggérèrent que des agrégations entre protéines (des interactions « protéine-protéine ») pouvaient avoir lieu, même en présence de facteurs de dissociation. Termine *et al.* (1979) montrèrent clairement que les amélogénines bovines pouvaient présenter une gamme de taille variant de 7 à 28 kDa ou de 10 à 40 kDa suivant la composition chimique des gels d'électrophorèse et leur capacité à dissocier les agrégats (l'amélogénine bovine entière fait 27 kDa). Ces interactions peuvent expliquer les confusions observées dans certains travaux à l'époque.

En 1980, Termine *et al.* introduisirent une nouvelle procédure d'extraction séquentielle et dissociative des protéines de l'émail en deux étapes. La première étape permettait d'extraire les acides aminés caractéristiques de l'amélogénine (dans les fractions dont la taille était comprise entre 5 et 27 kDa), alors que lors de la seconde étape on observait une préparation riche en asparagine, sérine, glycine (dans les fractions dont la taille était comprise entre 40 et 70 kDa). Eastoe proposa en 1979 d'appeler cette dernière « enaméline ». Cette méthode avait l'avantage d'être relativement simple et reproductible.

Pourtant, ces deux extraits étaient encore hétérogènes. En effet, la séparation n'était ni complète ni sélective. Aujourd'hui, les techniques actuelles montreraient que les extraits d'amélogénine de l'époque contenaient sans aucun doute d'autres protéines de la matrice de l'émail comme, par exemple, l'améloblastine. Il convient donc de bien faire la distinction entre les termes « amélogénine » et « enaméline » utilisés il y a quelques années pour désigner ces extraits de protéines, et les protéines de la matrice portant le même nom aujourd'hui.

Une grande avancée dans l'étude des protéines de la matrice de l'émail, et notamment l'amélogénine, a été obtenue avec la possibilité de connaître la séquence en acides aminés des protéines grâce à la technique de séquençage d'Edman (1970). Cette technique fut appliquée tout d'abord sur des protéines purifiées de dents de porc ou de vache qui étaient historiquement les premiers modèles utilisés pour les techniques d'extraction des protéines de l'émail. Chez le porc, la première séquence partielle date de 1979 (Fukae), mais les données étaient très incomplètes puisqu'elles concernaient seulement les treize premiers résidus. En revanche, d'autres données obtenues par le même auteur en 1980 et 1983 ont permis de connaître les 54 premiers acides aminés. Fukae constate ainsi que deux fractions protéiques de porc, de taille très différente (21 kDa et 6 kDa), possèdent les mêmes acides aminés. Il suggère alors qu'une enzyme protéolytique doit être présente dans l'émail immature et qu'elle dégrade l'amélogénine native de 26 kDa en fragments de taille plus petite. Il faudra cependant attendre 1989 et les travaux de Yamakoshi pour connaître la première séquence complète d'amélogénine porcine confirmant que l'amélogénine porcine possède bien une taille de 26 kDa. Chez la vache, la première étude de ce genre date de 1972 (Seyer) mais les résultats les plus intéressants ont été obtenus en 1983 par Fincham *et al.* Ces auteurs ont aussi comparé

des séquences d'acides aminés de l'amélogénine de vache, de porc et humaine. En plus de la séquence d'amélogénine humaine qui est présentée pour la première fois, ce travail est intéressant sur plusieurs plans. Non seulement les auteurs publient une partie de la séquence d'une amélogénine bovine (27 kDa), mais en plus, ils obtiennent la séquence des fragments LRAP de 7 kDa et TRAP de 5 kDa, qui etaient connus pour posséder des compositions globales en acides aminés différentes. Ils constatent ainsi que ces deux fractions possèdent effectivement une partie C-terminale complètement différente l'une de l'autre. Ne connaissant pas à l'époque les phénomènes d'épissage alternatif qui touchent les ARNm de l'amélogénine, ils ne purent en donner une explication, d'autant plus que la séquence complète de l'amélogénine bovine ne sera connue qu'un an plus tard, en 1984 (Takagi). Cela aurait permis de constater que la partie C-terminale de la protéine LRAP provient de l'extrémité de l'exon 6 (voir chapitre ci-après sur l'épissage alternatif) et celle de la TRAP du début de l'exon 6.

Les données sur l'amélogénine humaine ont été obtenues plus tard. Il faudra, en effet, attendre 1993 (Catalano-Sherman *et al.*) pour que la séquence complète de l'amélogénine humaine soit publiée. Mais cette séquence apparaît au moment où les techniques de séquençage de protéines sont déjà concurrencées par les techniques de biologie moléculaire et le séquençage d'ADN. D'ailleurs, dans leur article, Catalano-Sherman *et al.* (1993) présentent, en plus de la séquence protéique, une séquence d'ADN complémentaire (ADNc) d'amélogénine humaine. A partir de cet instant, on déterminera préférentiellement les séquences protéiques en les déduisant à partir de celle des ARNm. Entre-temps, sera publiée une des dernières études utilisant le séquençage de protéine; elle concerne l'amélogénine de lapin (Zeichner-David *et al.,* 1988), un modèle intéressant puisque les incisives de ce lagomorphe poussent (comme chez les rongeurs (rat, souris, …)) continuellement (dents hypsodontes). Cela permet une étude plus aisée de l'amélogénine à tous les niveaux (ADNc, ARN, protéine) car étant fabriquée en permanence, elle est présente chez l'adulte.

I.3. Les données génétiques

Dès lors, les données sur les séquences nucléiques d'amélogénine ont continué à s'accumuler chez les mammifères, comme le porc (Hu *et al.*, 1996a) et l'opossum (Hu *et al.*, 1996b), et divers rongeurs : le rat (Bonass, 1994a), le hamster (Lyaruu *et al.,* 1998) et le cochon d'inde (Cerny et Hammerström, 1998, 1999). Cependant, cela ne doit pas faire oublier les travaux pionniers en biologie moléculaire sur l'amélogénine, chez la souris (Snead, 1983, 1985) et chez la vache (Takagi, 1984). De plus, les séquences de rongeurs étant très proches les unes des autres, elles apportent assez peu d'informations nouvelles sur l'amélogénine elle-même, en dehors de quelques particularités propres aux rongeurs, par exemple l'insertion de 4 aa à la position 129 (Bonass, 1994a). Ces différences (insertions, duplications, délétions, …) qui peuvent exister entre les séquences d'amélogénine chez les différents groupes de mammifères soulèvent d'ailleurs chez Bonass (1994a) la question intéressante de la relation pouvant exister entre les résultats obtenus au niveau moléculaire et les altérations constatées dans la structure de l'émail. En revanche, comme chez le lapin, ces modèles rongeurs s'avèrent être très pratiques au niveau expérimental en raison de la croissance permanente des dents au cours de la vie de l'individu.

1.4. Approche évolutive

Les travaux sur l'amélogénine, et les autres protéines dentaires, n'ont concerné pendant longtemps que les mammifères. Ceci s'explique bien sûr par le fait que les chercheurs qui s'intéressent à l'amélogénine sont avant tout des "dentistes" et que les

modèles animaux doivent être proches du modèle humain. Cependant, la comparaison trouve très rapidement ses limites chez les quelques groupes de mammifères étudiés qui sont assez proches les uns des autres (comme la vache et le porc, la souris et le hamster, ...). De plus, les dents des mammifères présentent des caractéristiques tout à fait particulières chez les vertébrés. En effet, à de rares exceptions près (rongeurs, dauphins, ... , qui sont *monophyodontes*) ils sont tous *diphyodontes* (à une première dentition, appelée dentition de lait, succède une seconde dentition, appelée dentition définitive), leurs dents sont peu nombreuses (55 au maximum chez les marsupiaux, 44 au maximum chez les euthériens) et ils sont donc *oligodontes*. Ils sont également *hétérodontes* car leurs dents sont morphologiquement très différentes (incisives, canines, molaires). Chez les vertébrés non-mammaliens, contrairement aux mammifères, les dents sont généralement très nombreuses (*polyphyodontes*) et souvent très semblables (*homodontes*), et elles servent davantage à la préhension et la contention, mais jamais à la mastication. Les spécialisations importantes observées chez les mammifères se sont sans doute accompagnées d'adaptations ou de modifications des tissus dentaires. Il faut alors se demander si l'amélogénine n'a pas subi des pressions sélectives particulières en rapport avec ces phénomènes évolutifs. C'est pourquoi il était intéressant d'étudier l'amélogénine chez d'autres groupes de mammifères que ceux déjà connus et chez des vertébrés non-mammaliens afin d'essayer d'établir quelques modalités évolutives et fonctionnelles.

Une première approche a été réalisée par le séquençage d'ADNc d'amélogénine d'un marsupial (mammifère basal), l'opossum (*Monodelphis domestica*) (Hu, 1996). Cette séquence représente une étape intéressante avant de tenter d'obtenir des séquences dans des lignées non-mammaliennes. En effet, les marsupiaux (ou métathériens) sont séparés des mammifères placentaires depuis environ 135 MA (Figure 1), ce qui représente une distance évolutive importante au niveau de l'information génétique.

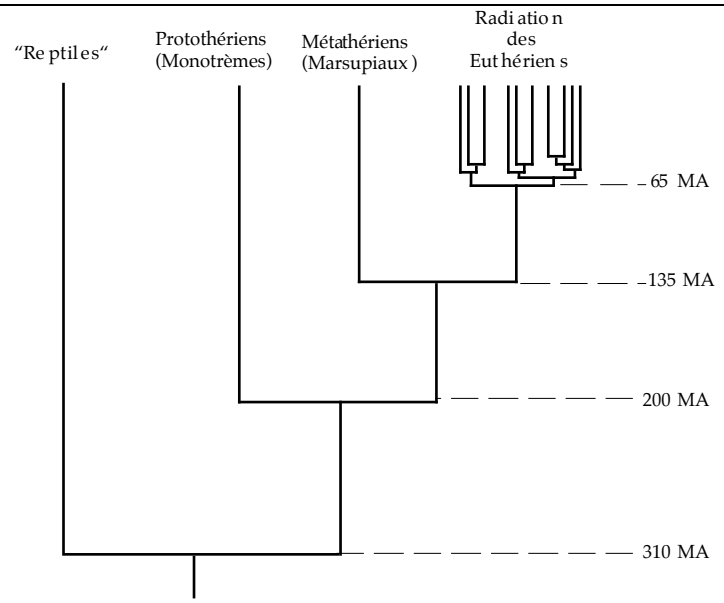


Figure 1. Relations évolutives des trois grands groupes de mammifères. Arbre enraciné avec les "reptiles". D'après : Kumar et Hedges, 1998; Rowe, 1999; Lecointre et Le Guyader, 2001.

La connaissance de l'amélogénine de l'opossum n'apporte que des informations sur le mode d'évolution des répétitions présentes dans la région centrale de l'exon 6 du gène (voir partie sur la structure du gène), mais elle permet d'identifier des régions très conservées au sein des mammifères. Depuis 1998, grâce au séquençage du gène de l'amélogénine chez un

bon nombre d'espèces, je dispose maintenant d'informations chez les primates, rongeurs, artiodactyles, périssodactyles, marsupiaux et monotrèmes, ce qui m'a permis de réaliser l'étude évolutive et phylogénétique de l'amélogénine chez les mammifères (voir chapitre 2).

En 1998, trois équipes de recherche présentent, pour la première fois, des séquences d'amélogénine non-mammaliennes. La première équipe, de l'institut Max-Planck en Allemagne, a obtenu la séquence d'un caïman (*Paleosuchus palpebrosus*), ainsi que deux séquences d'un amphibien anoure, le xénope (*Xenopus laevis*), qui présenterait une duplication du gène (Toyosawa *et al.,* 1998). Cet amphibien constitue, encore aujourd'hui, le taxon le plus basal (relativement aux mammifères) chez lequel on a pu séquencer le gène de l'amélogénine. Ces auteurs présentent également les séquences partielles de deux mammifères monotrèmes, l'échidné d'Australie (*Tachyglossus aculeatus*) et l'ornithorynque (*Ornithorhynchus anatinus*). Ce groupe, ordre unique des protothériens, est séparé du reste des mammifères depuis 200 millions d'années (Figure 1) et constitue de ce fait la branche la plus basale. Une deuxième équipe de l'Université de Niigata au Japon, a obtenu la séquence d'un serpent japonais (*Elaphe quadrivirgata*), un reptile appartenant à la lignée des squamates (Ishiyama *et al.*, 1998). Enfin, j'ai commencé ma thèse en obtenant la séquence partielle de l'amélogénine chez un autre squamate, le lézard *Gallotia stehlini* (Lacertidae) (Delgado *et al.*, 1998).

Depuis 1998, grâce à une collaboration avec une équipe japonaise (Dr. M. Ishiyama), j'ai pu avancer dans la connaissance de l'amélogénine chez un certain nombre de familles de squamates. Le gène et sa protéine sont maintenant connus, entièrement ou partiellement, chez diverses espèces de Scincidae, Lacertidae, Serpentes, Agamidae, Geckkonidae, Iguanidae et Varanidae (voir chapitre 2).

Grâce aux données moléculaires, nous savons maintenant que l'amélogénine est présente au moins dans toutes les lignées de Tétrapodes à l'exception des tortues et des oiseaux (Girondot et Sire, 1998).

(Toutes les séquences disponibles dans Genbank sont présentées en annexe 1).

I.5. Utilisation d'anticorps pour rechercher l'expression de l'amélogénine chez les vertébrés non-mammaliens

Dès 1980, les auteurs ont utilisé des anticorps élevés chez les mammifères pour rechercher l'expression de la protéine dans divers groupes de vertébrés non-mammaliens. Herold *et al*. (1980) démontrent la présence d'amélogénine au niveau des dents de reptiles, d'amphibiens, de téléostéens* et de chondrichtyens* en utilisant des anticorps anti-amélogénine bovine. Chez ces derniers, ils détectent aussi l'expression de l'amélogénine au niveau des denticules dermiques (= odontodes, homologues des dents) localisés à la surface de la peau de *Squalus acanthias*. Ils en concluent à une répartition très générale de l'amélogénine au sein des vertébrés* et à une origine très ancienne de cette protéine. En 1983, Slavkin *et al.* confirment ces résultats en détectant l'amélogénine chez la même espèce de requin (*Squalus acanthias*); la même année Slavkin *et al.* (1983b) détectent l'amélogénine chez une myxine*. Les myxines, crâniates dépourvus de mâchoires, sont actuellement considérées comme le groupe frère des vertébrés (janvier, 1996). Elles sont dépourvues de dents, mais possèdent des denticules "cornés". La mise en évidence de l'amélogénine au niveau des dents cornées, bien que surprenante, montre que l'amélogénine serait présente depuis très longtemps, et même en dehors des vertébrés. Par la suite, Krejsa *et al.* (1984) détectent de l'amélogénine dans les écailles et les dents de la carpe, *Cyprinus carpio* (un

* Voir figure 2.

téléostéen*). Pourtant en 1989, Herold *et al.* qui avaient précédemment détectés l'amélogénine chez les chondrichtyens (Herold *et al.,* 1980), remettent en question les précédents résultats concernant la présence d'amélogénine chez les vertébrés. En effet, en utilisant des anticorps anti-amélogénine et anti-énaméline bovine, les auteurs suggèrent que l'amélogénine ne serait apparue que chez les tétrapodes alors que l'énaméline serait très ancienne. L'amélogénine ne pouvait donc être détectée ni chez les chondrichtyens, ni chez les actinoptérygiens et les résultats précédents seraient dus à l'utilisation d'anticorps polyclonaux mal purifiés. Au contraire, une autre classe de protéines de la matrice de l'émail, l'énaméline serait présente chez les chondrichtyens et les actinoptérygiens. Pourtant, peu de temps après, Lyngstadass *et al.* (1990), en utilisant des techniques de biologie moléculaire, montrent qu'un ADNc de souris s'hybride sur le génome d'un téléostéen (*Anarhichas lupus*). Bien que ces auteurs n'obtiennent pas l'hybridation de leur sonde avec le génome du saumon *Salmo salar* (un autre téléostéen), ces résultats confirment les résultats précédents montrant la présence d'amélogénine chez les vertébrés non-tétrapodes. De plus, Zylberberg *et al.* (1997) ont mis en évidence l'expression de l'amélogénine chez un polyptère, *Calamoichthys calabaricus* (un actinoptérygien basal), au niveau de la ganoïde, un tissu minéralisé apparenté à l'émail et présent à la surface des écailles. Des études récentes ont montré la présence d'amélogénine au niveau des dents en formation chez *Neoceratodus forsteri* (un dipneuste*) (Satchell *et al*., 2000), et chez un lépisostée (un autre actinoptérygien basal) (Ishiyama *et al*., 2000). Tous ces résultats, même s'ils sont fondés sur des travaux utilisant des anticorps de mammifères, suggèrent que l'amélogénine est au moins présente chez l'ancêtre commun des actinoptérygiens et des sarcoptérygiens, ce qui ferait remonter son origine à plus de 400 MA (450 MA si l'on considère les chondrichtyens), contrairement aux conclusions de Herold *et al.* (1989).

Pour expliquer les résultats contradictoires obtenus à l'aide des anticorps, Slavkin (1984) a suggéré que l'amélogénine existe en effet chez les vertébrés non-tétrapodes mais que son expression est minoritaire et faible par rapport à celle des énamélines. Le rapport entre les deux classes de protéines s'inverserait à partir des tétrapodes, et l'amélogénine deviendrait alors la protéine majoritaire de la matrice de l'émail en formation. Cette théorie expliquerait les difficultés rencontrées pour détecter la protéine chez les non-tétrapodes, en particulier chondrichtyens et actinoptérygiens (Herold *et al.,* 1989*).* De plus, il faut noter que le gène de l'amélogénine n'a été cloné et séquencé chez des non-mammaliens qu'à partir de 1998, et que la plupart des auteurs a probablement surestimé la conservation de la protéine chez les vertébrés. En effet, on sait aujourd'hui à quel point l'amélogénine de taxons non-mammaliens (reptiles et amphibiens) présente des différences au niveau de certaines de ses régions. On peut penser que si l'amélogénine est présente chez les actinoptérygiens et les chondrichtyens, les différences accumulées au niveau de la protéine, au cours de l'évolution de ces lignées, sont bien plus grandes encore. Dans ces conditions, on peut supposer que les anticorps conçus à partir d'amélogénines de mammifères devraient être peu efficaces pour reconnaître des épitopes sur les amélogénines des non-tétrapodes. Toutefois, nous savons que l'amélogénine des lignées étudiées possède quelques régions très bien conservées. Il est probable que ces domaines, sans doute très importants pour la structure de la protéine, sont aussi présents chez les non-tétrapodes, ce qui expliquerait les résultats positifs obtenus avec les anticorps mammaliens, peut être moins sélectifs que les amorces d'ADN.

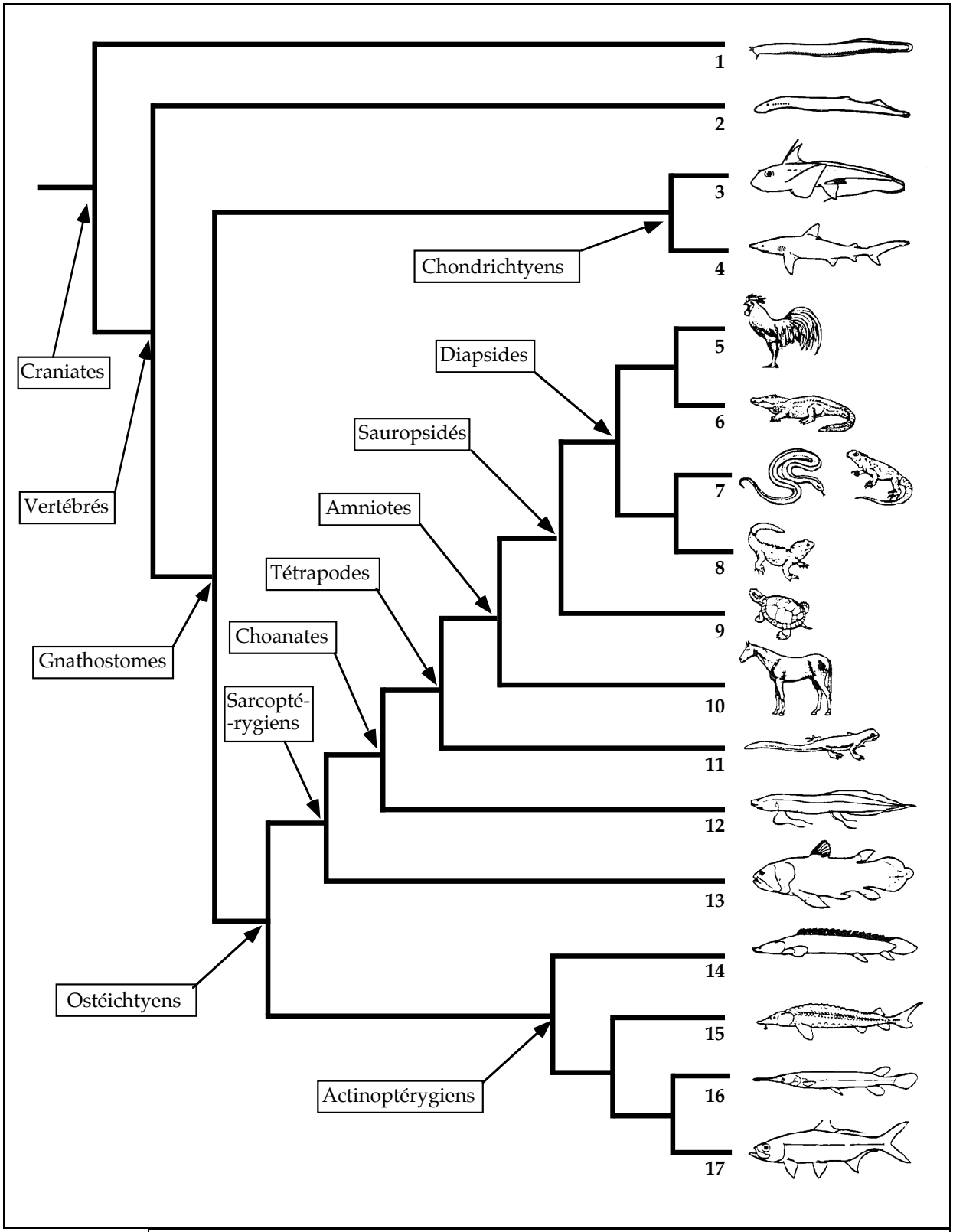


Figure 2. *Cladogramme résumant les relations de parenté des principaux groupes de vertébrés, d'après les connaissances actuelles. D'après Lecointre, 1993. Légendes :*

1 : Myxines ; 2 : Lamproies ; 3 : Chimères ; 4 : Requins et raies ; 5 : Oiseaux ; 6 : Crocodiles; 7 : Lézards et serpents ; 8 : Sphénodon ; 9 : Tortues ; 10 : Mammifères ; 11 : Amphibiens ; 12 : Dipneustes ; 13 : Cœlacanthes ; 14 : Polyptères ; 15 : Esturgeons ; 16 : Lépisostées ; 17 : Téléostéens.

II. BIOLOGIE MOLÉCULAIRE ET BIOCHIMIE DE L'AMÉLOGÉNINE

II.1. Structure du gène de l'amélogénine

Le gène de l'amélogénine est constitué de 9 exons (Figure 3). La région codante du gène de l'amélogénine commence à partir de la 12ème base de l'exon 2 par le codon d'initiation (ATG) et se termine à la 6ème base de l'exon 7 par le codon "stop" (TAA). Cela signifie que l'exon 1 est une région non codante dont le rôle reste à définir (s'il en possède un) et que l'exon 7 code pour un acide aminé seulement (Figure 3).

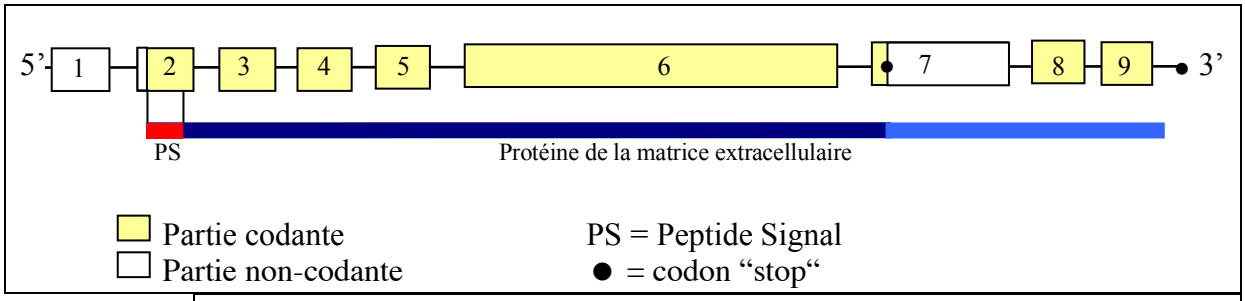


Figure 3. *Schéma général du gène et de la protéine de l'amélogénine représentant ses principales caractéristiques : les exons (1,2,...), proportionnels à leur taille, les introns symbolisés par un trait dont la longueur ne correspond pas à la taille réelle, la position des codons "stop", les parties codantes et non codantes.*

L'exon 4 n'est connu que chez l'humain, la vache, le porc et deux rongeurs (Figure 4), il n'a été détecté dans aucun des ADNc d'amélogénine isolés jusqu'à aujourd'hui et il semble bien, à ce titre, être caractéristique uniquement de ces groupes. De plus, chez l'humain et les bovins qui possèdent deux copies de l'amélogénine (sur les chromosomes X et Y), il semble que l'exon 4 ne soit présent que sur le chromosome X. Enfin, il faut noter que très récemment, un second exon 4 a été découvert chez le porc (Hu et al.*, 2001) (Figure 4).*

a.

Hamster	KSHSQAINTDRTAL
Cobaye	KSHSNAINIDRTAL
Souris	KSHSQAINTDRTAL
Homme X	NSHSQAINVDRTAL
Vache X	NSYFQGISIDETALVSLYFMQLN
Porc-a	KSGRWGARLTAFVSSVQ
Porc-b	DLYLEAIRIDRTAF

(Hu *et al.*, 2001)

b.

Exon 8	AFSPMKWYQGTARHP
Exon 9	LNMETTTEK

Figure 4. *Présentation des séquences protéiques connues de l'exon 4 (a) et des exons 8 et 9 de l'amélogénine de rat (b).*

Le rôle de l'exon 4 est pour l'instant inconnu. Il semble qu'il ait été perdu chez un certain nombre de mammifères alors qu'il s'est maintenu chez d'autres mais uniquement sur le chromosome X. Cependant, l'exon 4 étant présent chez un très petit nombre de transcrits d'amélogénine (Simmer, 1995), il se peut en réalité qu'il soit très difficile à détecter.

Les exons 8 et 9 ont été découverts assez récemment (Li *et al*., 1998) et pour le moment, leur séquence n'est connue que chez le rat (*Rattus norvegicus*). Cependant, ces exons ont été détectés par "southern blot" dans le génome de la souris et de l'homme (Li *et al*., 1998). De plus, l'utilisation d'anticorps dirigés contre la région de la protéine codée par l'exon 9 a donné des résultats positifs chez la souris et le porc (au niveau de dents qui n'ont pas encore percé) (Li *et al*., 1998). Les exons 8 et 9 représentent une fin "alternative" pour la protéine de l'amélogénine. En effet, la fabrication de la protéine s'arrête normalement au niveau de l'exon 7 qui contient un codon "stop" à l'extrémité de sa partie codante. Cependant, en raison d'un épissage alternatif, l'exon 7 est parfois exclu des ARNm et il y a alors production d'ARNm contenant les exons 8 et 9 se terminant également par un codon "stop". Toutes les données concernant les exons 8 et 9 suggèrent que ces produits d'épissage sont impliqués dans la formation de l'émail, même si leur présence est minoritaire (Li *et al*., 1998). En effet, jusqu'à ces travaux, les études à partir d'ADNc n'avaient jamais révélé la présence de ces deux exons. Cependant, leur rôle reste à définir ainsi que leur présence dans le génome d'autres mammifères, voire d'autres groupes de vertébrés. Il est aussi possible que la présence de ces exons soit un phénomène récent d'un point de vue évolutif, et qu'ils ne soient présents que chez quelques mammifères.

Les exons de l'amélogénine, à l'exception des exons 6 et 7, sont relativement courts (50 pb environ). L'exon 7 malgré sa taille (plus de 100 pb) ne code que pour un seul acide aminé, ce qui le rend peu intéressant pour définir des amorces de PCR ("Polymerase Chain Reaction"). L'exon 6 est le plus grand puisqu'il contient environ 400 pb; il est donc très intéressant car il rend possible l'amplification d'un fragment d'amélogénine à partir de l'ADN génomique (ADNg). En effet, il n'est pas possible de réaliser des PCR dans les "petits" exons sachant que les amorces de PCR mesurent environ 20 pb chacune. Il est aussi très difficile d'amplifier par PCR deux exons, car la taille des introns qui les séparent est très grande (de 1 à 2 kb), à l'exception de l'intron séparant les exons 5 et 6 qui est relativement court puisqu'il mesure de 200 à 300 pb (269 chez l'homme) (Figure 5).

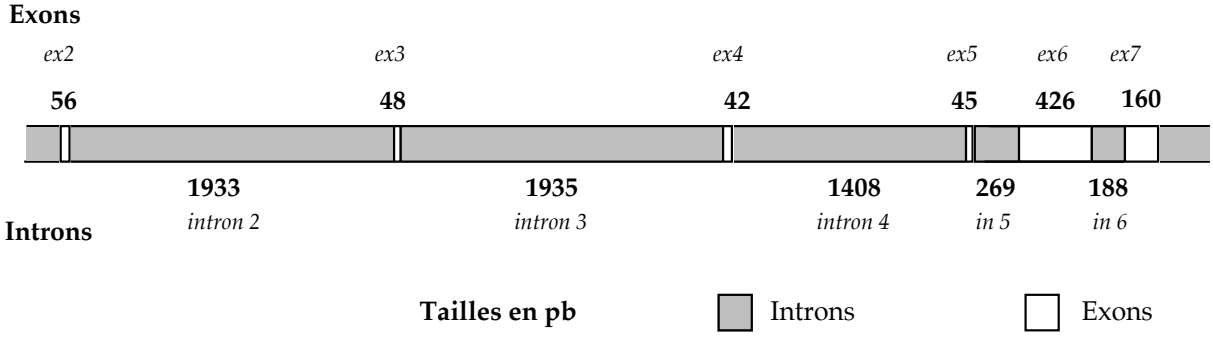


Figure 5. Schéma d'une partie du gène de l'amélogénine humaine (des exons 3 à 7) représentant les introns et les exons, en respectant l'échelle de taille.

Pour réaliser des PCR sur de l'ADNg, le choix s'est donc porté tout naturellement sur l'exon 6 et des amorces situées aux extrémités de cet exon

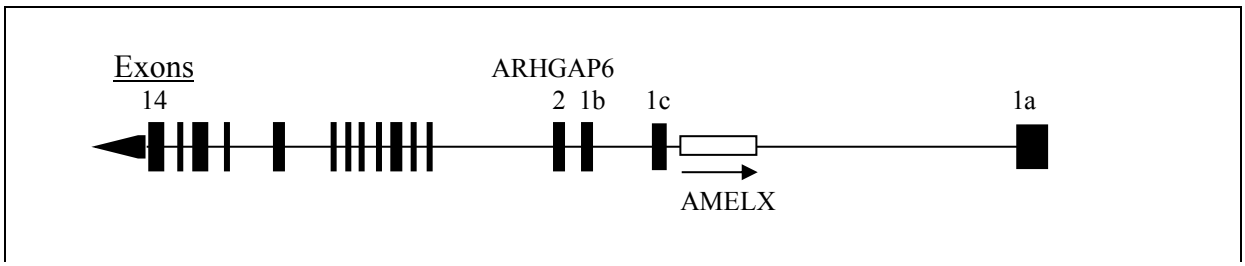
ont été confectionnées. Par la suite, pour obtenir les séquences des autres exons de l'amélogénine, il a fallu utiliser des ADNc fabriqués à partir des ARNm extraits de mâchoires entières ou de bourgeons de dents d'animaux fraîchement sacrifiés. En effet, les ARNm matures ne possèdent plus les introns ce qui permet en théorie d'amplifier directement l'amélogénine d'un bout à l'autre.

Localisation chromosomique

Chez les mammifères euthériens, le gène de l'amélogénine est présent sur le chromosome sexuel X et parfois sur le chromosome Y (homme, vache, chèvre, cheval, cerf, bouquetin) (Iwase *et al.,* 2001; voir aussi chapitre 2). En dehors des euthériens, le gène a seulement été localisé chez un marsupial, l'opossum, au niveau du chromosome 21 (Hu *et al.,* 1996). De plus, le gène de l'amélogénine peut être présent à deux exemplaires au sein d'un même génome. En effet, on trouve deux copies de ce gène chez le xénope (tétraploïde).

Par ailleurs, la présence du gène sur les chromosomes sexuels donne lieu à un phénomène évolutif intéressant, puisqu'une évolution darwinienne positive a été détectée sur les gènes de l'amélogénine humaine (Girondot et Casane, 1998). Cela signifie que les différences qui s'accumulent dans les séquences d'ADN d'amélogénine entre les deux chromosomes sont plus importantes que si une des deux séquences évoluait sous l'effet du hasard. Girondot et Casane (1998) ont proposé que ce phénomène pourrait être lié au fait que les protéines d'amélogénine s'associent pour former des agrégats (voir *§II.7.*). Ils ont supposé que les agrégats "homogènes" (formés uniquement par des protéines du chromosome X ou uniquement par des protéines du chromosome Y) seraient plus stables que les agrégats comportant un mélange des deux protéines d'amélogénine. Cela expliquerait l'existence d'un phénomène ayant tendance à séparer les deux protéines pour qu'elle ne puissent plus s'associer au cours de la minéralisation de l'émail.

En ce qui concerne la localisation précise du gène de l'amélogénine humaine, il est situé en position "Xp22.31-p22.1" sur le chromosome X et "Yp11.2" sur le chromosome Y. Sur le chromosome X, on a d'abord pensé qu'il était voisin du gène ARHGAP6* codant pour une protéine activatrice de GTPase de type *rho* (Schaefer *et al.,* 1997). Cependant, des travaux plus récents ont montré qu'en réalité, le gène de l'amélogénine serait présent dans l'intron 1 de ce gène (230 kb environ), mais codant dans le sens inverse d'ARHGAP6 (Prakash *et al.*, 2000). La confusion est venue du fait que toutes les régions codantes d'ARHGAP6 n'étaient pas connues en 1997. Il est apparu, à la suite de la découverte d'un nouvel exon (exon 1a) situé en amont du gène ARHGAP6 (Prakash *et al.*, 2000), que le gène de l'amélogénine se trouve localisé entre les exons 1a et 1c d'ARHGAP6.



* le gène ARHGAP6 humain n'a pas de copie sur le chromosome Y

Figure 6. Structure schématique du gène de l'ARHGAP6 situé sur le chromosome X humain. Le gène de l'amélogénine (symbolisé par un carré blanc) est situé dans l'intron 1 du gène ARHGAP6 mais s'oriente dans le sens inverse (l'orientation des deux gènes est indiquée par le sens des flèches).

Si l'amélogénine "cohabite" depuis longtemps avec ARHGAP6, il est alors possible d'envisager le séquençage du gène de l'amélogénine si l'on connaît la séquence complète de l'intron 1 d'ARHGAP6 (à condition évidemment qu'ARHGAP6 soit bien conservé au cours de l'évolution). Le séquençage en cours de certains génomes de vertébrés, et notamment des deux téléostéens, Fugu et Danio, devrait permettre de vérifier si l'amélogénine est bien présente en antisens dans l'intron 1 ce gène chez ces vertébrés non-tétrapodes. Cette découverte offre donc une nouvelle alternative dans la recherche du gène de l'amélogénine chez les vertébrés non-tétrapodes.

II.2. Phénomènes post-transcriptionnels : l'épissage alternatif

Comme pour la plupart des gènes, l'ADNg de l'amélogénine est transcrit en ARN. Par la suite, au cours de la maturation des ARN, ceux-ci sont soumis au mécanisme d'épissage des introns qui consiste en l'excision des introns et la ligation des exons du produit de transcription primaire. Au cours de cet épissage des introns, il se produit parfois un phénomène d'épissage alternatif qui se manifeste par l'élimination de régions codantes en même temps que les introns.

L'amélogénine subit également ce type d'épissage. L'élimination de parties d'exons, ou d'exons entiers, ne constitue pas ici une erreur conduisant à une amélogénine non fonctionnelle mais, au contraire, ce phénomène produit des amélogénines ayant des propriétés différentes et s'exprimant à des stades particuliers de la synthèse de l'émail. L'épissage alternatif de l'amélogénine est connu chez les bovins (Gibson et al.*, 1991), chez l'humain (Salido* et al.*, 1992), chez les porcins (Yamakoshi* et al.*, 1994), et chez les rongeurs tels que la souris (Simmer* et al.*, 1994) et le rat (Bonass* et al.*, 1994b ; Li* et al.*, 1995) et enfin, chez l'opossum (Hu* et al.*, 1996). Ces phénomènes n'ont pas encore été rapportés en dehors des mammifères, mais certains indices nous laissent supposer qu'ils existent également. Le phénomène d'épissage alternatif a été tout d'abord proposé pour expliquer l'existence des fragments LRAP-1 et LRAP-2 de 42 et 46 aa, isolés de l'émail d'un bovin pour la première fois en 1981 (Fincham* et al.*, 1981). Ces molécules furent ensuite séquencées et des études (Gibson* et al.*, 1991) montrèrent qu'un fragment LRAP de 59 aa correspond à la traduction d'un ARN messager plus court codant les exons 4, 5 et une partie (terminale) de l'exon 6. Les fragments LRAP-1 et LRAP-2 de Fincham* et al. *(1981) correspondent en fait au fragment LRAP isolé par Gibson* et al. *(1991), qui a subit un début de protéolyse. Initialement, il était supposé que ces fragments provenaient uniquement de ces coupures protéolytiques réduisant la protéine initiale en morceaux plus petits comme c'est le cas pour le fragment TRAP. L'épissage alternatif de l'amélogénine et l'existence de fragments LRAP n'a pas été observé chez l'humain : le site d'épissage est présent mais il ne produit pas de fragments LRAP. Cela signifierait que tous les messagers de l'amélogénine contiendraient l'exon 6 en entier. Il est aussi possible que ces isoformes particulières de l'amélogénine n'aient pas encore été détectées, peut-être parce qu'elles sont rares dans la matrice extracellulaire. La connaissance du phénomène d'épissage a beaucoup progressé grâce aux travaux de J. Simmer (1995) (Figure 7).*

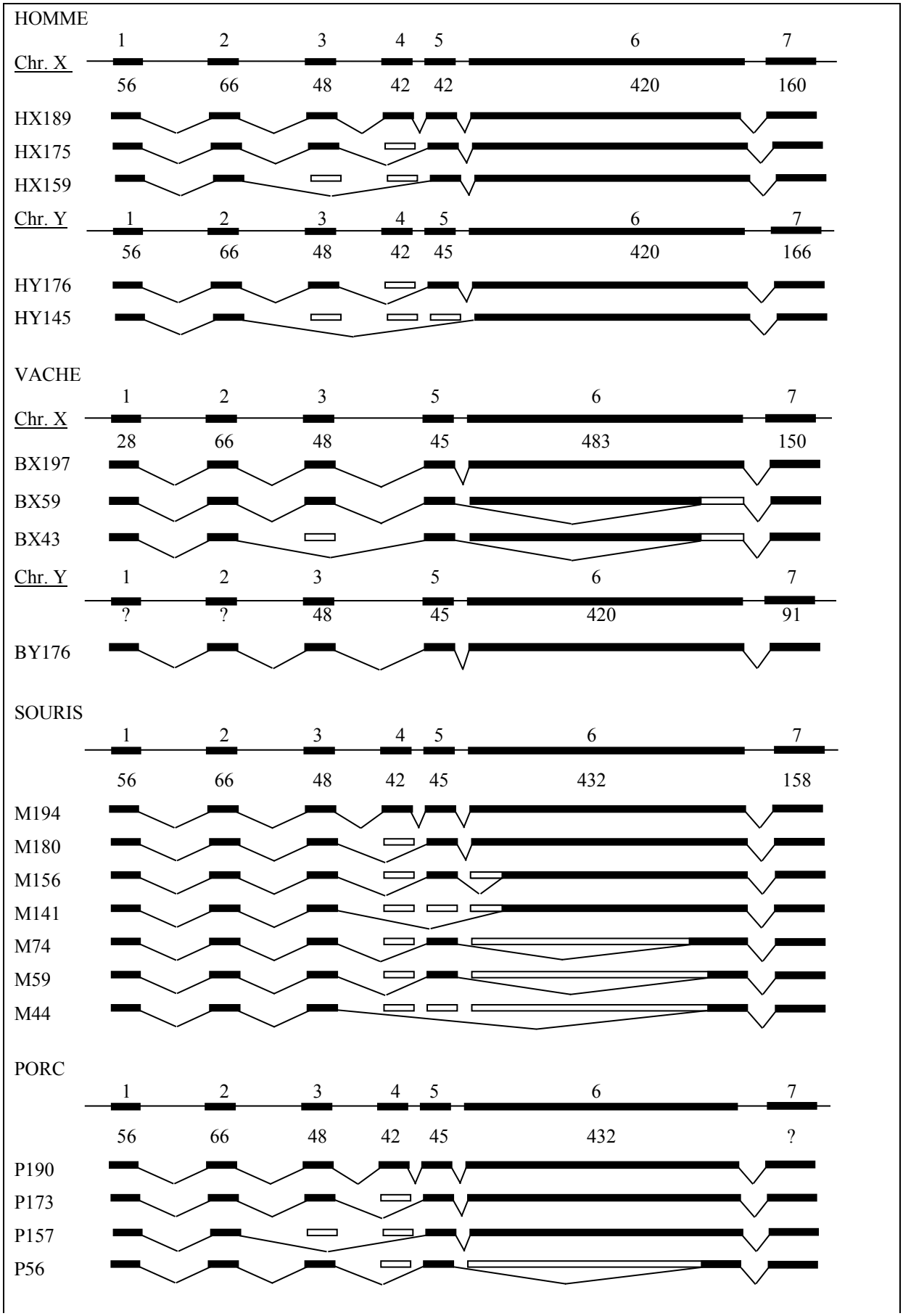

HOMME
1 2 3 4 5 6 7
Chr. X
56 66 48 42 42 420 160
HX189
HX175
HX159
Chr. Y
1 2 3 4 5 6 7
56 66 48 42 45 420 166
HY176
HY145
VACHE
1 2 3 5 6 7
Chr. X
28 66 48 45 483 150
BX197
BX59
BX43
Chr. Y
1 2 3 5 6 7
? ? 48 45 420 91
BY176
SOURIS
1 2 3 4 5 6 7
56 66 48 42 45 432 158
M194
M180
M156
M141
M74
M59
M44
PORC
1 2 3 4 5 6 7
56 66 48 42 45 432 ?
P190
P173
P157
P56

Figure 7. Résumé des connaissances sur l'épissage alternatif des exons de l'amélogénine. Ce schéma ne tient pas compte de l'épissage des exons 8 et 9 dont les données sont trop récentes et incomplètes pour être utilisées (d'après Simmer, 1995).

▭ = Régions subissant l'épissage alternatif

II.3. Structure primaire de l'amélogénine.

Sur la séquence protéique de l'amélogénine humaine (Figure 8), on peut observer l'emplacement des sites de protéolyse TRAP et du site d'épissage alternatif décrit précédemment. On observe également les différences existant entre les amélogénines X et Y (voir chapitre 2 - "Evolution du gène de l'amélogénine"), notamment la délétion d'une méthionine dans la séquence de l'exon 5 de l'amélogénine X. Il est également possible de remarquer la richesse en proline (P) et glutamine (Q) de l'exon 6, une caractéristique importante de l'amélogénine qui sera aussi étudiée dans le chapitre 2. Cette présentation de la séquence protéique va permettre également de se familiariser avec les acides aminés ou les groupes d'acides aminés dont on va discuter dans la suite.

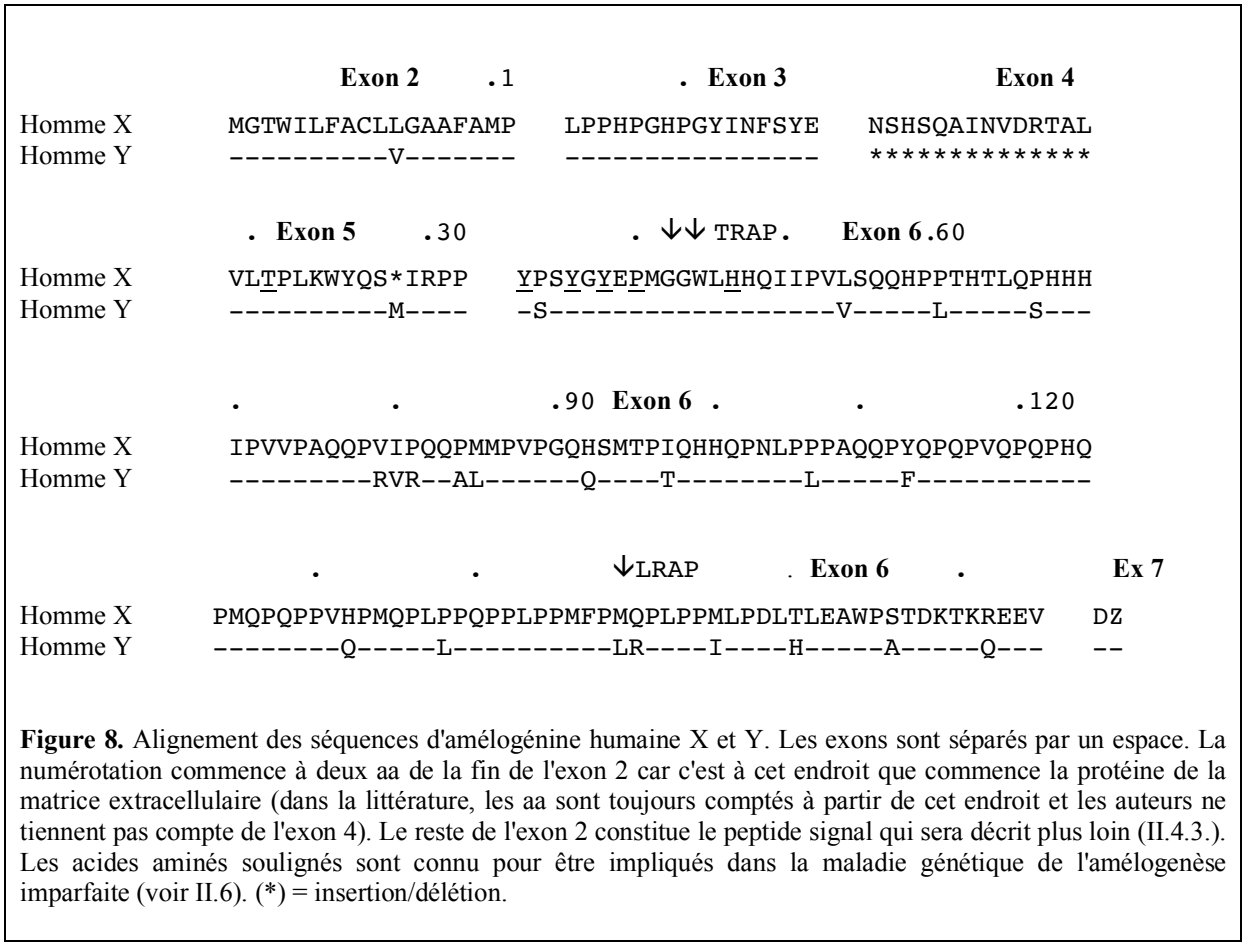


Figure 8. Alignement des séquences d'amélogénine humaine X et Y. Les exons sont séparés par un espace. La numérotation commence à deux aa de la fin de l'exon 2 car c'est à cet endroit que commence la protéine de la matrice extracellulaire (dans la littérature, les aa sont toujours comptés à partir de cet endroit et les auteurs ne tiennent pas compte de l'exon 4). Le reste de l'exon 2 constitue le peptide signal qui sera décrit plus loin (II.4.3.). Les acides aminés soulignés sont connu pour être impliqués dans la maladie génétique de l'amélogenèse imparfaite (voir II.6). (*) = insertion/délétion.

II.4. Phénomènes post-traductionnels

II.4.1. Phosphorylation

Takagi *et al.* (1984) montrèrent que l'amélogénine de vache est phosphorylée sur un seul résidu, la serine-16. Cette étude était en contradiction avec les travaux précédents dans lesquels il était indiqué que l'amélogénine bovine contenait trois résidus phosphorylés (Seyer

et Glimcher, 1977; Fincham *et al.*, 1981). Cependant, les recherches réalisées par Fincham et Moradian-Oldak en 1993 montrèrent que la masse calculée de la protéine d'amélogénine correspond à la protéine phosphorylée sur une seule position, ce qui semble confirmer les résultats de Takagi *et al.* (1984). Les études réalisées ensuite chez d'autres mammifères confirmèrent que la masse calculée des différentes formes de l'amélogénine n'est cohérente que si l'on suppose l'ajout d'un seul groupement phosphate.

II.4.2. Glycosylation

Les travaux de Ravindranath *et al.* (1999) ont montré qu'il existe un motif de liaison pour des N-acétyl-glucosamines (GlcNAc) constitué de 13 aa situé au début de l'exon 6. Grâce à des expériences de mutations dirigées sur un polypeptide de synthèse, les chercheurs ont pu caractériser les résidus nécessaires à la fixation des GlcNAc. Ces travaux ont montré que la présence de 3 tyrosines et d'une proline est essentielle à la liaison des N-acétyl-glucosamines sur l'amélogénine (en gras sur les 2 séquences suivantes). Chez les mammifères, cette portion de la protéine est très conservée et se présente ainsi : *N***Y***N*S**Y**G**Y**E**P**MGGW. Le motif de liaison a donc été défini comme étant : *X***Y***XX***Y***X***Y***X***P***. On peut également remarquer la ressemblance (50%) avec le motif de liaison aux GlcNAc d'une autre protéine, la WGA (Wheat Germ Agglutinin) :

Amélogénine mammalienne	…*NYN***S*YGY**EP**MGG**W…
Wheat Germ Agglutinin	…CC**S**Q**YGY**CG**MGG**D…

Au cours de la maturation de l'émail, l'amélogénine est soumise à des actions protéolytiques qui aboutissent à la formation de deux fragments TRAP. Un fragment TRAP-45 résultant de la coupure qui se produit après le tryptophane (W) (voir *§II.4.4.*), et un fragment TRAP 43 après la coupure entre les deux glutamines (GG). Ravindranath *et al.* (1999) suggèrent que la coupure protéolytique pourrait ne pas avoir lieu sans la présence de GlcNAc sur l'amélogénine. Ils pensent que le site de liaison aux GlcNAc pourrait être constitué par la partie C-terminale du peptide TRAP-45, c'est à dire –MGGW. Ce phénomène semble être extrêmement important pour la fonction de la protéine puisqu'il existe chez l'homme un cas d'*Amelogenesis imperfecta* qui est la conséquence d'une mutation touchant la proline (P40 sur l'amélogénine X) considérée comme essentielle à la liaison des GlcNAc (Collier *et al.*, 1997).

II.4.3. Peptide « signal »

Un peptide signal est un segment de 15 à 30 acides aminés présent à la partie N-terminale d'une protéine, et qui indique à la machinerie cellulaire que cette protéine doit être exportée ou sécrétée. Ce peptide, qui permet le passage de la protéine à travers une membrane, est généralement clivé au cours du processus de sécrétion ou d'exportation par une protéase spécifique. Il n'est donc pas présent dans la protéine mature (Figure 9).

Chez les eucaryotes, dès le début de la synthèse de ce type de protéines, la chaîne polypeptidique est insérée dans la lumière du réticulum endoplasmique. Cette opération complexe fait intervenir une reconnaissance spécifique de la séquence signal par des récepteurs situés sur la membrane externe du réticulum. Après cette reconnaissance, la séquence signal se retrouve à l'intérieur de la lumière et la phase d'élongation de la synthèse protéique (au niveau des ribosomes) se poursuit avant que le peptide signal soit éliminé (Figure 9). Le peptide signal de l'amélogénine est codé par une grande partie de l'exon 2; il commence au premier codon (12ème base de l'exon 2) et se termine 6 bases avant le début de

l'exon 3. Au cours de son passage dans le réticulum, l'amélogénine subit des modifications chimiques (glycosylation, phosphorylation, etc.) avant d'être libérée par exocytose dans la matrice extracellulaire.

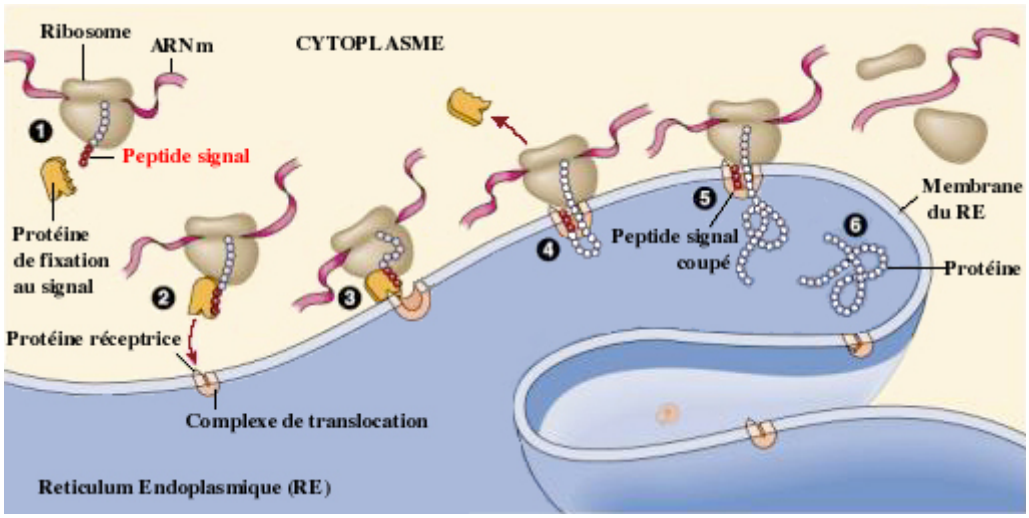


Figure 9. *Schéma montrant le rôle du peptide "signal" dans le transport des protéines à l'intérieur de certains compartiments cellulaires comme le réticulum endoplasmique.*

II.4.4. Protéolyse

Les fragments d'amélogénine sont soumis à des coupures protéolytiques au cours de la maturation de l'émail. Ces coupures réduisent la taille de l'amélogénine depuis la partie C-ter, aboutissant à la formation des fragments appelés TRAP 45 et 43. La protéolyse semble être un événement capital dans le cycle et la fonction de la protéine tant elle se retrouve chez tous les groupes étudiés. Les travaux de Fincham (1990) sur les phénomènes post-sécrétoires chez la souris post-natale, propose un schéma résumant ces processus (Figure 10).

De plus, Fincham et Moradian-Oldak (1995) ont indiqué, chez le porc, les emplacements exacts des coupures protéolytiques. C'est ainsi que l'on passe successivement d'une protéine de 173 résidus à un polypeptide de 43 acides-aminés, le TRAP-43 (Figure 11). Ces données sont intéressantes pour vérifier si les coupures protéolytiques se font de manière spécifique. En effet, les séquences sont fortement conservées chez les mammifères et l'on peut se demander si les sites de coupures le sont également.

D'après l'étude de Fincham et Moradian-Oldak (1995) on constate une grande diversité des emplacements des sites de coupures. Comme les séquences protéiques et nucléiques sont bien conservées, on peut penser que les coupures protéolytiques ne doivent pas nécessairement se produire sur un acide aminé précis ou sur une combinaison d'acides aminés. Ceci est confirmé par le fait que les emplacements des sites de coupures changent, chez une même espèce, pour les différentes protéines résultant de l'épissage alternatif (Figure 12 : souris/rat). Chez d'autres mammifères étudiés en médecine dentaire, les emplacements exacts des sites de coupures sont également connus (Fincham et Moradian-Oldak, 1996) (Figure 12).

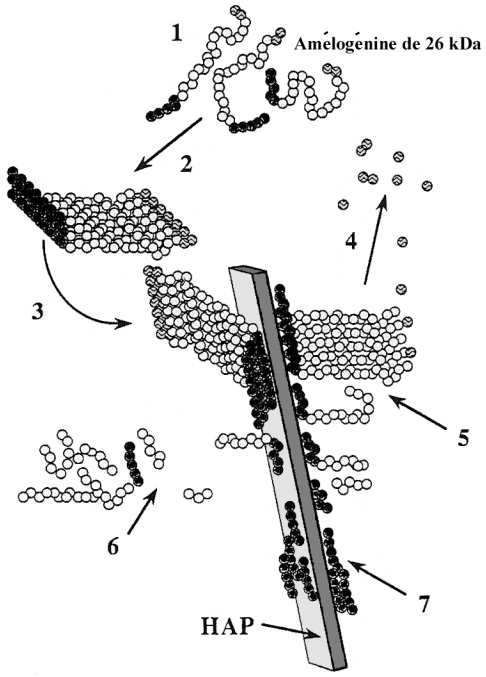


Figure 10. *Modèle des phénomènes post-sécrétoires que subit la protéine d'amélogénine durant la minéralisation de l'émail, d'après Fincham et* al. *(1991).* ***1.*** *Amélogénine de 26 kDa synthétisée par les améloblastes.* ***2.*** *Formation d'agrégats d'amélogénine suivant une polarité hydrophobe-hydrophile.* ***3.*** *Les agrégats d'amélogénine s'associent aux cristaux d'hydroxyhapatite (HAP).* ***4.*** *La partie hydrophile C-terminale est clivée.* ***5.*** *Les régions C-terminales de l'amélogénine subissent un clivage séquenciel et contrôlé.* ***6.*** *Les polypeptides résultants du clivage progressif deviennent "solubles" et se dispersent avant d'être totalement résorbés.* ***7.*** *La partie N-ter de l'amélogénine correspondant aux peptides TRAP persiste dans la matrice en cours de minéralisation jusqu'à une étape avancée du phénomène grâce à ses propriétés d'agrégation et son affinité pour l'hydroxyapatite.*

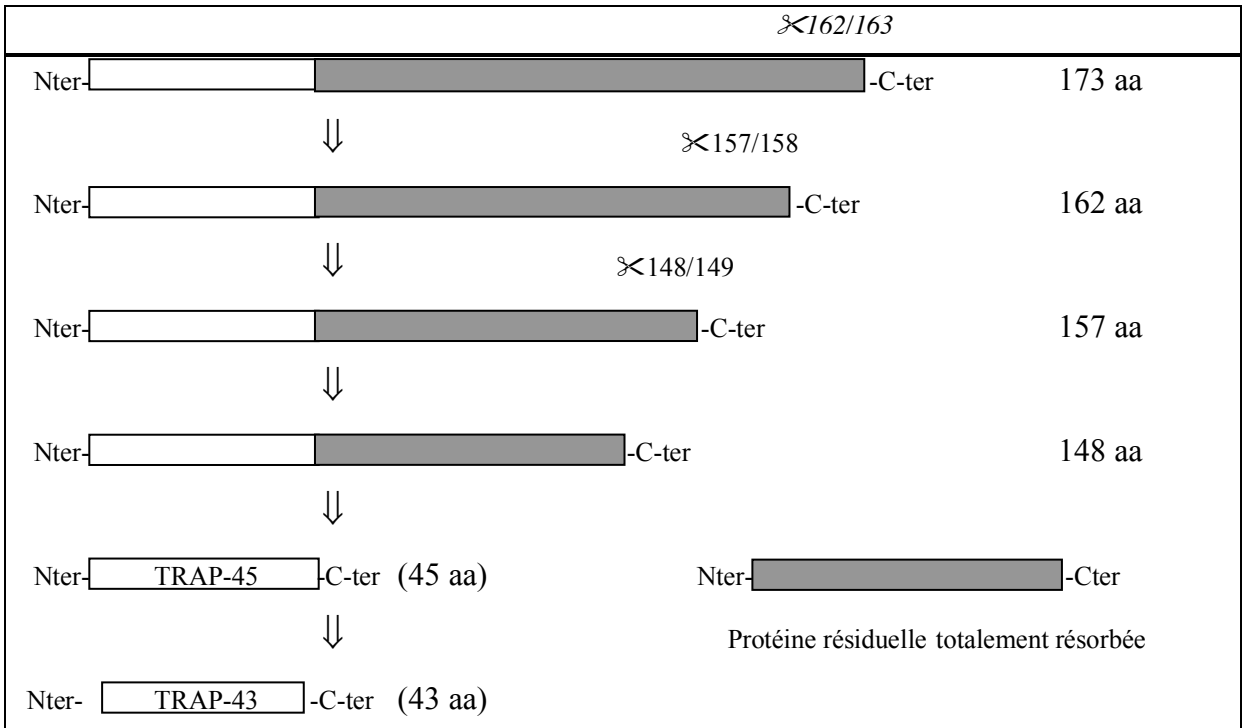


Figure 11. Schéma des processus protéolytiques concernant l'amélogénine porcine de 173 acides aminés. Le fragment TRAP est indiqué en clair (d'après Fincham et Moradian-Oldak, 1994).

Vache (X-197)
P Q P P V H P I Q P L P P Q P P L P P I F P M Q P L P P M L P D L P | L E A W P | A T D K T K R E E V D
Homme (X-175)
P Q P P V H P M Q P L P P Q P P L P P M F P M Q P L P P M L P | D L T | L E A W P | S T D K T K R E E V D
Cochon (173)
P Q S P M H P I Q P | L L P Q P P L P P M F | S | M Q S L - - - L | P D L P | L E A | W P | A T D K T K R E E | V D
Souris (180)
P Q S P L H P M Q P L A P Q P P L P P | L | F S M Q | P L | S P I | L P E | L P L | E A W | P | A T D K T K R E E | V D
Rat (180)
P Q S P L H P M Q P L A P Q P P L P P | L | F S M Q P L S | P | I | L | P | E L P L E A W P A T D K T K R | E E | V D
Souris (156)
P Q S P L H P M Q P L A P Q P P L P P L | F S M Q P L S P I | L P E | L P L E A W P A T D K T K R E E | V D
Rat (156)
P Q S P L H P M Q P L A P Q P P L P P L | F S M Q P L S P I L | P | E L P L E A W P A T D K T K R | E E | V D

Figure 12. Protéolyse supposée de la partie C-ter de l'amélogénine chez quelques mammifères. Les séquences souris/rat-180 et souris/rat-156 correspondent à des isoformes résultant du phénomène d'épissage alternatif (voir II.2.) (d'après Fincham et Moradian-Oldak, 1996). Les emplacements des coupures sont différents pour les deux isoformes de la souris et du rat.

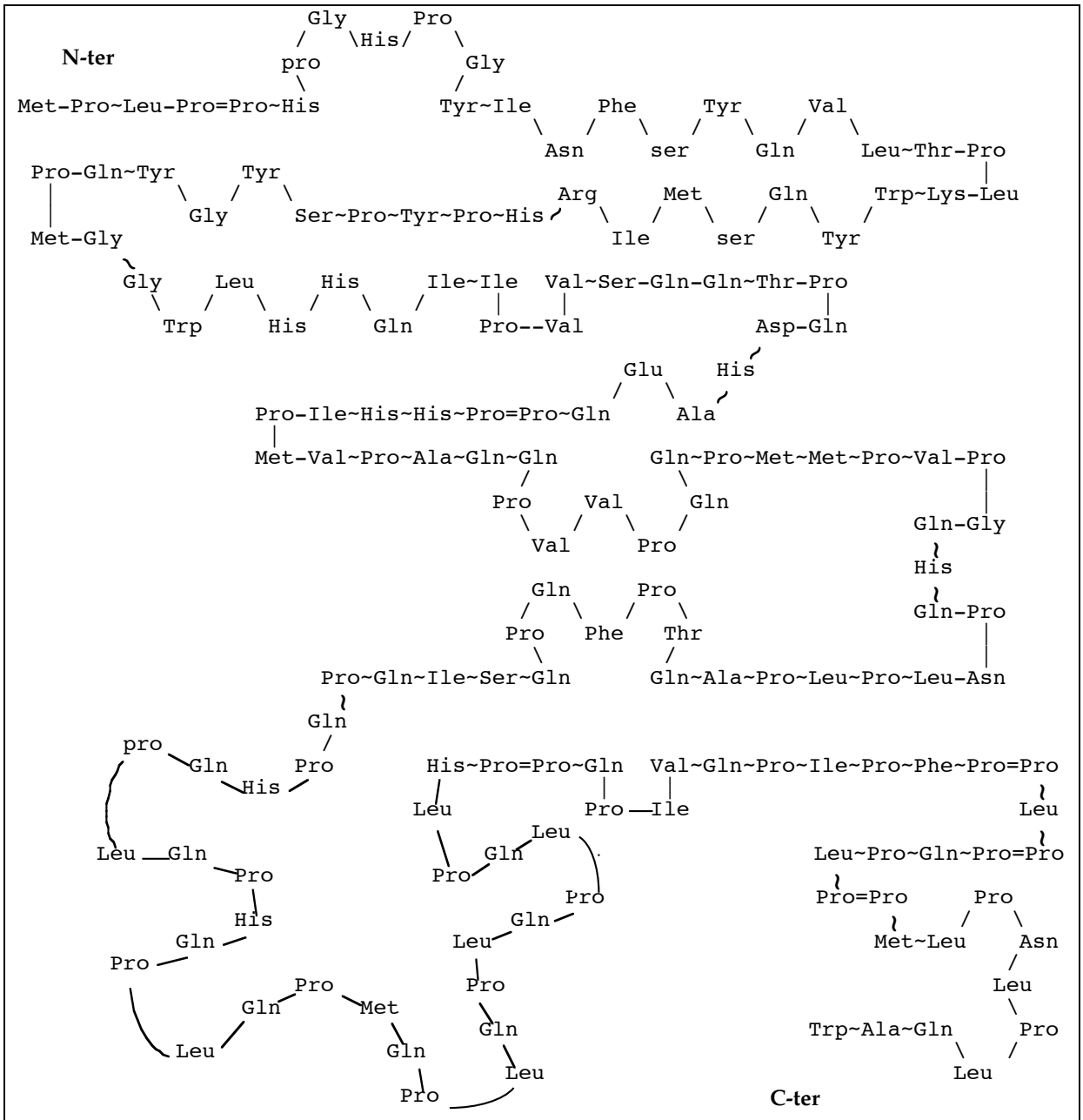
En revanche, les emplacements des deux derniers sites de coupure qui permettent d'obtenir les fragments TRAP-45 et TRAP-43 (coupure après les résidus 43 et 45 de l'amélogénine) sont toujours très conservés chez les mammifères. Ces sites se trouvent dans une région bien conservée de l'ADN, chez tous les vertébrés dont on connaît l'amélogénine. On peut donc penser à un rôle très important de ces deux coupures dans la fonction de la protéine et supposer l'existence de fragments TRAP dans l'émail en formation des autres vertébrés. Cependant, le rôle des TRAP reste peu connu, même si Akita *et al.* (1992) proposent que ce motif est impliqué soit dans des interactions avec les glycoprotéines de la matrice de l'émail, comme l'enaméline, la tufteline ou l'améloblastine, entraînant une stabilité structurale de la matrice de l'émail, soit comme molécule de signalisation en se fixant sur les glycoprotéines de la surface cellulaire.

II.5. Structure secondaire et tertiaire de l'amélogénine

Pour mieux comprendre le rôle et la fonction de l'amélogénine (voir *§II.7.*), certains auteurs ont tenté de mieux connaître la structure secondaire de l'amélogénine, c'est-à-dire les différentes liaisons qui ont lieu au sein de la protéine et qui lui donnent une structure particulière en deux dimensions. Grâce à une technique de spectroscopie infrarouge couplée à des calculs statistiques (Renugopalakrishnan *et al.*, 1986), il a été possible de prédire la structure secondaire de l'amélogénine bovine (d'après la séquence de Takagi *et al.*, 1984) (Figure. 13).

D'après ces travaux, la protéine serait constituée d'un certain nombre de feuillets-ß et de "ß-turn", lui donnant une structure inhabituelle appelée "spirale-ß". Cependant, les auteurs précisent qu'à ce stade des connaissances cette conformation doit être considérée comme une « tentative » de modélisation de la structure secondaire. Renugopalakrishnan *et al.* (1989a, 1989b), ont tenté de préciser la structure et de localiser plus précisément cette "spirale-ß",

mais sans succès. Toutefois, ils proposent que l'amélogénine constitue un canal membranaire permettant le passage de cations (comme le Ca^{2+}) de manière différente des autres canaux ioniques connus à cette époque. Ils indiquent également avoir identifié des interactions possibles de l'amélogénine avec des phospholipides membranaires chargés négativement, ce qui laisserait supposer une localisation membranaire de la protéine. Cette hypothèse, en désaccord à l'époque avec une localisation extracellulaire de la protéine, est aujourd'hui rejetée. Plus tard, Gibson *et al.* (1992) ont réalisé des calculs mathématiques sur l'amélogénine bovine pour tenter de comprendre les différences fondamentales entre les deux protéines des chromosomes X et Y. Cette étude a permis de relever certaines différences notables dans la structure secondaire des deux protéines X et Y, mais elle n'a pas pu associer ces différences à un fonctionnement précis de la protéine (Figure 14).



(/ \ / \ / \) = Hélice ß ; (|_|) = "ß-Turn" ; (~·~) = structure non-ordonnée; (=) = liaison polyproline.

Figure 13. Structure secondaire de l'amélogénine bovine déphosphorylée, prédite d'après la structure primaire rapportée par Takagi *et al.* (1984) en utilisant l'algorithme de Chou-Fasman. (D'après Renugopalakrishnan *et al.*, 1986).

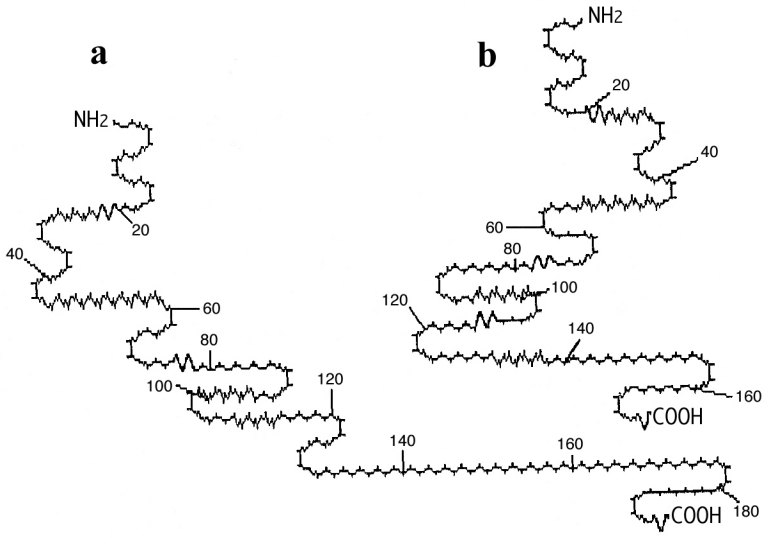


Figure 14. Structure secondaire de l'amélogénine bovine calculée par l'algorithme Chou-Fasman.
a. Amélogénine X ; **b**. Amélogénine Y. D'après Gibson *et al.* (1992).

De plus, la structure secondaire de l'amélogénine bovine déterminées par le même algorithme de calcul est assez différente dans ces deux études (comparer les Figures 13 et 14). Cela pose le problème de la fiabilité de cette méthode de prédiction : peut-elle tenir compte de tous les paramètres biochimiques et n'est-ce pas réducteur que de réaliser un modèle en 2 dimensions alors que la fonction d'une protéine est très souvent déterminée par sa structure en 3 dimensions ?

Dernièrement, Renugopalakrishnan (2002) a montré qu'une région centrale de l'exon 6 de l'amélogénine bovine (chromosome X) prend une conformation en Spirale-ß lui conférant des propriétés uniques. Cette découverte intéressante doit être relativisée sachant que cette région comprend une insertion d'acides aminés répétés qui n'existent pas chez de nombreux mammifères dont l'homme.

II.6. Amélogenèse imparfaite [Amelogenesis imperfecta]

Pour mieux connaître les régions qui jouent un rôle important dans la fonction d'une protéine, une des méthodes les plus employée est d'observer l'effet (le phénotype) d'un arrêt de son expression. Parfois c'est l'inverse qui se produit : un phénotype particulier ou une maladie génétique sont *a posteriori* associés à une mutation d'un gène particulier. Dans le cas de l'amélogénine, c'est ce qui s'est produit : il existait historiquement une maladie, l'amélogenèse imparfaite (AI), que l'on a tenté de lier à des modifications intéressant un ou plusieurs gènes.

L'amélogenèse imparfaite est une maladie génétique rare caractérisée par une coloration marron de la dent résultant d'un manque de calcium (hypocalcification) ou d'un sous-développement (hypoplasie) de la couche externe dure couvrant la dent (émail). Ce désordre est divisé en plusieurs formes basées sur la sévérité du défaut de l'émail (par exemple, une absence complète [agenesis], un mauvais développement, etc.). Les individus

atteints d'amélogenèse imparfaite peuvent être enclin à des pertes précoce et/ou des maladies des structures qui entourent et supportent la dent (maladies parodontales). Sa transmission génétique peut être liée au chromosome X, autosomique dominante ou autosomique récessive. Les différentes formes de la maladie sont :

- L'hypoplasie brune héréditaire de l'émail
- Le type hypocalcifié (Hypominéralisé)
- Le type hypomature ("Snow-Capped Teeth " en anglais)

Cette maladie a été étudiée au niveau de la composition minérale de l'émail, de manière à comprendre la différence qui pouvait exister entre une dent "malade" et une dent normale d'un point de vue chimique (Wright *et al*., 1995). La composition protéique de la matrice de l'émail a ensuite été observée pour les différentes formes de la maladie (Wright *et al*., 1996).

L'origine génétique des formes autosomiques est encore mal connue, bien que l'on sache qu'une des formes autosomiques dominantes de l'amélogenèse imparfaite est la conséquence d'une mutation dans le gène de l'énaméline (Rajpar *et al.*, 2001) (voir *§III.2.*) et qu'une des formes autosomique récessive étudiée chez trois familles suédoises (Forsman *et al.*, 1994) semble être liée au gène de l'améloblastine (voir *§III.1.*).

En revanche, la forme de la maladie liée au chromosome X est définitivement associée à des mutations dans le gène de l'amélogénine. L'emplacement exact de ces mutations reste mal connu, même s'il a été possible d'identifier une substitution C-T dans l'exon 5 de l'amélogénine (Thr21 remplacée par une Ile) (Lench *et al.*, 1995) (voir Figure 8). Cependant, Collier *et al.* (1997) décrivent un cas d'amélogenèse imparfaite localisé précisément au niveau de l'acide aminé 40 (Pro40 remplacée par une Thr) dont le phénotype unique, présent au niveau de l'émail, a ensuite été analysé par Ravassipour *et al.* (2000) et Hart *et al.* (2000a, b). Comme on l'a vu précédemment (*§II.4.2.*), cette région (TRAP) semble importante pour la fonction de la protéine. Elle a d'ailleurs fait l'objet d'une étude détaillée afin de mieux comprendre son rôle (Ravindranath *et al.*, 1999). Très récemment, Hart *et al.* (2002) ont rapporté une mutation au niveau de l'histidine 47 (His47 remplacée par une Leu), proche de la mutation Pro40Thr décrite par Ravassipour *et al.* (2000), confirmant l'idée que cette partie de l'amélogénine est soumise à d'importantes contraintes.

Les autres mutations identifiées correspondent à des "mutations non-sens", c'est-à-dire soit la création d'un codon "stop" au milieu de la séquence par mutation ponctuelle, soit un changement de phase de lecture résultant d'une délétion ou d'une insertion. Ce type de mutation supprime soit toute la partie de la protéine située après le codon "stop", soit modifie les propriétés chimiques de la protéine (Aldred *et al.*, 1992 ; Lench *et al.*, 1994 ; Kindelan *et al.*, 2000; Greene *et al.*, 2002). Ces mutations ne permettent pas de mettre en évidence le rôle d'un acide aminé précis ou d'une partie de la protéine. Cependant, Hart *et al.* (2002) ont identifié une mutation non-sens supprimant une vingtaine de résidus du coté C-terminal. Or, comme on le verra dans le chapitre II ("Evolution de l'amélogénine"), cette région est particulièrement conservée et semble jouer un rôle important pour la fonction de la protéine. Par ailleurs, cette région est la seule qui soit hydrophile dans la partie C-ter de l'amélogénine (Moradian-Oldak *et al*., 1994). Il existe également une mutation décrite dans le peptide signal (voir *§II.4.3.*) entraînant la perte de 2 acides aminés. Cette perte perturbe la translocation de la protéine au cours de sa synthèse et son exportation vers l'extérieur de la protéine.

II.7. Rôle de l'amélogénine

Comme nous venons de le voir dans les paragraphes précédents, la structure du gène et celle de la protéine, les sites de protéolyse et d'épissage, et de nombreuses mutations sont maintenant bien connus. Cependant, cela contraste avec le fait que les connaissances sur le

rôle de la protéine dans son ensemble et de ses différentes régions sont encore très imprécises. En effet, le rôle de l'amélogénine a beaucoup été discuté et différentes théories ont été proposées au fur et à mesure de l'avancée des connaissances. Il a tout d'abord été proposé que l'amélogénine jouait un rôle important dans l'organisation structurale des prismes d'émail (Travis et Glimcher, 1964), ou qu'elle était impliquée dans la nucléation de la partie minérale de l'émail (Glimcher, 1979), ou encore dans le transport du calcium (Renugopalakrishnan *et al.*, 1986). Actuellement, l'amélogénine est supposée avoir un rôle dans le contrôle de la croissance (en longueur et en épaisseur) et de l'orientation des cristaux d'émail durant le développement des dents. En 1996, Aoba a étudié, chez le porc, la cinétique de croissance de l'émail ainsi que la structure des cristaux en microscopie électronique et a tenté de comprendre les mécanismes de croissance des cristaux d'émail. D'après cet auteur, l'amélogénine servirait à « contrôler le nombre et l'orientation des fins "rubans" qui précipitent dans la région la plus externe de l'émail, à prévenir la prolifération aléatoire des cristaux d'apatite et enfin à réguler la cinétique de croissance de l'apatite pour éviter une sur-prolifération au niveau du "précurseur[1]" » (traduit de Aoba, 1996).

De plus, le rôle de l'amélogénine semble intimement lié à sa structure quaternaire qui permet l'association de plusieurs molécules (Fearnhead et Kawasaki, 1991). L'amélogénine formerait des agrégats, nanosphères de 20 nm dont le comportement *in-vitro* dépend principalement du pH et de l'hydrophobicité de la molécule (Moradian-Oldak *et al.*, 1995).

[1] Molécule d'octa-calcium phosphate qui sert de support à la croissance des cristaux.

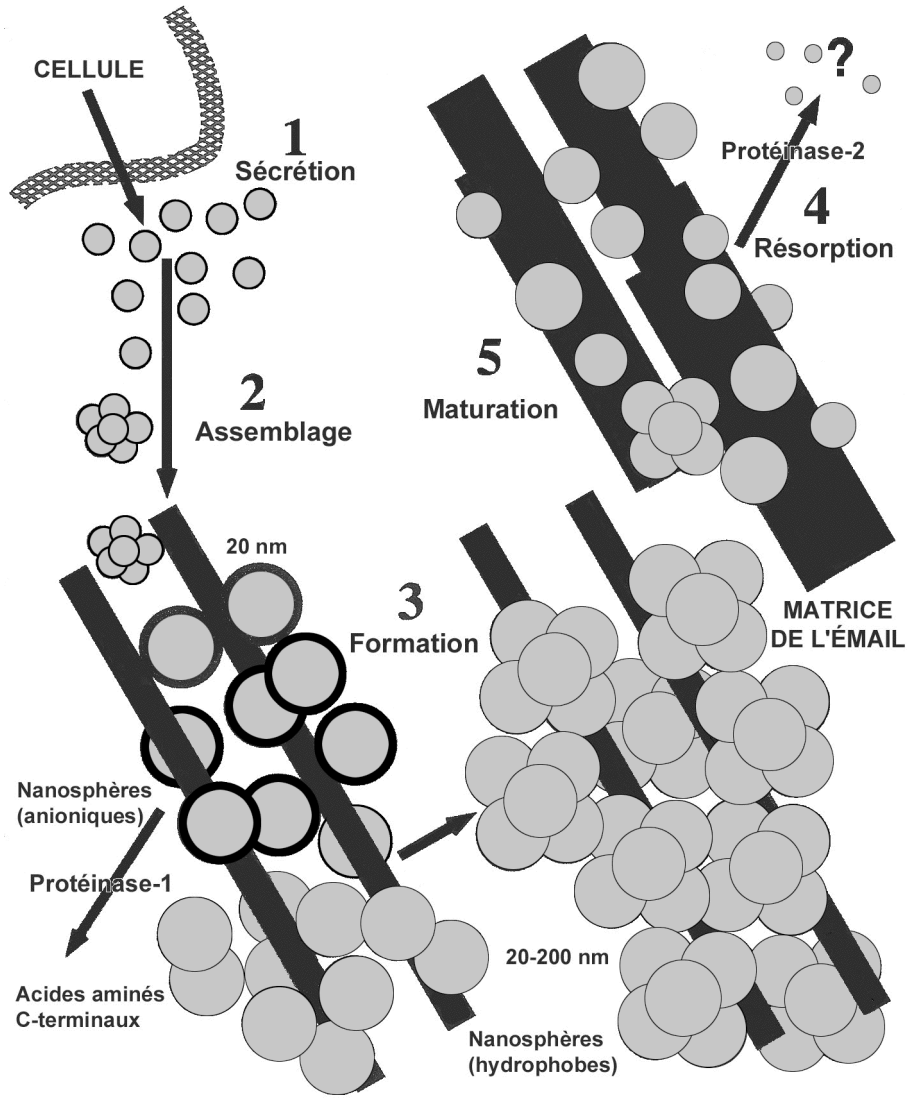


Figure 15. Schéma montrant le rôle de l'amélogénine dans le processus de minéralisation de l'émail (d'après Fincham *et al.*, 1999).

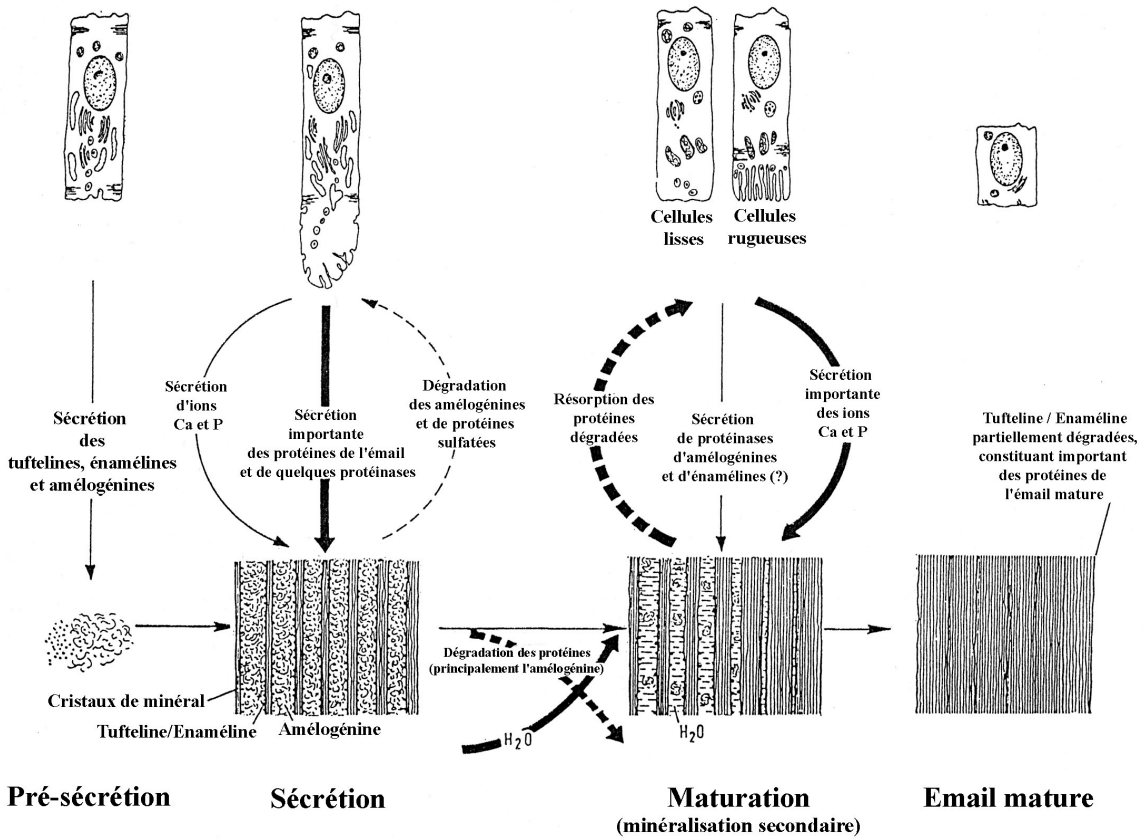
Il semble, d'après ces derniers travaux que la perte, au cours de la protéolyse, de la partie C-terminale hydrophile de la protéine changerait complètement les capacités d'assemblage des molécules. Wen *et al.* (1999) ont montré que les nanosphères d'amélogénine possèdent une structure hiérarchique : les nanosphères de 8-20 nm ont tendance à s'assembler pour former des structures sphériques de 40-70 nm de diamètre qui peuvent elle-mêmes former de larges assemblages de 70-300 nm de diamètre, le passage d'un état à l'autre dépendant essentiellement des conditions du milieu (température, pH). Tout cela semble suggérer que les interactions hydrophobes entre les molécules d'amélogénine sont importantes pour déterminer la structure des différents assemblages d'amélogénine et, par conséquent, celle de la matrice de l'émail. Fincham *et al.* (1999) intègrent toutes les données concernant l'amélogénine pour proposer un modèle de biominéralisation de l'émail en prenant en compte les différentes structures quaternaires de l'amélogénine (Figure 15).

L'accumulation de nombreuses données issues d'études biochimiques et de l'analyse génétique d'un plus grand nombre de mutations apporteront sans aucun doute des éléments

essentiels à une meilleure connaissance du rôle que tient l'amélogénine dans la formation de l'émail. A cet égard, on peut citer les travaux récents de Veis *et al.* (2000) qui ont montré que certains produits d'épissage de l'amélogénine sont des molécules signalantes jouant un rôle dans la maturation des préodontoblastes.

II.8. Expression de l'amélogénine

L'amélogénine est une molécule spécifique de l'émail et elle ne s'exprime donc que dans les cellules spécialisées dans la synthèse de la matrice de l'émail, les améloblastes. Les améloblastes sont des cellules différenciées de la couche basale de l'épithélium dentaire qui initie les premiers stades de développement de la dent en interagissant avec les cellules du mésenchyme sous-jacent. A la suite de cette interaction, l'épithélium dentaire s'invagine dans le mésenchyme qui forme, à son contact, une condensation cellulaire appelée la papille dentaire au sein de laquelle se différencient les odontoblastes, cellules spécialisées dans la production de la matrice de la dentine. Les tissus qui composent la dent (dentine et émail) se forment à l'interface entre ces deux types cellulaires, les améloblastes et les odontoblastes. Grâce à l'hybridation *in situ* et l'immunohistochimie, l'expression du gène et de la protéine d'amélogénine ont été étudiées durant la formation de l'émail et de manière intensive chez les mammifères, en particulier chez les rongeurs dont les dents poussent et fabriquent de l'émail continuellement. Les résultats obtenus sur l'expression de divers gènes lors de la morphogenèse et la différenciation des dents de rongeurs sont d'ailleurs résumés dans un grand projet de l'Institut de Biotechnologie de l'Université d'Helsinki (Finlande), appelé « Gene expression in tooth » et disponibles sur internet (http://bite-it.helsinki.fi/). De plus, les travaux de Bleicher *et al.* (1999) montrent, chez le rat, l'expression à différents stades de développement de plusieurs gènes de la matrice de l'émail et de la dentine. Chez les bovins, on peut consulter les travaux de C. Gibson (Gibson *et al*., 1995 ; Gibson, 1999), qui ont mis en évidence des différences d'expression entre le gène provenant du chromosome X et celui du chromosome Y. Chez l'homme, l'expression de l'amélogénine est observée chez un embryon *in-utero* dès le 3ème mois et chez l'embryon de souris les ARNm de l'amélogénine ont été détectés dès le 15ème jour alors que l'émail ne commence à se former qu'au 18ème jour (Gibson, 1999). Chez tous les mammifères étudiés, l'expression de l'amélogénine commence avant la minéralisation initiale de la dentine et se termine au stade précoce de la maturation de l'émail. Cette "maturation" est un phénomène qui combine, d'une part, la protéolyse de la matrice de l'émail par diverses protéinases (voir *§III.4.*) et, d'autre part, une minéralisation croissante qui aboutira à un émail mature contenant 95% de minéral. Ces protéinases sont secrétées par les améloblastes dès que la phase de minéralisation commence. Ces différentes étapes sont présentées dans le schéma ci-après.



Principales étapes de la formation et de la minéralisation de l'émail. Ce schéma illustre les modifications structurales des améloblastes en fonction de leur activité de synthèse, et les changements qui en résultent pour l'émail (d'après Deutsch *et al.*, 1995).

Récemment, Oida *et al.* (2002) ont montré, chez le porc, l'expression du gène de l'amélogénine au niveau des odontoblastes, bien que ces cellules sont spécialisées dans la fabrication des protéines de la dentine. C'est la première fois que l'on observe des ARNm d'amélogénine en dehors des améloblastes. Cependant, le niveau d'expression de l'amélogénine dans les odontoblastes est 1000 fois inférieur à celui des améloblastes. Les auteurs, qui se sont assurés de l'absence de contaminations, ont observé que les odontoblastes expriment d'autres protéines de l'émail comme l'améloblastine, l'énaméline ou l'énamélysine. D'après eux, ces protéines de l'émail, secrétés dans les odontoblastes, pourraient être impliquées dans la différenciation des améloblastes et jouer quand même un rôle dans la biominéralisation de l'émail.

III. LES AUTRES PROTÉINES DE LA MATRICE EXTRACELLULAIRE

Bien que l'amélogénine représente 90% des protéines de la matrice extracellulaire de l'émail chez les mammifères, il existe d'autres protéines indispensables à la fabrication de l'émail dentaire puisque l'on peut les associer à des cas d'amélogenèse imparfaite. Dès que les techniques ont permis de mettre en évidence la présence ces protéines, des études ont immédiatement tenté de les caractériser. Aujourd'hui, un certain nombre de ces protéines est connu, mais les données qui les concernent sont plus fragmentaires que celles sur l'amélogénine car historiquement plus récentes.

III.1. L'améloblastine (ou améline ou sheathline)

Le terme "améloblastine" fut utilisé pour la première fois en 1996 par Krebsbach *et al.* Avant cela, les parties C-ter et N-ter de l'améloblastine avaient été isolées séparément au cours de recherches des protéines de l'émail dentaire de porc. La partie C-ter fut découverte lors d'une recherche des protéines de l'émail qui se lient au calcium grâce une purification progressive de bandes protéiques marquées par Stain All Dye (Fukae et Tanabé, 1987). La partie N-ter fut découverte durant une recherche des protéines de l'émail dentaire "non-amélogénines" et ressemblant à l'énaméline (Fukae et Tanabé, 1985, 1987). Cette région N-ter de l'améloblastine se présente, dans la matrice de l'émail, comme un groupe de peptides de faible poids moléculaire variant de 13 à 17 kDa résultant du clivage d'une protéine "native", avec des propriétés agrégatives.

La composition globale en acides aminés de l'améloblastine en faisait une protéine distincte de l'amélogénine et de l'énaméline (Fukae et Tanabé, 1987). Plus tard, des expériences d'immunomarquage (Uchida *et al.*, 1991 ; Fukae *et al.*, 1993) ont montré que toutes les protéines de 10 à 20 kDa et de 20 à 30 kDa (dont l'améloblastine) sont présentes sur toute l'épaisseur de l'émail en formation excepté dans la partie la plus superficielle (à 30 μm des améloblastes). De plus, ce marquage montre un patron d'expression en forme de "nid d'abeille". Par la suite, des expériences d'adsorption ont montré que les protéines de 13 et 17 kDa non-amélogénines (dont l'améloblastine) présentaient une affinité faible pour l'hydroxyapatite (Akita *et al.*, 1992).

En 1995, Uchida obtient la première séquence partielle en acides aminés d'améloblastine par séquençage direct des protéines (méthode d'Edman, 1970). Il construit à partir de ces connaissances des anticorps contre ce morceau de la protéine et observe le même patron de marquage en forme de "nid d'abeille" que précédemment obtenu pour les protéines de 10-20 kDa et 20-30 kDa. Ce patron d'expression est dû, en réalité, à la présence d'améloblastine dans l'espace interprismatique ou "sheath space" en anglais, ce qui suggéra à Uchida *et al.* (1995) de baptiser cette nouvelle famille de protéines de l'émail les "Sheath" protéines.

En 1996, deux groupes de recherche (Cerny *et al.* (1996) et Krebsbach *et al.* (1996)) ont criblé au hasard des banques d'ADNc de dents de rat et ont obtenu indépendamment la séquence de l'ADNc de l'améloblastine de rat. Chacun a donné un nom différent à cette nouvelle protéine : "améloblastine" pour le groupe de Krebsbach "améline" pour celui de Cerny. La même année, les expériences d'hybridation *in situ* réalisées chez le rat (Fong *et al.,* 1996) font également référence au nom "améline". Plus tard (Hu *et al.*, 1997), l'ADNc de l'améloblastine de cochon fut cloné et la protéine appelée "sheathline" en référence aux

“sheath“ protéines d’Uchida *et al.* (1995). Cependant, depuis ces travaux, seul le nom “améloblastine“ a été conservé et les nouvelles séquences décrites le sont sous ce nom, comme celle de l’ADNc de l’améloblastine de vache introduite dans GenBank par Machule *et al.* (1999) (GenBank N° d’accession AF157019) et celle de l’ADNg de l’améloblastine humaine (Toyosawa *et al.*, 2000). Très récemment, l’améloblastine a été caractérisée chez un crocodilien : le caïman (Shintani *et al.*, 2002). Ces auteurs constatent que la similarité entre l'améloblastine du caïman et celle des autres mammifères n'est pas très élevée (elle varie de 38 à 47%). Cependant, les acides aminés de l'améloblastine de caïman sont très conservés dans la partie N-terminale de la protéine (entre la position 83 et 104) car il n'y a que deux substitutions et une insertion par rapport à l'améloblastine de l'homme, du cochon, du rat et de la souris. Pour confirmer que ce gène du caïman est bien homologue à l'améloblastine des mammifères, les auteurs ont également montré qu'il s'exprime uniquement au niveau des mâchoires et qu'il n'existe pas d'autres loci d'améloblastine dans le génome du caïman.

Après avoir obtenu les premières séquences d’améloblastine, plusieurs travaux ont tenté de situer ce nouveau gène dans le génome de différents mammifères. Ainsi, le locus du gène de l’améloblastine est localisé sur le chromosome 5 de la souris (Krebsbach, 1996) et sur le chromosome 4 (4q12) de l’homme (MacDougall, 1997), dans une région chromosomique qui a été préalablement liée à une forme locale récessive de “l’amélogenèse imparfaite“ dans trois familles suédoises (Forsman *et al.*, 1994).

Le phénomène d’épissage alternatif ayant été découvert chez l’amélogénine, des isoformes résultant du même mécanisme ont été recherchées pour l’améloblastine. Ainsi, deux isoformes ont été isolées qui diffèrent par la délétion de 15 acides aminés (Hu *et al.*, 1997). Par ailleurs, ces 15 acides aminés sont parfaitement conservés entre le rat et le cochon, ce qui montre qu’ils ont une fonction dans la protéine puisqu’ils n’évoluent pas à la manière de séquences non-codantes.

Dans un premier temps, les recherches en biologie moléculaire sur l’améloblastine n’ont concerné que les ADNc. Or, les séquences des ADNc de l’améloblastine ne permettent pas de connaître les frontières exons-introns. Mais, grâce aux travaux de Toyosawa *et al.* (2000) qui ont travaillé sur l’ADNg, on sait maintenant que le gène de l’améloblastine humaine est constituée de 13 exons (Figure 16).

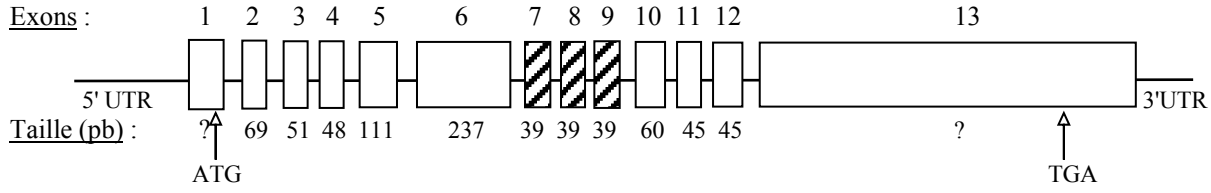


Figure 16. Structure du gène de l’améloblastine humaine d’après les travaux de Toyosawa *et al.* (2000). Les exons hachurés représentent les séquences issues de duplications internes. La taille des exons 1 et 13 n'est pas connue car les auteurs n'ont pas réussi à déterminer les limites des régions non codantes. L'exon 1 code pour 5 aa et l'exon 13 pour 181 aa.

L’améloblastine est, comme l'amélogénine, une protéine spécifique des dents. Elle est exprimée par toutes les cellules de la couche basale de l'organe de l’émail (gaine épithéliale d’Hertwig, améloblastes en présécrétion, en sécrétion, en maturation) et, de façon transitoire, au niveau des préodontoblastes. De plus, les parties C-ter et N-ter montrent des localisations différentes au cours du développement de la matrice de l’émail. En effet, contrairement à ce

qui a été observé pour la partie N-ter dans les expériences d'immunomarquage d'Uchida *et al.* (1991) et de Fukae *et al.* (1993) (voir plus haut), la partie C-ter est fortement concentrée à l'intérieur des 2 μm des processus de Tomes. On la retrouve ensuite de moins en moins concentrée sur une profondeur de 50 μm avec un patron d'immunomarquage en forme de "nid d'abeille" inversé. Elle ne montre aucun marquage dans l'émail plus profond.

Le rôle de l'améloblastine reste pour l'instant mal connu, mais certains auteurs ont suggéré que la protéine pourrait jouer un rôle important dans le contrôle de la croissance des cristaux d'émail et dans la détermination de la structure prismatique (Robinson *et al.*, 1998). Hu *et al.* (1997) pensent que l'améloblastine prévient la croissance des cristallites dans le manteau de l'émail durant la phase de sécrétion et maintient ouvert un chemin par lequel les protéines de l'émail profond peuvent s'échapper durant la phase de maturation.

III.2. L'énaméline

Enaméline est le nom donné par Termine *et al.* (1980) aux protéines de l'émail "non-amélogénines" qui sont fortement adsorbés aux cristaux d'émail et ne sont pas libérées dans la matrice tant que les cristallites ne sont pas dissous. Dans les études précédentes (Mecanics *et al.*, 1971), le terme « énaméline » était appliqué à toutes les protéines "non-amélogénines" de la matrice de l'émail, mais cet usage ne fut pas retenu. Le terme est actuellement utilisé pour désigner une protéine de l'émail caractérisée par son immunolocalisation au cours de la formation de l'émail ainsi que par ses propriétés de liaison aux cristaux d'hydroxyapatite. L'énaméline est une protéine spécifique de l'émail, d'une masse apparente de 186 kDa (la plus grande protéine de l'émail connue jusqu'à présent), dont l'ADNc a été cloné et caractérisé chez le cochon (Hu *et al.,* 1997b), la souris (Hu *et al.*, 1998) et l'homme (Hu *et al.*, 1998). Ces travaux ont permis de constater que la protéine de la souris se distingue de celle de l'homme et du cochon car elle possède 14 copies d'un segment de 11 acides aminés répétés en tandem.

Récemment, l'ADNg de l'énaméline de l'homme et de la souris a été cloné et caractérisé (Hu *et al.*, 2001b), permettant de constater qu'ils mesurent tous les deux environ 25 kb. L'énaméline possède 10 exons interrompus par 9 introns. Jusqu'à aujourd'hui, aucune isoforme résultant d'un épissage alternatif n'a été détectée dans les ARNm de l'énaméline, ce qui n'exclu toutefois pas l'existence de ce phénomène.

Comme toutes les protéines de l'émail, l'énaméline est soumise à une protéolyse après avoir été sécrétée. Les techniques biochimiques ont permis d'isoler les produits du clivage et de les caractériser. Un premier clivage produit des protéines de poids moléculaire apparent de 155, 142 et 89 kDa, puis la protéine de 89 kDa s'accumule avant d'être à nouveau coupée pour générer des fragments de 32 et 25 kDa (Fukae *et al.*, 1996). La protéine de 32 kDa est composée de 106 acides aminés (résidus 174 à 279) qui incluent deux phosphosérines et trois asparagines glycosilées (Yamakoshi, 1995 ; Fukae *et al.*, 1996 ; Yamakoshi *et al.*, 1998).

Les anticorps dirigés contre les fragments de 89 kDa montrent que, au cours de l'étape de sécrétion de l'émail, l'énaméline est spécifiquement localisée depuis la jonction dentine-émail jusqu'à la surface de la dent, et qu'elle disparaît ensuite rapidement durant les étapes précoces de la maturation de l'émail (Uchida *et al.*, 1991 ; Fukae *et al.*, 1993). Le patron d'immunomarquage de l'énaméline forme un "nid d'abeille" inversé qui signifie que la protéine est restreinte à l'émail prismatique et inter-prismatique et qu'elle est absente de l'espace de la gaine dentaire ("sheath space" en anglais). D'après ce patron d'expression et des études *in vivo*, on peut supposer que l'énaméline serait liée aux cristaux d'émail (Tanabe *et al.*, 1990).

Le rôle et la fonction de la protéine sont encore inconnus, mais on pense qu'elle participe à la nucléation et l'extension des cristaux d'émail et à la régulation de "l'apparence"

de l'émail (Hu *et al.*, 2000). On sait également que le gène humain est localisé sur le bras long du chromosome 4 (4q11-q21), dans une région précédemment liée à une forme autosomique dominante de l'amélogenèse imparfaite qui est la conséquence d'une mutation dans le gène de l'énaméline (Rajpar *et al.*, 2001).

III.3. La tufteline

La tufteline est une protéine très controversée, et tout d'abord par sa découverte puisqu'elle n'était pas la protéine recherchée. En effet, l'histoire de cette protéine est liée à la recherche d'une molécule pouvant jouer le rôle de nucléateur pour les cristaux d'émail, c'est-à-dire servant d'initiateur de la cristallisation. Le processus de nucléation s'étend de manière ininterrompue depuis la jonction énamélo-dentinaire (JED) jusqu'à la surface de la dent. Aussi, partant de l'hypothèse que l'élément nucléateur est constitué d'une seule protéine dont la composition globale en acides aminés est connue (Robinson *et al.*, 1975), les recherches d'un facteur de nucléation se sont concentrées dans la région de la JED. La protéine principale de la JED a une taille apparente de 66 kDa et a été appelée "tuft" protéine. Des anticorps dirigés contre cette protéine ont été construits et utilisés pour cribler une banque d'expression d'ADNc d'émail dentaire bovin en formation (Deutsch *et al.*, 1987). Une protéine fabriquée par l'un des clones, ayant répondu positivement aux anticorps, son ADNc a été ensuite séquencé (Deutsch *et al.*, 1989, 1991). La protéine déduite de cette séquence, appelée "tufteline", était nouvelle et ne montrait pas d'homologie avec les autres protéines connues. Sa masse moléculaire prédite (43,8 kDa) était compatible avec la protéine majeure de la JED de 66 kDa, en supposant la présence d'un site unique de glycosilation. Sa composition globale en acides aminés ressemblait à celle de la "tuft" protéine de 66 kDa. Par ailleurs, cette séquence protéique avait toutes les qualités biochimiques que l'on pouvait espérer d'un nucléateur, c'est-à-dire des propriétés acides et de possibles feuillets ß, le tout pouvant interagir avec les groupements phosphates des cristaux d'émail naissant. Pourtant, il apparaît aujourd'hui que la tufteline n'est pas la protéine de 66 kDa de la JED qui a été identifiée plus tard comme étant l'énaméline (Deutsch *et al.*, 1991 ; Zeichner-David *et al.*, 1997).

Une autre controverse est liée à la séquence de la tufteline elle même. En effet, le gène de la tufteline bovine a été cloné et séquencé (Bashir *et al.*, 1998), mais cette séquence d'ADNg diffère de la séquence de l'ADNc, précédemment découverte (Deutsch *et al.*, 1989, 1991); cette dernière présente une délétion d'un nucléotide (G) ce qui change la phase de lecture et modifie les 92 derniers acides aminés qui sont remplacés par 42 autres, car la présence d'un nouveau codon "stop" réduit la séquence protéique déduite auparavant (Deutsch *et al.*, 1989, 1991). La phase de lecture corrigée a aussi été trouvée dans l'ADNc de tufteline de souris et de cochon (MacDougall *et al.*, 1998). Il semble évident que l'ADNc bovin original a été parfaitement cloné, mais l'origine de la délétion reste inconnue. Il pourrait s'agir d'un phénomène "d'édition d'ARN" ("RNA éditing") qui change de manière très spécifique la séquence de certains ARNm en ajoutant ou supprimant des nucléotides à des sites précis. C'est un niveau supplémentaire de régulation de l'expression des gènes et certaines protéines, chez les végétaux ou les trypanosomes, ne seraient pas fabriquées si ce phénomène n'existait pas. Quelle que soit l'origine de la délétion, les discussions en résultant ont jeté un doute sur la séquence protéique réelle de la tufteline. D'ailleurs, Deutsch *et al.* (1989) ont reconnu que d'autres codons de démarrage étaient possible, ce qui ne permet pas de connaître, pour le moment, la séquence exacte en acides aminés de la tufteline, d'autant que la protéine n'a jamais été, à ce jour, isolée de la matrice de l'émail.

De plus, aucun peptide “signal“ n’a été trouvé dans la séquence de la tufteline, quelle que soit la phase de lecture choisie. Or, sans ce type de séquence, il est impossible à la tufteline d’être exportée vers la matrice de l’émail (voir II.4.3.), ce qui ne correspond pas à une protéine de la matrice extracellulaire.

Enfin, l'information la plus importante concernant la tufteline vient du fait que les analyses en Northern blot ont détecté les ARNm de cette protéine dans plusieurs tissus non minéralisés, ce qui supprime son caractère spécifique des tissus dentaires et son rôle supposé prépondérant dans la minéralisation.

III.4. Les protéinases

Le processus de minéralisation de l’émail dentaire se traduit par une disparition progressive de la matrice organique au profit des cristaux de minéral. Toutes les protéines de la matrice extracellulaire sont rapidement protéolysées après leur fabrication. Les protéinases sont donc des composantes essentielles de la matrice de l’émail, d’autant que la protéolyse est capitale pour le fonctionnement des protéines de l’émail, comme l’amélogénine (voir *§II.4.4.*).

Deux protéines ayant une activité protéolytique ont été récemment clonées et caractérisées. Il s’agit d’une métallo-protéinase appelée “enamélysine“ (MMP-20) identifiée chez le porc (Bartlett *et al.*, 1996) et chez l’homme (Llano *et al.*, 1997) et d’une “sérine-protéinase“ appelée “EMSP-1“ (Simmer *et al.*, 1998). Des expériences *in vivo* utilisant un extrait protéique enrichi en EMSP-1 ont montré que la serine-protéinase est capable de dégrader totalement l’amélogénine (Moradian-Oldak *et al.*, 1998). Sans toutefois en avoir les preuves expérimentales, on pense que ces deux protéinases sont capables d’agir, en dehors de l’amélogénine, sur d’autres protéines de l’émail comme l’énaméline, l’améloblastine, etc. De la même manière, on suppose que l’enamélysine est la protéine responsable de la dégradation progressive de l’amélogénine, générant la protéine TRAP qui a une fonction essentielle pour le développement de l’émail (Ravindranath *et al.*, 1999). La fonction de la sérine-protéinase serait d'éliminer la phase organique de la matrice de l'émail grâce à son activité protéolytique et de libérer de l'espace pour que les cristaux d’hydroxyapatite puissent croître rapidement durant la maturation de l’émail.

IV. L'AMÉLOGÉNINE COMME MARQUEUR DANS LES ANALYSES D'ADN

1V.1. Applications en médecine légale, archéologie et anthropologie

Comme cela a déjà été dit précédemment (voir *§II.1.*), dans le génome humain et chez d'autres mammifères, il existe deux copies du gène de l'amélogénine sur les chromosomes X et Y. Ces deux copies diffèrent par un certain nombre de substitutions et notamment, chez l'homme, par une insertion de 6 nucléotides dans l'intron 1 du gène (Sullivan *et al.*, 1993).

Ces différences peuvent être utilisées en biologie moléculaire pour distinguer les deux copies. Ainsi, par l'analyse des séquences d'ADN de l'amélogénine, il devient alors possible de savoir si une personne de sexe inconnu possède un chromosome Y ou pas (Figure 17). L'amélogénine est donc devenu un outil pratique en médecine clinique et en médecine légale pour déterminer le sexe d'une personne à partir d'un simple échantillon d'ADNg (Falconi *et al.*, 2001). On peut ainsi réaliser un diagnostic précoce du sexe d'un embryon, ou déterminer le sexe d'une personne dont les restes sont trop détériorés pour permettre des analyses morphologiques, comme lors d'incendies, d'explosions ou de catastrophes aériennes.

L'amélogénine a été utilisée dans une enquête à la suite à d'une catastrophe aérienne où les corps étaient brûlés (Soares-Vieira *et al.*, 2000a), dans le cas d'un viol pour déterminer si l'ADNg prélevé provenait de l'agresseur ou de la victime (Cina *et al.*, 2000) et dans le cas d'une fraude à l'assurance (Soares-Vieira *et al.*, 2000b).

Le gène de l'amélogénine est ainsi devenu un auxiliaire important pour de nombreuses cours de justice et bureaux d'enquête dans le monde entier (FBI, Interpol, etc.) (Voir : http://www.interpol.int/Public/Forensic/dna/loci.asp). En France, ce gène est l'objet d'un texte de loi publié le 19 mai 2000 dans le Journal Officiel de la République Française. Il s'agit de l'arrêté du 18 mai 2000, fixant la liste des segments d'ADN sur lesquels portent les analyses génétiques pratiquées aux fins d'utilisation du fichier national des empreintes génétiques (voir page suivante).

«
Le garde des sceaux, ministre de la justice, le ministre de l'intérieur et le ministre de la défense,
Vu le code de procédure pénale, et notamment ses articles 706-54 et R. 53-13 ;
Vu l'avis de la commission d'agrément des personnes habilitées à effectuer des missions d'identification par empreintes génétiques dans le cadre des procédures judiciaires en date du 9 avril 1999,
Arrêtent :

« Art. 1er. - Il est inséré dans le livre IV du code de procédure pénale (quatrième partie : Arrêtés) un titre XIX ainsi rédigé :

« Titre XIX
« DE LA PROCEDURE APPLICABLE AUX INFRACTIONS DE NATURE SEXUELLE ET AUX MINEURS VICTIMES
« Chapitre Ier
« Du fichier national automatisé des empreintes génétiques

« Art. A. 38. - La liste des segments d'ADN sur lesquels portent les analyses destinées à l'identification génétique figure dans le tableau ci-après :

Vous pouvez consulter le tableau dans le JO n° 116 du 19/05/2000, pages 7545 à 7546

« Les analyses portent également sur le gène de ***l'amélogénine****. »*

« Art. 2. - Le directeur des affaires criminelles et des grâces, le directeur général de la police nationale et le directeur général de la gendarmerie nationale sont chargés, chacun en ce qui le concerne, de l'exécution du présent arrêté, qui sera publié au Journal officiel de la République française.
»

Cette utilisation particulière du gène de l'amélogénine s'est étendue à d'autres domaines scientifiques comme l'archéologie et l'anthropologie. En effet, les techniques de biologie moléculaire permettent aujourd'hui d'obtenir de l'ADN à partir de restes anciens, ce que l'on a appelé "l'ADN fossile", mais qu'il convient mieux d'appeler "ADN ancien". En effet, la qualité de l'ADN décroît avec le temps et toute analyse devient impossible au-delà de 450 000 ans (Hänni, 2001). Les analyses d'ADN peuvent donc fournir de précieuses indications sur l'histoire de l'humanité comme les liens familiaux, la consanguinité dans certaines populations et le typage sexuel peut être intéressant par exemple pour mesurer la mortalité enfantine en fonction des sexes ou le sexe ratio à une époque donnée. Toutes ces données permettent de mieux comprendre le mode de vie des populations et des civilisations étudiées.

Ainsi, des analyses utilisant le gène de l'amélogénine ont été réalisées sur de l'ADN extrait de nombreux restes humains anciens. Le cas le plus célèbre a été l'identification du sexe des restes de la famille du Tsar Nicolas II (Gill *et al.*, 1994). Mais, il faut également citer les études spectaculaires réalisées sur 38 squelettes d'un cimetière allemand du moyen-age (450-750 AD) (Scholz *et al.*, 2001), sur 37 squelettes égyptiens d'une nécropole à l'ouest de Thèbes (2120-500 BC) et 4 squelettes de la nécropole d'Abydos (3000 BC) (Zink *et al.*, 2001), sur 13 squelettes provenant de la maison de Caius Iulius Polybius à Pompéi ensevelis

par l'éruption du Vésuve (79 A.D) (Cipollaro *et al.*, 1999) et sur 20 squelettes d'un cimetière amérindien du centre de l'Illinois (1300 AD) (Stone et Stoneking, 1999).

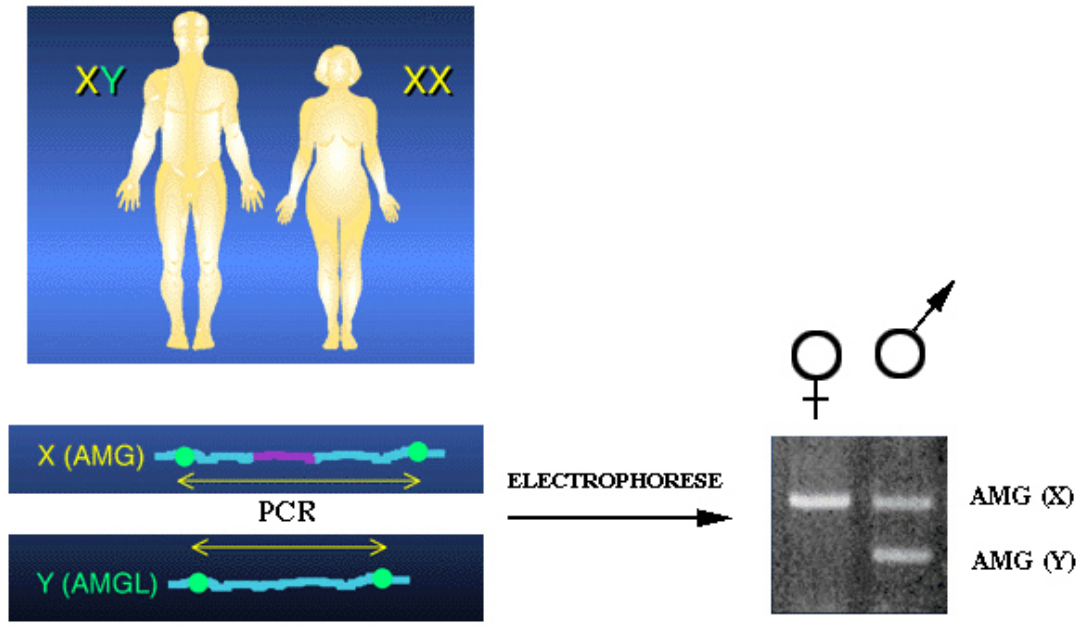


Figure 17. Méthode du "typage sexuel" à partir du gène de l'amélogénine. D'après "le Département de Médecine Légale de l'Université de Kansai", Japon.

IV.2. Applications en médecine vétérinaire et en biologie des organismes

L'amélogénine Y est connue chez d'autres espèces de mammifères : la vache, le cheval, la chèvre, le bouquetin, le cerf et l'ours. Ainsi, en utilisant des différences (ces différences ne sont pas les mêmes que chez l'homme) existant entre les deux copies de ces espèces, il est possible de réaliser des tests d'ADN pour déterminer le sexe de ces animaux. Ce genre de test a été mis au point chez le cheval dans le but de détecter des anomalies chromosomiques liées aux chromosomes sexuels comme le cas des juments XY (Hasegaw *et al.*, 2000). Des analyses d'ADN ont été réalisées chez le cerf japonais (*Cervus nippon*) à partir de matières fécales provenant de 34 à 37 individus (Yamauchi *et al.*, 2000) et le même genre d'étude a été effectué à partir d'échantillons de sang et de poils chez l'ours noir japonais (Yamamoto *et al.*, 2002).

Ces derniers travaux permettent d'envisager l'utilisation de l'amélogénine pour l'étude écologique de populations animales sans avoir à réaliser de captures ni de prélèvement de tissus et même sans avoir besoin de les observer dans leur milieu naturel, ce qui se révèle particulièrement intéressant dans le cas de populations particulièrement difficiles à approcher.

V. CHOIX DE L'AMÉLOGÉNINE COMME SUJET DE RECHERCHE

L'amélogénine est l'objet de nombreuses recherches dans plusieurs domaines de la biologie et de la médecine depuis de nombreuses années. J'ai choisi d'étudier ce modèle au cours de ma thèse pour plusieurs raisons.

V.1. Intérêts de l'amélogénine en médecine dentaire

En se reportant à l'historique que j'ai développé (voir partie I.), on constate que le gène de l'amélogénine, et la protéine qu'il code, sont la cible d'un vaste programme de recherche dont le but est de comprendre comment se construit l'émail et comment se mettent en place ses différentes composantes organiques et minérales. Les objectifs à long terme de ces recherches étant que l'on puisse un jour réparer la dent en agissant sur les différents gènes des protéines qui la constituent ou bien "guérir" certaines maladies génétiques liées aux tissus dentaires. Malgré tous les efforts investis depuis une dizaine d'année, les gènes des protéines de l'émail ne livrent pas facilement leurs secrets. De plus, les protéines de l'émail sont particulièrement difficiles à manipuler en raison de leurs propriétés hydrophobes. Les techniques actuelles ne permettent plus de progresser rapidement dans les connaissances des mécanismes de construction de l'émail et des relations qui peuvent exister entre la structure de l'émail et celle des protéines.

A ce jour, on ne connaît toujours pas le rôle que jouent chacune des différentes parties de l'amélogénine dans la mise en place de l'émail des mammifères. Aucune modélisation en 3D de cette protéine n'a pu être réalisée et la signification de la présence de tel ou tel acide aminé dans la séquence primaire reste très incertaine, voire inconnue. Pourtant, d'après les données obtenues en immunohistochimie (voir chapitre I.4.), le gène de l'amélogénine existe depuis plus de 400 millions d'années et, d'après mes données moléculaires récentes (voir chapitre II-"Origine de l'amélogénine"), son origine remonterait au Précambrien (630 MA), ce qui suggère la présence du gène de l'amélogénine dans toutes les lignées de vertébrés, et cela dès leur apparition (Figure 2). Or, comme je l'ai expliqué dans le chapitre I.3, les connaissances de l'amélogénine ont été principalement obtenues chez les mammifères et, de plus, chez un nombre réduit d'espèces non représentatives de toutes les lignées mammaliennes. Cette recherche "ciblée" se justifie car les auteurs espèrent qu'en étudiant les modèles les plus proches de l'homme, ils pourront appliquer rapidement les résultats obtenus. Il n'existe donc aucune vue globale de l'évolution du gène de l'amélogénine. Dans les autres lignées de vertébrés, on ne sait pas si l'amélogénine représente aussi 90% des protéines de la matrice de l'émail comme c'est le cas pour les mammifères, ou si, au contraire, elle est minoritaire par rapport à d'autres protéines de l'émail comme cela a déjà été suggéré pour les non-tétrapodes (Slavkin, 1984). Il est aussi impossible de dire si, dans les différentes lignées de vertébrés, l'amélogénine est exprimée de la même manière dans le temps et dans l'espace ni même si elle joue le même rôle au cours de l'amélogenèse que chez les mammifères. Voilà pourquoi les études comparatives menées dans d'autres lignées de vertébrés, et sur des modèles bien choisis, peuvent s'avérer extrêmement importantes dans la compréhension de ce processus biologique qu'est la formation de l'émail.

Les comparaisons de séquences obtenues dans chacune des lignées permettront de connaître les régions les mieux conservées au cours de l'évolution et ainsi mettront en évidence les parties fondamentales pour la fonction de la protéine. Actuellement, nous connaissons quelques mutations que l'on a associé à la maladie génétique appelée amélogenèse imparfaite (voir Introduction - II.6.). Ces mutations sont-elles toutes réellement

la cause de l'AI ? Quid des non-mammaliens qui possèdent de l'émail ? Or, il est certain que des mutations dans d'autres régions sont susceptibles de modifier la structure de l'amélogénine et il serait précieux pour la médecine de dresser une carte des régions "sensibles" du gène.

Les études comparatives permettent aussi de mettre en lumière des régions de la protéine caractéristiques de certaines lignées, c'est-à-dire qui seraient bien conservées uniquement à l'intérieur de certaines lignées de vertébrés. Ces conservations de motifs protéiques à un niveau plus "local" de la phylogénie peuvent être dues à des "effets phylogénétiques" (par exemple, la similarité des séquences protéiques dans une lignée peut être le résultat d'une spéciation rapide qui aurait fait hériter tous les descendants des mêmes acides aminés) ou à des différences de fonctionnement de la protéine dans certaines lignées qu'il peut être intéressant de comprendre.

Ces raisons expliquent la nécessité de connaître le gène de l'amélogénine dans toutes les lignées de vertébrés qui possèdent de l'émail (ou de l'émailloïde). Pour réaliser ce travail, j'ai choisi de commencer par la lignée mammalienne (seul groupe connu au début de mes recherches), puis de m'intéresser à la lignée de vertébrés la plus proche parente des mammifères, c'est-à-dire les reptiles, dont les représentants actuels possèdent de l'émail (à l'exception des tortues et des oiseaux). Ce choix s'explique par la difficulté rencontrée pour obtenir le gène de l'amélogénine chez des taxons très éloignés comme les actinoptérygiens ou les chondrichtyens séparés des mammifères depuis 400 à 500 millions d'années.

V.2. Intérêt de l'amélogénine pour établir des phylogénies moléculaires

Après avoir réalisé un certain nombre de phylogénies à partir de l'amélogénine des mammifères, comme par exemple l'étude d'une séquence de "myxine" (voir chapitre 2) et l'étude de la position phylogénétique de la séquence de l'ADNc de hamster découverte par Cerny *et al.* (1998) (travaux réalisés dans le cadre d'une collaboration avec Radim Cerny), il est apparu que l'amélogénine était un bon marqueur moléculaire capable de reconstituer des relations phylogénétiques. Il était alors tentant de tester la puissance de ce gène pour obtenir des phylogénies fiables dans diverses lignées comme les mammifères, les squamates, voire les amniotes, les tétrapodes et même les vertébrés dans leur ensemble.

Le gène de l'amélogénine (j'ai eu l'occasion de le démontrer, voir chapitre 2) possède des régions n'évoluant pas toutes à la même vitesse : l'exon 6 par exemple évolue plus vite que les autres régions. Ces différences peuvent sembler constituer un problème pour réaliser des phylogénies utilisant des groupes très éloignés (reptiles, mammifères, amphibiens), mais elles peuvent s'avérer intéressantes pour comparer des espèces plus proches (au sein d'un même genre ou d'une même famille, par exemple). Les régions les plus conservées du gène pourraient alors s'avérer utiles pour les phylogénies établies à un niveau plus élevé de l'évolution (grandes lignées de vertébrés par exemple).

Il est maintenant devenu possible de tester le gène de l'amélogénine en phylogénie moléculaire car le nombre de séquences connues pour ce gène est aujourd'hui assez important, d'une part, grâce à mon travail au cours de cette thèse et de la collaboration avec une équipe de recherche de l'université de Niigata (japon) et, d'autre part, aux séquences qui sont déjà connues dans la littérature.

Les outils phylogénétiques (gènes nucléaires, gènes mitochondriaux, etc.) ne manquent pas, mais tous n'ont pas le même pouvoir discriminatoire. De plus, les résultats obtenus ont parfois étaient critiqués : par exemple, les relations de parentés au sein des mammifères restent mal résolues aujourd'hui. Un outil supplémentaire peut donc être intéressant dans ce domaine de la biologie.

V.3. Intérêt de l'amélogénine dans l'étude de l'évolution du squelette dermique

Le squelette dermique est composé de différents éléments minéralisés incluant les odontodes[2], les dents, les denticules branchiaux[3] (ou branchiospines), les os dermiques du crâne et de la ceinture pectorale, les rayons des nageoires[4] ainsi que les écailles et autres éléments de type "écailleux" (scutes, ostéodermes[5], plaques osseuses). On considère généralement que tous ces organes sont dérivés du squelette dermique qui existait chez les premiers vertébrés, il y a près de 500 MA (revue dans Huysseune et Sire, 1998). L'histoire évolutive des différents tissus composant les éléments du squelette dermique a été longuement discutée car un certain nombre de tissus, susceptibles d'être dérivés du dermosquelette ancestral, ont subi tellement de changements morphologiques au cours de l'évolution que les comparaisons deviennent difficiles et qu'il en résulte des interprétations différentes et des controverses. Comme on va le voir dans la suite, l'émail dentaire aurait des origines communes avec d'autres tissus dérivés du squelette dermique ancestral. L'amélogénine étant une protéine spécifique de l'émail, elle pourrait aider à comprendre quels tissus sont homologues de l'émail.

Origine évolutive du dermosquelette

Selon le modèle évolutif généralement accepté, les odontodes qui étaient présents chez les premiers vertébrés sans mâchoires seraient à l'origine des dents et de toute une variété d'éléments dermosquelettiques situés à l'extérieur de la cavité buccale des vertébrés actuels.

Odontodes et dents ont relativement peu dérivé l'un de l'autre d'un point de vue morphologique ce qui permet d'établir des similitudes de façon relativement claire à la fois au niveau des organes et des tissus: tous les odontodes sont constitués de dentine qui entoure une cavité pulpaire et qui est couverte par un tissu hyperminéralisé et dépourvu de collagène (Figure 18).

Ce tissu est de l'émail ou de l'émailloïde[6]. Les relations évolutives entre ces deux tissus sont encore l'objet de vives controverses, certains auteurs pensant que la condition ancestrale était la présence d'émail vrai dont l'émailloïde serait dérivé, d'autres défendant une position opposée (voir revue dans Huysseune et Sire, 1998). La recherche des différentes protéines de l'émail dans l'émailloïde en formation, et notamment l'amélogénine, pourrait apporter des éléments de réponse sur l'origine évolutive de ces deux tissus et sur le rôle de ces

---

[2] Les odontodes sont constitués d'un cône de dentine (ou d'un tissu proche de la dentine) couvert ou non par une couche hyperminéralisée d'émail ou d'émailloïde et dont la base vascularisée consiste en un tissu osseux qui fonctionne comme un tissu d'attache.

[3] Petits denticules (odontodes) présents dans la muqueuse couvrant les arcs branchiaux.

[4] Encore appelés lépidotriches, ils sont des éléments minéralisés du squelette supportant les nageoires.

[5] Plaques osseuses minéralisées situées dans le derme de certains amphibiens anoures, de certains sauropsidés et d'un mammifère, et qui sont particulièrement bien développées chez les crocodiliens. Chez certains lézards, ils consistent en une épaisse plaque basale de tissu osseux formé par ossification métaplasique et d'une couche superficielle dépourvue de collagène.

[6] Tissu hyperminéralisé qui se distingue de l'émail par la présence, lors de sa formation, de fibres de collagène (toutefois beaucoup moins nombreuses que dans le tissu osseux et qui pourraient y jouer un rôle différent).

protéines dans la formation de l'émailloïde. En effet, la structure et la chronologie de formation de ce dernier sont différentes de celles de l'émail. Les odontodes sont généralement associés à un support osseux auquel ils sont attachés de manière variée. On trouve des odontodes actuellement chez les chondrichtyens (requins, raies, chimères) sous la forme d'écailles placoïdes[7], chez les sarcoptérygiens sur les os du crâne et les écailles du cœlacanthe, ainsi que chez les actinoptérygiens sur de nombreux os du crâne de *Denticeps* (Clupeïforme), sur le rostre de *Xiphias* (Perciforme) et sur la totalité des éléments du squelette dermique des Siluriformes cuirassés (Callichthyidae, Loricariidae).

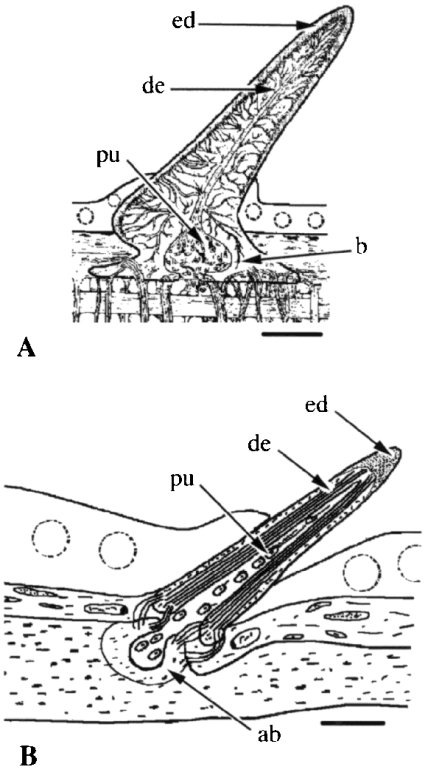


Figure 18. Exemples d'odontodes chez un chondrichtyen *Scyliorhinus canicula* (A) et un ostéichtyen *Corydoras aeneus* (B). ed: émailloïde ; de: dentine ; pu: cavité pulpaire ; b: os ; ab: os d'attache. D'après Huysseune et Sire (1998) *(Echelle= 50 µm).*

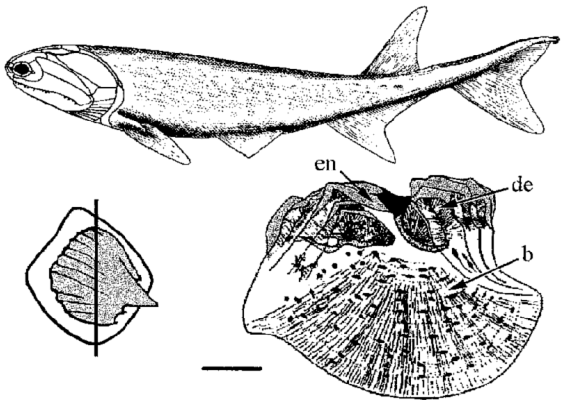


Figure 19. Le plus vieil actinoptérygien connu, *Cheirolepis* du Dévonien*;* vue superficielle et coupe transversale dans une de ses écailles dites rhombiques. en: émail ; de: dentine ; b: plaque basale osseuse. D'après Huysseune et Sire (1998) *(Echelle= 100 µm).*

Au cours de l'évolution, et dans certains groupes d'ostéichtyens, des odontodes ont fusionné pour former des structures appelées odontocomplexes[8]. Cette fusion procède soit par

[7] Petits denticules dermiques (odontodes) présents dans la peau des chondrichtyens. Elles consistent en une couche d'émailloïde couvrant une couronne de dentine qui entoure une cavité pulpaire. Elles se forment comme les dents et sont remplacées quand elles sont perdues.

[8] Agglomérations ou regroupements d'odontodes formés directement sur ou à côté des autres durant les étapes successives de croissance. Les odontocomplexes incluent les organes dérivés des odontodes et les organes possédant des tissus dérivés des odontodes.

superposition de différentes générations d'odontodes les unes sur les autres, soit par accrétion. On retrouve ces phénomènes de façon nette chez des fossiles du Dévonien (sarcoptérygiens et actinoptérygiens), dont le plus vieil actinoptérygien connu, *Cheirolepis* (Figure 19). Ce dernier possède des écailles dites rhombiques[9] composées d'une base osseuse sur laquelle se déposent des odontodes formés de dentine recouverte d'un tissu minéralisé dépourvu de collagène, de l'émail. On y distingue également des appositions successives d'odontodes. L'écaille rhombique est considérée comme l'ancêtre des différents types d'écailles des ostéichtyens.

Dans la faune actuelle, les polyptéridés possèdent des écailles ganoïdes (Figure 20A), le modèle le plus proche de l'écaille rhombique ancestrale (Figure 19). En effet, ce type d'écaille présente le plus de caractères plésiomorphes par rapport au modèle ancestral supposé, l'écaille rhombique (Figure 19) (Sire, 1990): c'est-à-dire, une couche superficielle hyperminéralisée et dépourvue de collagène appelée ganoïne et semblable à l'émail dentaire (Sire *et al.*, 1987 ; Zylberberg *et al.*, 1997), une couche de dentine particulière appelée ostéodentine, un tissu particulier appelé élasmodine, organisé en "contre-plaqué" et une épaisse plaque basale osseuse.

L'origine de l'élasmodine n'est pas très claire pour le moment, mais ce tissu, qui n'est comparable à aucun des tissus squelettiques existants, dériverait soit de la dentine soit de l'os, tous deux présents dans l'écaille rhombique ancestrale. L'écaille ganoïde existe aussi chez les lépisostéidés avec cependant quelques différences (Figure 20B): la ganoïne est déposée directement sur la plaque basale osseuse et il n'y a plus ni de dentine ni d'élasmodine. Il s'agirait ici d'un état dérivé, caractérisé par la disparition de ces deux tissus.

En dehors de ces exemples qui montrent clairement la similitude entre la ganoïne des écailles rhombiques, celle des écailles ganoïdes et l'émail dentaire, la comparaison est particulièrement difficile à établir pour les autres éléments du squelette dermique. C'est le cas notamment pour les écailles élasmoïdes présentes chez quelques sarcoptérygiens (cœlacanthes, dipneustes), mais surtout abondamment représentées chez les actinoptérygiens, et en particulier chez les téléostéens qui constituent la majorité des espèces de vertébrés aquatiques.

La structure d'une écaille élasmoïde est présentée sur la Figure 20C. Elle est caractérisée par une couche superficielle constituée de deux parties appelées couche limitante et couche externe, recouvrant une plaque basale lamellaire d'élasmodine. La couche limitante est un tissu constitué d'une trame minéralisée dépourvue de collagène mais traversée par des fibres de collagène minéralisées. Il est moins minéralisé que la ganoïne et il n'est pas directement en contact avec l'épiderme mais séparé de celui-ci par une mince couche de derme. Sa structure et son organisation ont donné lieu à des discussions sur l'origine de cette couche : les molécules qui la composent sont-elles sécrétées par l'épiderme, comme la ganoïne et l'émail dentaire, ou sont-elles sécrétées par le derme ? Dans cette dernière hypothèse, il n'y aurait pas d'homologie entre cette couche limitante et les autres tissus superficiels hyperminéralisés. Les données les plus récentes indiquent cependant que l'épiderme participerait, au moins en partie, à l'élaboration de cette structure (Sire *et al.,* 1997). La couche limitante se dépose à la surface d'une couche "externe" qui est constituée d'une trame fibreuse collagénique, elle même située au-dessus de la plaque basale d'élasmodine qui présente une organisation typique du collagène en "contre-plaqué".

[9] Ecailles épaisses et juxtaposées des ostéichtyens primitifs. Elles consistent en une épaisse plaque basale osseuse recouverte d'une couche superficielle d'os vasculaire, elle-même recouverte d'une couche de dentine et, d'une substance hyperminéralisée appelée ganoïne, un émail vrai.

A la suite de différents travaux sur le développement comparé des écailles ganoïdes d'un polyptère et des écailles élasmoïdes d'un téléostéen cichlidé, une homologie entre l'élasmodine de l'écaille élasmoïde et celle de l'écaille ancestrale a été proposée (Figure 20). La couche externe serait alors homologue de la dentine et la couche limitante homologue de la ganoïne (Sire, 1989). Ceci signifierait que la plaque basale osseuse aurait disparu lors de la transition entre l'écaille rhombique et l'écaille élasmoïde, et rendrait caduque les hypothèses selon lesquelles l'élasmodine pourrait aussi être dérivée de la plaque basale osseuse (Meunier, 1983 ; Schultze, 1996). L'origine commune des écailles ganoïde et élasmoïde est soutenue par le fait que l'écaille ganoïde, chez les polyptères juvéniles, possède la structure d'une écaille élasmoïde (Sire 1990). La structure de l'écaille élasmoïde serait alors le résultat d'un processus évolutif de paedomorphose: la présence chez le descendant, à l'âge adulte, d'une structure connue chez le juvénile de l'ancêtre. L'écaille élasmoïde, présente chez le juvénile, serait conservée à l'âge adulte et la possibilité d'une transformation en écaille ganoïde aurait été perdue.

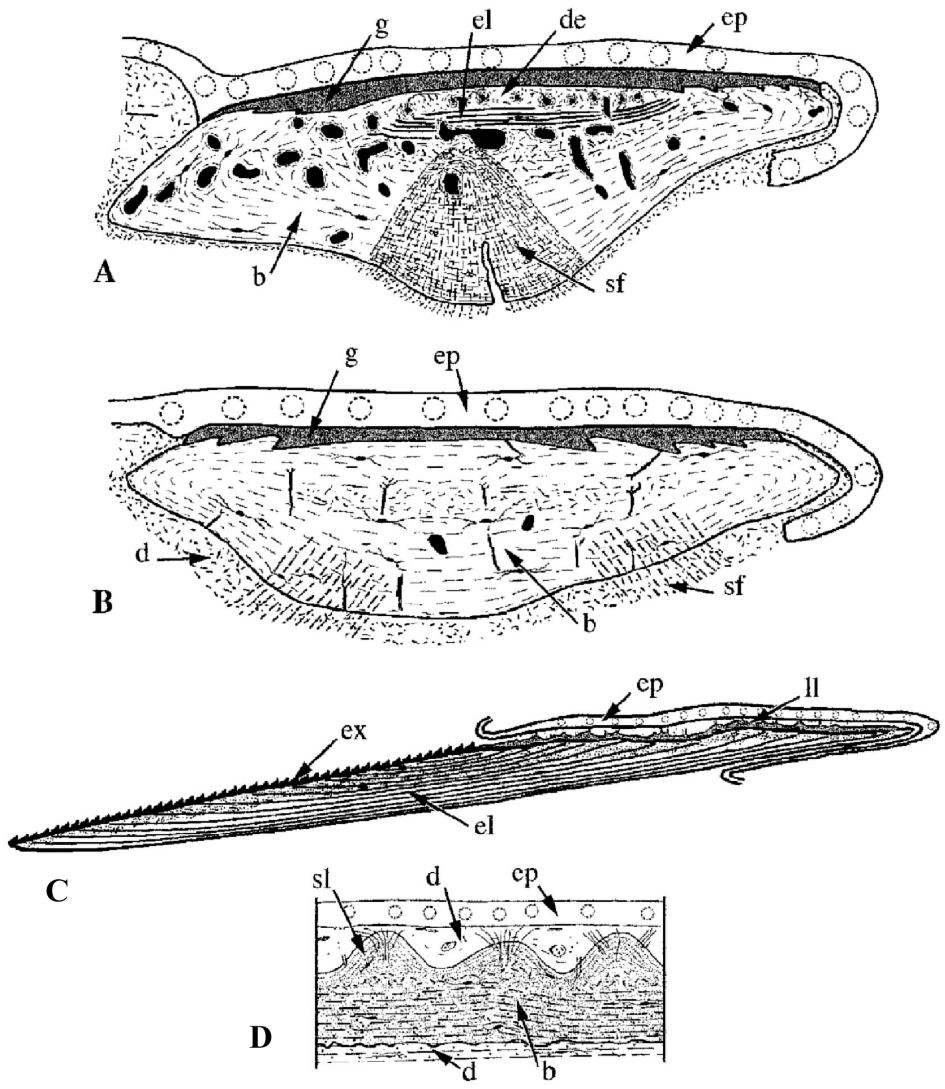


Figure 20. Différents éléments dermosquelettiques des ostéichtyens actuels ; **A**: Ecaille ganoïde de polyptéridé ; **B**: Ecaille ganoïde de lépisostéidé, sur laquelle la ganoine est déposée directement sur l'os ; **C**: Ecaille élasmoïde d'un téléostéen cichlidé ; **D**: Ostéoderme de squamate (Scincidé). b: os; d: derme; de: dentine; ep: épiderme; el: élasmodine; ex: couche externe; g: ganoïne; ll: couche limitante; od: odontode; sf: fibres de Sharpey; sl: couche superficielle. Modifiés d'après Huysseune et Sire (1998).

L'origine évolutive des ostéodermes chez certains tétrapodes (anoures, crocodiles et squamates) est encore plus difficile à comprendre (Figure 20D). Les ostéodermes sont composés d'une plaque basale osseuse recouverte d'une couche externe minéralisée et dépourvue de collagène. Contrairement aux écailles élasmoïdes, ils se forment plus profondément dans le derme et leur surface est séparée de l'épiderme par une couche plus épaisse de derme lâche (Figure 20D) et par des cellules pigmentaires.

Toutes les structures dermosquelettiques décrite dans la figure 20 ont un élément commun : la couche la plus externe comporte une matrice non-collagénique minéralisable. Cependant, il existe des différences, dans le taux et les processus de minéralisation par exemple mais aussi dans la présence de faisceaux de fibres de collagène traversant cette couche (ostéodermes de squamates et écailles élasmoïdes). Il n'en reste pas moins que ce tissu minéralisé ne présente pas de fibres de collagène dans la trame minéralisée alors que c'est le cas pour tous les autres types de tissus squelettiques de type osseux. En fait, pendant longtemps, le seul exemple connu de tissu dépourvu de collagène et hyperminéralisé était l'émail dentaire.

Les méthodes utilisées jusqu'à présent (histologie, microscopie électronique à balayage et à transmission) ainsi que les études de biologie du développement n'ont pas permis de répondre aux questions suivantes :

- Quelles sont les relations évolutives qui existent entre les différentes formations dermo-squelettiques qui sont présentes dans la nature ?
- Existe-t-il des homologies entre les tissus qui les composent ?
- La couche supérieure des écailles élasmoïdes et la couche supérieure des ostéodermes de certains tétrapodes dérivent-elles de la ganoïne ? La réponse à cette dernière question permettrait d'apporter des éléments de réponse aux deux premières.

Les méthodes histo-morphologiques ayant montré leurs limites, les techniques de biologie moléculaire ont été utilisées avec cette idée: si le tissu qui compose la couche la plus externe de certains éléments du squelette dermique dérive de l'émail, alors on devrait pouvoir en trouver des preuves au niveau moléculaire. Le choix d'une première molécule s'est porté sur l'amélogénine car cette protéine représente 90% des protéines sécrétées dans l'émail des dents en formation chez les mammifères (Termine *et al.*, 1980) et elle n'existe dans aucun autre tissu que l'émail. Par ailleurs, le gène est indétectable chez des groupes zoologiques ayant perdu l'émail au cours de l'évolution (tortues, oiseaux) ce qui indique que le seul rôle de la protéine serait de participer à sa formation (Girondot et Sire, 1998).

Le premier but de ma thèse de doctorat était donc de rechercher le gène de l'amélogénine chez les squamates possédant des ostéodermes et ensuite d'essayer de mettre en évidence son expression dans les cellules responsables de la formation de la couche superficielle des ostéodermes afin de montrer une similitude moléculaire entre la couche externe de ces organes et l'émail. En effet, si cette protéine, exclusivement d'origine épidermique et typique de l'émail chez les mammifères, pouvait être caractérisée comme entrant dans la composition de la couche superficielle des ostéodermes, cela pourrait apporter quelques éclaircissements sur l'origine évolutive de ces éléments. Bien évidemment, il resterait ensuite à déterminer si cette similitude provient d'un ancêtre commun, auquel cas il s'agirait d'une homologie, ou s'il s'agit de l'utilisation indépendante par deux organismes d'une machinerie génétique préexistante, ce qui serait une analogie.

C'est cette voie de recherche qui a déterminé le choix du modèle expérimental de départ, *Chalcides*, un lézard des iles Canaries qui possède des ostéodermes.

Cependant, en raison des difficultés rencontrées, au départ, pour cloner l'amélogénine de ce scincidé, l'objectif initial, c'est-à-dire la recherche de l'amélogénine dans la couche superficielle des ostéodermes, n'a pas été atteint. J'ai, en effet, dû concentrer mes efforts sur l'étude de l'évolution de l'amélogénine chez les mammifères, puis les squamates. Le gène de l'amélogénine de *Chalcides* ayant été obtenu assez tardivement, je n'ai pas pu en étudier l'expression très détaillée au niveau des tissus recouvrant les ostéodermes et au cours de leur développement. Toutefois, grâce à une collaboration avec M.L. Couble (Lyon) et A. Nanci (Montréal), j'ai pu apporter quelques résultats sur l'expression du gène et de la protéine d'amélogénine lors du développement des dents chez ce lézard (voir chapitre III), et l'étude du développement des ostéodermes et actuellement en cours.

AVERTISSEMENT

La partie “Résultats“ de ma thèse est constituée d'un ensemble de publications scientifiques, déjà publiées, sous presse, soumises ou en préparation. Ma volonté a été de rendre la présentation de ma thèse la plus homogène possible et j'ai donc adopté la même présentation pour tous les chapitres. Les articles publiés, sous presse ou soumis à des revues scientifiques de langue anglaise ont été traduits en français et les articles en préparation ont été présentés de la même façon. Le fait de conserver les textes originaux a conduit à des redites inévitables et inhérentes à la présentation choisie. De même, certaines informations détaillées dans le chapitre 1 “Introduction“ et dans le chapitre 6 “Matériel et Méthodes“ se retrouvent dans, mais seulement très partiellement, l'un ou l'autre des chapitres. Il était malheureusement impossible d'éviter ces répétitions. Seules les références bibliographiques de tous les chapitres ont pu être regroupées.

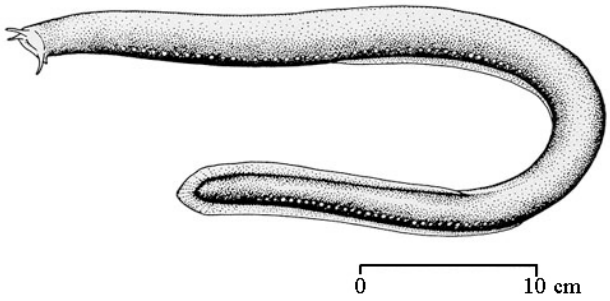
Je suis conscient de ces imperfections, mais j'espère que les lecteurs seront compréhensifs.

CHAPITRE 2
-
L'AMÉLOGÉNINE

Origine évolutive, évolution moléculaire
et utilisation en phylogenèse

Première Partie

ÉTUDE D'UNE SÉQUENCE D'AMÉLOGÉNINE DE MYXINE



La première tache que j'ai dû réaliser lorsque j'ai commencé mes travaux était de rechercher toutes les séquences connues d'amélogénine. Cette étape préalable était indispensable pour réaliser chacune des parties traitées dans ce manuscrit, comme celle d'obtenir des séquences d'amélogénine de squamates par des méthodes moléculaires ou réaliser des études comparatives et phylogénétiques.

C'est ainsi que je me suis immédiatement intéressé à une séquence protéique d'amélogénine supposée provenir d'un crâniate sans mâchoire, une myxine (Slavkin et Diekwish, 1996). Cela paraissait d'autant plus important que cette séquence concernait un taxon dans le groupe-frère des vertébrés et que c'était la seule séquence connue en dehors des mammifères; ceci était évidemment très intéressant pour tenter de comprendre l'histoire évolutive de l'amélogénine et essayer d'amplifier ce gène chez des lézards.

C'est ainsi que sous la direction de Marc Girondot (à l'époque MCU à Paris VII et membre de notre équipe), j'ai étudié, au cours de mon DEA, cette séquence d'amélogénine supposée être de myxine.

ANALYSE ÉVOLUTIVE DE «L'AMÉLOGÉNINE DE MYXINE»

Traduction de l'anglais à partir de l'article publié dans

Anatomical Record, 1998, **252**(4): 608-611,

Evolutionary Analysis of « Hagfish Amelogenin »

par Marc Girondot, **Sidney Delgado** et Michel Laurin

UMR 8570, Evolution et Adaptations des Systèmes Ostéo-musculaires, CNRS et Université Paris 7, Case 7077, 2 place Jussieu, 75251 Paris cedex 05, France.

Remerciements: Nous remercions Patricia Lai pour sa correction du manuscrit et Jean-Yves Sire et Armand de Ricqlès (UMR 8570) pour leur lecture critique et leurs nombreuses suggestions. Michel Laurin était financé par la fondation Alexander von Humboldt.

Résumé

Les myxines ont perdu leurs tissus minéralisés et leurs dents. La partie d'un brin d'ADNc, prétendu être de l'amélogénine, le gène majeur impliqué dans la formation de l'émail chez les mammifères, a récemment été clonée chez une myxine (Slavkin et Diekwish, Anat. Rec., 1996 ; 245 :131-150). Ce clonage est d'un grand intérêt puisqu'il pourrait changer la vision actuelle de l'évolution des tissus minéralisés, mais aucune analyse phylogénétique de ce morceau d'ADN n'a été faite par les auteurs.

L'analyse phylogénétique de ce morceau d'ADNc a été menée en employant à la fois les méthodes de la phénétique et de la cladistique. L'ADNc amplifié chez la myxine ne s'accorde pas avec une origine non mammalienne, mais correspond bien à une séquence dégradée de rongeur.

Le gène cloné chez la myxine est probablement d'origine mammalienne à la suite à d'une contamination survenue pendant la PCR.

Mots-clefs : Amélogénine, dent, évolution, myxine, contamination de PCR.

Les myxines sont des crâniates sans mâchoire, anguilliformes, appartenant aux Hyperotreti, le groupe frère des vertébrés (Janvier, 1996). Toutes les myxines existantes et tous les fossiles connus d'Hyperotreti ont perdu leurs tissus minéralisés ; leur squelette est cartilagineux et leur bouche est armée de dents cornées (odontoïdes ou denticules). L'absence des tissus minéralisés, qui sont par ailleurs trouvés chez tous les gnathostomes et chez beaucoup de groupes de fossiles de vertébrés sans mâchoire, a été différemment interprétée. Alors que les précédentes phylogénies ont impliqué que les myxines avaient perdu la capacité de produire des tissus minéralisés, le consensus actuel veut que les ancêtres des myxines n'aient jamais possédé cette capacité (Janvier, 1993). Le clonage récent d'un morceau d'ADNc de myxine par RT-PCR utilisant les amorces d'amélogénine (Slavkin et Diekwish, 1996 ; Anatomical Record, 245) est d'un grand intérêt dans la perspective de l'évolution des tissus minéralisés, mais est aussi tout à fait surprenant. En effet, l'amélogénine a été connue pendant quelques années comme étant impliquée dans la formation de l'émail des dents des mammifères (revue dans Deutsch, 1996), et récemment nous avons montré que le gène de l'amélogénine est probablement absent chez les sauropsidés édentés, comme les tortues et les oiseaux (Girondot et Sire, 1998). Cela renforce l'idée que le seul rôle de l'amélogénine chez les amniotes est de contribuer à la formation de l'émail et que celui-ci est perdu dans le taxon auquel il manque la pression sélective pour maintenir son intégrité (il y a moins de 200 Ma dans le cas des tortues et moins de 100 Ma pour les oiseaux). La présence d'amélogénine chez les myxines implique soit que ce gène a une autre fonction (au moins chez les myxines) soit, si nous supposons que les ancêtres des myxines avaient autrefois un squelette minéralisé (comme le suggèrent la plupart des phylogénies précédentes), que ce gène apparemment inactif a été conservé pendant plus de 300 Ma (âge de la plus ancienne myxine connue).

L'existence du gène de l'amélogénine chez les myxines a été utilisée par Slavkin et Diekwish (1996) pour valider la réactivité croisée, observée dans les denticules de myxine, des anticorps polyclonaux anti-amélogénine de mammifères (Slavkin et al., 1982, 1983, 1991). Cependant, la séquence publiée de 50 acides aminés de l'amélogénine de myxine ressemble beaucoup aux séquences connues des euthériens. Les auteurs n'ont pas réalisé de mesure quantitative de divergence entre ces séquences, alors qu'une telle analyse aurait permis soit la confirmation de l'origine non mammifère de la séquence, soit aurait détecté une quelconque source de contamination pendant la phase de PCR du protocole de clonage. Par conséquent, pour déterminer l'origine de l'amélogénine de la myxine, nous avons exécuté un test phylogénétique.

Récemment nous avons démontré que le gène de l'amélogénine peut être utile pour étudier de la phylogénie des mammifères. Il est possible de démontrer l'origine non-mammalienne de n'importe quelle séquence en établissant sa position fondamentale dans la phylogénie, bien qu'aucun groupe externe ne soit disponible puisque seules les séquences des mammifères sont connues. Les rapports phylogénétiques entre les myxines, les métathériens (les marsupiaux) et les euthériens (les mammifères placentaires), au sein des crâniates, sont connus sans aucune ambiguïté et cette phylogénie peut être considérée comme une phylogénie "vraie" (Janvier, 1996). La phylogénie de l'amélogénine des mammifères doit comprendre une première divergence entre les séquences des opossums et des euthériens, et une séquence d'amélogénine de myxine devrait se placer en dehors des mammifères.

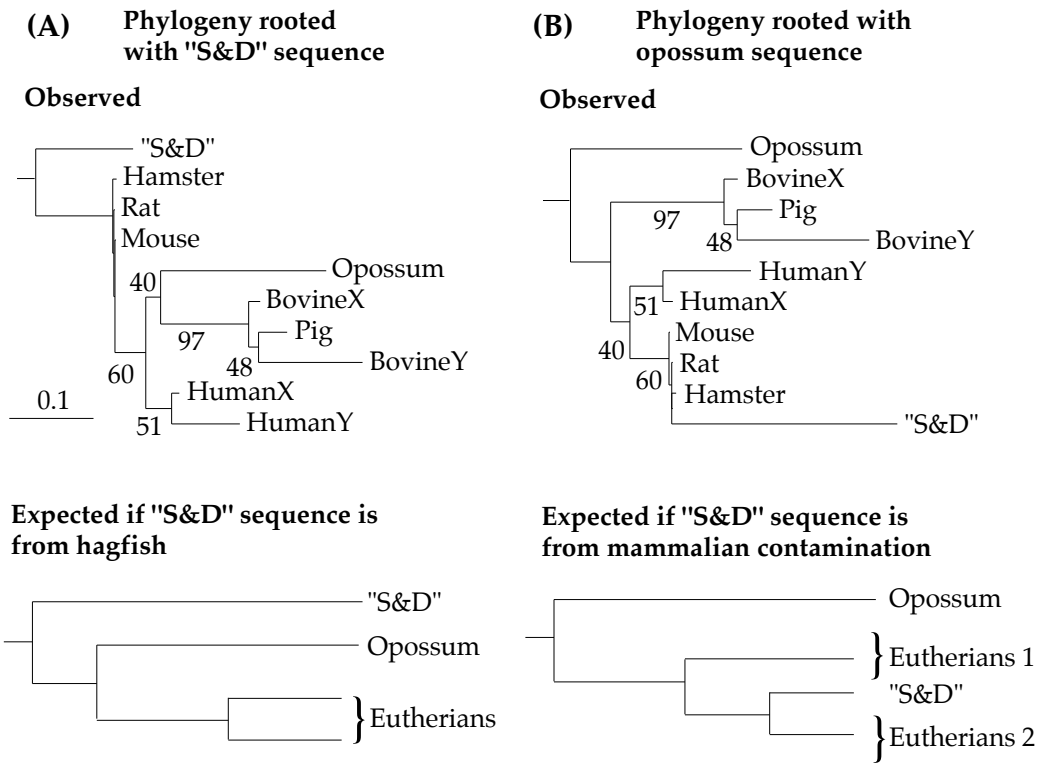


Figure 1. Arbres obtenus grâce à l'algorithme BIONJ avec la matrice PAM 350 pour les distances d'acides aminés et visualisés grâce à l'emploi de TreeView 1.4 (Page, 1996). La séquence supposée de myxine est indiquée par "S&D". Les arbres sont enracinés sur la divergence entre "S&D" et les autres séquences en (A) ou enracinés sur la divergence des métathériens avec les autres séquences en (B). Au-dessous de chaque arbre est indiquée la topologie attendue si la séquence "S&D" est une séquence de myxine (A) ou une séquence d'euthérien (B). Les valeurs de Bootstrap en pourcentage ont été obtenues grâce à 1000 ré-échantillonnages.

Cependant, étant donné qu'aucun groupe externe n'est disponible pour cette analyse, plusieurs options d'enracinement sont possibles mais seulement deux sont biologiquement significatives. Pour commencer, les arbres peuvent être enracinés entre les myxines et les mammifères (Figures 1A, 2A).

Si les séquences présentées appartiennent vraiment aux myxines, la dichotomie « correcte » entre les métathériens et les euthériens devrait être trouvée. Deuxièmement, les arbres peuvent être enracinés entre les métathériens (l'opossum) et les autres taxons (Figures. 1B, 2B). Si les séquences présentées sont le résultat d'une contamination par un euthérien, les séquences de « myxine » devraient former un clade avec l'espèce à partir de laquelle l'ADN a été réellement produit.

L'alignement a été réalisé en employant Clustal X 1.64b (Thompson *et al.*, 1994), mais les séquences peuvent être alignées sans aucune ambiguïté à la main.

Pour l'alignement, un seul « gap » est requis dans le gène X de l'homme (le gène de l'amélogénine est placé sur la partie hétérosomale des chromosomes sexuels chez les euthériens (Girondot et Sire, 1998)). Les analyses ont été exécutées en employant à la fois une méthode de distance (phénétique) et une méthode de parcimonie (cladistique) (voir les légendes des figures pour les procédures de calcul). Pour la méthode de distance, un algorithme récent (BIONJ (Gascuel, 1997)) a été employé pour réduire au minimum l'effet des différents taux de substitution dans les lignées chromosomiques X et Y de l'amélogénine des mammifères.

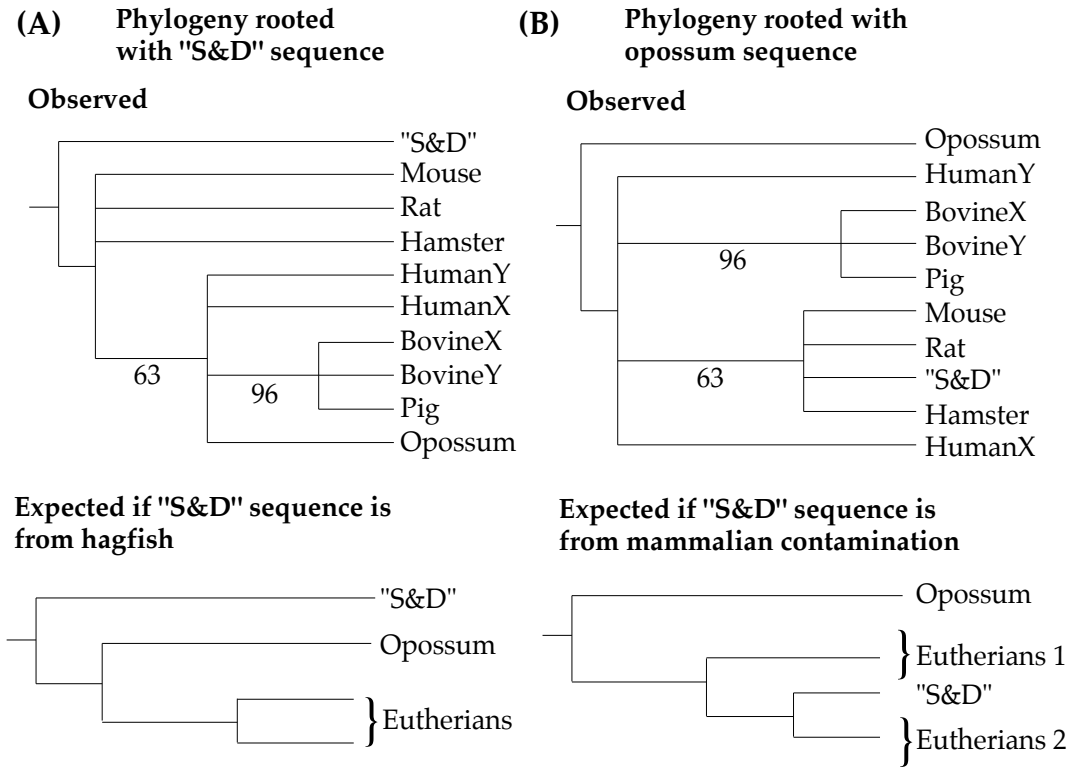


Figure 2. Arbres de consensus strict des 22 arbres les plus courts (41 pas) obtenus par la méthode de parcimonie mettant en œuvre une recherche exhaustive (2 027 025 arbres analysés) dans PAUP 3.1.1 (Swofford, 1993), avec l'option collectant les branches de taille zéro. La supposée séquence de myxine est indiquée par "S&D". L'arbre est enraciné à la divergence entre "S&D" et les autres séquences en (A) ou à la divergence entre les métathériens et les autres séquences en (B). La topologie attendue est indiquée au-dessous de chaque arbre si la séquence "S&D" est une séquence de myxine (A) ou si la séquence "S&D" est une séquence d'euthérien (B). Les valeurs de Bootstrap en pourcentage ont été obtenues grâce à 1000 ré-échantillonnages avec l'algorithme "branch-and-bound". Le consensus (50% "majority-rule") des arbres obtenus a la même topologie que le consensus strict des arbres les plus courts obtenu par la recherche complète.

En utilisant à la fois un arbre de distance de BIONJ (Figure 1) et un arbre consensus de parcimonie (Figure 2), la typologie de la phylogénie inférée n'est jamais compatible avec une origine non euthérienne des séquences supposées de myxine présentées par Slavkin et Diekwish (1996). Dans les deux cas, si les arbres sont enracinés entre les séquences présumées de myxine et les autres taxons, l'opossum est placé au sein des euthériens. Ce résultat suggère que la séquence de « myxine » représente en réalité un contaminant euthérien.

Pour évaluer la plausibilité de l'origine contaminante de la séquence de « myxine » et identifier la source réelle de la séquence d'ADN, nous avons ré-enraciné les arbres entre l'opossum et les autres taxons (Figure 1B et 2B). Dans ce cas, les séquences de « myxine » sont groupées avec les séquences de rongeur (rat, souris et hamster) (Figures 1B, 2B). Les vingt-deux arbres les plus courts (41 pas) sont obtenus par la méthode de parcimonie, et dans tous ces arbres, la séquence de la « myxine » est groupée avec celles des rongeurs. La topologie basale « correcte » est obtenue dans seulement trois des 118 arbres qui exigent un pas supplémentaire, et un consensus strict de ces arbres donne une topologie inconsistante avec les connaissances actuelles sur l'évolution de l'amélogénine basées sur la séquence

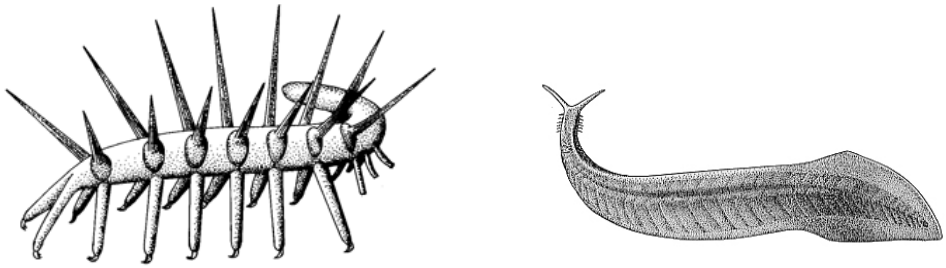
entière du gène (Girondot et Sire, 1998). En outre, la séquence de la « myxine » est toujours groupée avec au moins une séquence de rongeur dans les 115 arbres restant sur les 118. Cela suppose que la position de la myxine comme groupe-frère de toutes les séquences placentaires dans trois de ces 118 arbres n'est pas significative. Finalement, 63% et 60% des 1000 ré-échantillonnages de bootstrap, utilisant les méthodes de parcimonie ou de BIONJ, lient la myxine aux séquences de rongeurs, alors que les séquences analysées sont relativement courtes (Figures 1B, 2B). La méthode classique du neighbor-joining (Saitou et Nei, 1987) donne exactement la même topologie d'arbre que la méthode BIONJ (non montré).

Le temps de divergence entre les myxines et les autres espèces analysées ici est d'environ 470 millions d'années, tandis que le temps de divergence entre les métathériens (les marsupiaux) et les euthériens (les mammifères placentaires) est seulement de 120 millions d'années. Ainsi, on ne peut pas expliquer les résultats obtenus ici par des événements de spéciation rapide. Plus probablement, la séquence de myxine publiée est une séquence de mammifère obtenue par contamination pendant l'amplification par PCR. Cet artefact peut résulter de l'utilisation de certaines amorces potentiellement dégénérées, d'une température d'hybridation basse ou d'un nombre important de cycles de PCR (la séquence des amorces, la température d'hybridation, et le nombre de cycles de PCR ne sont pas décrits dans l'article original). Une telle contamination est fréquente lorsque l'on clone les gènes d'espèces éloignées en employant la PCR (voir, par exemple, le champignon et l'angiosperme à l'origine des gènes ribosomiques supposé de dinosaure (Wang *et al.*, 1997) ou l'origine humaine du gène mitochondrial supposé du dinosaure (Collura et Stewart, 1995), et pourrait être détectée avant publication par une analyse phylogénétique de la séquence produite. Les différences entre la séquence supposée de myxine et celle du mammifère le plus proche pourraient être dues au séquençage d'un gène d'amélogénine de mammifère précédemment non-séquencé ou plus probablement au séquençage d'un contaminant, le gène dégradé d'une souris.

Notre conclusion que la séquence supposée du gène de l'amélogénine de la « myxine », publié par Slavkin et Diekwish (1996) provient probablement d'un contaminant mammifère, n'infirme pas les résultats obtenus grâce à l'utilisation d'anticorps de mammifère (Slavkin *et al.*, 1982, 1983, 1991), mais cette séquence ne peut pas être employée pour les valider. La séquence primaire d'ADNc d'amélogénine chez la myxine reste donc à être trouvée.

Deuxième Partie

L'ORIGINE DE L'AMÉLOGÉNINE



GOULD, S.J.,1991, La vie est belle, Editions du Seuil.
Quelques exemples fossiles de la faune de Burgess illustrant "l'explosion cambrienne".

En 1998, j'effectuais mon stage de DEA dont le sujet était déjà l'amélogénine. A cette époque, je m'étais déjà rendu compte des ressemblances qui existaient entre les exons 2 de l'amélogénine et de la SPARC (ostéonectine). L'idée qu'il était possible de dater l'apparition du gène de l'amélogénine est venue en observant un arbre phylogénétique incluant ces deux exons. Cependant, le mode de calcul utilisé au cours de mon stage de DEA était très simple puisqu'il consistait uniquement à calibrer les distances évolutives sur le cladogramme construit à cette occasion avec la date connue de divergence entre sauropsidés et mammifères. De plus, le calcul des distances était basé sur l'alignement de 50 à 60 paires de bases, ce qui était un trop petit nombre pour obtenir une estimation solide des distances. D'ailleurs, les critiques (justifiées) n'ont pas manqué lors de l'évaluation de mon travail de DEA.

Malgré les difficultés de cette entreprise, l'estimation de cette date n'en restait pas moins intéressante. En effet, comme je l'ai indiqué dans le chapitre 1, "Introduction", de nombreux auteurs se demandent toujours si l'amélogénine est apparue chez les tétrapodes ou antérieurement. De plus, en raison de la spécificité de l'amélogénine (voir dans cette partie), une telle datation constituerait une bonne estimation de l'apparition de l'émail. Enfin, l'origine de l'amélogénine apporte un argument important dans la controverse sur la date d'apparition d'un tissu minéralisé chez les vertébrés, et donc sur l'explosion cambrienne.

Ces travaux ont été relancés quand les techniques de calcul sont devenues beaucoup plus fiables, grâce notamment au maximum de vraisemblance et lorsqu'il a été possible d'utiliser plus de 1000 pb pour le calcul (voir dans "matériel et méthodes"). La date d'apparition minimum de l'amélogénine obtenue dans ce travail est inférieure de presque 100 millions d'années à celle que j'avais présentée lors de ma première étude en 1998, mais surtout cette estimation est très fiable.

Ce travail est également la preuve que les études d'évolution moléculaire peuvent apporter des informations sur des questions anciennes, n'intéressant *a priori* que les paléontologues ou les morphologistes.

PREUVE MOLÉCULAIRE DE L'ORIGINE PRÉCAMBRIENNE DE L'AMÉLOGÉNINE, LA PROTÉINE MAJEURE DE L'ÉMAIL DES VERTÉBRÉS

Traduction de l'anglais à partir de l'article publié dans

Molecular Biology and Evolution, 2001, 18(12):2146–2153,

Molecular Evidence for Precambrian Origin of Amelogenin, the Major Protein of Vertebrate Enamel,

par **Sidney Delgado***, Didier Casane†, Laure Bonnaud*‡, Michel Laurin*, Jean-Yves Sire*, et Marc Girondot*§.

* UMR 8570, Evolution et Adaptations des Systèmes Ostéomusculaires, Paris, France;
† UMR 7622, Biologie Moléculaire et Cellulaire du Développement, Equipe ''Phylogénie, Bioinformatique et Génome,'' Université Pierre et Marie Curie, Paris, France;
‡ Institut Jacques Monod, Equipe ''Evolution du Développement des Nématodes,'' Université Denis Diderot, Paris, France;
§ UPRESA 8079, Ecologie, Systématique et Evolution, Université Paris-Sud, Orsay, France.

Remerciements: Nous remercions le Professeur Armand de Ricqlès et le Professeur Ann Huysseune pour leur lecture attentive du manuscrit et leurs nombreuses et précieuses

suggestions. Nous sommes aussi redevables à Matthew H. Godfrey pour sa correction de l'anglais.

Résumé

Si la datation des évènements cladogénétiques est possible, aucune méthode moléculaire n'a été décrite pour dater l'acquisition de différents tissus. En prenant en compte la spécificité de la protéine majeure de l'émail en formation (l'amélogénine), nous avons été capables de développer une telle méthode pour l'émail. En effet, puisque l'amélogénine est exclusivement associée à la formation et à la minéralisation de l'émail et parce qu'elle n'a pas d'effets pléiotropiques, cette protéine est un bon candidat pour estimer la date d'acquisition de ce tissu hautement minéralisé. Nous avons cherché dans les banques d'ADN des ressemblances entre l'amélogénine et d'autres séquences. Des ressemblances ont été trouvées seulement avec l'exon 2 de la SPARC (ostéonectine) chez deux protostomiens et chez huit deutérostomiens, et avec l'exon 2 de trois gènes de deutérostomiens apparentés à la SPARC (SC1, hevin et QR1). Les autres exons de l'amélogénine ne révélaient pas de ressemblances significatives avec d'autres séquences. Dans ces protéines, l'exon 2 encode surtout le « peptide signal » qui joue le rôle essentiel de permettre à la protéine d'être excrétée dans la matrice extracellulaire. Nous avons testé la signification des ressemblances de l'exon 2. Les valeurs observées étaient toujours significativement plus élevées que les ressemblances attendues par l'effet du hasard. Cela a démontré une origine commune au cours de l'évolution de cet exon. L'analyse phylogénétique de l'exon 2 a indiqué qu'il avait été dupliqué vers l'amélogénine à partir d'une séquence ancestrale de SPARC dans la lignée des deutérostomiens, avant la duplication de la SPARC des deutérostomiens et de SC1/hevin/QR1. Nous avons été capable de dater l'origine de la dernière duplication, il y a approximativement 630 millions d'années. Donc, l'exon 2 de l'amélogénine a été acquis avant cette date, dans le Protérozoïque, longtemps avant l'explosion Cambrienne. L'explosion cambrienne correspond à la soudaine apparition de plusieurs phyla de Bilateralia, dans le registre fossile, à la transition Protérozoïque-Phanérozoïque. Cette soudaine apparition a souvent été supposée être le reflet d'une cladogénèse intensive à cette époque. Cependant, la datation moléculaire de la divergence protostomien-deutérostomien et de la cladogénèse entre plusieurs clades majeurs de Bilateralia amène à plusieurs conclusions : beaucoup de clades de Bilateralia étaient déjà présents pendant le Protérozoïque tardif. Il a été précédemment proposé que ces Bilateralia n'étaient pas minéralisés et qu'ils avaient un faible potentiel de fossilisation. Nos résultats suggèrent fortement que des fossiles du Protérozoïque tardif possédant un tissu minéralisé homologue de l'émail pourraient être trouvés dans le futur.

Introduction

> *L'amélogénine est quantitativement la protéine majoritaire de l'émail en formation des mammifères, dans laquelle elle représente approximativement 90% du contenu organique (Termine* et al.*, 1980; Fincham* et al.*, 1982; Sasaki et Shimokawa, 1995). Plusieurs autres produits de gène ont aussi été trouvés dans la matrice de l'émail : diverses protéases et protéines anioniques (connues comme non-amélogénines). Certaines de ces protéines ont été caractérisées chez des mammifères, par exemple la tufteline (Deutsch* et al.*, 1995) et l'amélobastine (aussi nommée améline ou Sheathline) (Fong, Slaby, et Hammarström, 1996 ; Krebsbach* et al.*, 1996).*
>
> *Une fois que la minéralisation a commencé, les protéines de l'émail sont progressivement dégradées par les protéases et sont enlevées de la matrice ; ce processus dure jusqu'à la maturation complète de l'émail et aboutit à un tissu fortement minéralisé (Deutsch, 1989). La matrice minérale de l'émail est composée de cristaux d'hydroxyapatite, orientés perpendiculairement à la surface des améloblastes. Cet arrangement particulier, bien que légèrement variable (par exemple, prismatique ou pas) dans l'émail et l'émailloïde des diverses lignées des vertébrés, est caractéristique et permet de distinguer ces tissus des autres tissus minéralisés. Avant qu'elle ne soit dégradée, l'amélogénine est censée jouer un rôle important dans cette organisation des cristaux minéraux en empêchant la prolifération aléatoire des nucléations de cristaux et en réglant la cinétique de croissance, l'orientation et la taille de ces cristaux d'émail (Aoba, 1996; Lyngstadaas* et al.*, 1996).*

Toutes les caractéristiques précédentes sont spécifiques à la formation de l'émail des vertébrés et ne se rencontrent pas dans d'autres tissus minéralisés (même dans des lignées de non-vertébrés). Un autre caractère marquant, propre à la plupart de ces protéines de l'émail, est qu'elles n'ont pas été trouvées exprimées dans d'autres tissus minéralisables ou non-minéralisables ; ces protéines sont spécifiques à l'émail, autrement dit, elles n'ont aucun effet pléiotropique.

Nous nous sommes intéressés à l'origine évolutive des tissus composant le squelette dermique et la spécificité des protéines de l'émail a attiré notre attention. Puisque les gènes non-amélogénines ne sont connus que chez quelques espèces mammifères, ils n'ont pas pu être utilisés jusqu'à présent pour étudier l'évolution. En revanche, les séquences d'amélogénine sont disponibles dans plusieurs lignées de vertébrés (Girondot et Sire, 1998; Ishiyama *et al.*, 1998; Toyosama *et al.*, 1998).Par conséquent, nous avons choisi de travailler avec cette molécule. La présente étude a été entreprise pour dater l'apparition du gène de l'amélogénine chez les vertébrés. Actuellement, on pense que l'amélogénine est un gène orphelin sans aucun gène homologue connu. La connaissance de la date de duplication du gène de l'amélogénine, ou d'une partie de ce gène, et son écart-type fournirait une gamme d'âges possibles pour l'acquisition de cette protéine et du tissu, semblable à l'émail, dans lequel elle aurait été exprimée. Puisque le gène de l'amélogénine code une protéine qui joue un rôle important dans la minéralisation de l'émail et qu'il ne montre aucun effet pléiotropique, c'est un excellent candidat pour évaluer indépendamment la date d'acquisition d'un tissu minéralisé chez les vertébrés. En effet, cette date est actuellement débattue.

D'une part, selon le registre fossile, les vertébrés les plus anciens connus sont du Cambrien Ancien de la Chine (pendant la période que l'on appelle "l'explosion cambrienne"), mais ces derniers ne possédaient pas de tissus minéralisés (Chen, Huang et Li, 1999; Shu *et al.*, 1999). La minéralisation est probablement apparue plus tard chez les vertébrés. En effet, le plus

ancien présumé vertébré connu pour posséder un squelette minéralisé est l'*Anatolepis* trouvé dans le Cambrien Supérieur (520 millions d'années) (Repetski, 1978). Son squelette dermique était orné de petits tubercules composés d'un tissu identifié comme une sorte de dentine. Ce squelette était recouvert d'un tissu minéralisé d'homologie inconnue (Smith, Sansom et Repetski, 1996). De l'émail et d'autres tissus d'homologie incertaine, peut-être de l'os, de la dentine et du cartilage calcifié ont aussi été décrits chez les euconodontes du Cambrien Supérieur, mais ces interprétations ont été discutées (Schultze, 1996). Ces animaux aux propriétés controversées sont maintenant considérés comme des crâniates par beaucoup d'auteurs (Janvier 1996a; Aldridge et Donoghue, 1998). L'émail, ou le tissu émailloïde couvrant les tubercules, a été quant à lui clairement décrit chez les vertébrés ordoviciens (450 millions d'années) comme les placodermes (Ørvig, 1989). D'autre part, selon des données moléculaires, on estime que les Bilateralia trouvent leur origine pendant le Protérozoïque, entre 1000 et 830 millions d'années (Bromham *et al.*, 1998; Pagel, 1999; Bromham et Hendy, 2000).

Les tissus minéralisés des vertébrés sont-ils apparus dans le Cambrien, comme le suggère le registre fossile, ou se sont-ils développé lors du Protérozoïque ? Dans le présent article, nous avons essayé de discriminer les deux hypothèses en employant le gène de l'amélogénine. Nous avons cherché dans des banques d'ADN des ressemblances entre l'amélogénine et d'autres séquences, nous avons testé la signification de ces ressemblances et, en employant plusieurs méthodes, nous avons été capables de dater l'origine de l'exon 2 de l'amélogénine.

Matériel et méthodes

Les gènes de l'amélogénine

Les gènes de l'amélogénine ont été cloné (ADN ou ADNc) chez neuf espèces de mammifères (y compris un marsupial et deux monotrèmes), un squamate, un crocodile et un lissamphibien anoure. De plus, les gènes de l'amélogénine n'ont pas pu être détectés chez les tortues et les oiseaux, deux taxa qui ont perdu leurs dents il y a respectivement 150 et 60 millions d'années. Cela confirme, au moins chez les sauropsidés, la récente découverte qui montre que le gène de l'amélogénine n'a pas d'effets pléiotropiques qui lui auraient permis d'être maintenu en absence d'émail. Les séquences du gène de l'amélogénine ne sont pas disponibles chez les vertébrés non-tétrapodes. Cependant, de récentes études immuno-cytochimiques ont révélé des protéines semblables à l'amélogénine dans les dents d'un dipneuste, un sarcoptérygien basal (Satchell, Shuler et Diekwisch, 2000) et dans la ganoïne d'un polyptère, un actinoptérygien basal. La ganoïne qui recouvre les écailles des actinoptérygiens basaux vivants est homologue à l'émail (Sire et al., *1987; Huysseune et Sire, 1998) et cela est probablement vrai pour la ganoïne des plus anciens actinoptérygiens éteints (Janvier, 1996a). Ces résultats suggèrent que le gène responsable de la production d'amélogénine était présent au moins chez l'ancêtre commun des deux lignées sarcoptérygienne et actinoptérygienne. La présence d'amélogénine et/ou de protéines non-amélogénines dans les dents des chondrichtyens est toujours largement controversée (voir la discussion dans Zylberberg, Sire et Nanci, 1997). Cependant, tous ces résultats confirment l'absence d'effets pléiotropiques pour l'amélogénine chez les vertébrés et sa participation exclusive dans la formation de l'émail.*

Le gène de l'amélogénine n'est pas connu en dehors des vertébrés. Cependant, en 1996, Slavkin et Diekwisch ont publié une séquence partielle en

acides aminés d'amélogénine d'une myxine du Pacifique, pour soutenir l'immunodétection de protéines semblables à l'amélogénine dans la cavité buccale de ce crâniate sans mâchoire (Hyperotreri sensu *Janvier, 1996a). Les myxines vivantes n'ont pas de tissus minéralisés et cela correspond probablement à une condition primitive (Janvier, 1996b). Donc la présence d'amélogénine chez la myxine indiquerait que ce gène était exprimé dans des tissus non-minéralisés chez un ancêtre commun aux crâniates et aux vertébrés, résultat « inconsistant » avec notre hypothèse fondée sur le rôle essentiel de l'amélogénine dans la minéralisation de l'émail. Une analyse phylogénétique de l'ADN de la myxine utilisant des gènes d'amélogénine de mammifère a démontré que cette séquence supposée était probablement la conséquence d'une contamination par l'ADN d'un rongeur durant la PCR (Girondot, Delgado et Laurin, 1998). Cette découverte a été confirmée quand les séquences d'amélogénine récemment disponibles chez le caïman, le serpent et le xénope, ont été ajoutées à l'analyse (voir* matériel supplémentaire*, EMBL ALIGN_000059). L'arbre obtenu (ln L = -710.6) si la séquence de myxine est contrainte d'être le groupe frère des amniotes, grâce à la méthode ProtML de ré-échantillonnage de l'estimation du log de vraisemblance (RELL) (Hasegawa et Kishini 1994), se trouve être pire que l'arbre de maximum de vraisemblance (10000 ré-échantillonnages) dans 97,75% des cas. En outre, une analyse de parcimonie employant PAUP 3.1.1 (Swofford, 1993) a indiqué que quatre pas supplémentaires étaient nécessaires pour placer la séquence de la myxine à l'extérieur des amniotes.*

A la recherche de similarités de séquences

Une première recherche avec BLAST dans GenBank avec différents exons de l'amélogénine n'a pas révélé de ressemblances significatives avec d'autres séquences. Néanmoins, notre attention a porté sur l'observation suivante : les améloblastes de rat (les cellules épithéliales qui produisent l'amélogénine) montrent un signal positif quand ils sont hybridés *in situ* avec une sonde d'ARN de SPARC (Liao *et al.,* 1998). Cette protéine ("Secreted Protein, Acidic and Rich in Cystein"), aussi appelée ostéonectine, BM-40 et protéine 43K, est une glycoprotéine qui est impliquée dans la médiation des interactions de la matrice cellulaire, mais qui ne joue pas de rôles structuraux (Brekken et Sage, 2000). Liao *et al.* (1998) ont suggéré que ce signal était probablement dû à la présence de 10 bases identiques dans les séquences de SPARC et d'amélogénine de rat. En effet, l'alignement de SPARC et des séquences d'amélogénine a révélé que la traduction complète du second exon des deux gènes (le premier exon n'est pas traduit) montre des ressemblances. Nous avons cherché d'autres séquences de SPARC dans des banques d'ADN en employant les outils Nentrez ou Blast. Un total de 13 séquences a été trouvé : 8 séquences complètes de SPARC chez les deutérostomiens (dSPARC), 2 séquences complètes chez les protostomiens (pSPARC), un nématode et une drosophile, et les séquences partielles (c'est-à-dire, sans l'exon 2) de trois autres deutérostomiens. Trois autres gènes apparentés à la SPARC ont été inclus dans l'analyse : SC1 des rongeurs, hevin, son homologue chez l'homme et QR1 leur homologue chez la caille. Dans la suite de cet article, ces trois gènes apparentés à la SPARC seront collectivement appelés SC1 (le premier identifié). De plus, au cours de notre recherche nous avons trouvé 19 séquences exprimées (EST) de poisson zèbre (*Danio rerio*) montrant quelques ressemblances avec la SPARC. Après une vérification étendue, nous avons reconstruit la séquence complète d'ARNm de SPARC chez le poisson zèbre (voir matériel supplémentaire, section 2) et cette séquence a été ajoutée à notre analyse (sur le site Internet

de MBE, voir *matériel supplémentaire*, section 3, EMBL Align_000006).

Les séquences d'amélogénine ont aussi été recherchées dans les banques d'ADN en employant les mêmes procédures et 20 séquences ont été trouvées. Seulement 12 séquences d'amélogénine étaient complètes (avec l'exon 2) et ont ainsi été retenues pour l'analyse. Dans la SPARC, SC1 et l'amélogénine, l'exon 2 est composé de trois régions : la première est constituée de 12 nucléotides non traduits, la deuxième de 48 nucléotides traduits constituant le peptide signal, et la troisième de 6 nucléotides qui constituent le début de la protéine elle-même. Les régions 2 et 3 de dSPARC, de SC1 et de l'exon 2 de l'amélogénine (un total de 19 séquences) peuvent être alignées avec seulement un seul ensemble de 3 délétions pour la totalité des 54 nucléotides (sur le site Internet de MBE, voir *matériel supplémentaire*, section 4, EMBL Align_000005). Le peptide signal est fortement hydrophobe et permet à la protéine en cours de synthèse (qui est fortement hydrophile) de passer à travers la membrane du réticulum endoplasmique rugueux. Une fois localisée dans le *lumen* du réticulum endoplasmique, la protéine est transportée dans l'appareil de Golgi, puis à l'extérieur de la cellule au moyen de vésicules intracellulaires.

Evaluation des ressemblances

Nous avons eu besoin de nous assurer que le dSPARC, SC1 et les séquences d'amélogénine partageaient un ancêtre commun, c'est-à-dire, que la similitude observée reflétait une homologie réelle. Ainsi, nous avons testé les ressemblances de dSPARC, de SC1 et de l'exon 2 de l'amélogénine.

Les proportions d'identité des différents couples de séquences ne peuvent pas être directement comparées parce que ces valeurs ne sont pas indépendantes. En effet, nous supposons que toutes les amélogénines, d'une part, et dSPARC et SC1, d'autre part, ont une histoire phylogénétique commune. Donc, la plupart des séquences ancestrales les plus parcimonieuses pour le second exon de chaque gène ont été calculées séparément par MacClade 3.07 (Maddison et Maddison, 1992). En raison des ambiguïtés dans la séquence ancestrale, la position *i* dans une séquence *j* est entièrement décrite par les cinq fréquences f_{Aij}, f_{Gij}, f_{Cij}, f_{Tij} et f_{Iij}, qui sont respectivement les probabilités que A, G, C, T ou un gap soient présents dans cette position pour cette séquence (sur le site Internet de MBE, voir *matériel supplémentaire*, section 5). Puis, la probabilité moyenne que deux bases identiques soient présentes dans un alignement de *l* bases est estimée grâce à :

$$\Phi = \sum_{i=1}^{l}\left(f_{Ai1}\cdot f_{Ai2} + f_{Ti1}\cdot f_{Ti2} + f_{Gi1}\cdot f_{Gi2} + f_{Ci1}\cdot f_{Ci2}\right) \div \left(l - \sum_{i=1}^{l}\left(f_{Ii1}\cdot f_{Ii2}\right)\right).$$

Pour chaque couple, la distribution de Φ sous l'hypothèse nulle (c'est-à-dire, que la valeur observée Φ n'est pas significativement différente de ce qui est attendu pour des séquences aléatoires) a été établie en utilisant 1000 copies de deux séquences aléatoirement générées par ordinateur et partageant les mêmes caractéristiques que les séquences dans le test (même longueur, même niveau d'ambiguïté et même composition en bases). Une simple permutation a également été utilisée pour générer les séquences aléatoires. Chaque paire de séquences générées aléatoirement a été alignée de manière à maximiser la valeur estimée de Φ sous H_0 (Φ_0). Ce test a été réalisé avec les trois bases des codons aussi bien qu'avec les bases 1 et 2 pour ne pas tenir compte des substitutions synonymes qui ont lieu principalement sur la troisième base de chaque codon.

Recherche de l'origine de l'exon 2 de l'amélogénine

Des arbres BIO-NJ (Gascuel, 1997), utilisant soit les distances de Kimura à 2 paramètres (Kimura, 1980) soit les distances de Tajima et de Nei à 4 paramètres (Tajima et Nei, 1993), ont été produits avec seulement la partie traduite de SPARC, de SC1 et de l'exon 2 de l'amélogénine (Figure 1). Plusieurs méthodes ont été appliquées car la séquence (54 nucléotides) est courte et nous avons eu besoin de nous assurer que le résultat était significatif. Quand ces arbres ont été enracinés avec les deux séquences pSPARC, trois groupes monophylétiques ont toujours été observés : dSPARC, SC1 et l'amélogénine. Cependant, la position relative de ces trois groupes était incertaine.

La méthode des covarions a été appliquée pour éliminer une grande part de la saturation de substitution (Lopez, Forterre et Philippe, 1999). Les positions avec plus d'un changement inféré par la parcimonie dans n'importe lequel des quatre groupes monophylétiques (pSPARC, dSPARC, SC1 et amélogénine) ont été changées en état de caractère manquant ("?") dans les séquences de ce groupe (équivalent de la matrice H1 de Lopez, Forterre et Philippe, 1999) et les gaps ont été considérés comme des "NEWSTATES" pour l'analyse par parcimonie.

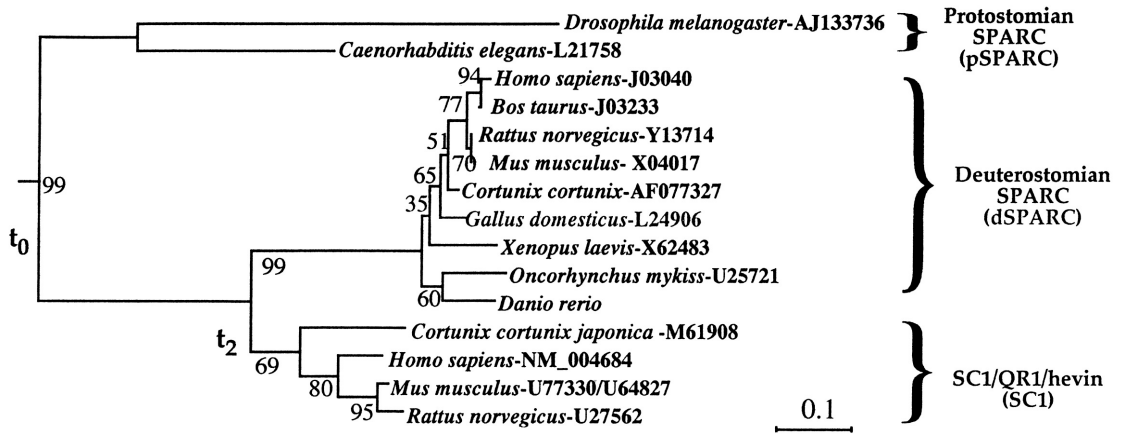


Figure 1. Arbre phylogénétique de SPARC et des gènes liés à SPARC produit en utilisant les séquences en acides aminés. Les distances de Gonnet ont été appliquées avec la méthode de construction d'arbre de neighbor-joining. Les valeurs de bootstrap sont indiquées aux nœuds correspondants. Les dates t_0 et t_2 sont, respectivement, les temps de divergence entre les protostomiens et les deutérostomiens, et la duplication qui a produit dSPARC et SC1 chez les deutérostomiens.

Ce nouvel alignement a été analysé en employant la parcimonie et les bootstraps relatifs aux trois positions possibles de dSPARC, de SC1 et de l'amélogénine, ont été établis grâce à 100 copies avec une recherche heuristique dans PAUP 3.1.1 (Swofford, 1993). Finalement, les trois positions relatives possibles de dSPARC, de SC1 et de l'amélogénine ont été évaluées grâce à la méthode de RELL (Hasegawa et Kishimo, 1994) avec des options par défaut. La position phylogénétique des séquences parmi chaque groupe monophylétique (dSPARC, SC1 et l'amélogénine) a été choisie en accord avec la phylogénie des vertébrés couramment acceptée (Kumar et Hedges, 1998; Murphy *et al.*, 2001).

La datation de l'origine de l'exon 2

Le petit nombre de nucléotides dans l'exon 2 ne permettait pas de dater son origine. En revanche, dater l'origine de la duplication de dSPARC/SC1 (plus de 1000 nucléotides disponibles) était possible.

Les fonctions différentes qui ont été attribuées à l'amélogénine et aux protéines SPARC et SC1 soulèvent la possibilité de différences dans les taux évolutifs pour ces gènes paralogues ; donc, une méthode qui pourrait prendre en compte cette potentialité est nécessaire. Rambaut et Bromham (1998) ont proposé une méthode de quartet, dans laquelle les taux évolutifs spécifiques d'une lignée sont calculés en utilisant le maximum de vraisemblances. Cette méthode exige que quatre espèces soient groupées en paires dans une phylogénie et que deux dates internes de divergence doivent être connues. Les dates estimées de divergence suivantes ont été utilisées : sauropsides *vs*. synapsides, 310 millions d'années ; lissamphibiens *vs*. amniotes, 360 millions d'années ; et actinoptérygiens *vs*. sarcoptérygiens, 410 millions d'années (Kumar et Hedges, 1998). Cette méthode emploie des références internes pour les dates de divergence et cela est plus approprié pour notre étude que l'utilisation de dates de divergence externes, comme la date de divergence entre les protostomiens et les deutérostomiens qui est fortement controversée (Ayala et Rzhetsky, 1998; Bromham *et al.*, 1998; Gu, 1998; Bromham et Hendy, 2000). De plus, pour cette phase de l'évolution des métazoaires, le registre fossile est extrêmement pauvre (Valentine, Jablonski, et Erwin, 1999).

> *Pour estimer les dates de duplication de SPARC et SC1, toutes les combinaisons de quartet possibles (n=84) impliquant deux séquences dSPARC, d'une part, et deux séquences SC1, d'autre part, ont été produites. Pour chacun de ces quartets, le taux de substitution a été calculé en utilisant le modèle de substitution à 2 paramètres de Hasegawa, Kishino et Yano (1985). Le taux de substitution estimé a été testé par le modèle de substitution à 5 paramètres. Si une différence significative est observée, cela signifie que le modèle de substitution à 2 paramètres n'est pas suffisant pour produire une évaluation fiable du taux de substitution car le modèle de substitution à 5 paramètres altère de façon significative l'estimation ; aucun des 84 tests n'a été rejeté au niveau de 5%.Par conséquent, des dates minimale et maximale ont été évaluées à la limite de confiance de 95%, pour chacun des 84 quartets, en utilisant le modèle de substitution à 2 paramètres. Ces 84 valeurs ne sont pas indépendantes les unes des autres puisque ces séquences partagent une histoire phylogénétique commune. Pour cette raison, nous avons présenté tous les résultats des dates évaluées de duplication entre SPARC et SC1. L'emploi de la méthode de l'horloge moléculaire permet ici de calculer un taux évolutif spécifique d'une molécule (dSPARC ou SC1/QR1/Hevin) (Rambaut et Bromham, 1998). La cohérence parmi les 84 évaluations suggère que les taux évolutifs spécifiques des espèces pour une molécule donnée (dSPARC ou SC1/QR1/Hevin), ne sont pas largement divergents.*

Résultats

L'arbre SPARC révèle qu'une duplication d'un gène SPARC ancestral dans la lignée des deutérostomiens a créé SC1/HEVIN/QR1 et dSPARC, et que cet événement est arrivé avant la divergence entre les actinoptérygiens et les sarcoptérygiens (Figure 1). Donc, SC1, hevin et QR1 sont considérés comme orthologues les uns aux autres et ils sont paralogues à dSPARC. Cela est confirmé par l'absence de SC1/hevin et de QR1 dans le génome complètement séquencé de deux protostomiens, *Caenorhabditis* et *Drosophila*.

L'origine de SPARC, SC1 et de l'exon 2 de l'amélogénine

La valeur Φ est de 0,4787 pour dSPARC *versus* SC1, de 0,5064 pour dSPARC *versus* amélogénine et de 0.5447 pour l'amélogénine *versus* SC1, quand les 3 bases des codons sont analysées, et ces valeurs sont 0,5625, 0,6038 et 0,6476 respectivement, quand on ne tient compte que des deux premières bases. En évaluant les ressemblances des séquences de l'exon 2 des deutérostomiens, les valeurs observées sont toujours significativement plus grandes que les similarités attendues lorsqu'elle sont produites aléatoirement (Φ_{H0} ; $P < 0{,}001$ pour dSPARC *vs.* SC1, dSPARC *vs.* amélogénine et amélogénine *vs.* SC1 quelle que soit la méthode utilisée pour les générer ; voir *Matériels et Méthodes*) (Figure 2).

Ces ressemblances démontrent clairement une origine évolutive commune de cet exon. Un test interne de la méthode (c'est-à-dire, que la méthode ne surestime pas les similarités) est obtenu quand l'exon 2 de pSPARC est comparé à n'importe quel exon des deutérostomiens. En effet, l'organisation génomique des exons de SPARC des protostomiens et des deutérostomiens, la fonction du peptide signal, et le profil d'hydrophobicité de la protéine, suggèrent que l'exon 2 de dSPARC et de pSPARC sont réellement homologues, mais la similarité n'est pas statistiquement significative. Cela indique que les mutations ont effacé une grande part du signal phylogénétique. Il doit aussi être noté que les insertions/délétions de 3 bases requis pour un alignement correct des séquences des deutérostomiens (dSPARC, SC1 et amélogénine) ne changent pas la phase ouverte de lecture.

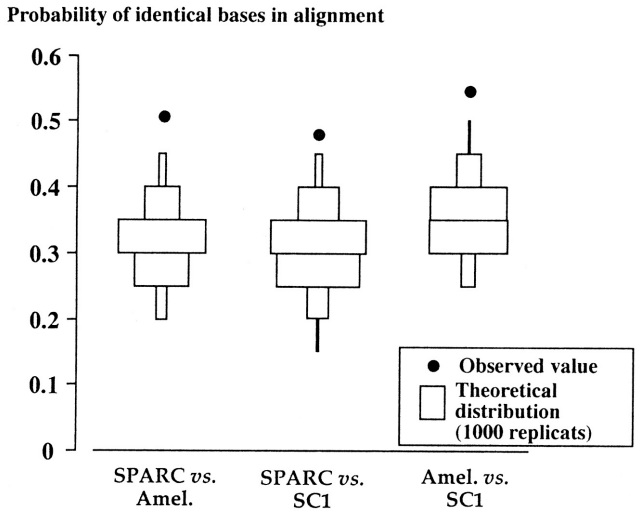


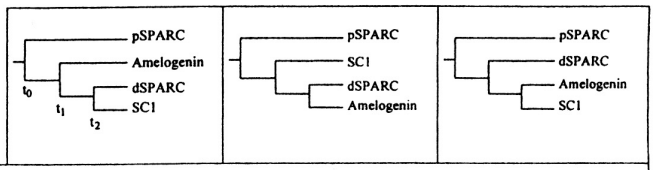
Figure 2. Comparaison et test de signification des probabilités de bases identiques dans l'alignement de l'exon 2 de SPARC, SC1 et l'amélogénine quand les 3 bases du codon sont utilisées. Les points noirs représentent les valeurs de la probabilité d'avoir des bases identiques dans l'exon 2 pour des comparaisons entre paires de séquences ancestrales (reconstituées en utilisant la parcimonie). Les histogrammes montrent la distribution de cette valeur sous l'hypothèse nulle, c'est-à-dire que les séquences ne sont pas homologues, en utilisant des séquences générées par ordinateur.

Un autre argument en faveur d'une histoire évolutive commune de l'exon 2 est obtenu par l'observation de l'organisation génomique. En effet, l'extrémité 3' de la zone de similitude dans les séquences d'amélogénine et de SPARC/SC1 correspond exactement à la fin de l'exon 2.

Une date pour la présence de l'exon 2 dans l'amélogénine

L'arbre de l'exon 2 montre que le peptide signal qu'il code a été dupliqué à partir d'une séquence de SPARC vers l'amélogénine avant la duplication entre dSPARC et SC1. Cette conclusion est soutenue par toutes les analyses phylogénétiques de ces séquences (Figure 3). Cette duplication s'est faite avant la séparation entre les actinoptérygiens et les sarcoptérygiens car l'exon 2 de l'amélogénine se trouve à l'extérieur du groupe formé par les SPARC des actinoptérygiens (*Danio rerio* et *Oncorhynchus mykiss*) et des sarcoptérygiens (qui comprend les séquences des tétrapodes).

De plus, l'acquisition de l'amélogénine précède la duplication de SPARC/SC1 qui s'est produit très probablement autour de 630 millions d'années (Figure 4). Donc, cette date de duplication peut être considérée comme une sous-estimation de la date d'acquisition de l'exon 2 dans le locus de l'amélogénine. Cette date se trouve être bien antérieure à l'explosion cambrienne et à l'apparition soudaine de minéralisation dans le registre fossile (543 millions d'années) et plus antérieure que le Vendéen (610 millions d'années), à partir duquel plusieurs fossiles présumés de métazoaires sont connus (Valentine, Jablonski et Erwin, 1999).L'origine des autres parties de la séquence de l'amélogénine est toujours inconnue. La fonction de l'exon 2 (le peptide signal), qui consiste dans le passage des protéines à travers la membrane du réticulum endoplasmique est essentielle pour transférer ces mêmes protéines hors des cellules et dans la matrice extracellulaire. L'émail ne pourrait pas exister si l'amélogénine n'était pas produite extra-cellulairement par les améloblastes, puisque cette protéine joue un rôle crucial dans la minéralisation de l'émail (Deutsch, 1989; Schwarzbauer et Spencer, 1993).



RELL method (Hasegawa, and Kishino 1994)			
Ln *L*	-760.9	-762.8	-763.1
AIC	1611.7	1615.7	1616.2
S.E. from ML	ML tree	3.2	3.0
p	-	$p<0.001$	$p<0.002$

Figure 3. Trois relations de parenté possibles entre dSPARC, SC1 et l'exon 2 de l'amélogénine, évaluées grâce à la méthode de ré-échantillonnage des log de vraisemblance (Hasegawa et Kishino 1994). L = vraisemblance; AIC = Critère d'Information Akaike (Akaike, 1974); S.E. de ML = erreur standard du maximum de vraisemblance. Cela représente la distance (c à d, le nombre standard d'erreurs de vraisemblance) entre un arbre et le meilleur des arbres (l'arbre ML).La valeur de *P* associée est la probabilité qu'une telle vraisemblance soit observée avec l'arbre ML. Les dates t_0, t_1 et t_2 sont respectivement, le temps de divergence entre les

protostomiens et les deutérostomiens et les deux événements de duplication qui ont produit (1) l'amélogénine et dSPARC et (2) SC1 chez les deutérostomiens.

Théoriquement, la période maximale pendant laquelle une séquence non-exprimée peut persister dans un génome, en restant potentiellement fonctionnelle, a été évaluée comme n'excédant pas 10 millions d'années (Marshall, Raff et Raff, 1994). Au-delà de cette durée, le gène de l'amélogénine serait perdu en raison de l'accumulation des mutations.

Donc, nous concluons que la date d'acquisition de l'exon 2 pour l'amélogénine coïncide probablement avec la date d'apparition de l'amélogénine. Cela signifie probablement que l'émail des vertébrés (ou de ses précurseurs) n'est pas apparu plus tard que 630 millions d'années, même en tenant compte de l'écart type de l'évaluation (Figure 4).

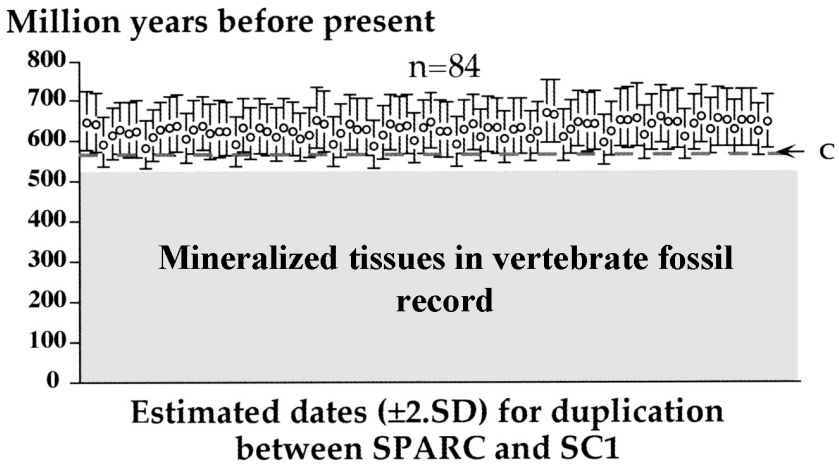


Figure 4. Dates évaluées de duplication entre SPARC et SC (t_2 dans les figures 1 et 3) en utilisant la méthode de quartet (Rambaut et Bromham, 1998). "C" indique la limite entre le Précambrien et le Cambrien.

Discussion

En développant une méthode de datation utilisant un gène (l'amélogénine) impliqué dans la minéralisation de l'émail et sans aucun effet pléiotropique connu, nous avons démontré que l'exon 2 de ce gène a été recruté à partir d'une duplication de l'exon 2 du gène SPARC. Cette duplication est survenue très probablement lors du Précambrien, il y a 630 millions d'années, et presque 100 millions d'années avant l'explosion cambrienne, une période caractérisée par l'apparition soudaine de la minéralisation dans le registre fossile. L'origine de l'exon 2 de l'amélogénine est proche du début du Vendéen (610 millions d'années), dans lequel plusieurs fossiles présumés de métazoaires sont connus (Valentine, Jablonski et Erwin, 1999). La méthode d'horloge moléculaire employée dans la présente étude (Rambaut, et Bromham, 1998) tient compte du taux évolutif spécifique de la molécule (dSPARC ou SC1). Une méthode tenant compte du taux évolutif spécifique des espèces pourrait être plus fiable, mais aucune n'est actuellement

valable (Pagel, 1999). La cohérence entre les 84 évaluations suggère que les taux évolutifs spécifiques des espèces pour une molécule donnée (dSPARC ou SC1) ne sont pas largement divergents.

L'acquisition de l'exon 2 dans le locus de la « pré-amélogénine » est légèrement antérieure à la date de la duplication de dSPARC et de SC1, survenue il y a presque 630 millions d'années (t_2 dans la Figure 1). On peut considérer la date calculée pour la duplication de dSPARC et de SC1 comme une évaluation fiable pour la date limite possible d'acquisition de l'exon 2 dans le locus de l'amélogénine.

Le manque de fossiles minéralisés dans le Protérozoïque peut être dû à la très petite taille des individus ou à la fragilité du squelette minéralisé (par exemple, les écailles d'*Anatolepis* ont moins de 0,1 mm d'épaisseur). Les fossiles de métazoaires ont été trouvés dans seulement quelques sites Protérozoïques. Nous ne serions pas étonnés si des fossiles minéralisés étaient trouvés dans les derniers dépôts du Protérozoïque.

L'histoire du phylum des Bilateralia est caractérisée par leur apparition soudaine dans le registre fossile dans le Cambrien Ancien à la transition Protérozoïque-Phanérozoïque (il y a environ 543 millions d'années), au cours de la supposée « explosion Cambrienne » (Knoll et Carroll, 1999; Morris, 1997; Valentine, Jablonski et Erwin, 1999). Les vertébrés n'ont pas été une exception à cette règle et les premiers fossiles du Cambrien Ancien appartenant indiscutablement à la lignée des vertébrés ont été récemment décrits dans une localité chinoise (Chen, Huang et Li, 1999; Shu *et al.*, 1999). Un de ces fossiles peut être rapproché des lamproies et l'autre est probablement un vertébré basal. L'apparition soudaine de clades bilatéraliens, pendant les 10 millions d'années qu'a duré « l'explosion Cambrienne », a longtemps été interprétée comme le résultat d'une radiation évolutive grande et rapide (voir par exemple, Valentine, Jablonski et Erwin, 1999). Cependant, cette découverte a été mise en doute par plusieurs études moléculaires. En employant différents modèles d'horloge moléculaire avec plusieurs jeux de gènes, l'origine de nombreux clades de Bilateralia est systématiquement évaluée comme étant plus ancienne que « l'explosion Cambrienne » (Dortey, Briggs, et Wills, 1997) et la divergence entre les Protostomiens et les Deutérostomiens a été estimée comme ayant eu lieu entre 1000 et 830 millions d'années, pendant le Protérozoïque (Wray, Levinton et Shapiro, 1996; Bromham *et al.*, 2000; Gu, 1998). Toutefois, toutes ces estimations sont soumises à caution car la plupart d'entre-elles utilise une horloge moléculaire qui est supposée être régulière au cours de la période étudiée. Cette hypothèse a récemment été testée et la même conclusion s'y applique (Bromham et Hendy, 2000).

Cependant, cette hypothèse a récemment reçu l'appui des données fossiles : un animal semblable à un mollusque (Fedonkin et Waggoner, 1998), de nombreux terriers et quelques embryons de spiraliens (comme des Annélides ou des Mollusques) (Bengton, 1998; Xiao, Zhang et Knoll, 1998) ont été décrits comme appartenant au Protérozoïque tardif (Valentine, Jablonski et Erwin, 1999). Ainsi, la grande augmentation de l'abondance de fossiles de métazoaires dans le Cambrien pourrait être un artefact taphonomique dû aux conditions environnementales ne conduisant pas à la fossilisation ou au manque de Bilateralia dans le Protérozoïque, à leur petite taille, ou à leur manque de minéralisation. Quoi qu'il en soit, cela a rendu les chances de fossilisation moins importantes et a conduit à la conclusion incorrecte que les Bilateralia se sont diversifiés en un temps très court. L'hypothèse d'un manque de minéralisation est soutenue par l'observation d'une acquisition simultanée de la minéralisation dans la plupart des phyla animaux pendant le Cambrien (Bengston, Conway et Morris, 1992). Le peu de fossiles du Protérozoïque qui ont été interprétés comme appartenant au clade bilatéralien manquent de tissus minéralisés. L'exemple le plus révélateur est *Kimberella* qui a été réinterprété comme un mollusque (placophore ?) sans coquille minéralisée (Fedonkin et

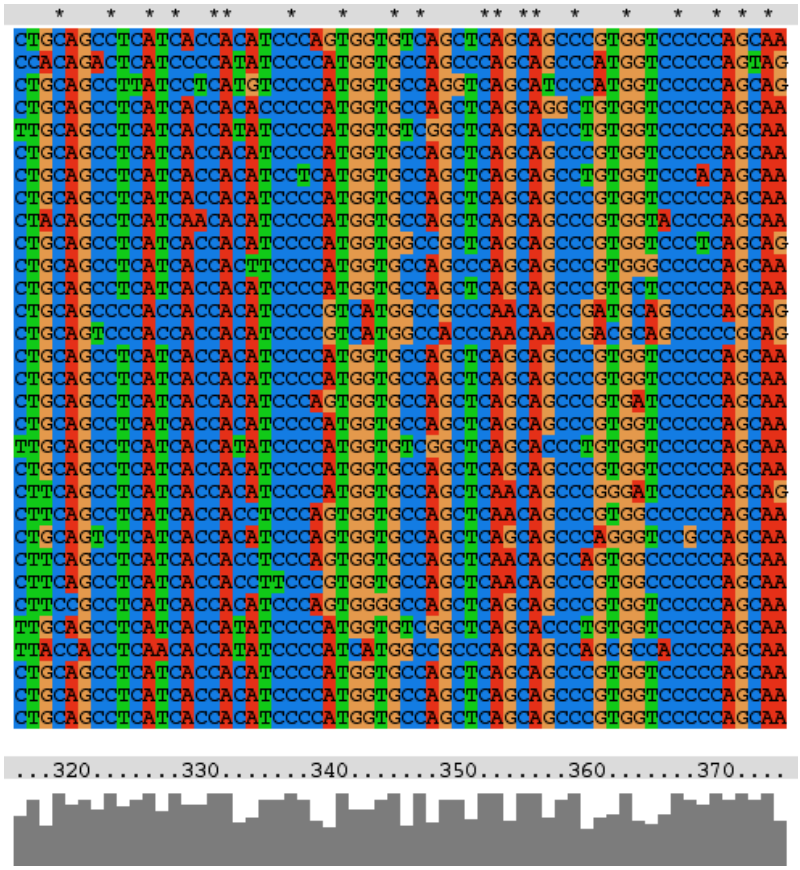
Waggoner, 1998).Quelques fossiles Bilateralia du Cambrien Ancien, qui sont les groupes frères de clades qui ont par la suite acquis la minéralisation, ne sont eux-mêmes pas minéralisés. Par exemple, les lobopodes du Cambrien Ancien peuvent être le groupe frère des arthropodes, mais ils possèdent une cuticule non-minéralisée (Min, Kim et Kim, 1998).

La phylogénie des crâniates présentée par Janvier (1996b) suggère que l'absence de tissus minéralisés chez la myxine et les lamproies est primitive. Si cette conclusion est acceptée, nos résultats suggèrent que les vertébrés basaux (comme *Anatolepis* ou les Thélodontes) étaient déjà présents il y a 630 millions d'années.

Notre étude est la première tentative visant à discriminer les hypothèses d'une radiation évolutive tardive et rapide de métazoaires (soutenue par une interprétation littérale du registre fossile) et une radiation évolutive plus ancienne et plus lente (soutenue par les dates moléculaires de divergence des taxa de métazoaires), grâce à l'emploi d'une date moléculaire d'acquisition d'une protéine, l'amélogénine, impliquée dans la minéralisation de l'émail.

Troisième Partie

L'ÉVOLUTION DU GÈNE DE L'AMÉLOGÉNINE



Exemple d'un alignement de séquences d'ADN d'amélogénine illustrant le travail effectué en évolution moléculaire.

L'intérêt de ce travail est apparu tout naturellement lorsque j'ai commencé à obtenir de nouvelles séquences d'amélogénines chez les mammifères et chez les reptiles. Immédiatement, de nouvelles questions sont apparues en constatant que les différentes parties du gène n'évoluent pas au même rythme : pourquoi certaines régions sont-elles plus conservées que d'autres ? Pourquoi certains acides aminés ne varient jamais ? Quelles sont les conséquences fonctionnelles de ces mutations dans la protéine ?

Pour tenir compte des spécificités des deux lignées, mammalienne et reptilienne, j'ai décidé d'étudier l'amélogénine dans chacun des deux groupes, séparément, en commençant par les mammifères, car il s'agit du groupe chez lequel l'amélogénine est la plus étudiée et à partir duquel j'ai pu définir les premières amorces pour la PCR.

ÉVOLUTION MOLÉCULAIRE DE L'AMÉLOGÉNINE CHEZ LES MAMMIFÈRES

Ce travail a été réalisé par :

Sidney Delgado*, **Marc Girondot†, et Jean-Yves Sire*.**

* UMR 8570, Évolution et Adaptations des Systèmes Ostéomusculaires, Équipe "Evolution et Développement du Squelette Dermique", Paris, France.
† UPRESA 8079, Laboratoire "Écologie, Systématique et Évolution", Équipe de Conservation des Populations et des Communautés, CNRS et Université Paris-Sud, Paris XI, , Bâtiment 362, 91405 Orsay Cedex.

Remerciements: Je suis grandement reconnaissant auprès des collègues suivants qui ont fourni de l'ADN ou des tissus congelés ou conservés dans l'alcool: Dr F. Catzeflis (Université Montpellier 2), Dr. G. Véron (Laboratoire de Zoologie, mammifères et oiseaux, Muséum national d'Histoire naturelle, Paris), Dr. W. Dabin (Muséum de la Rochelle) et les Drs F.

Ollivet et A. Lécu (Zoo de Vincennes, Muséum national d'Histoire naturelle, Paris).

Résumé

L'amélogénine est la protéine majeure de l'émail en formation chez les mammifères et, en tant que telle, joue certainement des rôles importants dans la détermination de la structure de ce tissu protecteur fortement minéralisé. Cependant, la fonction exacte de cette protéine reste toujours peu comprise. Une des approches à mettre en œuvre pour comprendre le rôle de ses différentes régions consiste à comparer des séquences dans une perspective évolutive. En effet, les régions qui ont été conservées pendant une longue période d'évolution indiquent que ces dernières peuvent avoir un rôle dans la structure ou dans la fonction de la protéine.

Premièrement, nous avons cherché des séquences d'amélogénine de mammifères dans des bases de données ; 23 séquences appartenant à 16 espèces ont été trouvées. Certaines de ces séquences étaient des amélogénines liées au chromosome Y et nous les avons analysées au regard des connaissances actuelles sur l'évolution des chromosomes sexuels. Deuxièmement, nous avons cloné et séquencé les gènes de l'amélogénine dans 14 autres espèces pour réaliser notre analyse évolutive sur des séquences d'amélogénine représentatives des lignées principales de mammifères. Un total de 26 séquences d'amélogénine ont été incluses dans cette analyse. La structure primaire de l'amélogénine montre des résidus fortement conservés dans les régions terminales hydrophiles N-(exons 2, 3, 5 et région 5' de l'exon 6) et C-(région 3 ' de l'exon 6 et l'exon 7). La région centrale hydrophobe de la protéine (la plus grande partie de l'exon 6) est caractérisée par des répétitions de trois acides aminés (des triplets incluant souvent une proline et une glutamine) et montre de nombreuses substitutions, des délétions et des insertions. Il est clair que cette région a évolué à une vitesse plus rapide que les autres régions bien conservées qui sont structurellement et/ou fonctionnellement contraintes. En effet, la comparaison des taux de substitution et de la richesse en CpG a démontré que les régions fortement conservées de l'amélogénine sont soumises à de fortes pressions sélectives. Les acides aminés localisés à des sites importants (phosphorylation, liaison au N-acétyle-glucosamine, épissage alternatif, protéolyse) sont présents dans toutes les séquences. Les résidus connus, quand ils sont substitués, pour mener à l'amélogenèse imparfaite sont aussi conservés dans toutes les espèces. Néanmoins, le rôle d'un certain nombre de résidus fortement conservés reste à découvrir.

Au contraire, la variabilité observée dans l'exon 6 suggère une région moins contrainte. Cependant, pendant l'évolution de l'amélogénine cette région a conservé une grande quantité de proline et de glutamine indiquant que ces résidus jouent certainement un rôle dans la structure de la protéine. Les tests sur les répétitions ont montré que les insertions de répétitions observées chez quelques espèces s'étaient produites récemment et indépendamment dans les différentes lignées. D'autres répétitions étaient plus anciennes et sont probablement apparues comme une insertion de répétitions chez un ancêtre commun des mammifères. L'analyse évolutive de l'amélogénine actuellement entreprise chez les lépidosauriens et les crocodiliens pourra confirmer et préciser ces données.

Introduction

L'amélogénine est le composant principal de l'émail en formation des mammifères, dans lequel elle représente approximativement 90% du contenu en protéine (Termine *et al.*, 1980; Fincham *et al.*, 1983; Sasaki et Shimokawa, 1995). Étant une protéine largement prédominante, l'amélogénine joue certainement des rôles essentiels dans la structuration de l'émail en développement, mais son organisation et le rôle exact de ses différents domaines sont toujours peu compris.

L'émail bien formé des mammifères est un tissu épais et fortement minéralisé qui fonctionne dans la protection des dents. Cela représente un rôle crucial, quand on considère i) l'importance des dents pour l'alimentation et ii) le fait que les dents des mammifères ne sont pas renouvelées durant la plus grande partie de la vie (diphyodontie et monophyodontie), à la différence d'autres lignées de vertébrés qui remplacent leurs dents durant toute leur vie (polyphyodontie). Étant donnée la forte prédominance de cette protéine dans l'émail en formation, on peut facilement s'imaginer ce que pourraient être les conséquences des défauts dans la structure de l'amélogénine pour la formation de l'émail. Chez l'homme, on a relevé quelques mutations (des délétions ou des substitutions de nucléotides) dans la séquence du gène de l'amélogénine pouvant être responsables de différents types d'amélogénèse imparfaite (Lagerstrom, *et al.*, 1991; Aldred *et al.*, 1992; Lench *et al.*, 1994; Collier *et al.*, 1997; Gibson *et al.*, 2001). Un tel rapport entre une malformation observée de l'émail et un défaut dans la structure du gène de l'amélogénine pourrait mettre en évidence des domaines (ou même des acides aminés particuliers) essentiels pour que cette molécule puisse jouer sa bonne fonction.

Ces dix dernières années un certain nombre de d'approches biochimiques et *in vitro* ont apporté de nouveaux éclairages dans la connaissance de la structure de l'amélogénine (voir revue dans Fincham et Moradian-Oldak, 1995; Gibson *et al.*, 2001). Les différents domaines de la protéine de l'amélogénine ont été supposés être impliqués dans différentes fonctions du développement de l'émail, comme le contrôle du processus de minéralisation à travers des interactions avec les cristaux d'hydroxyapatite (Aoba *et al.*, 1989; Fincham *et al.,* 1991; Deutsch *et al.,* 1995; Aoba, 1996; Paine *et al.,* 2000), la réalisation de rôles régulateurs ou encore la signalisation intercellulaire (Veis *et al.*, 2000; Gibson *et al.*, 2001). Moradian-Oldak *et al.* (1995) ont montré que l'amélogénine native forme des agrégats sphériques (nanosphères). Ravindranath *et al.* (1999) ont examiné les interactions possibles de l'amélogénine avec d'autres glycoprotéines de la matrice de l'émail (tufteline, énaméline, améloblastine) et ils ont montré qu'une région de l'amélogénine était liée aux N-acétyle-D-glucosamines. Plus récemment, Veis *et al.* (2000) ont montré que certains produits d'épissage sont des molécules de signalisation qui jouent un rôle dans la maturation des pré-odontoblastes. Clairement ces travaux ont ajouté de nouvelles pièces à la connaissance de l'amélogénine, mais de nouvelles données sont toujours nécessaires pour résoudre l'énigme des rapports entre la structure et la fonction.

Une approche qui pourrait fournir une information importante sur la structure de l'amélogénine (et qui permettrait par conséquent de concentrer l'attention sur les fonctions biologiques possibles) consiste dans la comparaison de séquences dans une perspective évolutive. En effet, parmi les vertébrés, les variations observées dans la structure de l'émail résultent probablement de changements évolutifs des composants de la matrice de l'émail en développement. Des études morphologiques ont montré que la structure de l'émail est grossièrement semblable chez les mammifères et les amniotes non-mammaliens, comme les squamates et les crocodiles (Sander, 2001). Si la fonction de l'amélogénine est aussi importante qu'on le suppose cela signifierait que sa structure est semblable chez toutes les espèces mammaliennes. Des études précédentes ont en effet montré qu'un certain nombre

d'acides aminés sont bien conservés dans l'amélogénine du petit nombre d'espèces de mammifères examinées (Bonass *et al.,* 1994; Hu *et al.*,1996; Toyosawa *et al.*, 1998).

Jusqu'à présent, les gènes de l'amélogénine des mammifères étaient partiellement, ou entièrement séquencés dans 16 espèces, deux protothériens, un métathérien et 13 euthériens, mais plusieurs taxa sont étroitement apparentés. Par exemple, on connaît l'amélogénine chez quatre rongeurs et six artiodactyles, tous étant des espèces fréquemment employées comme modèles en médecine dentaire. Ces 16 espèces appartiennent à seulement sept ordres, tandis qu'il y a 26 ordres pour un total actuel de 4629 espèces de mammifères (Wilson et Reeder 1993). Un si petit nombre d'espèces examinées n'est pas représentatif des lignées principales de mammifères et elles ne peuvent pas être utilisées pour conclure que la structure primaire de l'amélogénine est conservée chez les mammifères.

Dans cette présente étude, notre objectif principal était de trouver des indicateurs de l'importance de certains domaines de l'amélogénine à travers la conservation de leur structure moléculaire au cours de l'évolution des mammifères. Une telle analyse évolutive a exigé un certain nombre d'espèces représentatives de toutes les lignées principales mammaliennes et cette condition n'a pas été remplie par les espèces disponibles dans les bases de données. Ainsi, nous avons cloné et séquencé le gène de l'amélogénine chez un certain nombre d'espèces pour obtenir des séquences bien distribuées dans la phylogénie des mammifères. Dans ce travail, nous avons voulu nous concentrer principalement sur la région la plus variable du gène, c'est-à-dire l'exon 6. De plus, nous avons analysé les séquences d'amélogénine X et Y disponibles dans les bases de données à la lumière des découvertes récentes sur l'évolution des chromosomes sexuels.

Matériel et méthodes

Source des ADN

Vingt-trois séquences de gènes de l'amélogénine de mammifères (incluant sept gènes de l'amélogénine liés au chromosome Y) ont été obtenues dans GenBank. Elles appartiennent à 16 espèces distribuées dans 7 ordres et 10 familles (Tableau 1A). De plus, nous avons séquencé le gène de l'amélogénine (l'exon 6 et une partie de l'exon 5) à partir d'ADN génomique dans 14 autres espèces, représentant 10 ordres et 14 familles (Tableau 1B).
Nous avons aussi inclus dans notre étude un cobaye (*Cavia porcellus*) pour vérifier une donnée antérieure qui semblait étonnante.

L'ADN génomique d'un phoque gris (*Halichoerus grypus*) et d'un tamanoir (*Tamandua tetradactyla*) ont été fournies par Dr. F. Catzeflis (UMR 5554, Université de Montpellier 2, France).

L'ADN génomique des autres espèces a été isolé à partir de tissus mous (rein, foie, rate et muscle) par extraction avec du phénol-chloroforme ou en employant le kit « Dneasy tissue » (Qiagen). La source de ce matériel était :

L'éléphant d'Asie (*Elephas Maximus*), le fennec (*Vulpes zerda*), le guépard (*Acinonyx jubatus*), l'hippopotame nain (*Hexaprotodon liberiensis*) et le zèbre de Grévy (*Equus grevyi*) ont été fournis sous forme de tissus congelés par Dr. F. Ollivet et le Dr. A. Lécu (Zoo de Vincennes, Muséum national d'Histoire naturelle, France).

Le grand dauphin (*Tursiops truncatus*) a été fourni sous forme de tissus congelés par Dr. W. Dabin (Muséum de la Rochelle, France).

Tableau 1. Gènes de l'amélogénine mammaliennes étudiés. A- 23 séquences (16 espèces) trouvées dans GenBank. *: Séquences d'ADN partielles ; B- Nouvelles données (14 espèces).
X, Y : Gènes de l'amélogénine localisés sur les chromosomes X et Y.

A

Nom	Genre espèce	Ordre	Famille	N° d'accession
Homme (X)	*Homo sapiens*	Primates	Hominidae	M86933
Homme (Y)	*Homo sapiens*	Primates	Hominidae	M86933
Souris	*Mus musculus*	Rodentia	Muridae	D31769
Rat	*Rattus norvegicus*	Rodentia	Muridae	U67130
Hamster	*Mesocricetus auratus*	Rodentia	Muridae	AF005245
Cochon d'inde	*Cavia porcellus*	Rodentia	Muridae	AJ012200
Vache (X)	*Bos taurus*	Cetartiodactyla	Bovidae	M63499
Vache (Y)	*Bos taurus*	Cetartiodactyla	Bovidae	M63500
Bouquetin jap. (X)*	*Capricornis crispus*	Cetartiodactyla	Bovidae	AB073741
Bouquetin jap. (Y1)*	*Capricornis crispus*	Cetartiodactyla	Bovidae	AB073742
Bouquetin jap. (Y2)*	*Capricornis crispus*	Cetartiodactyla	Bovidae	AB073743
Chèvre (X)	**Capra hircus**	Cetartiodactyla	Bovidae	AF215889
Chèvre (Y)	*Capra hircus*	Cetartiodactyla	Bovidae	AF215890
Cerf Sika (X)*	*Cervus nippon*	Cetartiodactyla	Cervidae	AB028027
Cerf Sika (Y)*	*Cervus nippon*	Cetartiodactyla	Cervidae	AB028028
Cochon	*Sus scrofa*	Cetartiodactyla	Suidae	U43406
Cheval (X)	*Equus caballus*	Perissodactyla	Equidae	AB032193
Cheval (Y)	*Equus caballus*	Perissodactyla	Equidae	AB032194
Loup	*Canis lupus*	Carnivora	Canidae	AB080687
Chien	*Canis familiaris*	Carnivora	Canidae	AB080686
Opossum	*Monodelphis domestica*	Marsupiala	Didelphidae	U43407
Échidné Australien	*Tachyglossus aculeatus*	Monotremata	Tachyglossidae	AF095567
Ornithorynque	*Ornithorhynchus anatinus*	Monotremata	Ornithorhynchidae	AF095566

B

Nom	Genre espèce	Ordre	Famille
Lémur volant de Malaisie	*Cynocephalus variegatus*	Dermoptera	Cynocephalidae
Toupaye commun	*Tupaia glis*	Scandentia	Tupaiidae
Grand dauphin	**Tursiops truncatus**	Cetartiodactyla	Delphinidae
Hippopotame nain	*Hexaprotodon liberiensis*	Cetartiodactyla	Hippopotamidae
Zèbre de Grévy	*Equus grevyi*	Perissodactyla	Equidae
Chauve-souris géante	*Cynopterus brachyotis*	Chiroptera	Pteropodidae
Fennec	*Vulpes zerda*	Carnivora	Canidae
Phoque gris	**Halichoerus grypus**	Carnivora	Phocidae
Guépard	*Acinonyx jubatus*	Carnivora	Felidae
Pangolin de Malaisie	*Manis javanica*	Pholidota	Manidae
Hérisson d'Europe de l'ouest	*Erinaceus europaeus*	Insectivora	Erinaceidae
Éléphant d'Asie	*Elephas maximus*	Proboscidea	Elephantidae
Tamanoir	*Tamandua tetradactyla*	Edentata	Myrmecophagidae
Paresseux	*Bradypus infuscatus*	Edentata	Bradypodidae

Le toupaye commun (*Tupaia glis)*, la chauve-souris géante (*Cynopterus brachyotis*), le lémur volant (*Cynocephalus variegatus*), le hérisson européen (*Erinaceus europaeus*), le pangolin (*Manis javanica*), le tamanoir (*Tamandua tetradactyla*) et le paresseux (*Bradypus infuscatus*) ont été fournis sous forme de tissu conservé dans l'éthanol par le Dr. G. Veron (Laboratoire de Zoologie, mammifères et oiseaux, Muséum national d'Histoire naturelle).

Le cobaye, quant à lui, est mort naturellement chez un particulier.

En prenant ensemble des séquences des bases de données et notre nouveau matériel, nous avons pu travailler avec 30 espèces représentatives (14 ordres, 22 familles) des lignées principales reconnues dans les phylogénies récentes des mammifères (Figure 1).

Brève présentation de l'amélogénine chez les mammifères

Chez un certain nombre de mammifères, les gènes de l'amélogénine transcriptionnellement actifs sont localisés sur les deux chromosomes sexuels (par exemple, chez l'homme, la vache, la chèvre, le cheval, ...), tandis qu'il y a seulement une copie sur le chromosome X chez d'autres espèces (par exemple, chez les rongeurs). Le gène de l'amélogénine est généralement connu pour être composé de 7 exons. L'exon 1 n'est pas traduit et l'exon 2 contient le peptide signal. La séquence codante commence à partir de la 12ème base de l'exon 2 par le codon d'initiation (ATG) et se termine à la 6éme base de l'exon 7 par le codon "stop". De plus, deux exons supplémentaires, les exons 8 et 9, existent au moins chez l'homme et les rongeurs (Li *et al.,* 1998). L'amélogénine est soumise un important épissage alternatif. L'exon 4 n'est pas inclus dans les ARNms de la vache, tandis qu'il est parfois trouvé chez l'homme et le porc, et qu'on le rencontre fréquemment chez la souris (Salido *et al.*, 1992; Yamakoshi *et al.*, 1994; Hu *et al.*, 1997, Yuan *et al.*, 1997; 2001). Cet exon n'est pas présent dans les gènes de l'amélogénine liés au chromosome Y de l'homme et de la vache, mais l'existence de deux copies des exons 1, 2, 4 et 7 a été démontrée récemment chez le porc (Hu *et al.,* 1996; 2001). Étant donné que l'exon 4 est inconnu chez la plupart des mammifères, il n'a pas été inclus dans notre analyse. L'exon 6 possède plusieurs sites 5' d'épissage alternatif chez la souris, mais seulement un seul a été trouvé dans les autres espèces (voir revue dans Simmer, 1995; Gibson *et al.*, 1991). Les exons 1 à 5 ont moins de 60 pb tandis que l'exon 6 contient plus de 400 pb ; les introns sont grands (> 1kb) sauf l'intron 5. Les seuls exons accessibles par PCR sur l'ADN génomique sont l'exon 6 et l'exon 5, en profitant de la petite taille de l'intron 5.

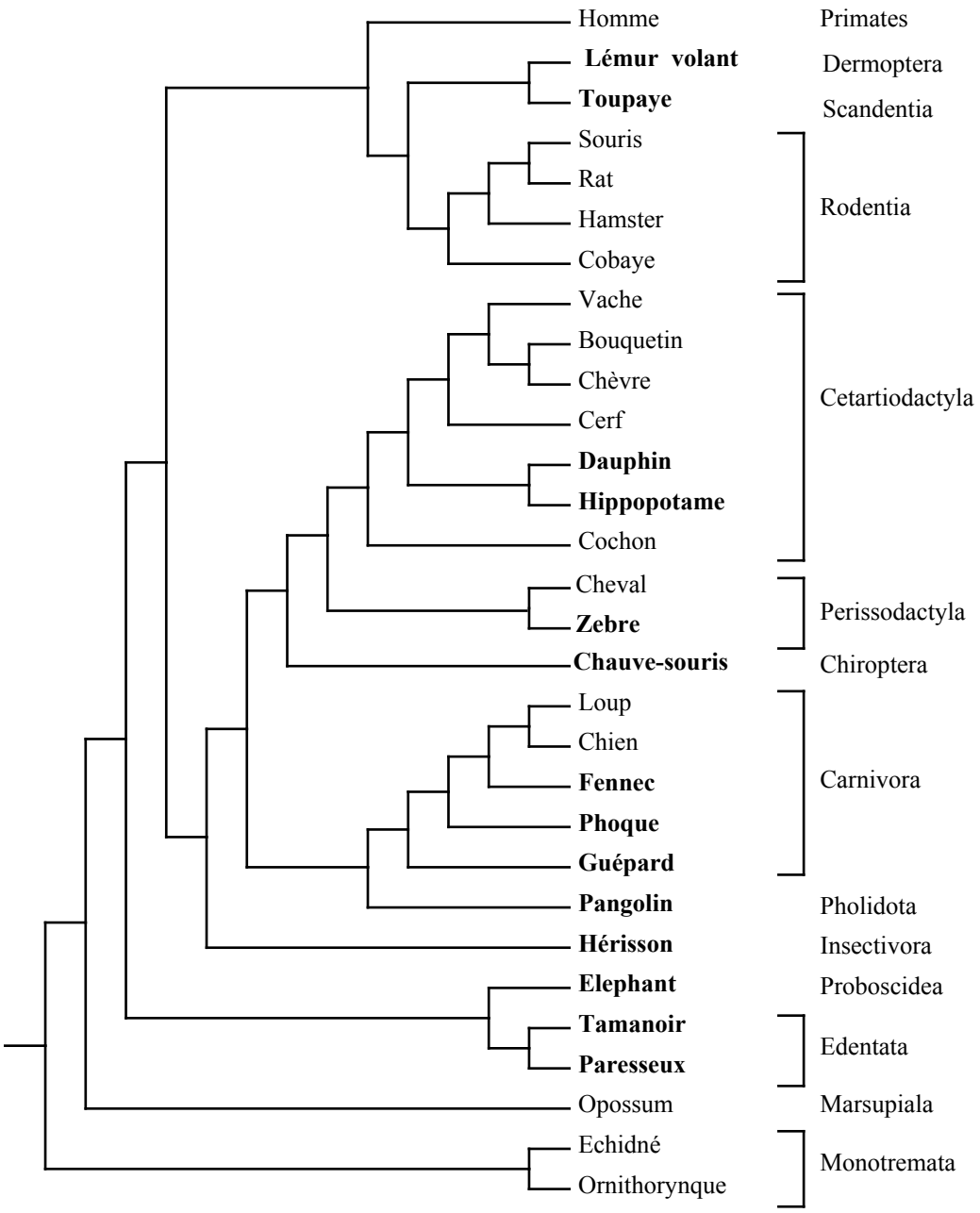


Figure 1. Relations de parenté des 30 espèces de mammifères utilisées dans cette étude. D'après Rowe (1999) et Madsen *et al.* (2001). Mes données sont en gras.

Amplification par PCR

Les gènes de l'amélogénine ont été amplifiés en utilisant les amorces suivantes basées sur une comparaison des séquences nucléotidiques déjà connus des gènes de l'amélogénine des mammifères:

Mam1 (sens ; 5'- TACGAACCATGGGTGGATGGCTGC -3') et

Mam3 (*sens* ; 5'- TACCCTTCCTATGGTTACGAG -3)
ont été conçues pour s'hybrider à la région 5' de l'exon 6.

Mam2 (antisens ; 5'- CACTTCCTCCCGCTTGGTCTT -3') et
Mam4 (*antisens* ; 5'- GCCAAGCTTCCAGAGTCAGAT -3)
ont été conçues pour s'hybrider à la région 3' de l'exon 6.

Mam5 (*sens* ; 5'- ACCCCTCTGAAGTGGTACCAG -3)
a été conçue pour s'hybrider à la région 5' de l'exon 5.

Mam6 (*antisens* ; 5'- GCACGGGAATGATTTGGTGGT -3)
a été conçue pour s'hybrider à la région 5' de l'exon 6.

L'ADN génomique (1 µl) a été amplifié par PCR dans un mélange composé de 5 µl de tampon de Taq (10x) pH 8,8, 3 µl de $MgCl_2$ 2 mM, et 1 µl de dNTP 10 mM, en présence des amorces sens et antisens et de 0,3 µl de polymérase Red Hot (Advanced Biotechnologies Ltd.). L'amplification s'est déroulée dans un thermocycleur (Genius Techne) en 38 cycles, chaque cycle consistant en : 1 minute de dénaturation à 94°C, 1 minute d'hybridation à 59°C et 1 minute de synthèse d'ADN à 72°C. La synthèse finale d'ADN a été réalisée pendant 30 minutes à 72°C.

Clonage

Un microgramme de produit de PCR a été isolé, inséré dans le plasmide vecteur pCR 2.1-TOPO (Invitrogen) par la méthode de TA-cloning, qui a ensuite été utilisée pour transformer les bactéries compétentes E. coli TOP10F. Les bactéries transformées ont été cultivées toute la nuit à 37°C dans un milieu Luria-ampiciline broth, puis elles ont été lysées à 0° C pendant 5 minutes avec 200 µl de NaOH 0,2 M - SDS 1%, puis 150 µl d'AcK 3 M ont été ajoutés à 0°C durant 5 minutes pour précipiter les protéines. Les plasmides ont été purifiés grâce à un mélange de phénol/chloroforme. Le séquençage a été fait par Génome Express S.A.

Analyses moléculaires

Les séquences d'acides aminés ont été traduites à partir des séquences en nucléotide grâce au logiciel DNA Strider 1.2. Les séquences protéiques inférées ont été alignées par rapport à l'amélogénine humain du chromosome X grâce à Clustal X 1.81 (puis l'alignement a été vérifié à la main).

Les séquences nucléotidiques liées au chromosome Y ont été ont été uniquement comparées à leurs homologues liées au chromosome X.

Taux de substitution et richesse en CpG

Les 26 séquences d'ADN retenues ont été alignées grâce à Clustal X. Le taux de substitution moyen a été calculé en employant l'algorithme proposé par Tajima et Nei (1993). Nous avons choisi ce modèle parce qu'il ne postule pas de fréquence à l'équilibre identique pour les quatre bases (ATGC) et l'amélogénine est bien connue pour être riche en GC. Nous avons réalisé ce calcul en employant une fenêtre de 20 nucléotides et un minimum de 10 positions. Le nombre maximal de gap a été fixé au nombre de séquences moins 2 pour

conserver un minium de deux séquences pour le calcul. Le pourcentage de CpG, que l'on sait être lié à un fort taux de substitution (Jamain *et al.*, 2001), a été calculé de la même façon.

Les répétitions dans l'exon 6

Déduction de la séquence ancestrale de l'amélogénine de mammifère

La comparaison des 26 séquences, nous a conduit à envisager l'hypothèse que la majeure partie de la structure primaire de l'exon 6 de l'amélogénine s'est constituée primitivement par l'insertion de plusieurs répétitions de 3 acides aminés (des triplets). On pense que de telles répétitions (9 pb) ont été générées par le glissement de l'ADN polymérase au cours de la division cellulaire (Elder et Turner, 1995; Liao, 1999; Ohta, 2000). Cependant, si certaines répétitions sont évidentes, particulièrement dans la région des insertions de certaines séquences, dans les autres régions, les limites exactes des répétitions sont parfois difficiles à trouver en raison de plusieurs substitutions, délétions ou insertions qui ont eu lieu au cours de l'évolution. Pour tester cette hypothèse, il fallait définir précisément les limites entre les répétitions situées de part et d'autres des régions d'insertion. La séquence ancestrale de mammifère, c'est-à-dire, la séquence d'amélogénine théorique de l'ancêtre de tous les taxa actuels de mammifères, apparaît être la plus appropriée pour tenter de définir ces limites. L'alignement des 26 séquences en acides aminés, qui sont bien représentatives des principales lignées mammaliennes, a permis de déterminer cette séquence ancestrale en utilisant PAUP 4.0 à partir d'une phylogénie récente des mammifères (Madsen *et al.*, 2001).

Tester l'origine des répétitions

Pour comprendre, sur une base statistique, l'origine de ces répétitions, nous avons développé une méthode qui s'appuie sur une fenêtre de 10 répétitions sur la séquence. Dans chaque fenêtre et pour chaque séquence de l'exon 6, le nombre de couples de répétitions avec 0 à 9 bases identiques a été calculé pour deux répétitions adjacentes (a0 à a9) ainsi que pour deux répétitions non-adjacentes (n0 à n9). Quand les distributions (ai) et (ni) n'étaient pas semblables, cela signifiait que deux répétitions adjacentes étaient plus semblables que deux répétitions éloignées; donc, les répétitions adjacentes étaient générées par le glissement de la polymérase.

La similitude pour les distributions (ai) et (ni) est mesurée grâce au test statistique χ^2 : plus le χ^2 était bas, plus (ai) et (ni) étaient semblables. Quand $p<0.05$, deux répétitions adjacentes sont plus semblables que ce qui était attendu sous l'effet du hasard ; une création de répétition est ainsi détectée.

Ce test a été utilisé pour caractériser la différence entre les distributions (ai) et (ni), et non pour évaluer sa signification.

La fréquence de proline et de glutamine

Le pourcentage de P et Q a été calculé dans la région variable de l'exon 6 (aa 46 à 210 dans l'alignement) pour chaque séquence d'amélogénine. Les graphes de P, Q, et la somme P+Q ont été ensuite tracés.

Résultats

Analyse de séquence, alignement et comparaison

Parmi les 23 séquences trouvées dans les bases de données, seulement 12 étaient utilisables pour notre analyse évolutive. En effet, les séquences d'amélogénine du chien, du loup gris et du fennec étant identiques (même au niveau des nucléotides), seule la séquence de fennec a été utilisée. Les séquences du gène de l'amélogénine du chromosome X du cerf et du bouquetin japonais n'ont pas été incluses parce qu'elles étaient trop courtes. Cependant, leurs séquences nucléotidiques ont été utilisées pour l'étude de l'évolution du gène de l'amélogénine du chromosome Y (voir ci-dessous). Les sept séquences du chromosome Y ont été exclues de l'analyse évolutive générale parce que leur taux évolutif était incertain (voir discussion).

L'amplification par PCR de l'ADN génomique de 14 espèces par des combinaisons d'amorces sens et antisens (*Mam1* ou *Mam3 / Mam 2* ou *Mam4*), a produit des fragments compris entre 380 et 420 nucléotides. L'analyse des séquences a révélé que ces fragments étaient homologues à la majeure partie de l'exon 6 du gène de l'amélogénine. De plus, l'amplification par PCR de ce même ADN en utilisant la paire d'amorce *Mam1 / Mam6* a produit des fragments de 350 nucléotides. L'analyse des séquences a révélé qu'ils étaient homologues à la région 3' de l'exon 5, à l'intron 5 et à la région 5' de l'exon 6 du gène de l'amélogénine. Nous avons donc obtenu la majeure partie de la séquence de l'exon 6 dans les 14 espèces. Ces séquences nucléotidiques ont été transformées en séquences d'acides aminés puis elles ont été ajoutées aux séquences primaires liées aux chromosomes X obtenues dans les bases de données. Un total de 26 séquences (12 à partir des bases de données et 14 nouvelles) ont été alignées et comparées (Figure 2).

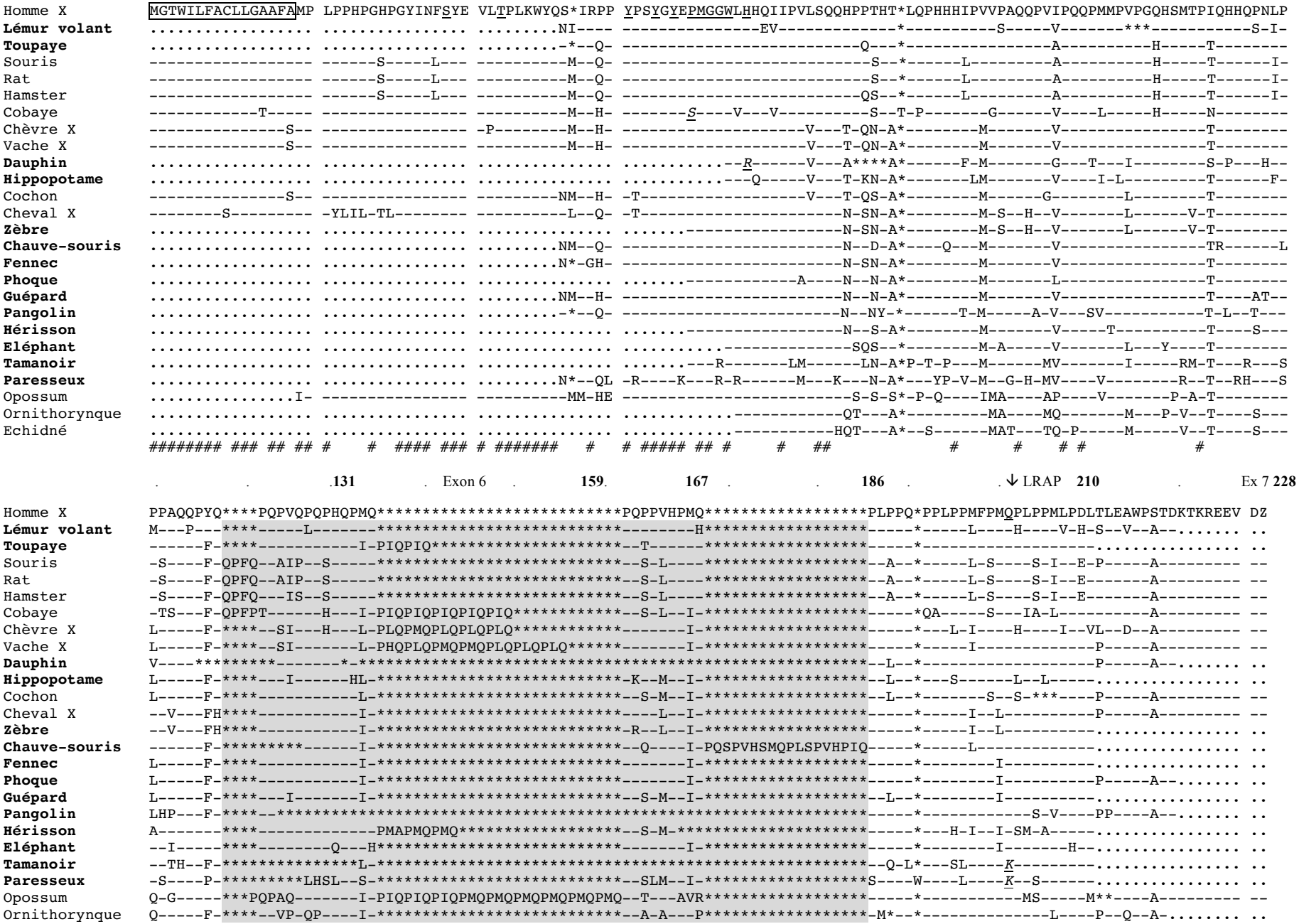
Les acides aminés de l'alignement ont été numérotées à partir de la méthionine N-ter (M1), après clivage du peptide signal de l'exon 2, jusqu'au second codon C-ter de l'exon 7 (Z228) (Figure 2).

Deux résultats principaux se dégagent de la comparaison des séquences en acides aminés (aa). D'abord, les séquences primaires d'espèces étroitement liées sont très semblables. Comme il est indiqué ci-dessus, dans la famille Canidae, les séquences nucléotidiques du chien, du loup et du fennec sont 100% identiques.

Dans la famille Equidae, la similitude nucléotidique entre le cheval et le zèbre est de 99,7% (seulement une substitution). Dans la famille Muridae, la similitude entre les séquences nucléotidiques de souris et de rat est de 99,15% (seulement une substitution). Cependant, cette similitude diminue quand la distance évolutive augmente comme entre la souris et le hamster (92,7%), la souris et le cobaye (81,7%), etc.

Figure 2 (page suivante). Alignement en acides aminés des 26 séquences d'amélogénines analysées. La présentation suit l'ordre de l'arbre phylogénétique (voir Figure 1). Les séquences du cerf Sika et du bouquetin japonais n'ont pas été utilisées parce qu'elles étaient trop courtes (elles sont utilisées dans l'analyse des gènes de l'amélogénine liés au chromosome Y, Figure 6). Les séquences d'amélogénine liées au chromosome Y ne sont pas utilisées à cause d'un taux d'évolution incertain (mais voir la Figure 6 pour la comparaison). Les séquences du chien, du loup et du fennec sont identiques, donc seul le fennec est montré. Les jonctions des exons sont indiquées par un espace. Les résidus remarquables sont soulignés. *S*: sérine substituée à la proline 41 chez le cobaye (cochon d'inde). (.) = résidu inconnu; (-) = identique au résidu de la séquence humaine X; (*) insertion

ou délétion; (#) résidus identiques chez toutes les espèces étudiées. TRAP: les deux sites de protéolyse (double flèche) donne naissance au peptide “Tyrosine Rich Amelogenin Peptide“. LRAP: la localisation du site intra-exonique 3' d'épissage alternatif donne naissance au “Leucine Rich Amelogenin Peptide“. Cadre = peptide signal; partie grisée = région soumise à de nombreuses insertions et délétions.


Homme X
Lémur volant
Toupaye
Souris
Rat
Hamster
Cobaye
Chèvre X
Vache X
Dauphin
Hippopotame
Cochon
Cheval X
Zèbre
Chauve-souris
Fennec
Phoque
Guépard
Pangolin
Hérisson
Eléphant
Tamanoir
Paresseux
Opossum
Ornithorynque
Echidné
131
Exon 6
159
167
186
LRAP
210
Ex 7
228

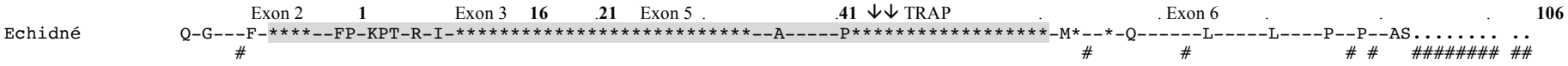


Figure 2. (légende page précédente).

Le second résultat est que la similitude de la séquence primaire varie selon les domaines de l'amélogénine. En effet, les résidus identiques (marqués # sur la Figure 2) sont surtout présents dans les régions hydrophiles (aa 1-45 et aa 211-228), tandis que la région hydrophobe (aa 46-210) est caractérisée par de nombreuses substitutions, des insertions et des délétions (Figure 2).

Un haut degré de conservation de séquences dans les régions hydrophiles

Les résidus conservés sont particulièrement nombreux dans l'exon 2 (83,4%), l'exon 3 (56,3%), l'exon 5 (60%), la région 5' de l'exon 6 (aa 34 à 45 dans l'alignement) (83,4%) et la région 3' des exon 6 et 7 (aa 211 à 228) (66,6%), même dans les lignées basales des mammifères (métathériens et protothériens). On considère ainsi que ces régions de l'amélogénine évoluent à une vitesse lente. Elles contiennent non seulement des résidus "remarquables", qui sont supposés jouer des rôles importants dans la fonction de la protéine, mais aussi plusieurs domaines qui ont été conservés dans toutes les séquences, mais pour lesquelles aucune fonction particulière n'est connue. Par exemple, dans la région C-terminal (aa 211-228) les 10 derniers acides aminés sont inchangés dans les séquences examinées.

Dans la région N-terminal (aa 1 à 45), des résidus remarquablement conservés sont : dans l'exon 3, le site de phosphorylation au niveau de la serine 16 est conservé dans toutes les espèces (100%); dans l'exon 5, la région des hélices alpha (aa 19 à 28) est conservée à 80%; dans la région 5 ' de l'exon 6, le motif tyrosine (aa 33 à 45) qui se lie au N-acétyle-D-glucosamine est conservé à 83,4% avec les trois résidus tyrosine présents dans toutes les séquences examinées. De plus, plusieurs sites de mutations qui sont impliqués dans les différents types d'amélogénèse imparfaite (c'est-à-dire thréonine 21 et proline 41) sont conservés dans toutes les espèces. Cependant, la proline 41 est substituée par une serine dans l'amélogénine du cobaye. Étant donné l'importance de ce site et sa conversation dans les autres espèces examinées, nous avons mis en doute ce résultat et nous avons séquencé le gène de l'amélogénine chez un autre cobaye. Chez cet individu P41 n'est pas substitué par une serine. D'autres résidus importants dans la région 5' de l'exon 6 concernent les sites protéolytiques de la glutamine 43 et du tryptophane 45 qui donne naissance à TRAP43 et TRAP45 (Tyrosine Rich Amelogenin Peptide); on les trouve dans toutes les séquences analysées.

Variabilité et répétitions dans la région hydrophobe

La région variable de l'exon 6 du gène de l'amélogénine (aa 46 à 210) est soumise à de nombreuses substitutions d'acides aminés. En comparant l'alignement des 26 séquences seulement 6% de résidus restent non substitués (Figure 2). Cela suggère que cette région est seulement peu contrainte. La région centrale de l'exon 6 (aa 112 à 188) est aussi caractérisée par de nombreuses délétions et par des insertions (Figure 2).

La majorité de cette région de l'amélogénine semble composée de répétitions de 3 résidus ("des triplets"). Dans la région des insertions, les triplets sont reconnaissables. Il sont surtout constitués par PxQ, c'est-à-dire qu'ils commencent par une proline (P), qu'ils finissent par une glutamine (Q) et que le troisième acide aminé (x) varie dans les différents taxa, mais il est généralement soit une isoleucine (I), une méthionine (M) ou une leucine (L). L'existence de tels triplets explique pourquoi cette région est plus riche en P et Q que les autres régions.

Les répétitions et insertions

Dans les sept séquences possédant des insertions (toupaye, cochon d'inde, chèvre, vache, chauve-souris, hérisson et opossum), la succession des triplets PxQ (Figure 2) est évidente. Dans ces triplets, le nombre des substitutions du résidu central (x) indique que cette position n'est pas aussi importante pour le triplet que les résidus P et Q. Les domaines de l'amélogénine placés des deux côtés de cette région (c'est-à-dire, aa 46-111 et aa 189-210), possèdent aussi des triplets, mais soit ils sont reconnus par le seul motif répété Pxx, soit ils ne sont pas identifiés du tout (Figure 2). Puisque ces séquences peuvent-être facilement alignées, cela suggère que ces triplets ont été insérés dans un ancêtre commun à la lignée mammalienne.

L'existence de répétitions claires de PxQ dans la région des insertions et des délétions suggère qu'elles ont été acquises plus récemment que celles des régions voisines de l'amélogénine. De plus, plusieurs séries indépendantes de création de répétitions sont nécessaires pour expliquer la répartition de triplets fortement semblables trouvés dans les séquences du toupaye, du cobaye, de la chèvre, de la vache, du hérisson et de l'opossum. De plus, puisque ces triplets ne sont pas trouvés dans tous les taxa et puisque que les taxa qui les partagent sont moins liés d'un point de vue évolutif que les taxa qui ne les possèdent pas, nous concluons que les insertions dans l'exon 6 ont été acquises indépendamment dans ces lignées et ne sont donc pas dérivées d'un ancêtre commun.

Pour savoir si les triplets ont été ou pas produits par la copie successive de triplets ("glissement de l'ADN polymérase"), nous avons testé l'origine de ces répétitions sur les 26 séquences des aa 46 à 210. En raison du nombre de substitutions, délétions et insertions qui ont eu lieu indépendamment dans chaque lignée mammalienne conduisant aux taxa actuels, il est souvent difficile de définir les limites entre les triplets. Pour essayer d'obtenir une meilleure définition des limites ces triplets, la séquence ancestrale des mammifères a été déterminée (Figure 3).

Exon 2 1 **Exon 3** 16 21 **Exon 5**
MGTWILFACLLGAAFAMP LPPHPGHPGYINFSYE VLTPLKWYQSMIRHP

TRAP
41 ↓↓ 47 **Exon 6**
YPSYGYEPMGGWLHHQII PVLSQQQT PTHALQ PHHHI PVMAAQQ PMQ

Exon 6
PQQ PMM PM PGQHSVT PTQHHQSNL PQ PAQQ PFQ PQL PQQP PHQ

LRAP
Exon 6 ↓
PIQ PQA PVH PMP PMP PQP PLP PMF PMQ PLP PLL PDL PLE PWP

Exon 6 **Ex 7**
ATDKTKREEV DZ

Figure 3. Séquence ancestrale mammalienne de l'amélogénine obtenue avec PAUP 4.0 à partir de l'alignement des 26 séquences de mammifères. Les jonctions entre les exons sont indiquées par un espace large. L'espace

simple délimite les triplets supposés. Légendes dans la Figure 2.

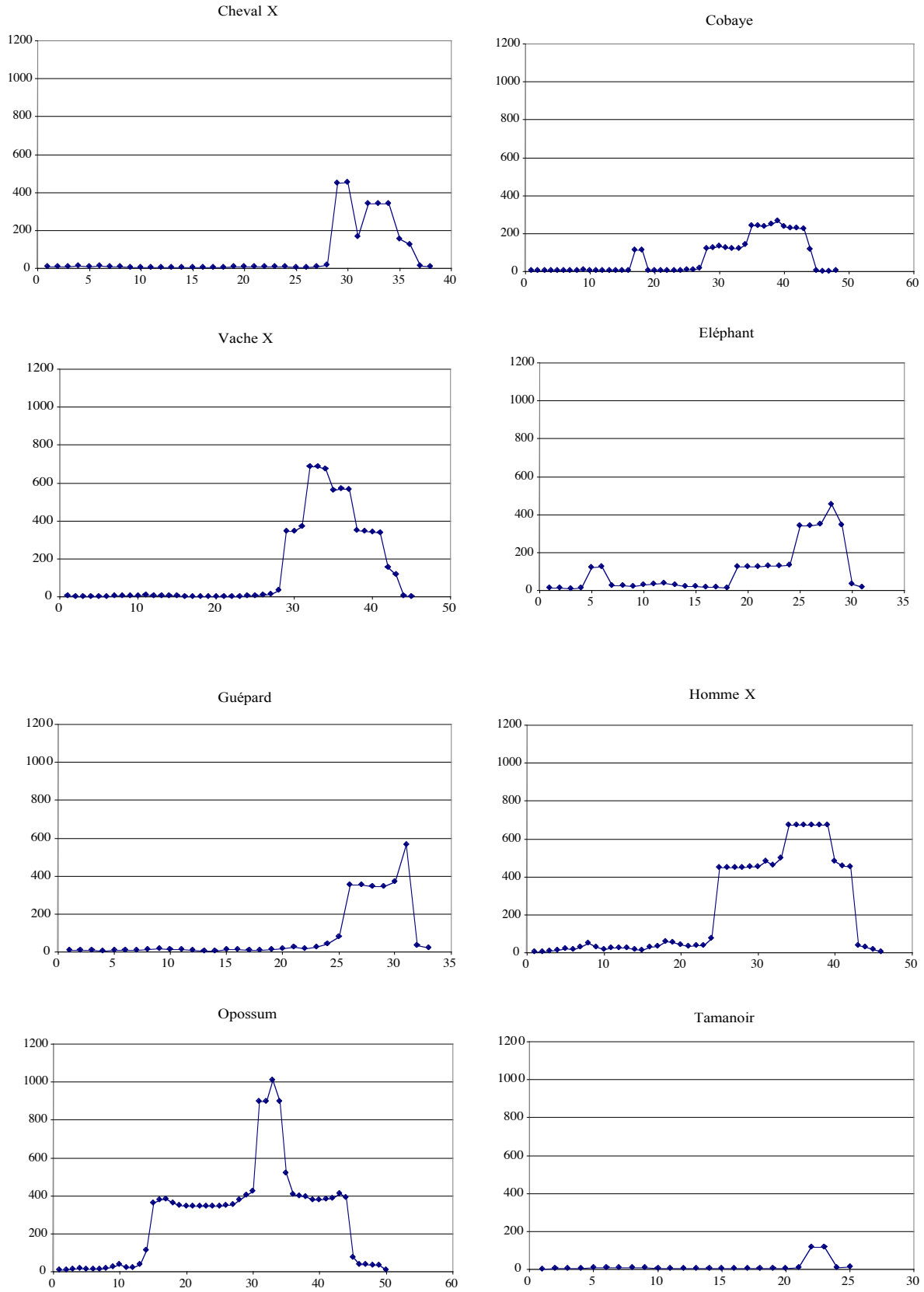


Figure 4. Analyse des triplets de répétitions dans l'exon 6 de certaines séquences d'ADN de mammifères. La valeur du test de χ^2 est comprise entre 0 (non significatif) à 1200 (la plus haute signification). L'axe des abscisses représente la position sur la séquence d'amélogénine : le test analyse une fenêtre de 10 répétitions de 9

nucléotides qui se déplace le long de la séquence d'ADN (par exemple, la position 1 représente l'analyse des répétitions 1 à 10).

La séquence primaire de l'amélogénine de l'ancêtre des mammifères pourrait avoir comporté 177 résidus. Toutes les positions importantes étaient déjà présentes (TRAP, LRAP, …). Les triplets étaient déjà présents, mais ils sont plus ou moins visibles selon les régions car des substitutions se sont déjà produites. Cela suggère que les triplets ont été insérés plus tôt dans l'évolution de l'amélogénine. Les limites des triplets ont été définies à partir de cette séquence ancestrale et utilisées pour définir les limites des triplets dans toutes les séquences de l'alignement. Quelques exemples de cette analyse sont présentés dans la Figure 4. Le test n'a pas réussi à reconnaître des triplets voisins plus semblables que des triplets éloignés dans le domaine 5' de la région variable (aa 46 à 111), mais il a mis en avant une haute fréquence de l'identité dans des triplets voisins ($p < 0.05$) dans la région centrale (aa 112 à 188) avec la plus grande similitude pour la région d'insertion (Figure 4). Ainsi, nous concluons que ces répétitions ont été auto-générées, probablement par un mécanisme de glissement de la polymérase pendant la copie de l'ADN. Le dernier triplet est alors copié plusieurs fois. Dans la séquence de la chauve-souris, l'insertion a été produite par la copie de trois triplets et cela par deux fois.

La région entre les aa 112 et 188 est caractérisée par de nombreuses insertions et délétions. L'observation détaillée du lieu d'insertion indique que toutes les insertions sont placées entre un codon CAG (glutamine) et un codon CCN (proline).

La richesse en P et Q

La région variable de l'exon 6 (aa 46 à 210) est particulièrement riche en proline et en glutamine, cela représente, par exemple, respectivement 33 et 20% dans la séquence d'amélogénine de l'homme, contre 19 et 9% dans les autres régions. Étant donné le nombre de substitutions dans cette région, nous avons vérifié si le pourcentage important de PQ a été conservé dans toutes les espèces étudiées (Figure 5).

Dans toutes les séquences étudiées, le pourcentage de proline est compris entre 25 et 33% (en moyenne, 29,84) et celui de glutamine entre 15 et 23% (en moyenne, 19,57), la somme P+Q varie de 40 à 55% (en moyenne, 49,41). Seulement deux taxa, le paresseux et le tamanoir (édentés), affichent un pourcentage inférieur pour P+Q (40%) comparé à ce qui est observé chez les autres mammifères. Les insertions de triplets (chacun contenant un P et un Q) augmentent le pourcentage P+Q chez la vache et l'opossum par rapport aux valeurs moyennes. Il est clair, cependant, que ces insertions contribuent à augmenter plus le pourcentage des Q que celui des P (Figure 5). Le pourcentage bas de P+Q chez l'édenté est lié aux délétions qui ont intéressé cette région.

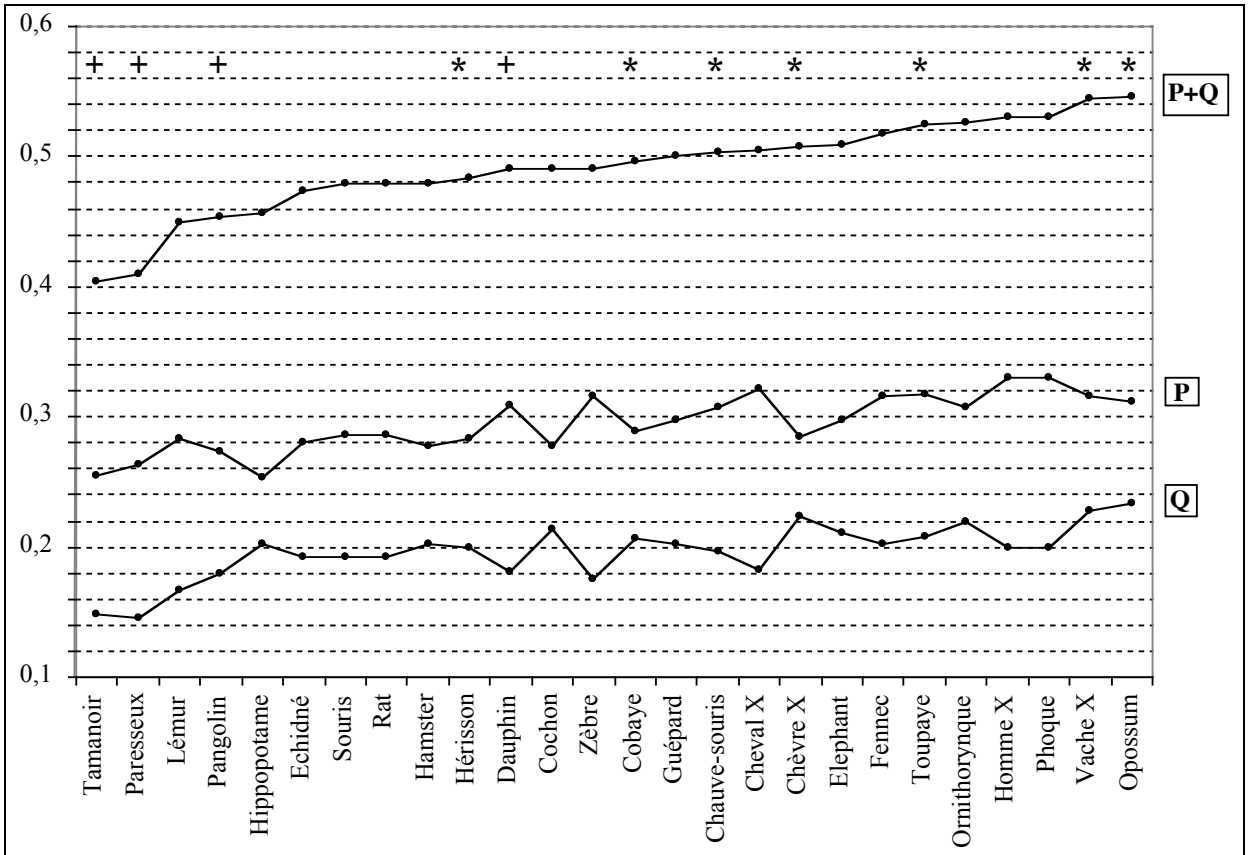


Figure 5. Pourcentage de proline (P) et glutamine (Q), et leur somme (P+Q) dans la région variable de l'exon 6 (aa 46 à 210 de l'alignement) pour les 26 séquences analysées. (*) = présence d'insertions; (+) = présence de délétions. L'axe des abscisses représente les mammifères étudiés classés par ordre croissant de P+Q.

Le site d'épissage de LRAP

Le site d'épissage alternatif exonique, qui est responsable de la présence de fragments d'amélogénine appelés LRAP (Leucin Rich Amelogenin Peptide) est remarquablement conservé chez la plupart des espèces. L'épissage se produit entre les codons codant pour les résidus Q201 et P202 dans l'alignement. Cette glutamine est conservée dans toutes les séquences d'amélogénine des mammifères étudiés sauf chez les deux xénarthres (Edentata) où elle est substituée par une lysine (K), ce qui supprime le site d'épissage chez ces espèces (Figure 2).

Taux de substitution et pourcentage de CpG dans les séquences d'ADN

A partir des 26 séquences d'ADN, le taux de substitution moyen a été calculé pour chaque position nucléotidique depuis la région 5' du peptide signal jusqu'à la région codante 5' de l'exon 7 (678 positions dans l'alignement nucléotidique, excluant les insertions) (Figure 6B). Les grandes valeurs du taux de substitution moyen (> 0,2) ont été trouvées dans l'exon 2, dans l'exon 3, dans la région 5' de l'exon 5 et dans les régions 5 ' et 3 ' de l'exon 6 (Figure 6B).

Le taux de substitution moyen a été comparé à la richesse en CpG pour vérifier si les régions conservées et variables des séquences nucléotidiques de l'amélogénine sont liées ou pas à la présence de paires de nucléotides CG (=CpG) que l'on sait favoriser les substitutions. En effet, une région particulière peut être conservée pendant l'évolution parce qu'elle mute lentement ou parce qu'elle est fonctionnellement contrainte.

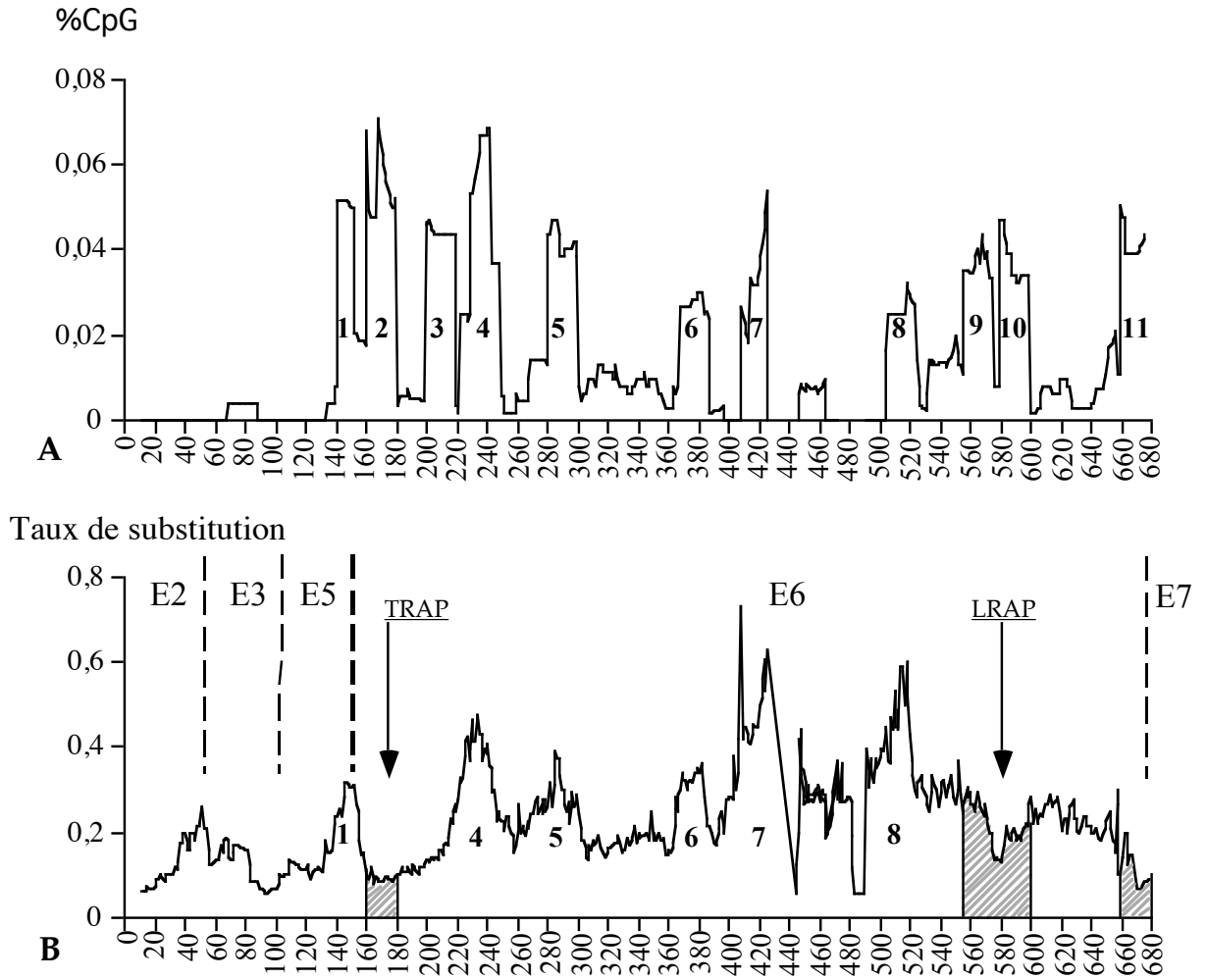


Figure 6. Taux de substitution moyen et pourcentage de CpG dans les 26 séquences d'amélogénine de mammifères, de l'exon 2 à l'exon 7. E2, E3, ... = exon 2, exon 3, ... Les flèches indiquent les pourcentages élevés de CpG liés à un faible taux de substitution. Calculs réalisés avec l'algorithme de Tajima et Nei (1993).

Le pourcentage de CpG a été calculé (Figure 6A) pour les mêmes positions nucléotidiques que dans la Figure 6B. La séquence nucléotidique de l'amélogénine est pauvre en CpG dans l'exon 2, l'exon 3, la région 5' de l'exon 5 et dans quelques régions de l'exon 6. Au contraire, la région 3' de l'exon 5 et la majeure partie de l'exon 6 montrent un fort pourcentage de CpG. La comparaison des deux graphiques révèle à la fois une corrélation entre un petit nombre de CpG et un faible taux de substitution dans quelques régions (et *vice-versa*), ainsi que l'absence de corrélation entre un grand nombre de CpG et un faible taux de substitution dans d'autres régions. Ce dernier résultat indique que certaines régions sont fortement contraintes (par exemple, la région 3' de l'exon 5 et les régions 5' et 3' de l'exon 6). Ce résultat pourrait signifier que les substitutions sont survenues au hasard dans le reste de l'exon 6.

La structure du gène de l'amélogénine du chromosome Y

Les sept séquences codantes des gènes de l'amélogénine liés au chromosome Y (désignés par *AMGY*) retrouvées dans les bases de données (les séquences d'ADN entières de l'homme, de la vache, de la chèvre et du cheval, et les séquences partielles du cerf et du bouquetin japonais; voir Tableau 1A pour les numéros d'accession), ont été alignées et comparées à leurs paralogues *AMGX* (Figure 7). Les séquences *AMGX* et *AMGY* de la chèvre n'ont pas été montrées car elles étaient identiques. Le bouquetin japonais possède deux copies (*AMGY1* et *AMG*Y2) du gène de l'amélogénine sur le chromosome Y. A la fois *AMGY1* et *AMGY*2 sont divergents d'AMGX (respectivement 5% et 3,6% de substitution), mais de façon intéressante la région 5' d'*AMGY1* est plus divergente qu'*AMGY2*, ce qui indique que les deux copies évoluent différemment sur le chromosome Y. Comparé au paralogue *AMGX*, l'*AMGY* humain montre une plus grande variation : 35 substitutions de nucléotides, plus un gap sur un codon (représentant 6% de substitution). Ces substitutions sont survenues dans l'exon 2, l'exon 5 et l'exon 6, et 23 substitutions ont changé les acides aminés. La mutation la plus importante dans l'*AMGY* humain consiste dans la substitution d'une glutamine par une arginine menant à la suppression du site d'épissage alternatif exonique (LRAP) dans l'exon 6. Une telle suppression d'épissage est aussi observée dans la séquence de la vache et dans celle du bouquetin japonais, chez lesquelles le dernier nucléotide a été changé. Cela n'a pas changé la glutamine, mais le troisième codon est crucial pour que l'épissage puisse se produire. L'*AMGY* de la vache montre 11 substitutions de nucléotides dans l'exon 6 (1,89%), parmi lesquels trois ont changé les acides aminés et un a changé le troisième codon de la glutamine au même site d'épissage exonique que chez l'homme. L'*AMGY* de la vache montre aussi une délétion entière de l'insertion nucléotidique qui existe dans *AMGX*. L'*AMGY* du cheval et de la chèvre sont identiques à leurs paralogues *AMGX* excepté une délétion de 24 pb dans l'exon 6 de l'*AMGY* du cheval. Les séquences publiées de l'*AMGX* et l'*AMGY* du cerf sont assez courtes, mais *AMGY* montre 4,6% de substitution.

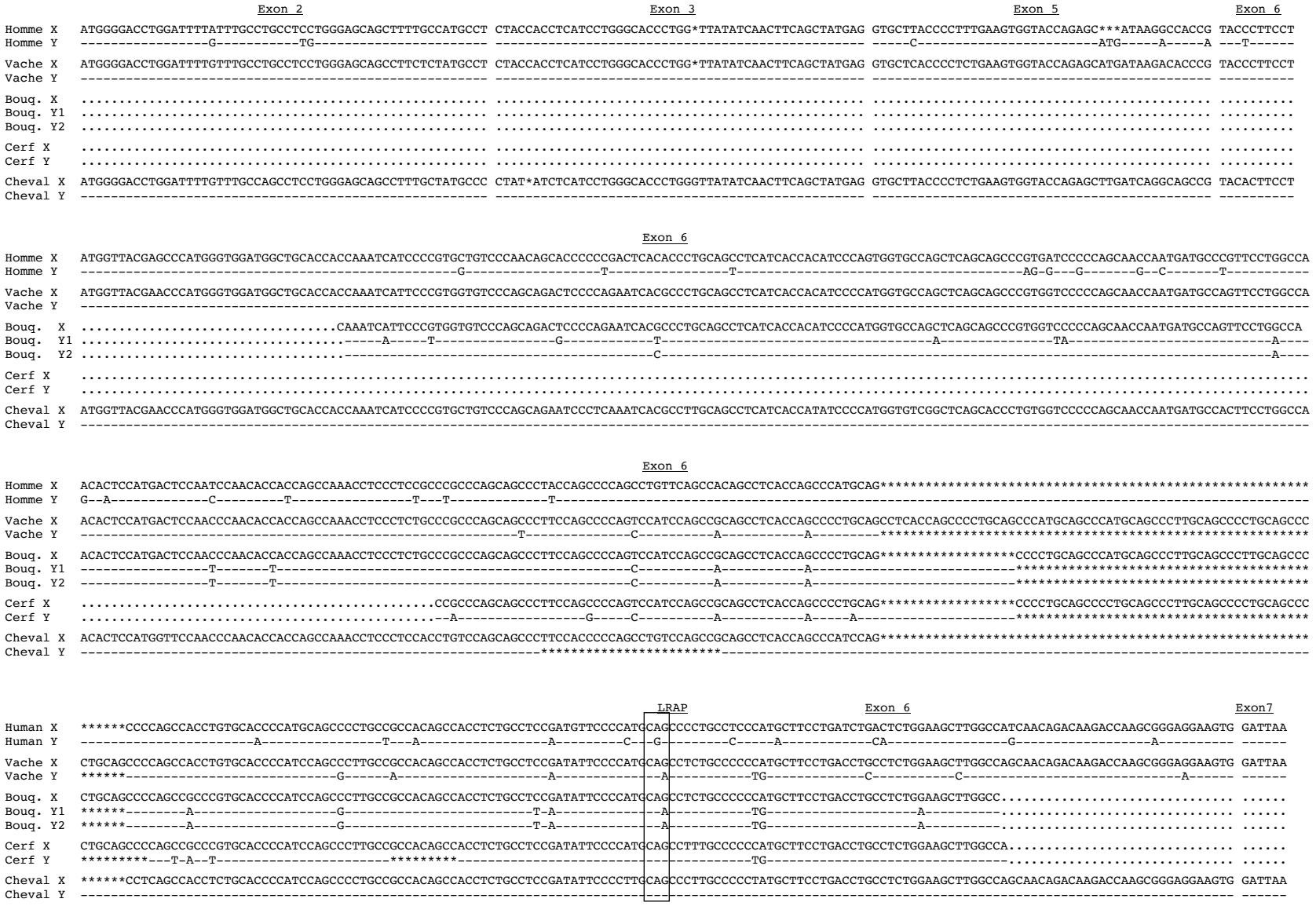


Figure 7. Alignement des séquences nucléotidiques *AMGX* et *AMGY* de mammifères obtenues dans les bases de données (5 espèces).
(-) = nucléotides identiques par rapport à la séquence X; (*) = insertions ou délétions. Bouq. = Bouquetin japonais. Encadré = site d'épissage alternatif.

Le nombre de variations observées dans *AMGY* indique que ces gènes ne sont pas soumis aux mêmes contraintes évolutives que leurs paralogues *AMGX* sachant tout de même qu'il existe plus de mutations dans les lignées germinales. *AMGY* a tendance à accumuler des mutations chez quelques espèces (c'est-à-dire, chez l'homme, la vache, le bouquetin japonais, le cerf), mais pas chez la chèvre, qui est apparentée à la vache, au bouquetin japonais et au cerf, tous des bovidés. Ces variations sont probablement plus liées à l'évolution particulière des chromosomes sexuels qu'à l'amélogénine elle-même (voir la discussion). Par conséquent, en raison de leur taux évolutif incertain, les séquences d'*AMGY* n'ont pas été prises en compte dans l'analyse évolutive.

Discussion

La comparaison des 26 séquences d'amélogénine distribuées dans les principales lignées mammaliennes indique qu'un certain nombre de résidus, situés dans les domaines hydrophiles de l'amélogénine, ont été fortement conservés pendant l'évolution de ce clade. Ces domaines sont localisés sur les régions N- et C- terminales: la région 5 ' riche en tyrosine (exons 2, 3, 5 et les 12 premiers résidus de l'exon 6) et la région 3 ' (les 18 derniers résidus de l'exon 6). Au contraire, la région centrale hydrophobe de l'amélogénine (la majeure partie de l'exon 6) est plus soumise aux variations. Elle se caractérise par un fort pourcentage de proline et de glutamine qui sont inclus dans les triplets répétés d'acides aminés. Cette découverte confirme et précise les conclusions d'études précédentes basées sur un nombre moins important d'espèces (Fincham *et al.*, 1983; Hu *et al.*, 1996; Toyosawa *et al.*, 1998). Cependant, l'ajout de nouvelles séquences d'amélogénine aux séquences déjà disponibles dans les bases de données a permis de travailler avec 30 espèces et de pouvoir, pour la première fois, réaliser une analyse évolutive des amélogénines mammaliennes. L'intérêt principal de cette étude est d'apporter de nouvelles informations sur l'évolution de la région la plus variable, l'exon 6.

Grande conservation des résidus dans les domaines hydrophiles

On peut proposer deux hypothèses pour expliquer la conservation des résidus au cours de l'évolution : une distance évolutive relativement courte entre les lignées ou l'influence de pressions sélectives importantes. La lignée des mammifères trouve son origine il y a environ 300-250 millions d'années. Les protothériens (monotrèmes) se sont séparés des autres mammifères il y a approximativement 200 millions d'années, puis les métathériens (marsupiaux), il y a approximativement 140 millions d'années et finalement les lignées principales euthériennes qui se sont progressivement différenciées il y a 100 millions d'années (voir Easteal, 1989; Bulmer *et al.*, 1991). La forte conservation observée dans les régions hydrophiles de l'amélogénine dans l'ensemble des lignées des mammifères, pourrait donc s'expliquer par l'existence d'importantes pressions sélectives en relation avec les rôles essentiels joués par quelques acides aminés dans la structure et la fonction de l'amélogénine. En effet, ces régions, et particulièrement la région N-terminal, contiennent des acides aminés qui sont soupçonnés être impliqués dans des fonctions diverses : un site de phosphorylation sur la sérine 16 (Fincham *et al.*, 1991), des hélices alpha se formant probablement dans la région des aa 19-28 (Toyosawa *et al.*, 1998) et un motif tyrosine (aa 33-40 qui se lient au N-acétyle-D-glucosamine) pouvant se lier à d'autres glycoprotéines de la matrice de l'émail comme la tufteline, l'énaméline et l'améloblastine (Ravindranath *et al.*, 1999). De plus, on sait que cette région N-ter a des propriétés agrégatives et une forte affinité pour

l'hydroxyapatite (Fincham *et al.*, 1991). Les enzymes protéolytiques coupent la protéine de l'amélogénine en enlevant d'abord la région C-terminal, puis, la région N-terminal, par une séparation entre les aa 43-44 ou entre les aa 45-46 pour produire deux peptides appelés TRAP43 et 45 (Terminal Tyrosine Rich Amelogenin Peptides) (Fincham *et al.*, 1991; Fincham et Moradian-Oldak, 1993; Moradian-Oldak *et al.*, 1994). Ces étapes dans la dégradation de la protéine de l'amélogénine sont supposées être importantes pour la minéralisation normale de l'émail et cela explique pourquoi ces régions sont si fortement conservées dans les 26 séquences mammaliennes examinées. On sait que la région C-terminal bien conservée joue un rôle crucial dans les propriétés agrégatives de l'amélogénine par sa contribution dans la stabilisation des nanosphères (Moradian-Oldak *et al.*, 1994).

Toutes ces fonctions probables pourraient expliquer le haut niveau de contraintes agissant sur la structure primaire de l'amélogénine dans ces domaines. Cette hypothèse est soutenue par la découverte qu'un fort pourcentage en CpG n'est pas lié à un fort taux de substitution dans la séquence nucléotidique de ces régions pour les 26 séquences d'amélogénine analysées. Ces résidus bien conservés sont soumis à une forte pression sélective. Cette similitude des séquences pourrait aussi expliquer la similitude de la structure de l'émail observée chez les mammifères.

Variations dans le domaine hydrophobe

Par opposition aux régions hydrophiles N- et C-terminales, la variabilité observée dans la région hydrophobe indique que la région centrale de l'exon 6 n'est pas aussi structurellement-fonctionnellement contrainte comme les autres régions du gène. Cette hypothèse est soutenue par le taux de substitution moyen qui varie selon le pourcentage de CpG. Cependant, la tendance générale de cette région (située au milieu de la protéine) est d'être hydrophobe et les résidus substitués se conforment à cette règle. La variabilité observée dans cette région entre les différentes espèces étudiées pourrait s'expliquer par la distance évolutive qui existe entre les lignées des monotrèmes, des marsupiaux et des euthériens. Le nombre de substitutions observées dans cette région peut être en effet lié au temps de divergence. Les substitutions sont évidemment moins nombreuses quand on compare l'amélogénine d'espèces relativement proches d'un point de vue évolutif plutôt que celle d'espèces éloignées les unes des autres. Par exemple, le loup (*Canis lupus*) et le chien (*Canis familiaris*) se sont séparés il y a 100 000 ans (Wayne *et al.*, 1987; Vila *et al.*, 1997) et aucune différence n'a été trouvée dans les séquences nucléotidiques. Cela enlève, par exemple la possibilité d'identifier à l'aide de l'amélogénine la peau de chien de celle de loup. De façon intéressante, chez le fennec (*Vulpes zerda*), qui appartient aussi à la famille des Canidae, la séquence du gène de l'amélogénine n'a pas montré de différence avec celle du chien et du loup, même au niveau des nucléotides. Un certain nombre d'exemples similaires peuvent illustrer cette règle, comme par exemple les rongeurs, chez lesquels la souris et le rat se sont séparés il y a approximativement 10 millions d'années, et qui montrent seulement une substitution de nucléotide dans la région variable de l'exon 6. La même identité de séquence (une substitution nucléotidique) a été observée entre le cheval et le zèbre, deux Equidae. Au contraire, le nombre de substitutions augmente en comparant des espèces de différentes familles, puis quand on observe différents ordres de mammifères, les plus grandes variations étant observées en comparant la séquence d'amélogénine des protothériens à celle des euthériens.

L'analyse détaillée de la région variable de l'exon 6 dans les 26 séquences indique cependant quelques particularités, c'est-à-dire, les répétitions de triplets, un fort pourcentage de résidus proline et glutamine, qui méritent davantage d'attention.

Les répétitions de triplets

La région centrale de l'exon 6 a révélé des répétitions composées d'un nombre variable de triplets, qui ont été insérés récemment et indépendamment dans plusieurs taxa. Ces insertions ont été auto-générées par un mécanisme de glissement de l'ADN polymérase pendant la copie de l'ADN génomique (Elder et Turner, 1995; Liao, 1999; Ohta, 2000). Trois résidus de la séquence, du côté 5' précédant les répétitions, ont été copiés plusieurs fois (par exemple 2 fois chez le toupaye et 9 fois chez l'opossum). Dans la séquence de la chauve-souris, l'insertion s'est produite par deux répétitions de la même succession de trois triplets. Aussi, dans d'autres taxa, des délétions d'un nombre variable de résidus se sont produites dans la même région de l'exon 6. Cela indique que la région est favorable à de telles insertions et délétions. La présence de répétitions de triplets dans la région située en amont du locus aurait pu permettre à la nouvelle insertion de se produire (Stephan, 1989). Dans tous les cas d'insertion observée, les copies de triplets ont été insérées entre les nucléotides codants pour Q et P, deux résidus qui sont fortement présents dans cette région de l'exon 6. Ainsi, la paire QP semble être attractive pour les insertions. Puisque ces triplets PxQ ont été insérés récemment et indépendamment dans diverses lignées mammaliennes, nous proposons que les résidus PMQ, PIQ et PLQ sont à la base d'insertion de triplets dans ces régions de l'amélogénine des mammifères. La comparaison avec les séquences d'amélogénine du crocodile (Toyosawa *et al.*, 1998) et du serpent (Ishiyama *et al.*, 1998) indique que, contrairement aux mammifères, l'amélogénine reptilienne n'a pas de triplets répétés qui auraient été récemment insérés dans l'exon 6.

Dans la combinaison des trois résidus PxQ récemment insérés, les substitutions intéressent surtout le résidu central (x), moins la glutamine (Q), alors que le premier résidu, proline (P), est souvent conservé. Cette particularité a permis de définir les régions de l'exon 6 situées des deux côtés des insertions composées aussi de répétitions de triplets. L'analyse du χ^2 des répétitions a confirmé l'existence de tels triplets, qui ont probablement été générés par un même mécanisme de glissement de l'ADN polymérase, ce que l'on a déjà montré pour les insertions. Cependant, l'origine de ces triplets reste à découvrir soit dans un ancêtre de la lignée mammalienne, soit, peut être, plus précocement dans l'histoire de l'amélogénine. Ces répétitions de triplets sont principalement reconnues par le signal Pxx. Si ces triplets ont été générés de la même façon que les triplets récents, nous pouvons formuler l'hypothèse qu'ils proviennent d'un motif PxQ ancestral transformé en Pxx après substitution de Q. De plus, des signaux Pxx similaires semblent exister dans la région qui suit immédiatement le TRAP, mais, étant donné le nombre élevé de substitutions, le test n'a pas permis de les reconnaître comme des répétitions de triplets.

Notre hypothèse concernant l'origine de la région centrale de l'exon 6 de l'amélogénine, au moyen d'insertions de répétition PxQ, implique qu'il y a eu plusieurs événements d'insertions pendant l'évolution, le dernier s'étant produit récemment dans un nombre limité de lignées de mammifères. Étant donné que cette région de l'exon 6 est soumise à plus de substitutions que les autres régions, les 600 millions d'années d'évolution de l'amélogénine (Delgado *et al.*, 2001) ont été suffisantes pour que de nombreuses substitutions aient eu lieu (PxQ $\Rightarrow$ Pxx $\Rightarrow$ xxx). Cela pourrait expliquer la difficulté rencontrée pour reconnaître les triplets quand la copie s'est produite il y a longtemps.

Dans l'exon 6 de l'amélogénine du crocodile et du serpent, la région centrale contient des signatures de triplets Pxx. Cela soutient notre hypothèse que 1) précocement dans l'évolution des vertébrés une grande partie de l'exon 6 a été générée à la suite des insertions de répétition de triplets et 2) que pendant l'évolution des principales lignées de vertébrés, ces triplets insérés ont été soumis à des substitutions qui les ont rendus difficile à reconnaître en tant que tels.

Jusqu'à présent, nous ne connaissons pas la fonction exacte de ces répétitions et on ne signale pas de mutation, dans cette région, conduisant à des défauts de l'émail. Cependant, Williamson (1994) a proposé que les répétitions de triplets pouvaient jouer un rôle, même secondaire, dans la conservation d'un nombre important de deux acides aminés, la proline et la glutamine, pour faciliter les interactions protéine-protéine.

Le pourcentage élevé de proline et de glutamine

Dans les 26 séquences examinées, la région centrale de l'exon 6 est caractérisée par un fort pourcentage de proline (P=29,84%) et de glutamine (Q=19,57%) comparé aux autres résidus, même dans les séquences qui ne possèdent pas d'insertion de triplets. Une telle richesse n'est pas surprenante si l'on considère l'existence de répétitions de triplets PxQ. Une telle quantité de P et de Q soutient aussi notre hypothèse sur l'origine de l'ensemble de la région centrale de l'amélogénine de l'exon 6 comme étant primitivement composée de telles répétitions. En partant avec 33% de P et de Q (répétitions de PxQ), il apparaît que le pourcentage de P a été plutôt bien conservé pendant l'évolution tandis que celui de Q a diminué. Cela est probablement à rapprocher d'une sélectivité dans les substitutions comme il a été noté ci-dessus, P étant moins sensible que Q. Cela pourrait aussi expliquer le signal durable Pxx dans cette région. L'insertion de répétitions de triplets dans la région centrale hydrophobe de l'amélogénine pourrait contribuer à maintenir un niveau élevé de proline et, à un moindre niveau, de glutamine, et pourrait donc faciliter les interactions avec d'autres protéines de l'émail (tufteline, améloblastine et énaméline). Cette quantité de P et Q pourrait améliorer la résistance de l'émail. En effet, il ne faut pas oublier que les mammifères ont perdu la polyphyodontie il y a environ 300-250 millions d'années et que l'émail doit protéger les dents (principalement par résistance à la rupture et à l'abrasion) pendant une période plus longue que, par exemple, les reptiles, chez lesquels les dents sont remplacées plusieurs fois pendant la vie. Étant donné que les dents ne sont pas remplacées, une plus forte pression sélective que chez les reptiles pourrait agir sur la structure de l'émail, et l'insertion de répétitions dans la région centrale de l'amélogénine pourrait soutenir cette hypothèse.

Conservation de la séquence et amélogenèse imparfaite

Parmi les résidus conservés dans la région N-ter, on sait que la substitution de deux acides aminés, thréonine 21 par isoleucine (dans l'exon 5) ou proline 41 par thréonine (au début de l'exon 6), conduit à une amélogenèse imparfaite, un défaut qualitatif et quantitatif de l'émail chez l'homme (Lench et Winter, 1994; Collier *et al.*, 1996; Ravindranath *et al.*, 1999). La proline 41 est présente dans toutes les séquences de mammifères examinées sauf chez le cobaye, un rongeur, chez lequel elle est substituée par une sérine (GenBank, numéro d'accession AJ012200). La participation probable de la substitution de P41 dans l'amélogénèse imparfaite chez homme et la conservation de P41 dans toutes les séquences mammaliennes examinées et même dans les séquences du crocodile, de la grenouille (Toyosawa *et al.*, 1998) et du serpent (Ishiyama *et al.*, 1998) supporte un rôle important possible de ce résidu pour la structure et/ou la fonction de l'amélogénine. La substitution dans la séquence publiée du cochon d'inde paraissait donc douteuse. Le séquençage de l'amélogénine d'un autre spécimen nous a permis de répondre à cette question : P41 n'est pas substituée. Donc, soit la substitution rapportée précédemment était une erreur pendant le séquençage, soit le cobaye souffrait d'une amélogenèse imparfaite. Puisque aucun défaut n'a été signalé dans la structure de l'émail de ce rongeur, nous pensons que cette substitution était

une erreur de séquençage. Quelques cas seulement de défaut de l'amélogénine ont été signalés jusqu'à présent chez l'homme, mais étant donné la forte conservation de quelques résidus dans toutes les lignées de mammifères étudiées, on peut imaginer que de nouveaux cas d'amélogenèse imparfaite seront signalés dans un avenir proche. À notre connaissance, l'amélogenèse imparfaite n'a pas été signalée jusqu'à présent dans d'autres espèces de mammifères. Cependant, étant donné les conséquences importantes dans la structure de l'émail de tels défauts, l'amélogenèse imparfaite serait probablement létale à court terme dans la nature.

Conservation de séquence et épissage alternatif de l'ARN

L'épissage alternatif, qui est la suppression de régions codantes pendant l'épissage des introns, est un mécanisme essentiel dans la fonction de l'amélogénine. On connaît ces phénomènes chez la vache (Gibson *et al.*, 1991), l'homme (Salido *et al.*, 1992), le porc (Yamakoshi *et al.*, 1994), les rongeurs (Simmer *et al.*, 1994; Bonnas *et al.*, 1994; Li *et al.*, 1995) et l'opossum (Hu *et al.*, 1996). Donc, il pourrait aussi exister dans d'autres espèces étudiées. De multiples ARNms sont synthétisées à partir d'un gène unique et le modèle d'épissage alternatif est différent pour chaque gène d'amélogénine. Même les dimensions des fragments épissés changent pendant la formation de la dent comme cela a été montré chez la souris (Couwenhoven et Snead, 1994; revue dans Simmer, 1995). Le principal site d'épissage alternatif connu jusqu'à maintenant est situé sur la région 3' de l'exon 6. Chez la vache, il produit des fragments LRAP (Leucin Rich Amelogenin Peptide). Chez la souris, le rat, la vache et le porc, le segment LRAP correspond aux exons 3, 5 et aux 25 derniers aa de l'exon 6. Les protéines LRAP ont été trouvées chez un certain nombre d'espèces mammaliennes étudiées sauf chez l'homme. L'épissage de l'exon 6 (c'est-à-dire, la suppression de la fin 5' de ce grand exon) est obtenu grâce à un site accepteur d'épissage alternatif 3' intra-exonique. Ce site est identifié comme un signal d'épissage, comme celui qui est situé aux frontières introns/exons. La séquence des introns commence toujours par le dinucléotide GT- et se termine par le dinucléotide -AG. Au niveau de la séquence LRAP de mammifère, le site (codé par PM**QP**LP) est reconnu comme étant un accepteur d'épissage alternatif intra-exonique 3'. En effet, la glutamine est codée par CAG dans tous les cas sauf dans le gène de l'amélogénine lié au chromosome Y de la vache, dans lequel il est codé par CAA et dans le gène de l'amélogénine lié au chromosome Y de l'homme, dans lequel elle est substituée par de l'arginine. Cela confirme les résultats précédents selon lesquels le chromosome Y de la vache ne subit pas d'épissage, mais ne répond pas à la question de l'absence de fragments LRAP dans la protéine de l'amélogénine humaine, puisque le codon CAG est bien présent (Simmer, 1994; Gibson *et al.*, 1994).

Notre analyse évolutive nous montre que le résidu glutamine précédant le site d'épissage est fortement conservé dans l'amélogénine mammalienne (sauf chez les xénarthres) et elle confirme le rôle essentiel que joue le dinucléotide AG dans l'épissage alternatif. Les particularités observées dans les séquences d'amélogénine des deux xénarthres (paresseux et tamanoir) sont peut-être à mettre en relation avec l'absence probable de dents ou d'émail (voir chapitre 4 "Discussion conclusion générales", *§III*). Cependant, un certain nombre de résidus de glutamine codés CAG pourraient être de potentiels sites accepteurs d'épissage 3', particulièrement ceux qui sont conservés dans toutes les séquences de mammifères étudiées.

L'évolution des gènes de l'amélogénine liés au chromosome Y

La plupart des euthériens possèdent une copie du gène de l'amélogénine sur les deux chromosomes sexuels X et Y (désignés par *AMGX* et *AMGY*). Au contraire, le gène de l'amélogénine est autosomique chez les protothériens (monotrèmes) et les métathériens (marsupiaux) bien qu'ils aient des chromosomes X et Y (Watson *et al.*, 1992; Sasaki et Shimokawa, 1995; Quintana-Murci *et al.*, 2001). Dans les deux lignées, l'hybridation par southern blot indique la présence de loci multiples d'amélogénine (Watson *et al.*, 1992). Mais des séquences d'amélogénine de deux loci ne sont disponibles ni chez un monotrème, l'ornithorynque (Toyosawa *et al.*, 1998) ni chez un marsupial, l'opossum (Hu *et al.*, 1996). Les chromosomes sexuels ont une origine indépendante dans de nombreuses classes de vertébrés et ainsi, la localisation chromosomique des gènes de l'amélogénine n'est pas attendue sur les chromosomes sexuels. De plus, deux gènes de l'amélogénine dérivés de deux loci sont présents dans le génome du xénope (un amphibien connu comme tétraploïde) (Toyosawa *et al.*, 1998). Les deux séquences montrent de nombreuses différences (des substitutions, des délétions et des insertions) qui s'expliquent soit par une duplication ancienne du génome soit par une divergence rapide, analogue à ce qui est observé pour l'*AMGX* et l'*AMGY* de l'homme.

Dans une étude précédente, nous avons montré que les relations phylogénétiques d'*AMGX* et d'*AMGY* sont différentes de ce que l'on pourrait attendre d'après l'évolution des chromosomes sexuels (Girondot et Sire, 1998). Bien que les chromosomes Y des euthériens soient homologues et aient divergé après la séparation des métathériens (Charlesworth, 1991; Quintana-Murci *et al.*, 2001), ils ne se sont pas placé dans le même groupe, avec par exemple, *AMGX* d'un côté et *AMGY* de l'autre. Le regroupement était en désaccord avec l'origine chromosomique du gène de l'amélogénine. Cette découverte nous a amené à tester et à proposer l'hypothèse que plusieurs duplications indépendantes du gène de l'amélogénine sur les deux chromosomes sexuels avaient eu lieu pendant l'évolution mammalienne (Girondot et Sire, 1998). La même conclusion a été obtenue par Toyosawa *et al.* (1998). Cependant, la publication récente des séquences *AMGY* chez le cheval (Hasegawa *et al.*, 2000) et la chèvre (soumission électronique 2001: tableau 1A) et la séquence partielle de l'ADN du cerf japonais (Yamauchi *et al.*, 2000) ont changé cette hypothèse. La comparaison entre ces séquences d'*AMGY* et leur paralogue *AMGX* indique clairement qu'*AMGY* a évolué différemment dans les cinq espèces de mammifères étudiées : de grandes différences sont observées chez l'homme, alors que quelques variations existent seulement chez la vache et le cerf, et qu'aucune différence n'a été relevée dans les séquences codantes AMGX et AMGY du cheval et de la chèvre. Ces variations sont aussi en désaccord avec les relations phylogénétiques entre les lignées mammaliennes : le bœuf, la chèvre et le cerf sont des bovidés artiodactyles tandis que le cheval est un équidé périssodactyle, deux clades éloignés dans les phylogénies récentes des mammifères (Madsen *et al.*, 2001). De plus, *l'AMGY* de la vache est plus différente de son paralogue *AMGX* que le sont l'*AMGX* de la vache et l'*AMGX* de la chèvre. Ces observations nous poussent à reconsidérer notre hypothèse sur les duplications indépendantes du gène de l'amélogénine sur les deux chromosomes sexuels pendant l'évolution mammalienne.

Les études récentes sur l'évolution des chromosomes sexuels chez les mammifères pourraient résoudre le problème de compréhension de l'évolution de l'*AMGX* et de l'*AMGY* (Lahn *et al.*, 2001; Quintana-Murci *et al.*, 2001; Iwase *et al.*, 2001). Brièvement, les chromosomes X et Y des mammifères se sont développés à partir d'une paire de chromosomes autosomiques d'un ancêtre commun, de type "reptile-mammalien" (un thérapside). Cette évolution s'est déroulée en plusieurs étapes. Il y a eu au moins quatre inversions multi-gèniques successives dans le chromosome Y, responsables de recombinaisons restreintes de

X-Y. Chaque inversion était ensuite suivie par la différenciation de nouvelles régions non-recombinantes, puis par une attrition autosomique progressive (Lahn et Page, 1999). La première inversion a eu lieu il y a 300 millions d'années, avant la séparation des protothériens. La recombinaison X-Y a été stoppée dans la région inversée (appelé strate 1), et la différenciation de cette strate a commencé dans les deux chromosomes sexuels. La deuxième inversion s'est produite il y a ~150 millions d'années (strate 2), avant la séparation des métathériens. La troisième inversion, ~100 millions d'années (strate 3), a eu lieu avant la séparation des principales lignées euthériennes et la quatrième inversion ~30-50 millions d'années (strate 4) est probablement survenue indépendamment dans chaque lignée euthérienne (Jegalian et Page, 1998; Lahn et Page, 1999; Lahn *et al.*, 2001; Quintana-Murci *et al.*, 2001). Chez les primates (homme et singes), Iwase *et al.* (2001) ont déduit, à partir de différences dans les séquences d'ADN de l'amélogénine (sur des sites silencieux), qu'*AMGX* est placé à la limite de la strate 3 et de la strate 4 et que la jonction 3-4 est proche du début du gène de l'amélogénine. Cette observation est déduite du fait que le promoteur, l'exon 1 et la partie amont de l'intron 2 d'*AMGX* présentent plus de différences que la partie localisée en aval de l'intron 2 et le reste du gène. Dans les régions proches des inversions, on observe une forte diminution de la recombinaison. L'inversion 4 ayant eu lieu près du gène de l'amélogénine, la zone d'inhibition de recombinaison s'est étendue jusqu'au milieu de l'intron 2.

L'hypothèse d'une duplication indépendante du gène de l'amélogénine sur les deux chromosomes sexuels pendant l'évolution mammalienne est donc caduque. Il y a eu une seule origine du gène de l'amélogénine sur les chromosomes X et Y dans la lignée euthérienne, quelque part après la divergence des métathériens. Puis, tout à fait récemment (30-50 millions d'années), *AMGX* et *AMGY* ont commencé à se différencier après que la strate 4 du chromosome Y se soit inversée. Notre étude comparative montre que cette différenciation d'*AMGX* et *AMGY* est indépendante chez les cinq mammifères étudiés. Les loci *AMGY* sont devenus non-recombinants indépendamment, c'est-à-dire, après la divergence entre les périssodactyles, les artiodactyles et les primates. Puis, la différenciation s'est déroulée différemment dans chaque lignée même chez les espèces étroitement liées comme la chèvre, le cerf et la vache. La différenciation a commencé relativement tôt chez l'homme, pour lequel *AMGX* et *AMGY* ont probablement arrêté de se recombiner il y a 30-50 millions d'années, tandis qu'elle a commencé plus récemment chez la vache et le cerf et ne s'est pas encore produite chez le cheval et la chèvre (au moins pour les séquences codantes). Cette hypothèse doit être confirmée quand les séquences d'ADN génomiques seront disponibles chez d'autres mammifères et particulièrement chez le cheval et la chèvre.

Lahn et Page (1999) ont remarqué que la période de 30-50 millions d'années de non-recombinaison est un temps suffisant pour la dégénérescence d'*AMGY* (parce qu'une telle dégénérescence est arrivée pour tous les autres homologues Y des gènes X de l'homme). Cela est en accord avec la dégénérescence génétique du chromosome Y (voir les revues récentes dans Charlesworth et Charlesworth, 2000; Lahn *et al.*, 2001). Dans leur récente revue Lahn *et al.* (2001) concluent que la divergence de séquences chez l'homme est *« plus en accord avec une évolution différentielle adaptative des protéines des homologues qu'une hasardeuse persistance d'une protéine fonctionnellement non-contrainte»*. Considérant que le site d'épissage exonique dans l'exon 6 a été éliminé dans l'*AMGY* de l'homme et de la vache et que les produits d'épissage d'*AMGX* (LRAP) sont probablement impliqués, comme signaux régulateurs, dans des stades précoces de l'éruption des dents, ils postulent qu'*AMGY* a divergé en fonction d'*AMGX* de manière bénéfique aux mâles grâce à un retard dans la dentition. Mais cette contrainte fonctionnelle a-t-elle été sélectionnée dans d'autres lignées mammaliennes étant donné que la recombinaison s'est probablement arrêtée indépendamment ?

Les rongeurs sont exceptionnels dans le fait que seul *AMGX* a été identifié dans les séquences d'ADN. *AMGY* a été soit supprimé soit altéré à un niveau indétectable (Snead *et al.*, 1985; Lau *et al.*, 1989; Bonass *et al.*, 1994). La dernière hypothèse pourrait s'expliquer par une différentiation plus précoce ou plus accélérée de la strate 4 dans la lignée des rongeurs. Chez l'homme, les défauts structuraux qui mènent à plusieurs amélogenèses imparfaites concernent seulement l'*AMGX* (défauts liés aux chromosomes X) (Aldred *et al.*, 1992; Lench *et al.*, 1994; Collier *et al.*, 1997).

Conclusion

Au moyen de la comparaison de séquences d'amélogénine dans 26 espèces représentatives des lignées principales des mammifères, nous avons montré que les régions terminales N- et C- de la protéine ont évolué lentement. Elles sont caractérisées par de nombreux résidus fortement conservés. Au contraire, la région centrale (exon 6) a évolué plus rapidement et montre de nombreuses substitutions qui pourraient représenter un fort signal phylogénétique. Cette région est caractérisée par l'insertion de répétitions de triplets et par un fort pourcentage de proline et de glutamine. Ces insertions, aussi bien que la richesse en P et en Q, sont propres à la lignée mammalienne. Elles ont probablement été sélectionnées pendant l'évolution mammalienne pour renforcer la résistance de l'émail des dents en raison de la perte du remplacement des dents dans la lignée mammalienne. Cependant, évolutivement parlant, les mammifères ne sont pas d'une grande aide pour étudier l'évolution des régions de l'amélogénine qui sont soumises à de fortes pressions sélectives. En effet, les lignées principales mammaliennes se sont séparées dans une période de temps assez court et cela a abouti à une grande similitude des différents gènes de l'amélogénine. Pour entrer en profondeur dans l'histoire de l'amélogénine et pour améliorer notre connaissance sur ses régions différentes, nous devons étudier l'évolution de cette protéine dans d'autres lignées de vertébrés, particulièrement chez les reptiles et les amphibiens.

ÉVOLUTION MOLÉCULAIRE DE L'AMÉLOGÉNINE CHEZ LES SQUAMATES ET LES CROCODILIENS

Ce travail a été réalisé par :

Sidney Delgado*, **M. Ishiyama‡, M. Mikami‡, A. Imai‡, H. Shimomura‡, Marc Girondot§, et Jean-Yves Sire*.**

* UMR 8570, Évolution et Adaptations des Systèmes Ostéomusculaires, Équipe "Evolution et Développement du Squelette Dermique", Paris.
§ UPRESA 8079, Laboratoire "Écologie, Systématique et Évolution", Équipe de Conservation des Populations et des Communautés, CNRS et Université Paris-Sud, Paris XI, Orsay.
‡ Department of Histology and Oral Microbiology, The Nippon Dental University School of Dentistry at Niigata, Niigata, Japon.

Remerciements: Je suis grandement reconnaissant à Nicolas Vidal (Muséum national d'Histoire naturelle, Paris) qui m'a fourni les tissus de serpents, ainsi qu'à M. Luc Fougeirol,

(Ferme aux Crocodiles, Pierrelatte) pour avoir donné les tissus de jeunes spécimens de caïman et de crocodile du Nil. Merci aussi à Olivier Babiar pour la gestion de l'élevage de lézards.

Résumé

L'amélogénine, la protéine majeure dans la formation de l'émail, joue un rôle important dans la détermination de la structure de ce tissu fortement minéralisé. Cependant, la fonction exacte de cette protéine reste peu comprise. Un moyen d'évaluer l'importance d'un seul acide aminé ou d'un groupe d'acides aminés pour la fonction de la protéine est de comparer un nombre de séquences représentatives des différentes lignées dans une perspective évolutive. En effet, les résidus qui restent inchangés durant une longue période d'évolution indiquent qu'ils jouent certainement des rôles importants.

Dans le présent travail, l'évolution de l'amélogénine a été analysée chez des reptiles qui possèdent des dents: les squamates (lézards et serpents) et les crocodiliens, les lignées les plus proches des mammifères. Seules une séquence de crocodile et une séquence de serpent ont été obtenues à partir des bases de données. J'ai cloné et séquencé le gène de l'amélogénine de 16 espèces représentatives des principales lignées (10 lézards, 4 serpents et 2 crocodiles). Après l'alignement et la comparaison des séquences, le patron de l'évolution de l'amélogénine a été déduit dans ces deux lignées.

L'amélogénine des reptiles est caractérisée par la conservation, dans toutes les séquences examinées, d'un certain nombre d'acides aminés dans les régions N-terminale (codées par les exons 2, 3, 5 et le début de l'exon 6) et C-terminale (codées par la fin de l'exon 6 et l'exon 7). Ces résidus semblent essentiels pour la fonction de l'amélogénine et sont sujets à de fortes pressions sélectives. Parmi ces acides aminés conservés, certains sont déjà suspectés de jouer un rôle important pour la fonction de l'amélogénine chez les mammifères (sites de phosphorylation, sites de liaison aux N-acétyle-D-glucosamines, sites d'épissage alternatif, sites protéolytiques). Il existe aussi des résidus conservés qui sont connus, quand ils sont substitués, pour conduire chez l'homme à l'amélogenèse imparfaite. De plus, il y a de nombreux résidus hautement conservés dont le rôle est inconnu.

Chez les squamates, la région centrale de la protéine (la majeure partie de l'exon 6) montre un très grand nombre de substitutions, d'insertions et de délétions comparée aux séquences des crocodiliens. Cette région se caractérise par la présence de nombreux triplets d'acides aminés et par un grand pourcentage de proline et de glutamine. Les séquences des crocodiliens sont très semblables bien qu'elles appartiennent à trois genres différents. Elles possèdent également de nombreux triplets d'acides aminés dans l'exon 6. L'analyse comparative des séquences des squamates, des crocodiles et des mammifères indique que l'amélogénine a évolué plus rapidement chez les squamates que dans les autres lignées. Tirant avantage de cette différence de vitesse d'évolution, il a été possible de réduire le nombre de résidus qui sont suspectés d'avoir une fonction importante dans l'amélogénine.

Une séquence ancestrale d'amélogénine des squamates a été définie et comparée à la séquence ancestrale des mammifères et aux séquences crocodiliennes. Cette comparaison permet de proposer qu'une série de triplets (Pro-x-x) est probablement à l'origine de la majeure partie de l'exon 6 chez un ancêtre commun aux amniotes.

Introduction

Au cours de l'étude comparative des différents éléments du squelette dermique, nous nous sommes intéressés à l'histoire évolutive des tissus dérivés des dents chez les vertébrés (Huysseune et Sire, 1998; Sire et Huysseune, 2002). Plus précisément, notre but est de montrer l'homologie entre l'émail, l'émailloïde et certains tissus fortement minéralisés tels que la ganoïne (Sire *et al.*, 1989; Zylberberg *et al.*, 1997) et d'autres couches superficielles dans les éléments du squelette dermique des poissons actinoptérygiens (Sire *et al.*, 1997; Sire, 1997; Huysseune et Sire, 1998). La cible principale est l'amélogénine, la protéine de la matrice prédominante dans le développement de l'émail dentaire chez les mammifères, et qui est considérée comme essentielle pour la fonction normale de l'émail.

L'amélogénine n'est pas seulement la protéine majeure de l'émail en formation, mais elle est aussi spécifique à ce tissu (elle n'a aucun effet pléiotropique connu). Chez les mammifères, l'amélogénine représente approximativement 90% du contenu protéique (Termine *et al.*, 1980; Fincham *et al.*, 1983; Sasaki et Shimokawa, 1995) et elle n'a jamais été trouvée exprimée dans d'autres tissus, excepté une expression transitoire dans les pré-odontoblastes (Oida *et al.*, 2002). En raison de cette présence dominante dans la matrice organique, avant et pendant la minéralisation de l'émail, l'amélogénine joue certainement un rôle important dans la structuration de ce tissu. Elle est supposée contrôler le processus de minéralisation de l'émail à travers des interactions avec les autres protéines de la matrice et les cristaux d'émail (Aoba *et al.*, 1989; Fincham *et al.*, 1991; Deutsch *et al.*, 1995; Aoba, 1996). L'organisation structurale et le rôle des différentes régions de l'amélogénine sont actuellement étudiés en profondeur chez des modèles mammaliens (à savoir les rongeurs, la vache et le cochon) en raison des possibilités d'expériences *in vitro*. Ces études sont aussi importantes au regard des malformations de l'émail (amélogenèse imparfaite) apparentées à la substitution ou à la délétion de certains acides aminés.

Deux raisons expliquent pourquoi notre intérêt s'est focalisé sur l'amélogénine. Premièrement, nous voulions définir les relations évolutives et développementales entre l'émail, l'émailloïde et d'autres tissus apparentés, en étudiant l'expression de l'amélogénine, dans des tissus dérivés des dents, chez des clades basaux de vertébrés, c'est-à-dire les actinoptérygiens et les chondrichtyens. Le gène de l'amélogénine et son évolution sont connus chez les mammifères (voir chapitre précédent “Evolution moléculaire de l'amélogénine chez les mammifères“), quelques données moléculaires sont disponibles chez les reptiles (squamates et crocodiles, Ishiyama *et al.*, 1998; Toyosawa *et al.*, 1998) et chez un amphibien (Toyosawa *et al.*, 1998), mais aucune information n'existe dans des taxa non-tétrapodes (sarcoptérygiens basaux, actynoptérygiens et chondrichtyens). Cependant, un certain nombre d'études immunohistologiques, bien qu'employant des anticorps de mammifères, ont suggéré que l'amélogénine est présente dans la plupart des lignées de vertébrés (Herold *et al.*, 1989) et même chez la myxine (Slavkin et Diekwish, 1996). Récemment, l'origine du gène de l'amélogénine a été estimée à 630 MA (durant la période Précambrienne) au moment où les restes de vertébrés minéralisés apparaissent dans le registre fossile. Durant cette période, la séquence codante du peptide signal de SPARC (ostéonectine) a été dupliquée et ajoutée à une séquence d'ADN dont l'origine est encore inconnue, pour construire un nouveau gène codant pour une nouvelle protéine, l'amélogénine (Voir chapitre 2 - Deuxième partie). Cependant, l'amélogénine étant connue seulement chez les tétrapodes, dont l'origine est estimée à 360 millions d'années (Kumar et Hedges, 1998), il y a donc une période d'environ 250 millions d'années dépourvue de données évolutives dans l'histoire de l'amélogénine.

Le second objectif, quelque peu lié au premier, était d'étudier l'évolution moléculaire de l'amélogénine chez les tétrapodes pour mieux comprendre le rôle des différents domaines de cette molécule. Il est apparu rapidement qu'une telle étude comparative était un pré-requis

pour réaliser le premier objectif qui nécessite de définir des amorces adéquates pour le clonage de l'amélogénine chez les vertébrés non-tétrapodes. Donc, nous avons commencé notre programme en étudiant l'évolution moléculaire de l'amélogénine dans la lignée mammalienne, pour laquelle cette protéine était bien connue. Dans deux travaux récents, nous avons analysé l'évolution de l'amélogénine dans la lignée mammalienne (voir chapitre précédent "Evolution moléculaire de l'amélogénine chez les mammifères") et nous avons aussi démontré le fort signal phylogénétique contenu dans l'exon 6 du gène de l'amélogénine pour étudier les relations mammaliennes (voir chapitre suivant). Un total de 26 séquences d'amélogénine mammalienne, réparties dans les principales lignées de mammifères, ont été comparées et la séquence ancestrale de l'amélogénine des mammifères a été déduite. Un grand nombre de résidus sont fortement conservés dans les régions hydrophiles N- et C-terminales, alors qu'un grand nombre de substitutions, d'insertions et de délétions caractérisent la région hydrophobe centrale de la protéine (voir le chapitre précédent "Evolution moléculaire de l'amélogénine chez les mammifères"). Ces découvertes indiquent que les régions très conservées sont sujettes à de fortes pressions sélectives (révélant une relation étroite entre la structure primaire de la protéine et sa fonction), alors que l'autre région, plus variable, subit moins de contraintes. Cependant, le rôle d'un certain nombre de résidus fortement conservés dans les régions hydrophiles est toujours inconnu. Dans la partie précédente ("Evolution moléculaire de l'amélogénine chez les mammifères") il a été conclu que les principales lignées de mammifères se sont séparées trop rapidement (entre 173 et 130 millions d'années (Kumar et Hedges, 1998)) pour déterminer avec précision si les résidus sont conservés pendant l'évolution parce qu'ils ont une fonction importante ou à cause à la relation de parenté étroite entre les taxa de mammifères. Donc, pour améliorer nos connaissances sur l'évolution moléculaire des différents domaines de l'amélogénine, de nouvelles études sont nécessaires sur d'autres clades de vertébrés. En étudiant l'évolution de l'amélogénine chez les reptiles [10], nous espérons mieux tester si les résidus sont importants ou non dans les différentes régions de la protéine de l'amélogénine, et si l'amélogénine a évolué ou non différemment dans deux lignées d'amniotes, les reptiles et les mammifères.

Chez les squamates (lézards et serpents) et chez les crocodiliens, l'émail est une couche épaisse et fortement minéralisée qui couvre la dent et joue un rôle important dans la protection de l'abrasion dentaire. Quand on considère l'importance des dents dans l'alimentation, on peut facilement comprendre le rôle crucial de ce tissu. Alors que l'émail des dents des crocodiles ressemble beaucoup à celui des mammifères (présence de prismes), les squamates possèdent un émail dépourvu de prismes et plusieurs types d'organisation des cristaux de minéral ont été décrits (Sander, 2001). Mais, contrairement aux dents des mammifères qui ne sont pas renouvelées pendant une grande partie de la vie (les mammifères sont diphyodontes ou monophyodontes), les crocodiles et les squamates remplacent leurs dents durant toute leur vie (ils sont polyphyodontes) suivant un patron régulier de remplacement (Edmund, 1969; Osborn, 1971; voir aussi chapitre 3 - 1ère partie). Cependant, excepté la présence ou l'absence de prismes, la structure de l'émail est approximativement semblable chez les mammifères et les reptiles. De plus, l'amélogénine ne semble pas impliquée dans la mise en place des prismes au niveau desquels la tufteline et l'énaméline serait abondantes (Deutsch *et al.,* 1995b*)*. Nous pouvons donc supposer que l'amélogénine est (i) prédominante dans le développement de l'émail autant chez les reptiles que chez les mammifères, (ii) qu'elle remplit les mêmes fonctions importantes, et (iii) que par conséquent la partie codante a une structure similaire.

[10] Nous utilisons ici le terme "reptile" pour regrouper les lézards, les serpents et les crocodiles. Ce terme comprend également les tortues et les oiseaux qui n'ont pas de dents, et chez lesquels les recherches du gène de l'amélogénine se sont avérées sans succès (Girondot et Sire, 1998). Jusqu'à maintenant, nous n'avons pas pu obtenir de tissus de *Sphenodon punctatus*, le taxon basal des lépidosauriens.

Avant ces travaux, le gène de l'amélogénine était connu chez un seul crocodile, un caïman (Toyosawa *et al.*, 1998) et chez un seul squamate, un serpent (Ishiyama *et al.*, 1998). Ces deux séquences, bien que très informatives, n'étaient évidemment pas représentatives de l'ensemble des lignées de squamates et de crocodiliens, et ne pouvaient pas être utilisées seules pour analyser l'évolution de l'amélogénine chez les reptiles. Ainsi, pour atteindre notre objectif, nous avons cloné et séquencé (à partir d'ADN génomique et d'ADNc) le gène de l'amélogénine, de 14 espèces représentatives des principales lignées de squamates et chez deux crocodiliens appartenant à deux genres différents.

Nous présentons ci-dessous l'analyse évolutive de l'amélogénine basée sur 18 séquences de reptiles. Ces résultats sont ensuite comparés aux données obtenues dans la partie concernant les mammifères ("Evolution moléculaire de l'amélogénine chez les mammifères").

Matériel et méthodes

Brève présentation du gène de l'amélogénine

Contrairement aux mammifères, chez lesquels deux copies du gène de l'amélogénine se trouvent sur les chromosomes sexuels X et Y, la localisation de ce gène est inconnue chez les reptiles. Cependant, il n'y a aucune raison de penser que le gène se trouve aussi sur les chromosomes sexuels car les origines des chromosomes sexuels sont différentes entre les mammifères et les reptiles (Quintana-Murci *et al.*, 2001). Chez un certain nombre de mammifères, le gène de l'amélogénine est composé de 7 exons, mais seulement 6 exons (1, 2, 3, 5, 6,7) sont connus à ce jour chez les reptiles. Chez les reptiles, l'exon 4 n'existe probablement pas sachant qu'il n'est présent ni dans le génome du caïman, ni dans celui du xénope (Toyosawa *et al.*, 1998). L'exon 1 et la majeure partie de l'exon 7 ne sont pas traduits. L'exon 2 contient le peptide signal. Chez quelques mammifères, des insertions de répétitions sont présentes dans la région centrale, mais de telles insertions n'existent pas dans les deux séquences d'amélogénine reptilienne connues à ce jour. Chez les reptiles, la séquence codante commence à partir de la 12ème base de l'exon 2 par le codon d'initiation (ATG) et se termine à la 6ème base de l'exon 7 par le codon "stop". Les exons 1 à 5 sont à moins de 60 pb, alors que l'exon 6 contient plus de 400 pb ; les introns sont larges (> 1kb) sauf l'intron 5. Les seuls exons accessibles à partir de l'ADN génomique par PCR sont les exons 6 et 5, quand on tire parti du petit intron 5.

Matériel

Squamates

La séquence de l'amélogénine d'un serpent japonais colubridé, *Elaphe quadrivirgata*, a été obtenue dans GenBank (Ishiyama *et al.*, 1998) (Tableau 1A).

Nous avons cloné et séquencé complètement le gène codant l'amélogénine (ADNc) ou la majeure partie de son exon 6 (ADNg), dans 14 espèces de squamates distribuées dans 10 familles (Tableau 1B).

L'origine du matériel était :

Chalcides viridanus, *C. sexlineatus*, *Gallotia stehlini*, *G. galloti*, *Psammodrommus algirus*, *Anguis fragilis* et *Lacerta muralis* ont été élevés dans notre laboratoire.

Gekko japonicus, *Iguana iguana* et *Leiolepis belliana* ont été élevés dans le laboratoire japonais à Nigata.

Varanus niloticus (sang) nous a été fournie par le Dr. V. de Buffrénil du Muséum national d'Histoire naturelle, Paris.

Micrurus lemniscatus, *Boa constrictor* et *Trimesurus purpureomaculatus* (tissus fixés dans 70% d'éthanol) provenaient d'un don généreux du Dr. N. Vidal du Muséum national d'Histoire naturelle, Paris.

Tableau 1. Nom vernaculaire, nom latin et famille des squamates et crocodiliens étudiés. **A**: Les 2 séquences d'amélogénine obtenues dans GenBank et leur numéro d'accession. **B**: Données personnelles (14 squamates, 2 crocodiliens). Voir Figure 1 pour les relations de parenté de ces espèces.

A

Nom vernaculaire	***Genre, espèces***	**Sous-famille**	**Famille**	**N° d'accession**
Couleuvre japonaise	*Elaphe quadrivirgata*	–	Colubridae	AF095568
Caïman nain	*Paleosuchus palpebrosus*	Alligatorinae	Crocodylidae	AF118568

B

Nom vernaculaire	***Genre, espèces***	**Famille**
Iguane commun	*Iguana iguana*	Iguanidae
Agame	*Leiolepis belliana*	Agamidae
Orvet	*Anguis fragilis*	Anguidae
Varan	*Varanus niloticus*	Varanidae
Serpent corail d'Amérique du sud	*Micrurus lemniscatus*	Elapidae
Vipère de mangrove d'Asie	*Trimeresurus purpureomaculatus*	Viperidae
Boa	*Boa constrictor*	Boidae
Lézard géant des îles Canaries	*Gallotia stehlini*	Lacertidae
Lézard des îles Canaries	*Gallotia galloti*	Lacertidae
Grand psammodrome	*Psammodromus algirus*	Lacertidae
Lézard des murailles	*Lacerta muralis*	Lacertidae
Scinque des îles Canaries	*Chalcides sexlineatus*	Scincidae
Scinque doré	*Chalcides viridanus*	Scincidae
Gecko japonais	Gekko japonicus	Gekkonidae
		Sous-famille
Caïman	*Caiman crocodylus*	Alligatorinae
Crocodile du Nil	*Crocodylus niloticus*	Crocodylinae

Les 15 séquences à notre disposition étaient représentatives des principales lignées de squamates et sont bien distribuées à l'intérieur des phylogénies récentes de squamates (cf. Heise *et al.*, 1995; Brochu, 1997; Harris *et al.*, 1999) (Figure 1). Quelques unes de ces espèces ont été choisies à l'intérieur d'un même genre (c'est-à-dire *Gallotia*, *Chalcides*) ou à l'intérieur d'une même famille (c'est-à-dire Lacertidae) pour tester le taux d'évolution intra-générique et intra-familial.

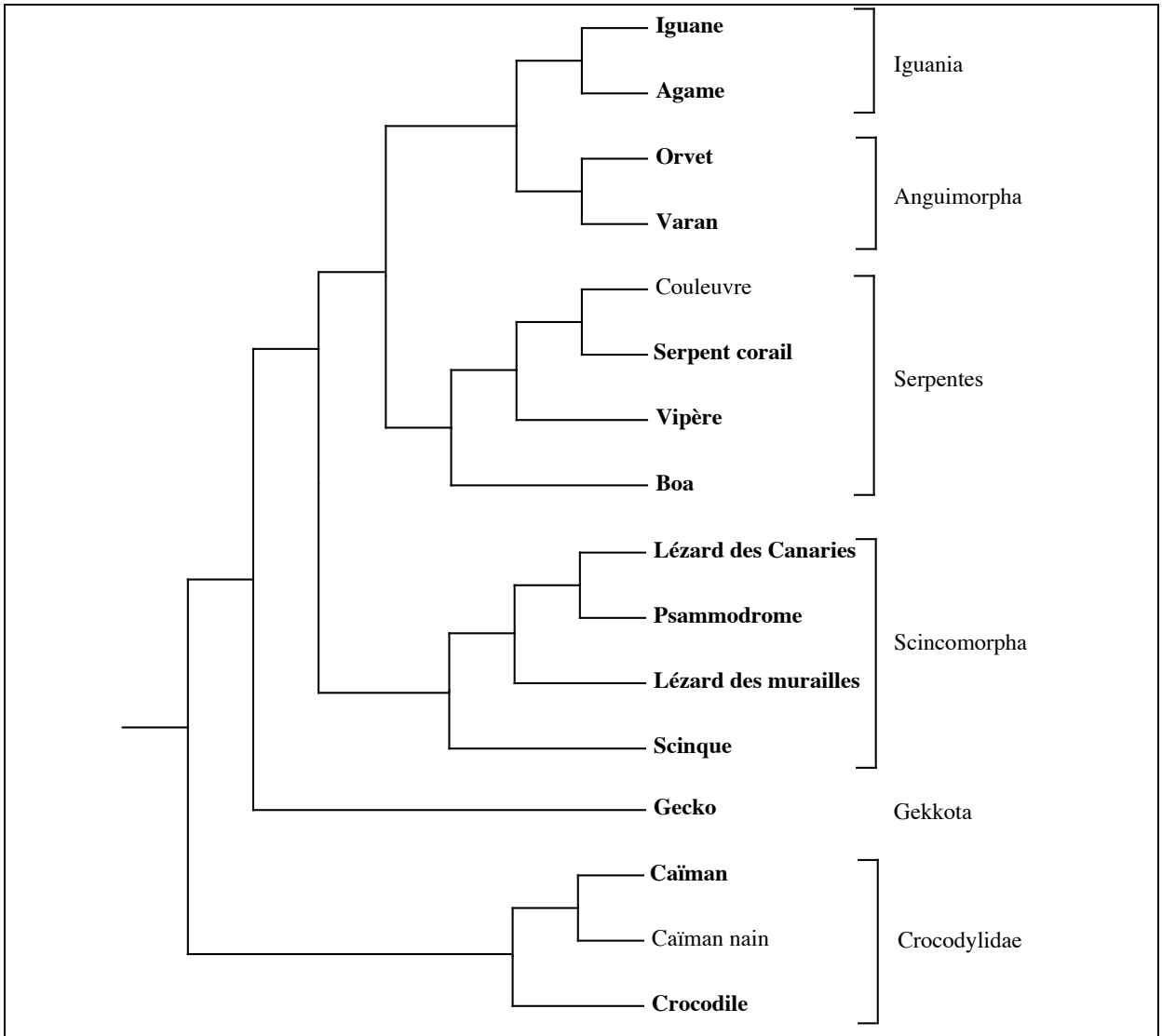


Figure 1. Relations phylogénétiques des 15 espèces de squamates et des 3 espèces de crocodiliens utilisées dans cette étude. Phylogénie des serpents d'après Heise *et al.* (1995), des squamates d'après Harris *et al.* (1999) et des crocodiliens d'après Brochu (1997). Les espèces dont nous avons obtenu les séquences d'amélogénine sont indiquées en gras.

Les crocodiliens

Nous avons séquencé le gène de l'amélogénine chez *Crocodylus niloticus* (un Crocodylinae) et chez *Caiman crocodylus* (un Alligatorinae) (Tableau 1B). Ce matériel provenait d'un don généreux de “la Ferme aux Crocodiles“ de Pierrelatte, France. Un total de trois séquences représentatives de trois genres et de deux sous-familles ont été comparées (Figure 1).

Extraction d'ADN

L'ADN génomique a été isolé à partir de petits morceaux de tissus “mous“ (muscle, foie) ou de sang en utilisant le kit “DNeasy Tissue“ (Qiagen) ou par la méthode d'extraction avec phénol-chloroforme. Ces tissus ont été prélevés à partir de spécimens fraîchement tués ou conservés dans l'éthanol.

Pour la préparation des ADNc, les spécimens ont été tués par une overdose d'anesthésiant (chloroforme). Les mâchoires entières (ou des fragments de mâchoire pour les spécimens de grande taille) contenant des dents en formation ont été immédiatement disséquées, congelées dans de l'azote liquide et réduites en fine poudre. Les ARN totaux ont été extraits en utilisant le kit “Rneasy Midi“ (Qiagen) et les ARNm ont été purifiés par la méthode “Oligotex“ (Qiagen).

Amplification par PCR, RACE 5' et 3'

Les gènes de l'amélogénine ont été amplifiés en utilisant les amorces suivantes définies à partir d'une comparaison de séquences nucléotidiques des gènes d'amélogénine de reptiles déjà connus.

- Pour les squamates :
AML2.2 (*sens*; 5' - GACTTTGGTTATGTGCCTCTT - 3') a été construite pour s'hybrider sur la région 5' de l'exon 2 sur l'ADNc;

AML 242 (*sens*; 5' - TATGAACCTATGGGAGGATGGATGC - 3'),
AML B2 (*sens*; 5' - CATGGGTGGNTGGYTNCAYCA - 3'),
AML A2 (*sens*; 5' - TATGGBTATGARCCHATGGGAGG - 3') et
AML tyr6 (*sens*; 5' - GTATCCAYSTTATGGYTATGA - 3') ont été construites pour s'hybrider sur la région 3' de l'exon 6 à la fois sur l'ADNc et l'ADNg.

AML 529 (anti*sens*; 5' - CACTTCCTCCCGCTTGGTCTT - 3') et
AML C (anti*sens*; 5' - CTTCCTCCCGYTTNGTYTTRT - 3') ont été construites pour s'hybrider à la région 3' de l'exon 6 à la fois sur l'ADNc et l'ADNg.

- Pour les crocodiles :
AML_CROC (*sens*; 5' - TATGAGCCCATGGGAGGATGG - 3') a été construite pour s'hybrider sur la région 5' de l'exon 6 sur l'ADNg et a été utilisée avec AML C (cf. ci-dessus).

La séquence non-codante de la fin de l'exon 6 et la séquence du début de l'exon 2 ont été obtenues chez plusieurs espèces grâce, respectivement, aux techniques RACE 3' et RACE 5'. Plusieurs PCR successives, utilisant une amorce poly(dT)18 et une amorce spécifique construite pour le gène cible, ont été nécessaires. L'ADN génomique ou l'ADNc (1 µl) ont été amplifiées par PCR dans une mixture composée de 5 µl de tampon de Taq (10x) pH 8.8, de 3

µl de Mg Cl_2 2 mM, et de 1 µl de dNTP 10 mM, en présence des amorces sens et antisens, et de 0,3 µl de polymérase "Red Hot" (Advanced Biotechnologies Ltd.). L'amplification a été réalisée dans un thermocycleur (Genius Techne) en 38 cycles, chaque cycle consistant en : 1 min de dénaturation à 94 ° C, 1 mn d'hybridation à 59° C et 1 mn de synthèse à 72° C. la synthèse finale a durée 30 mn à 72° C.

Clonage (voir partie précédente : "Evolution moléculaire de l'amélogénine chez les mammifères")

Analyses moléculaires

Taux de substitution et richesse en CpG (voir "Evolution moléculaire de l'amélogénine chez les mammifères")

Les répétitions dans l'exon 6

1- Déduction de la séquence ancestrale de l'amélogénine de squamate

Dans le chapitre précédent, consacré aux mammifères, nous avons postulé que la plus grande partie de la structure primaire de l'exon 6 provient de multiples insertions de répétitions de 3 aa (triplets). Ces insertions de triplet ont eu lieu plusieurs fois durant l'évolution de l'amélogénine, et probablement avant la divergence de la lignée mammalienne. De telles répétitions sont probablement générées par "dérapage" de la polymérase (Elder et Turner, 1995; Liao, 1999; Ohta, 2000). Nous avons utilisé ici les séquences des reptiles avec la même approche que pour l'analyse de l'amélogénine des mammifères pour vérifier si oui ou non il y a des répétitions dans les régions variables de l'exon 6 chez les squamates et les crocodiliens. Parce que les limites exactes entre les répétitions sont parfois difficiles à trouver à cause de nombreuses substitutions, délétions et insertions ayant eu lieu au cours de l'évolution de l'amélogénine chez les reptiles, nous avons construit une séquence ancestrale. En théorie, les limites entre les répétitions pourrait être mieux définies (si elles existent) dans la séquence de l'amélogénine de l'ancêtre de tous les squamates actuels. L'alignement des 15 séquences protéiques représentatives des principales lignée de squamates, a permis de déterminer cette séquence ancestrale en utilisant PAUP 4.0 après avoir placé les séquences dans une phylogénie récente (Heise *et al.*, 1995; Harris *et al.*, 1999)

2- Tests sur l'origine des répétitions (voir "Evolution moléculaire de l'amélogénine chez les mammifères")

Fréquence de proline et glutamine

Le pourcentage de P et Q a été calculé dans la région variable de l'exon 6 (des aa 49 à 169 sur l'alignement des squamates) pour chaque séquence d'amélogénine de reptile, et pour la séquence ancestrale des squamates.

Résultats

Comparaison de séquences

En utilisant différentes paires d'amorces, l'amplification par PCR d'ADNc (pour 8 squamates) ou d'ADNg (pour 6 squamates et deux crocodiles) a produit des fragments qui se

révèlent être homologues, respectivement, aux régions allant de l'exon 2 à l'exon 6 (ADNc) ou uniquement de l'exon 6 (ADNg) du gène de l'amélogénine. Les séquences nucléotidiques ont été transformées en séquences d'acides aminés (aa). Les deux séquences obtenues dans GenBank pour un crocodile, *Paleosuchus palpebrosus* et un serpent, *Elaphe quadrivirgata* ont été ajoutées à notre matériel. Un total de 15 séquences de squamates et 3 de crocodiliens ont été ensuite alignées et comparées.

Squamates

Quelques insertions/délétions ont été introduites dans la région variable de l'exon 6 pour obtenir le meilleur alignement de séquences possible (Figure 2). A partir de cet alignement nous avons déduit que l'organisation de la séquence primaire de l'amélogénine est la même dans toutes les espèces de squamates comme nous l'avons déjà décrit pour les séquences mammaliennes (voir "Evolution moléculaire de l'amélogénine chez les mammifères").

Cinq exons ont été identifiés à partir des produits de PCR (exons 2, 3, 5, 6, 7). L'exon 4, qui est connu que dans quelques espèces mammaliennes, n'a jamais été trouvé dans aucune des 8 espèces, pour lesquelles l'ADNc a été séquencé. De même, les exons 8 et 9 n'ont pas été identifiés dans la région C-ter. L'exon 2 est composé de 18 acides aminés, les 16 premiers aa étant identifiés comme appartenant au peptide signal. L'exon 3 est composé généralement de 16 aa mais trois ont été supprimés durant l'évolution des lacertidés, des serpents et des iguanes. L'exon 5 contient 15 aa. L'exon 6 est composé d'un nombre variable de résidus s'étendant de 129 (chez le colubridé *Elaphe quadrivirgata*) à 147 (chez le scincidé *Chalcides viridanus*). Cependant, 155 positions ont été obtenues dans l'alignement l'exon 6 à cause de l'insertion de plusieurs gaps (Figure 2). Dans l'exon 7 les deux premiers codons seulement sont traduits, le second étant le codon stop. Par conséquent, l'alignement des 15 séquences d'amélogénine de squamates est composé de 190 aa. Les régions 5' (aa 1-48) et 3' (aa 175 à 190) sont hydrophiles alors que la région centrale (aa 49 à 174) est hydrophobe.

La comparaison des séquences révèle que la majeure partie de la séquence primaire est conservée dans les régions N- et C-terminales hydrophiles alors que la région centrale de l'amélogénine est caractérisée par de nombreuses substitutions, délétions et insertions (Figure 2).

La variabilité intra-générique du gène de l'amélogénine dans le genre *Chalcides* (Scincidae, 2 espèces) et dans le genre *Gallotia* (Lacertidae, 2 espèces) est, respectivement, de 2,7% et de 4,4% au niveau nucléotidique. Dans ces deux genres, 7 et 9 substitutions ont observées dans l'exon 6, respectivement. La comparaison inter-générique chez les Lacertidae (3 genres, *Lacerta, Psammodromus* et *Gallotia*), a montré 8,85% de différences nucléotidiques pour *Lacerta* vs *Gallotia*, 8,06% pour *Lacerta* vs *Gallotia* et 5,95% pour *Psammodromus* vs *Gallotia*. Le pourcentage de substitution de nucléotides augmente quand on compare les séquences au sein d'une famille, la plus grande différence étant obtenue pour les séquences de gecko vs iguane, avec 30,54% de différences.

Figure 2. Alignement des 15 séquences d'amélogénines de squamates et des 3 séquences de crocodiliens analysées. Squamates et crocodiliens ne sont pas alignés entre eux. L'ordre des séquences suit celui des relations de parenté (Figure 1). La jonction des exons est indiquée par un espace. Les résidus remarquables sont soulignés. (.) = inconnu; (-) = identique à la séquence de l'iguane; (*) = insertions ou délétions; (#) = résidus identiques chez tous les squamates étudiés. TRAP = les deux sites protéolytiques (double flèche) qui donnent naissance aux

peptides TRAP (Tyrosine Rich Amelogenin peptide); LRAP = emplacement du site d'épissage intra-exonique qui donnent naissance au peptide LRAP (Leucine Rich Amelogenin peptide). Séquence encadrée = peptide signal.

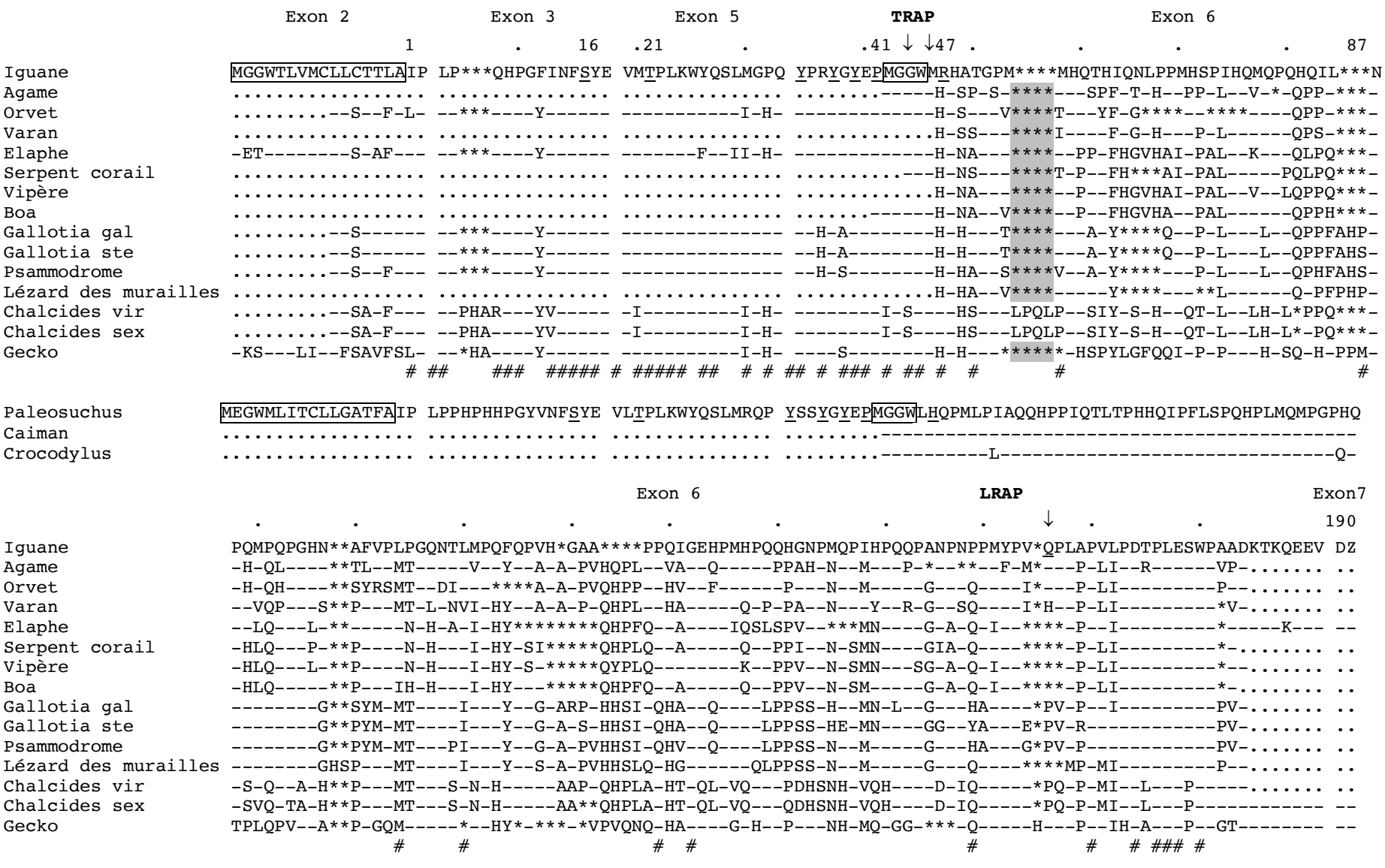

Exon 2
Exon 3
Exon 5
TRAP
Exon 6
Iguane
Agame
Orvet
Varan
Elaphe
Serpent corail
Vipère
Boa
Gallotia gal
Gallotia ste
Psammodrome
Lézard des murailles
Chalcides vir
Chalcides sex
Gecko
Paleosuchus
Caiman
Crocodylus
Exon 6
LRAP
Exon7

Figure 2.

Crocodiles

Les séquences primaires d'amélogénine des crocodiliens ont été plus faciles à aligner que les séquences de squamates (Figure 2). L'organisation de l'amélogénine est la même que celle qui a été décrite ci-dessus pour les 15 séquences de squamates, mais contrairement à ces dernières, les trois séquences crocodiliennes sont très semblables, avec seulement quelques substitutions de nucléotides. La protéine est composée de 185 résidus distribués à l'intérieur des 5 exons (exons 2, 3, 5, 6, 7), la séquence de l'exon 6 étant composée de 148 aa chez les trois espèces étudiées.

Les différences sont seulement de 0,8% au niveau nucléotidique en comparant les deux séquences d'Alligatorinae et de 2,3% entre *Crocodylus* et les deux autres genres, *Caïman* et *Paleosuchus*. Etant donnée la forte ressemblance des séquences d'amélogénine de crocodiliens, la relation entre le taux de substitution et le pourcentage de CpG n'a pas été analysée.

Dans les alignements des séquences de squamate et de crocodile, les acides aminés sont comptés à partir de la fin du peptide signal, c'est-à-dire, à partir des deux derniers aa de l'exon 2, jusqu'au second aa de l'exon 7. On obtient alors 190 aa pour les squamates et 185 aa pour les crocodiliens.

Un grand nombre de résidus est conservé dans les régions hydrophiles

Squamates

Dans les 15 séquences protéiques d'amélogénine des squamates étudiés (aa1-190), 54 résidus (marqués # sur la Figure 2) sont identiques et 50 autres positions sont substituées par seulement un résidu. La plupart des résidus inchangés (50) se trouve dans l'exon 3, l'exon 5, au début de l'exon 6 (aa 34 à 48) et à la fin de la séquence, exons 6 et 7 (aa 165 à 190 de l'alignement). Etant donné ce pourcentage important de similarité, on considère que ces régions de l'amélogénine ont évolué lentement chez les squamates. Parmi les résidus inchangés il y a quelques acides aminés "remarquables" dont le rôle chez les mammaliens est important (ou supposé important) pour la fonction de l'amélogénine.

Dans la région hydrophile N-terminale (aa 1 à 48), les résidus bien conservés au cours de l'évolution sont : (i) dans l'exon 3, la serine 16, un site de phosphorylation connu chez les mammifères (conservés dans les 15 espèces de squamates examinées); (ii) dans l'exon 5 la région des hélices alpha (aa 19-29 : conservés à 90%); (iii) dans l'exon 6 les 12 aa composant le "motif tyrosine" qui se lie aux N-acétyl-D-glucosamines (aa 34 à 45 : conservés à 90% et toutes les tyrosines présentes dans toutes les séquences); et (iv) quelques sites de mutations qui sont connus chez les mammifères comme étant impliqués dans quelques déficiences de l'émail (amélogenèse imparfaite) (par exemple la thréonine 21 et la proline 41 : conservées dans toutes les espèces).Cependant, l'histidine 47, un résidu impliqué dans la malformation de l'émail lorsqu'il est substitué par une leucine, se trouve être remplacé par une arginine (R) chez *Chalcides* et l'iguane. D'autres résidus importants, conservés dans l'exon 6, concernent les sites protéolytiques, glutamine 43 et tryptophane 45, qui donnent naissance aux peptides TRAP43 et TRAP45 (Tyrosine Rich Amelogenin Peptides). Ces résidus sont inchangés dans toutes les autres séquences analysées (Figure 2), sauf chez les deux *Chalcides* pour lesquels les acides aminés voisins, méthionine 42 et glycine 44, sont substitués par une isoleucine et une sérine, respectivement.

Chez les mammifères, la région C-terminale possède un site d'épissage alternatif remarquablement conservé qui est responsable de la formation d'un fragment d'amélogénine appelé LRAP (Leucin Rich Amelogenin Peptide) lors de la formation de la matrice de l'émail. D'après ce qui est connu pour l'amélogénine mammalienne, l'épissage du LRAP peut se produire dans les séquences de squamates entre les résidus Q163 (glutamine, soulignée sur la Figure 2) et P164 (proline). Cependant, ces glutamines (codées par CAG) et ces prolines sont présentes seulement dans 4 des 15 séquences étudiées. Dans les autres séquences, ces acides aminés sont absents ou substitués (Figure 2).

Nous pouvons aussi noter que le peptide signal, une région importante pour le transfert de la molécule à partir de la cellule jusqu'à la matrice extracellulaire, est sujet à des variations, et que ces substitutions sont spécialement nombreuses dans la séquence de l'amélogénine du gecko.

Crocodiles

Tous les résidus connus pour être importants dans le bon fonctionnement de l'amélogénine chez les mammifères sont conservés dans les régions hydrophiles N- et C-ter des trois séquences de crocodiliens, incluant le site d'épissage LRAP sur la position Q157 (Figure 2). La conservation de ces acides aminés, au cours de plus de 350 millions d'années d'évolution séparée de ces deux lignées, confirme la fonction clé de ces résidus pour l'amélogénine.

Pourcentage de CpG et taux d'évolution

Le pourcentage de CpG et le taux de substitution nucléotidique ont été calculés seulement pour les séquences de squamates (Figure 3). Pour certaines positions, le pourcentage important de CpG correspond à un niveau important de substitutions (par exemple, la fin de l'exon 2, la région centrale de l'exon 6). Dans quelques régions, le niveau important de CpG n'est pas synonyme d'une augmentation du niveau de substitution (par exemple, l'exon 3, le début de l'exon 6 (région TRAP) et la région suivant la région LRAP). Cela signifie que ces régions sont sujettes à de fortes contraintes fonctionnelles.

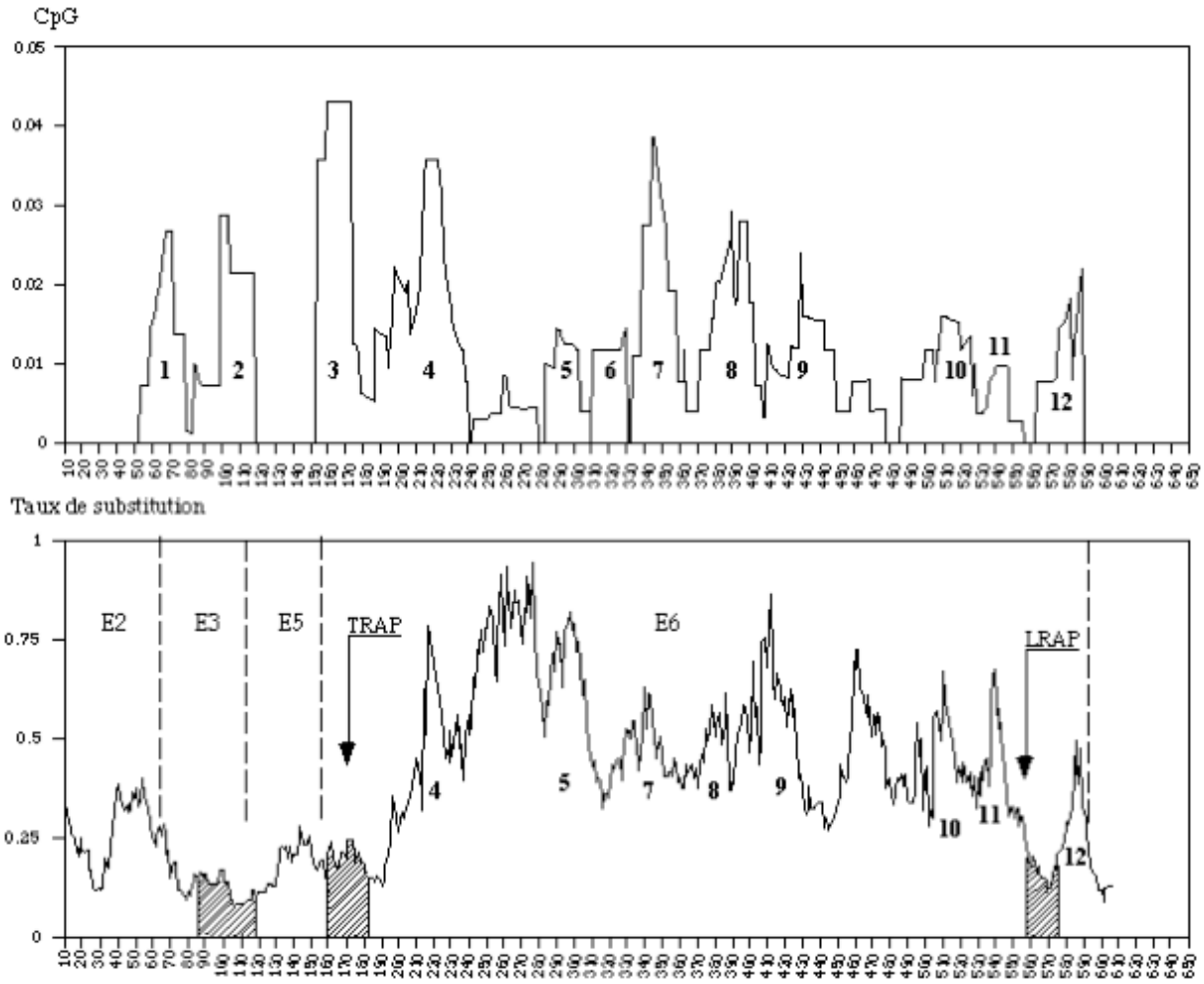


Figure 3. Pourcentage de CpG et taux de substitution dans les 15 séquences d'amélogénine des squamates, depuis l'exon 2 jusqu'à l'exon 7 (E2 = exon 2, ...). Les régions TRAP et LRAP sont indiquées par des flèches. Calculs réalisés avec l'algorithme de Tajima et Nei (1993). Les régions hachurées en gris indiquent les régions ayant un taux de substitution faible malgré un taux de CpG élevé.

Les variations dans la région centrale de l'exon 6

Squamates

La région centrale, hydrophile, de l'exon 6 (des aa 49 à 164) présente de nombreuses substitutions, délétions et insertions. Seuls 8 résidus sont inchangés dans l'alignement des 15 séquences. Cela indique une évolution rapide de cette région qui est probablement moins sujette aux contraintes que les régions hydrophiles.

Chez les mammifères, cette partie de l'exon 6 est supposée d'avoir été créée par l'insertion de nombreuses répétitions de trois acides aminés (triplets) (voir "Evolution moléculaire de l'amélogénine chez les mammifères"). Chez les squamates, étant donné le nombre important de variations dans la région centrale, nous avons vérifié la présence de tels triplets sur la séquence ancestrale de l'amélogénine des squamates qui a été déduite des 15 séquences disponibles en utilisant le logiciel PAUP 4.0 (Figure 4).

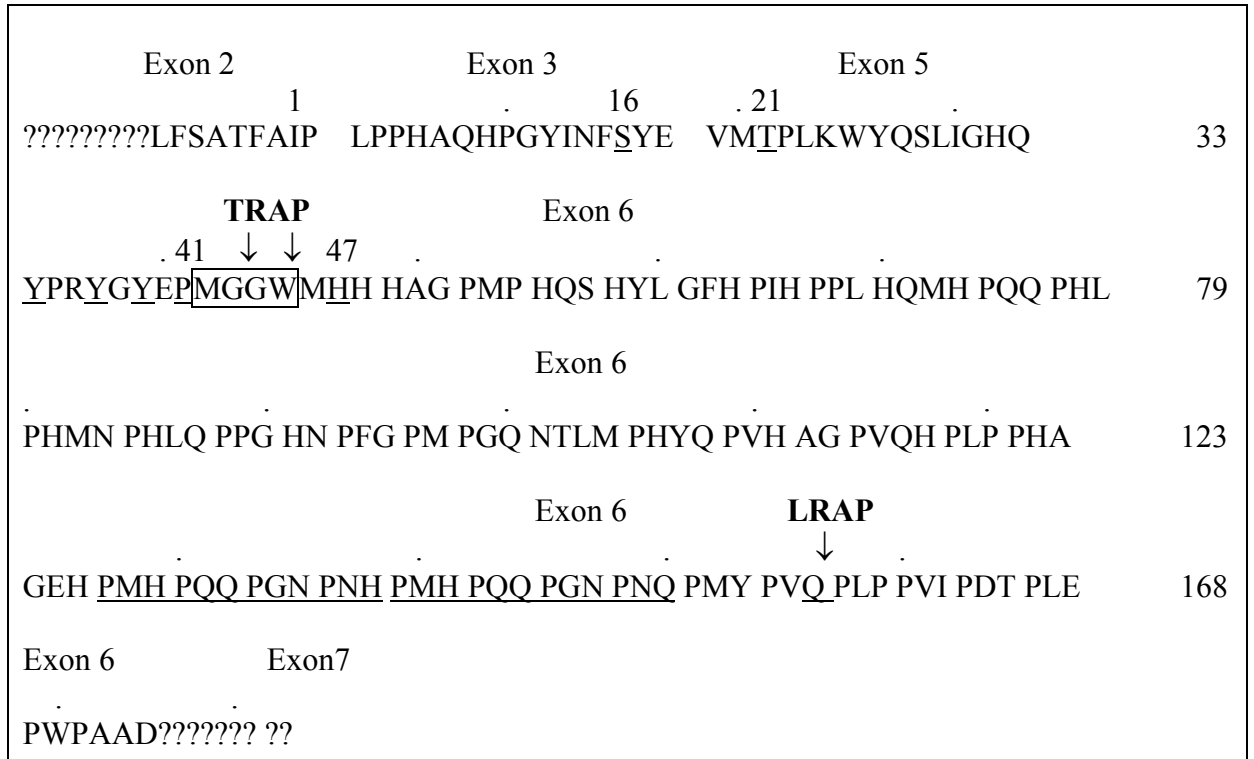


Figure 4. Séquence ancestrale de l'amélogénine des squamates. Cette séquence a été obtenue avec PAUP 4.0 à partir des 15 séquences protéiques représentatives des principales lignées de squamates. Recherche des triplets d'acides aminés dans la région variable de l'exon 6 de la séquence. Les jonctions entre les exons sont indiquées par un espace large. L'espace simple délimite les triplets supposés. La partie soulignée montre la répétition de 4 triplets. Légendes dans la Figure 2.

Les triplets sont plus faciles à identifier dans la région 3' de l'exon 6 que dans la région 5'. Dix-sept triplets ont été reconnus (aa 118-168). La plupart d'entre eux sont clairement basés sur un motif Pxx et seulement quatre étaient identifiables comme PxQ. Les triplets sont moins visibles dans la région 3' (aa 46-116), probablement parce que de nombreuses insertions, délétions et substitutions ont masqué le motif répété Pxx. Cependant, quelques uns de ces triplets ont été reconnus (Figure 4). Dans cette région, aucune répétition de triplets similaires à ceux décrits chez quelques mammifères n'a été observée et on ne trouve pas de grandes insertions. Cependant, une répétition de quatre triplets a été identifiée (aa 127-138 / aa 139-150) dans la séquence ancestrale (Figure 4).

La région de l'exon 6 (aa 49-168) est aussi riche en proline et en glutamine, et le pourcentage de ces deux résidus est à peu près semblable chez les diverses espèces (Figure 5). La proline varie de 22% chez les scincidés à 30,5% chez l'agame (en moyenne, 25%). La glutamine varie de 12% chez le gallotia à 19% chez les scincidés (en moyenne, 15%) (Figure 5). La somme de P+Q varie de 37,5% chez l'élaphe à 44,5% chez l'agame (en moyenne 40%) du total des résidus dans cette région. Dans la séquence ancestrale de l'amélogénine des squamates, le pourcentage de proline est élevé (29,4%) mais celui de la glutamine (11,7%) est plus bas que dans chez les autres espèces.

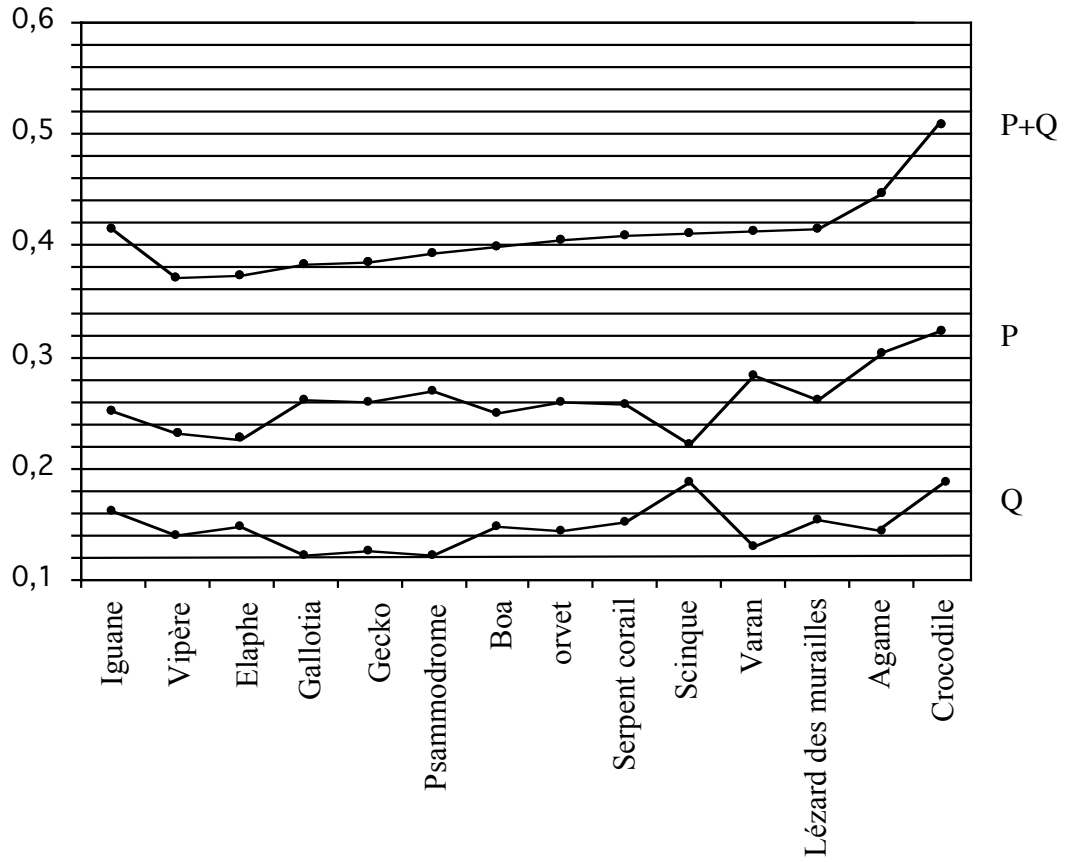


Figure 5. Taux de proline (P), de glutamine (Q) et de leur somme (P+Q) dans la région variable de l'exon 6 (aa 46-175 de l'alignement) pour les séquences analysées des 15 squamates et d'un crocodile. L'axe des abscisses représente les reptiles étudiés classés par ordre croissant de P+Q.

Crocodiles

La région de l'exon 6 (aa 49-169) est fortement conservée dans les trois séquences d'amélogénine de crocodile étudiées. Seules quatre positions d'acides aminés sont substituées dans cette région (Figure 2). Elle possède de forts pourcentages en proline (32,5%) et en glutamine (19,5%) (Figure 5). La somme P+Q est élevée (51%), les deux résidus représentant plus de la moitié des résidus dans cette région.
Vingt-quatre triplets ont été facilement identifiés dans la région 3' de l'exon 6 (aa 104-175) (Figure 6). La plupart de ces triplets sont basés sur un motif Pxx et sept ont été identifiés comme PxQ. Les triplets sont moins visibles dans la région 5' (aa 49-103), mais huit motifs Pxx ont été trouvés. Cette région centrale de l'exon 6 ne contient pas de répétitions évidentes de triplets comme celles qui ont été décrites chez quelques mammifères et aucune insertion n'a été trouvée. Cependant, quatre répétitions de quatre triplets ont été identifiées (aa 104-115 / aa 116-127 aa 128-139 / aa 140-151) (Figure 6).

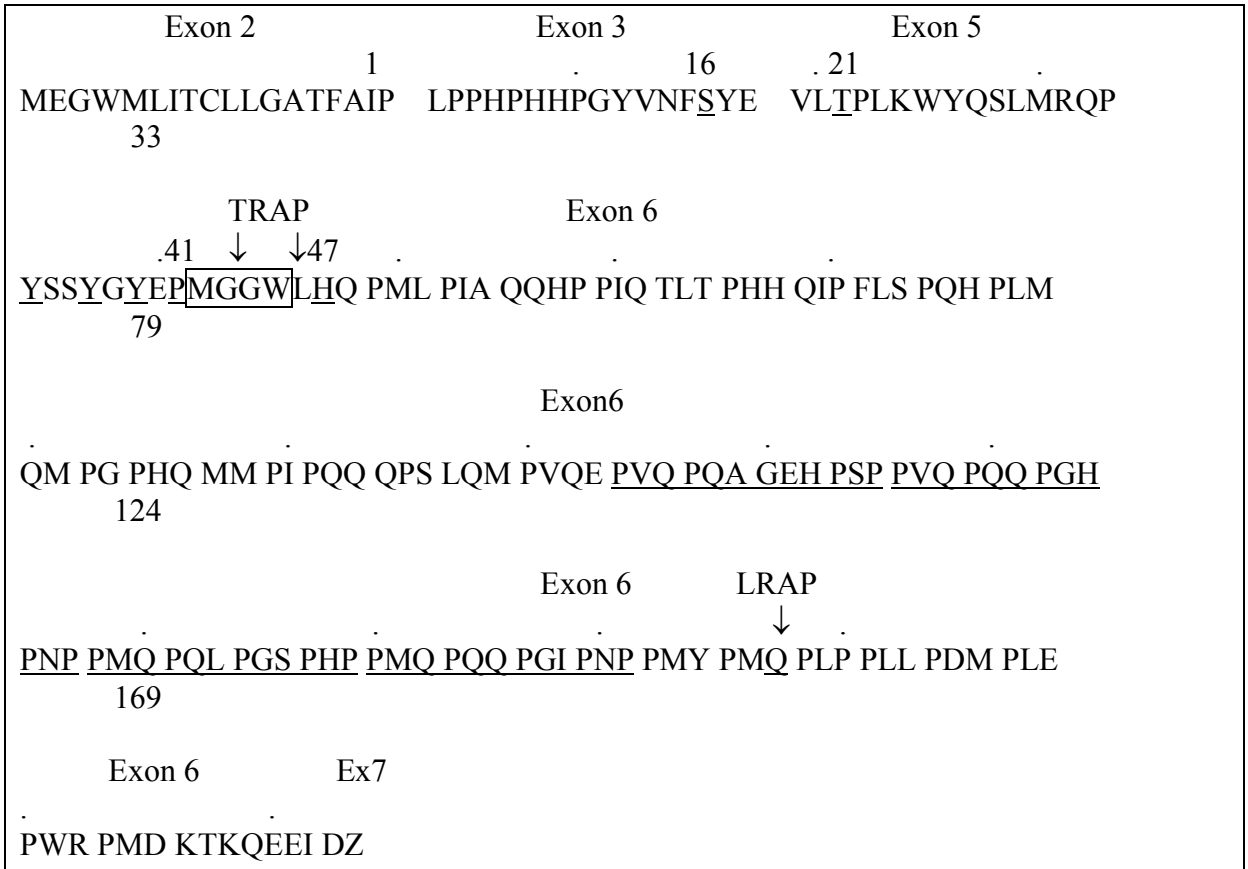


Figure 6. Séquence "consensus" des trois crocodiliens. Recherche des triplets d'acides aminés dans la région variable de l'exon 6 de l'amélogénine des crocodiles. Les jonctions des exons sont symbolisées par un grand espacement. Les petits espacements, présent dans l'exon 6 à partir de l'aa 49, délimitent les triplets supposés. Quatre répétitions de 4 triplets ont été identifiées (soulignés). Pour les autres légendes, voir la Figure 2.

Une comparaison des séquences de squamates, crocodiles et mammifères

Nos résultats ont permis de comparer la séquence ancestrale de l'amélogénine des squamates avec celles des crocodiliens et la séquence ancestrale de l'amélogénine des mammifères. Nous avons pu déduire ainsi la séquence d'amélogénine d'un ancêtre commun aux amniotes (mammifères + reptiles) (Figure 7). En l'absence de données sur l'amélogénine du tuatara (sphénodon) et pour un certain nombre d'amphibiens, une telle comparaison est seulement une tentative d'approche vers une séquence plus ancienne.

Dans la séquence ancestrale "amniotes", tous les résidus remarquables sont conservés, y compris le site d'épissage LRAP. L'alignement de séquences a permis de tester l'hypothèse de l'origine de l'exon 6 par insertion de répétitions de Pxx. Le meilleur alignement dans la région des triplets a été obtenu en introduisant plusieurs insertions/délétions dans les trois séquences (Figure 7). Ces gaps sont d'hypothétiques délétions ou insertions de résidus qui pourraient avoir eu lieu dans chaque lignée durant leur évolution. La comparaison indique que l'amélogénine des reptiles aurait connu plus de délétions que celle des mammifères. La succession de triplets Pxx est plus évidente que lorsque l'on prend chaque séquence séparément. La séquence de crocodile semble être plus proche de la séquence ancestrale de l'amélogénine des amniotes que ne le sont les deux autres lignées.

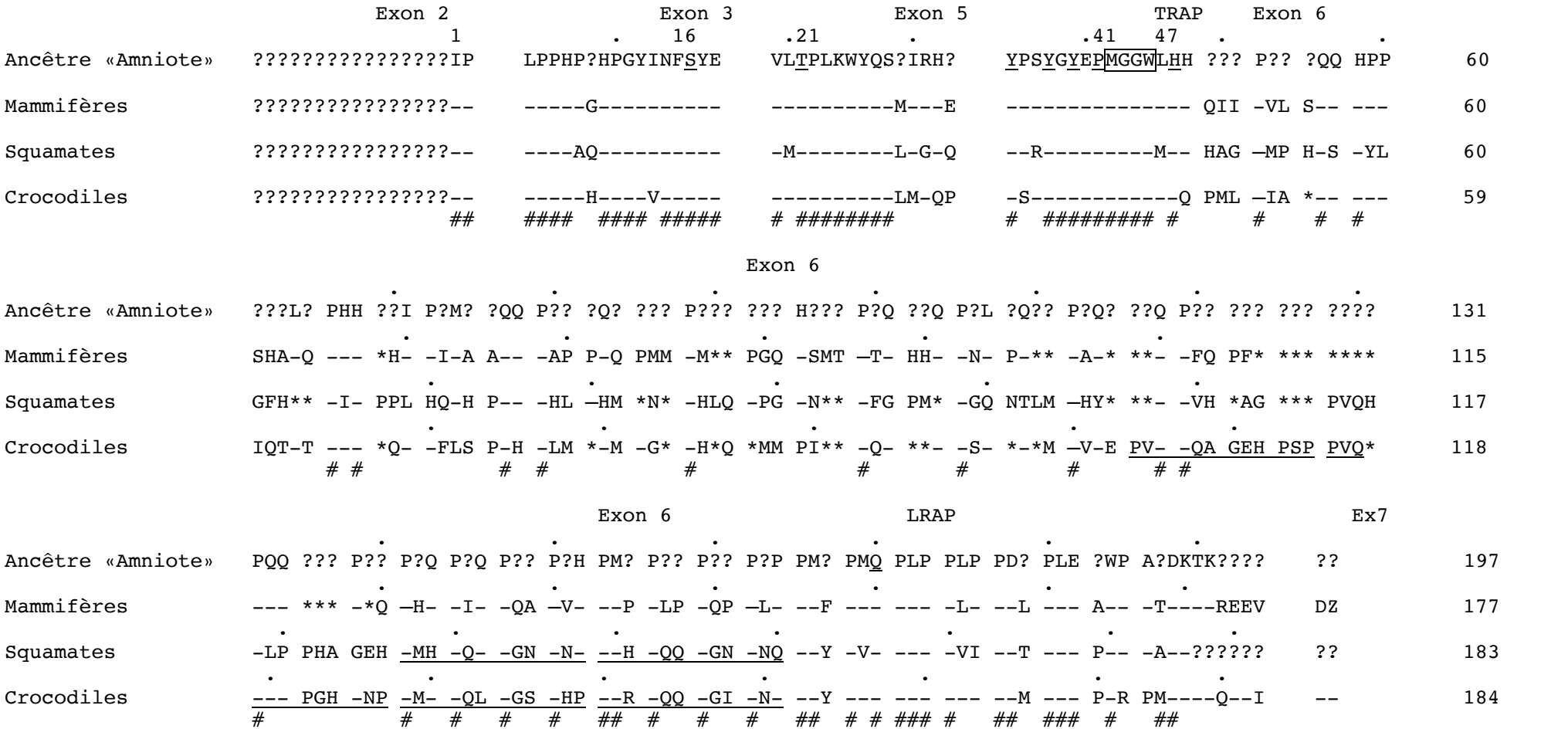


Figure 7. Séquence d'amélogénine de l'ancêtre commun aux amniotes. Cette séquence est déduite de l'alignement de la séquence ancestrale des squamates (présent travail), de la séquence ancestrale des mammifères (voir chapitre précédent) et de la séquence "consensus" des crocodiles (présent travail). Les jonctions des exons sont indiquées par un grand espacement. Les petits espacements dans l'exon 6, à partir de l'aa 49, délimitent les triplets supposés. Des insertions/délétions (*) ont été introduites pour mettre en évidence les triplets. Les répétitions de 4 triplets sont soulignées. (?) = données incomplètes. Autres légendes dans la figure 2. Le grand nombre de "?" dans l'exon 6 est lié à l'absence de données sur la séquence ancestrale des amphibiens, et dans un moindre degré sur celle des sphénodons.

Discussion

La présente étude apporte de nouvelles données sur l'évolution de l'amélogénine chez les squamates (lépidosauriens) et les crocodiles, basées sur un échantillon non exhaustif, mais plutôt bien représentatif, des espèces dans les deux lignées. Etant donné que les deux autres lignées actuelles de reptiles, les tortues et les oiseaux, ont perdu leurs dents respectivement il y a 60 et 200 millions d'années, et qu'aucun gène d'amélogénine n'a été trouvé (Girondot et Sire, 1998), nous pouvons considérer que ce travail concerne les principales lignées de reptiles. Malheureusement, nous n'avons pas eu la possibilité de travailler sur l'ADN de tuatara, *Sphenodon punctatus*, le taxon basal des lépidosauriens et cela nous aurait permis de compléter l'analyse évolutive de l'amélogénine chez les reptiles. La comparaison de 15 séquences de squamate et de 3 séquences de crocodile révèle que le gène de l'amélogénine reptilien et la protéine qu'il code, sont caractérisés par (i) des résidus bien conservés dans les régions N- et C-ter, (ii) une région centrale variable hydrophobe de l'exon 6, riche en proline et en glutamine, (iii) un plus grand nombre de variations chez les squamates que chez les crocodiles, (iv) l'absence de grandes insertions dans l'exon 6 comme celles qui ont été décrites chez quelques espèces de mammifères, mais aussi par (v) la présence de nombreux triplets d'acides aminés.

Ces résultats, ajoutés à ceux qui ont été obtenus pour l'amélogénine des mammifères, permettent de préciser non seulement l'évolution de l'amélogénine dans ces lignées mais aussi d'aborder l'évolution de l'amélogénine chez les amniotes.

Au cours de l'analyse évolutive de l'amélogénine mammalienne, deux hypothèses ont été proposées pour expliquer la conservation des résidus à travers l'évolution : soit une distance évolutive relativement courte entre les lignées soit l'influence d'importantes pressions sélectives (voir "Evolution moléculaire de l'amélogénine chez les mammifères"). L'amélogénine de squamate est celle qui présente la plus grande variation, comparée à l'amélogénine mammalienne et crocodilienne. Cette observation confirme les travaux moléculaires récents, qui ont montré un plus grand taux de substitution chez les squamates que chez les autres amniotes (Hughes et Mouchiroud, 2001). Cependant, ces variations pourraient être aussi liées à une plus longue histoire évolutive des différentes familles et genres de squamates que chez les crocodiles et les mammifères, pour lesquels les familles sont plus proches. Au cours de l'évolution des squamates, la divergence progressive des différentes familles a permis l'accumulation de mutations. Un taux aussi élevé de substitutions dans l'amélogénine des squamates est cependant très intéressant. En effet, nous pouvons supposer que les résidus qui sont conservés dans les 15 séquences de squamates sont réellement importants pour le bon fonctionnement de la protéine.

Une forte conservation de résidus dans les amélogénines de crocodile et de squamate

Grâce à de nombreux travaux antérieurs, la plupart réalisées chez les mammifères, un certain nombre de résidus dans les régions terminales NH_2 et CH_3 sont connus pour être essentiels au bon fonctionnement de l'amélogénine (par exemple, Sasaki et Shimokawa, 1995 ; Simmer, 1995 ; Ravindranath *et al.*, 1999 ; Veis *et al.*, 2000 ; Moradian-Oldak *et al.*, 1995, 2002). Les acides aminés localisés à des sites importants (site de phosphorylation, sites de liaison aux N-acétyl-glucosamines, site d'épissage alternatif, sites protéolytiques) sont présents dans toutes les séquences d'amélogénine examinées jusqu'ici, y compris les résidus qui sont connus pour conduire à l'amélogenèse imparfaite quand ils sont substitués (voir "Evolution moléculaire de l'amélogénine chez les mammifères"). Notre comparaison des

séquences d'amélogénines reptiliennes indique que la plupart de ces résidus sont conservés dans toutes les espèces de crocodiles et de squamates examinées, bien que ces deux lignées soient séparées l'une de l'autre depuis 276 millions d'années et de la lignée mammalienne depuis plus de 310 millions d'années (Kumar et Hedges, 1998).
Les résidus remarquables de la région N-ter sont conservés chez les squamates. Il s'agit de la sérine 16, connue comme un site potentiel de phosphorylation (Fincham *et al.*, 1991) ; des tyrosines 34, 37, 39, qui appartiennent à un motif supposé de se lier aux N-acétyle-D-glucosamines (Ravindranath *et al.*, 1999) ; de la glycine 43 et du tryptophane 45 les sites protéolytiques donnant les deux peptides appelés TRAP43 et TRAP45 (Tyrosine Rich Amelogenin Peptide) (Fincham *et al.*, 1991; Fincham et Moradian-Oldak, 1994). Cette dégradation de l'amélogénine est supposée être importante pour la minéralisation normale de l'émail. En effet, les fragments de TRAP, en raison de leurs propriétés agrégatives et leur affinité pour les cristaux d'hydroxyapatite, représentent la seule région de la protéine qui reste fixée au minéral quand la protéine a été hydrolysée (Fincham *et al.*, 1991). La fonction de ces TRAPs est toujours inconnue, mais elle semble être essentielle au contrôle de la minéralisation de l'émail (Akita *et al.*, 1992).

Quelques résidus de la région terminale NH_2 sont si importants que, lorsqu'ils sont substitués, la structure de la molécule est modifiée et mène à une déficience de la qualité et de la quantité d'émail, appelée amélogenèse imparfaite (Lench et Winter, 1995; Collier *et al.*, 1997 ; Ravindranath *et al.*, 1999). Quelques cas ont été rapportés dans la littérature comme la substitution de la thréonine 21 par une isoleucine dans l'exon 5 (Lench et Winter, 1995). Ce résidu est présent dans toutes les séquences de reptile examinées. Un autre exemple vient de la proline 41, située près du site protéolytique TRAP. Chez l'homme, une substitution de cette proline par une thréonine provoque également une amélogenèse imparfaite (Collier *et al.*, 1996). Récemment, Hart *et al.* (2002) ont rapporté qu'une substitution de l'histidine 47 par une leucine conduit aussi à un type d'amélogenèse imparfaite. L'histidine 47 est substituée par une arginine (R) chez deux genres de squamates, *Chalcides* et *Iguana*. Auparavant, nous avons montré que l'histidine 47 est bien conservée dans les amélogénines mammaliennes sauf chez le dauphin où His47 est aussi substitué par une arginine (voir partie précédente). Etant donné que l'émail chez le dauphin, les scinques et l'iguane est bien formé, nous pouvons conclure que la substitution His-Arg sur la position 47 est permise. Ces deux acides aminés appartiennent au même groupe et cela ne change pas la supposée fonction de cette région dans sa liaison aux glycoprotéines et aux cytokératines de la matrice de l'émail (Ravindranath *et al.*, 2000). Au contraire, c'est le cas pour une substitution His-Leu. Seuls quelques cas de défauts d'émail ont été rapportés à ce jour chez l'homme, mais étant donnée la très grande conservation de certains résidus dans les séquences de reptiles et de mammifères étudiées, on peut imaginer que de nouveaux cas d'amélogenèse imparfaite pourraient être décrits dans un futur proche. Des cas d'amélogenèse imparfaite n'ont pas été, jusqu'à maintenant, relevés chez les reptiles. Etant donné les modifications importantes de la structure de l'émail, et l'importance d'une dentition saine pour la survie dans la nature, les amélogénèse imparfaite seraient probablement létales.

Notre comparaison de séquences montre que la délétion de 3 résidus (Pro5, His6 et Ala7), dans l'exon 3 de l'amélogénine des lacertidés, ne provoque pas de défauts dans l'émail. De la même façon, la substitution de résidus Met42-Iso et Gly44-Ser dans les sites protéolytiques TRAP de l'amélogénine de *Chalcides* n'a pas d'effet négatif pour l'émail.

Dans la région N-ter de l'amélogénine, il existe aussi de nombreux résidus qui sont bien conservés (en raison d'importantes contraintes sélectives) mais pour lesquels la fonction est inconnue (par exemple, NFSYE à la fin de l'exon 3, PLKW dans l'exon 5, etc.). Il est clair que ces acides aminés jouent un rôle crucial dans la fonction de l'amélogénine et il serait intéressant d'essayer de comprendre leur rôle, comme cela a été fait, par exemple pour les

résidus situés au début de l'exon 6 (Ravindranath *et al.*, 1999). En effet, étant donné le taux élevé de substitution observé chez les squamates, la forte conservation de ces résidus signifie que chacun d'eux est sujet à de fortes pressions sélectives.
Dans la région hydrophile C-terminale le seul résidu connu pour avoir une fonction importante dans l'amélogénine mammalienne est la glutamine du site d'épissage alternatif intra-exonic LRAP (Leucin Rich Amelogenin Peptide. L'épissage alternatif, qui correspond à la suppression des régions codantes durant l'épissage des introns, est un mécanisme essentiel pour la fonction de l'amélogénine. Ce phénomène est connu pour l'amélogénine de la vache (Gibson *et al.*, 1991), l'homme (Salido *et al.*, 1992), le cochon (Yamakoshi *et al.*, 1994), les rongeurs (Simmer *et al.*, 1994 ; Bonass *et al.*, 1994 ; Li *et al.*, 1995) et l'opossum (Hu *et al.*, 1996). Cependant, ce mécanisme pourrait aussi se produire chez les autres espèces étudiées. De nombreux ARNm différents sont synthétisés à partir d'un seul gène et le patron d'épissage alternatif est unique pour chaque gène d'amélogénine, et même les proportions changent durant la formation des dents comme cela a été montré chez la souris (Couwenhoven et Snead, 1994 ; revue dans Simmer, 1995). Chez la vache, le cochon, la souris et le rat, le segment LRAP correspond aux aa 1 à 33 (exons 3, 5) et aux 25 derniers acides aminés de l'exon 6. Les protéines LRAP ont été observées chez un certain nombre d'espèces de mammifères étudiées, mais pas chez l'homme bien que le site d'épissage soit présent. L'épissage de l'exon 6 (c'est-à-dire, la suppression de l'extrémité 5' de ce grand exon) est obtenu grâce à un site 3' accepteur d'épissage alternatif intra-exonique. Ce site est identifié par un signal d'épissage connu aussi dans les frontières intron/exon. Les introns commencent toujours avec le dinucléotide GT- et se terminent par le dinucléotide –AG. De plus les bases qui précèdent –AG sont constituées de 75% de C ou de T, une condition qui favorise l'épissage (Rogan et Schneider, 2000). Chez les mammifères, la glutamine qui précède le site d'épissage (PM**Q**PLP), est conservée dans presque toutes les séquences d'amélogénine (voir "Evolution moléculaire de l'amélogénine chez les mammifères"). Cette glutamine est codée par le codon CAG reconnu comme site accepteur d'épissage alternatif intra-exonique 3'. L'analyse évolutive de l'amélogénine de reptile montre que la glutamine (localisée grâce à la comparaison des séquences reptiliennes avec les séquences mammaliennes) est conservée chez tous les crocodiles mais elle n'est pas aussi fortement conservée chez les squamates. Cela amène à s'interroger sur la présence ou non d'un épissage alternatif dans cette région de l'amélogénine, mais nous ne pouvons pas négliger la possibilité de l'existence d'autres sites d'épissage (motif CAG précédé par 75% de C/T). Cependant, les fragments de LRAP restent à trouver chez les reptiles.

Les 20 derniers résidus de l'exon 6 sont bien conservés chez tous les reptiles étudiés, mais leur fonction est toujours inconnue. Cependant, le domaine hydrophile de C-ter est le premier à être supprimé par la protéolyse et, de plus, il joue un rôle important dans la minéralisation de l'émail. La forte conservation observée pour ces résidus dans cette région est probablement liée à une fonction possible comme sites spécifiques de protéolyse. Mais ces sites restent à être définis. Cette région joue aussi un rôle crucial dans les propriétés agrégatives de l'amélogénine par sa contribution dans la stabilisation des nanosphères (Moradian-Oldak *et al.*, 1994). Ici aussi, les propriétés de chaque acide aminé dans cette région semble être aussi très importants car ces résidus ont été conservés durant plus de 300 millions d'années d'évolution, et cela malgré le fort taux de substitution observé chez les squamates.

La région variable dans l'exon 6

La région centrale de l'amélogénine (connue comme étant une région riche en répétitions de proline/glutamine chez les mammifères) est plus sujette à des variations chez les squamates que chez les crocodiles et les mammifères. Chez les squamates, cette région était impossible à aligner sans l'introduction de délétions/insertions ce qui signifie qu'un certain nombre de délétions et/ou d'insertions sont apparues dans chaque lignée de squamate. La forte variabilité de cette région est en opposition avec le fait que l'émail est bien structuré chez toutes les espèces étudiées. Cela indique que les acides aminés localisés dans cette région hydrophobe de l'amélogénine n'ont pas de contraintes fonctionnelles importantes. Cependant, la conservation du caractère hydrophobe de cette région pourrait être importante pour les inter-relations avec les autres protéines (Williamson, 1994). Cela pourrait expliquer la conservation d'un grand nombre de résidus de proline et de glutamine malgré le taux élevé de substitution de cette région. Cela indique qu'il y a probablement des contraintes pour conserver la structure générale et la composition de cette région en acides aminés spécifiques des régions hydrophobes. Chez les reptiles, cette région de l'amélogénine est plus riche en proline et plus pauvre en glutamine que chez les mammifères. Cette relative richesse en glutamine de l'amélogénine des mammifères pourrait être due à l'insertion de répétitions récentes (voir "évolution de l'amélogénine chez les mammifères") et de telles insertions n'existent pas dans les lignées reptiliennes. La présence de glutamine semble donc être moins importante pour l'amélogénine que ce qui avait été supposé précédemment dans l'étude de l'évolution de l'amélogénine chez les mammifères. En revanche, le nombre de prolines et leur présence répétée tous les trois résidus semblent être d'une grande importance pour le maintien de la structure typique de cette région. En effet, dans l'amélogénine des squamates, alors que le taux de substitution est très élevé, on observe seulement quelques substitutions de ce résidu comparé aux autres acides aminés.

L'insertion de répétitions de triplets n'existe pas dans les séquences d'amélogénine de reptiles comme cela a été décrit dans les séquences des mammifères (voir "évolution de l'amélogénine chez les mammifères"). A ce jour, nous ne connaissons pas le rôle que jouent ces répétitions dans l'amélogénine des mammifères. Elles contribuent probablement à une meilleure minéralisation de l'émail, ce qui lui permettrait de résister à l'abrasion durant de longues années. En effet, au contraire des reptiles qui remplacent leurs dents de nombreuses fois pendant leur vie, les mammifères ont perdu la polyphyodontie il y a 250 millions d'années. En revanche, l'émail mammalien est plus épais et plus solide que chez les reptiles.

Grâce à la comparaison des séquences ancestrales de mammifères et de squamates avec les séquences de crocodiles, il a été montré que des triplets d'acides aminés existent dans l'exon 6 de l'amélogénine des reptiles. Cela conforte l'hypothèse de la création de l'exon 6 (au moins dans sa région 3') par la répétition de triplets d'acides aminés. Ces triplets

sont basés sur le motif Pxx répété plusieurs fois. Notre analyse révèle que ces triplets ont été générés chez un ancêtre commun aux amniotes, probablement par l'insertion répétée d'une copie d'un Pxx isolé ou de la copie de plusieurs motifs Pxx, comme cela a été montré chez les crocodiles et les squamates. La meilleure conservation de ces triplets est trouvée dans l'amélogénine de crocodile qui a apparemment conservé plus de caractères ancestraux que dans la lignée des squamates. Le fait que la région 3' soit mieux conservée que la région 5' en tant que succession de triplets Pxx, pourrait être liée soit à des contraintes structurales soit à une insertion plus récente de ces répétitions.

Chez les squamates, la région centrale de l'exon 6 est la source de la plupart des variations phylogénétiques. Dans une même famille, les variations intra-génériques (2,7 à 4,4%) et inter-génériques (5,95 à 8,85%) dans les séquences nucléotidiques de l'exon 6, sont plus importantes que celles qui ont été observées chez les crocodiles et les mammifères. En effet, chez les trois crocodiles la comparaison inter-générique a révélé seulement 4 positions d'acides aminés substitués (0,8% de variations nucléotidiques et seulement 2,3% pour les comparaisons entre sous-familles). Chez les mammifères, les comparaisons intra-génériques n'ont révélé aucune substitution, même au niveau nucléotidique, et les variations inter-génériques dans une même famille étaient relativement faibles. Ces différences peuvent être expliquées soit par des relations de parenté plus proches entre les genres chez les crocodiles et chez les mammifères que chez les squamates, soit parce qu'il y a un taux plus élevé de substitution chez les squamates par rapport aux autres lignées, soit une superposition des deux phénomènes. En terme de distance évolutive, les lignées mammaliennes et archosauriennes/lépidosauriennes ont divergé il y a 310 MA alors que les squamates et les crocodiliens se sont séparés il y a 276 MA.

Cependant le peu de substitutions observé dans l'amélogénine des crocodiles est surprenant car les deux sous-familles, Alligatorinae et Crocodylinae, sont connues pour être séparées depuis plus de 66 millions d'années (Brochu, 1997). L'amélogénine de crocodile semble avoir conservée quelques caractères ancestraux suggérant que le taux de substitution, chez les crocodiles, est plus faible que dans les autres lignées. Cependant, nous avons à prendre en considération que ces données concernent seulement 3 espèces appartenant à 3 genres alors qu'il existe 30 espèces de crocodiliens appartenant à 10 genres. De nouvelles données sur les Gavialinae, l'autre sous-famille et sur d'autres genres, sont nécessaires pour compléter cette analyse.

Conclusion

La comparaison de 15 espèces représentatives des principales lignées de squamates et de 3 espèces de crocodiles a révélé que :

- les régions 5' et 3' du gène de l'amélogénine évoluent lentement et sont caractérisées par de nombreux acides aminés fortement conservés;
- la région centrale (exon 6) évolue plus rapidement et possède un pourcentage important de proline et de glutamine.

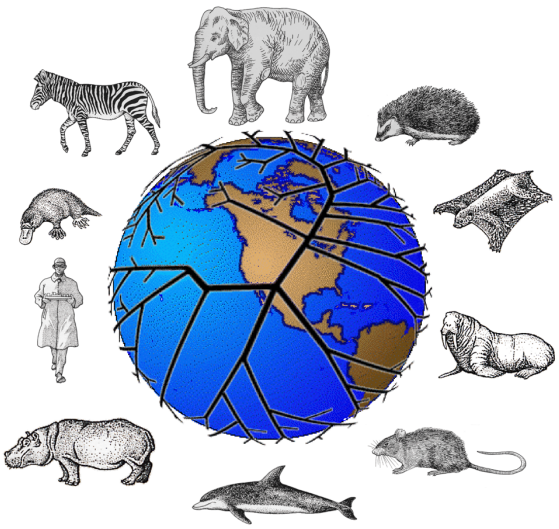
Une comparaison entre les séquences ancestrales de reptiles et de mammifères a montré que la région 3' de l'exon 6 est composée de nombreux triplets Pro-x-x, et que cette région pourrait avoir été créée grâce à l'insertion de tels triplets.

L'amélogénine de la lignée reptilienne partage avec la lignée mammalienne un certain nombre de caractéristiques qui indiquent une fonction similaire dans la formation de l'émail pour chaque domaine de la protéine. Cependant, la fonction d'un certain nombre de résidus,

fortement conservés, reste à connaître. Des études supplémentaires sur l'évolution de l'amélogénine dans la lignée des amphibiens anoures et urodèles et chez les sarcoptérygiens basaux (dipneustes) sont nécessaires pour apporter quelques réponses à ces questions.

Quatrième Partie

UN NOUVEL OUTIL MOLÉCULAIRE POUR L'ÉTUDE DE LA PHYLOGÉNÉSE DES MAMMIFÈRES ET DES SQUAMATES : L'AMÉLOGÉNINE



Les illustrations des animaux sont extraites du livre "Classification phylogénétique du vivant" Lecointre et le Guyader, 2001, Editions BELIN.
L'arbre phylogénétique provient du site "The Tree of Life" (http://tolweb.org/tree/phylogeny.html)

La première étude montrant que l'amélogénine possède un pouvoir phylogénétique concernait la séquence d'amélogénine de myxine (voir première partie). Par la suite, j'ai construit des phylogénies avec le gène de l'amélogénine pour vérifier l'origine des séquences obtenues dans la troisième partie ("Étude de l'évolution du gène de l'amélogénine") et éviter les contaminations d'ADN. Au cours de ces études phylogénétiques, il était évidant que l'amélogénine possède un fort signal phylogénétique. Encore restait-il à déterminer dans quel cadre il est possible d'utiliser ce gène. L'amélogénine peut-elle être utilisée pour résoudre les relations de parenté au sein des amniotes, des mammifères, des squamates ? Etant donné la divergence qui existe entre l'amélogénine des mammifères et des squamates, il était évident qu'il fallait traiter les séquences des deux lignées indépendamment.

UN NOUVEL OUTIL MOLÉCULAIRE POUR L'ÉTUDE DE LA PHYLOGÉNÉSE DES MAMMIFÈRES ET DES SQUAMATES : L'AMÉLOGÉNINE

Ce travail a été réalisé par :

Sidney Delgado*, **Marc Girondot†, Géraldine Véron§ et Jean-Yves Sire*.**

* UMR 8570, Évolution et Adaptations des Systèmes Ostéomusculaires, Équipe "Evolution et Développement du Squelette Dermique", Paris.
† UPRESA 8079, Laboratoire "Écologie, Systématique et Évolution", Équipe de Conservation des Populations et des Communautés, CNRS et Université Paris-Sud, Paris XI, , Bâtiment 362, 91405 Orsay Cedex.
§ Laboratoire de Zoologie, mammifères et oiseaux, Muséum national d'Histoire naturelle, Paris.

Remerciements: Je suis grandement reconnaissant auprès des collèges suivants qui ont fourni de l'ADN ou des tissus congelés ou conservés dans l'alcool: Dr F. Catzeflis (Université Montpellier 2), Dr. W. Dabin (Muséum de la Rochelle) et le Dr F. Ollivet et A. Lécu (Zoo de Vincennes, Muséum national d'Histoire naturelle, Paris). Je remercie également mes collègues japonais, M. Ishiyama, M. Mikami, A. Imai, H. Shimomura (Department of Histology and Oral Microbiology, The Nippon Dental University School of Dentistry at Niigata, Japon) qui nous ont donné les séquences du gène de l'amélogénine chez trois squamates.

Résumé

Si les outils moléculaires ne manquent pas pour la réalisation de phylogénies, en revanche tous n'ont pas la même spécificité et les résultats sont parfois très controversés. Certains gènes se sont avéré être de très bons outils pour résoudre les relations de parenté dans certains groupes, et au contraire inefficaces dans d'autres. En plus de l'ADN mitochondrial qui permet de travailler avec plus de 10 000 pb, de nombreux gènes nucléaires ont été utilisés à ce jour dans la réalisation de phylogénies moléculaires, comme le gène *C-mos*.

Nos récents travaux ont montré que le gène de l'amélogénine, codant pour la protéine majeure de l'émail, était très bien conservé au cours de l'évolution et qu'il possédait également un bon signal phylogénétique. Dans cette étude nous avons voulu tester ce signal à la fois chez les mammifères et les squamates pour lesquels nous possédions 30 et 14 séquences, bien réparties dans les principales lignées. Nous avons tenté de reconstituer les relations de parenté dans ces deux groupes à partir d'un fragment du gène de l'amélogénine d'environ 400 pb (situé dans l'exon 6). Chez les mammifères, les grandes lignées se sont séparées en un temps très court ce qui diminue l'information phylogénétique disponible dans les gènes. Chez les squamates, au contraire, les différentes familles se sont séparées depuis très longtemps et le génome accumule plus de mutations que chez les autres amniotes, les crocodiliens et chez les mammifères, ce qui a tendance à brouiller le signal phylogénétique.

Les résultats obtenus chez les mammifères confirment l'existence de deux groupes d'euthériens : les Laurasiathériens et les Afrothériens tels qu'ils ont été révélés récemment par des études moléculaires (Madsen *et al.*, 2001; Murphy *et al.*, 2001). En accord avec ces travaux, notre étude confirme également la proximité des primates et des rongeurs et permet de déterminer que les dermoptères et les scandentiens sont plus proches des primates que des rongeurs. Le résultat le plus inattendu et le rapprochement des périssodactyles (chevaux, zèbres) avec les xénarthres (paresseux, tamanoir), un groupe longtemps considéré comme basal. Ces résultats qui, d'une part, confirment certains des travaux les plus récents et, d'autre part, apportent des regroupements nouveaux, montrent que l'amélogénine possède un bon signal phylogénétique chez les mammifères.

Chez les squamates, malgré quelques incertitudes, nos résultats confirment un certain nombre de relations phylogénétiques connues, comme la position basale du gecko ou la proximité des varanidés et des serpents. Notre étude apporte également des résultats originaux comme les relations entre lacertidés et anguidés, et la séparation des scincidés et des

lacertidés. Ce dernier résultat est le plus étonnant car ces groupes sont réunis dans les Scincomorpha depuis longtemps. Dans ce cas également, l'amélogénine apparaît comme un gène pertinent au niveau phylogénétique.

Tous ces résultats doivent encore être complétés par l'apport de nouvelles données à la fois chez les mammifères pour confirmer le regroupement de certains clades, de certaines branches, et chez les squamates où l'échantillonnage est encore assez incomplet pour pouvoir déduire de nos résultats des affirmation irréfutables.

Introduction

Les dents et leurs tissus, dentine et émail, ont une longue histoire chez les vertébrés et, parfois, ils constituent les seuls restes de vertébrés rencontrés dans les gisements fossiles. Les dents des vertébrés actuels sont dérivées d'éléments dermiques, les odontodes, qui se trouvaient à la surface du squelette dermique extra-oral des premiers vertébrés sans mâchoires, il y a quelques 500 millions d'années (voir revues dans Smith et Hall, 1993; Huysseune et Sire, 1998; Smith et Coates, 2000; Sire et Huysseune, 2002). Une fois localisées dans la bouche des premiers vertébrés à mâchoires, les dents, et les tissus qui les composent, ont subi une très forte sélection en raison de leur fonction capitale pour la prise de nourriture. Cela explique pourquoi les dents, et surtout leur structure, ont été conservées chez la plupart des lignées de vertébrés actuelles. Chez les mammifères et les squamates (lézards et serpents) les dents sont recouvertes d'un émail épais et fortement minéralisé. L'amélogénine est la principale protéine de la matrice de l'émail en formation où elle représente, chez les mammifères, plus de 90% du contenu protéique.

En raison de son importance en tant que protéine majeure d'un tissu protégeant un organe aussi vital que la dent, on peut penser que l'amélogénine est sujette à très peu de variations. En fait, deux études récentes sur l'évolution de l'amélogénine chez les mammifères et les squamates (voir chapitres précédents) ont montré que cette protéine était très conservée dans certaines régions, hydrophiles, alors qu'elle présentait des variations dans une région plus centrale et hydrophobe, correspondant à une grande partie de l'exon 6 (environ 120 acides aminés). Ces variations sont plus importantes en nombre chez les squamates que chez les mammifères, et elles représentent un fort signal phylogénétique (voir chapitre "Analyse évolutive de l'amélogénine de la myxine"). Il était alors très tentant d'essayer de vérifier si cette région de l'amélogénine pouvait représenter un bon outil pour construire la phylogénie de ces deux grandes lignées de vertébrés amniotes que sont les mammifères et les squamates.

L'intérêt que peut représenter un nouvel outil en phylogénie moléculaire n'est pas négligeable dans la mesure ou les relations de parenté dans ces grands groupes sont encore le sujet d'un grand nombre de controverses. Chez les mammifères, au-delà de la division entre métathériens et euthériens (Simpson, 1945) la difficulté réside dans la radiation explosive, et récente (moins de 60 millions d'années), des euthériens qui a donné naissance à une phylogénie en étoile qui ne rassemble pas l'accord de tous les auteurs (Shoshani, 1986; Bulmer *et al.,* 1991). Plusieurs phylogénies utilisant les méthodes moléculaires fondées sur les séquences nucléotidiques de divers gènes (mitochondriaux, nucléaires, etc.) ont été proposées ces dernières années. Certains regroupements obtenus par ces méthodes rejoignent ceux qui ont été obtenus par des analyses morphologiques alors que la position de certains clades diffère fortement en fonction des différentes approches moléculaires. Ainsi, la monophylie de certains groupes a été remise en question, de même que les relations entre différents ordres de mammifères. Récemment, de nouvelles phylogénies proposées par Madsen *et al.* (2001) et par Murphy *et al.* (2001) ont conduit à un nouveau consensus mais, toutefois, les résultats diffèrent à propos de la position de certains clades. Il est clair que la rapide radiation des mammifères euthériens pose toujours un problème. Chez les squamates il existe aussi de nombreuses controverses et les relations de parentés entre les grandes lignées n'ont pas encore été bien définies (*e. g*., Harris *et al*., 1999). Cependant, les problèmes rencontrés sont différents de ceux relatifs aux mammifères. D'une part, l'origine des grandes familles de squamates remonte à une radiation probablement très ancienne (Kumar et Hedges, 1998) et, d'autre part, le génome de ce clade semble évoluer plus rapidement que chez les autres grandes lignées, ce qui provoque une saturation des méthodes d'analyses moléculaires (Hughes et Mouchiroud, 2001).

Dans le présent travail, le signal phylogénétique que pouvait représenter le gène de l'amélogénine a donc été testé chez les mammifères et les squamates. Cette étude phylogénétique a été réalisée à partir des 30 séquences disponibles chez les mammifères et des 14 séquences de squamates (voir chapitres précédents "Evolution moléculaire de l'amélogénine chez les mammifères et chez les squamates") à l'aide de différentes méthodes d'analyse (distance, parcimonie et maximum de vraisemblance).

Matériel et méthode

Matériel biologique

Mammifères

Voir chapitre "Evolution de l'amélogénine chez les mammifères, Tableau 1" pour la liste des espèces étudiées et la Figure 1, ci-après, pour les liens de parenté de ces espèces.

Définitions

Pour une meilleure compréhension de la discussion, il est nécessaire de définir certains des clades qui vont y être cités. Ces définitions s'appuient à la fois sur de récentes phylogénies moléculaires (Madsen *et al.*, 2001; Murphy *et al.*, 2001) et sur la littérature plus ancienne (résumée dans Lecointre et Le Guyader, 2001).

Afrothériens: ce groupe a été créé récemment (Madsen *et al.*, 2001; Murphy *et al.*, 2001). Il comprendrait les proboscidiens (éléphants), les siréniens (dugong, lamantin), les tubulidentés (oryctérope), les tenrecs (autrefois placés dans les insectivores), les chrysochloridés (taupe dorée) et les macroscélidés.

Archontes : signifiant "pénis pendant", ce clade inclut les primates, les scandentiens (toupaye), les dermoptères (lémur volant) et les chiroptères (chauves-souris).

Carnivores : groupe très diversifié comprenant les félidés (chat, guépard, ...), les canidés (fennec, chien, loup,...), les pinnipèdes (morse, phoque, etc.), les ursidés, etc.

Cétartiodactyles : regroupement des artiodactyles (ruminants, suinés et tylopodes) et des cétacés (dauphin, baleine, ...). Autrefois groupe-frère des artiodactyles, les cétacés y sont maintenant inclus comme groupe-frère des hippopotamidés.

Xénarthres : mammifères dépourvus de dents ou d'émail et comprenant les fourmiliers, les tatous et les paresseux. Autrefois, les xénarthres faisaient partie des "édentés" qui incluaient aussi les pholidotes (pangolin) avant que l'on découvre que les pangolins n'étaient pas apparentés aux xénarthres. Le mot "édentés" est encore utilisé dans la littérature comme synonyme de xénarthres.

Insectivores : ce groupe comprenait les hérissons, les musaraignes, les taupes et les tenrecs. Dans les phylogénies moléculaires récentes (Stanhope *et al.*, 1998; Madsen *et al.*, 2001; Murphy *et al.*, 2001), le groupe des insectivores n'est plus monophylétique car les tenrecs sont maintenant placés dans les afrothériens, le nom du groupe étant maintenant les "euinsectivores".

Laurasiathériens : ce groupe a été créé récemment (Madsen *et al.*, 2001; Murphy *et al.*, 2001). Il comprendrait les cétartiodactyles, les périssodactyles, les euinsectivores, les chiroptères et les carnivores.

Ongulés : clade ancien qui regroupait les cétartiodactyles, les périssodactyles, les siréniens et les proboscidiens.

Périssodactyles : ils incluent les chevaux, les zèbres, les rhinocéros et les tapirs ainsi que le groupe des hyracoïdes dont les représentants ressemblent à des marmottes.

Téthythériens : ils incluent les proboscidiens et les siréniens.

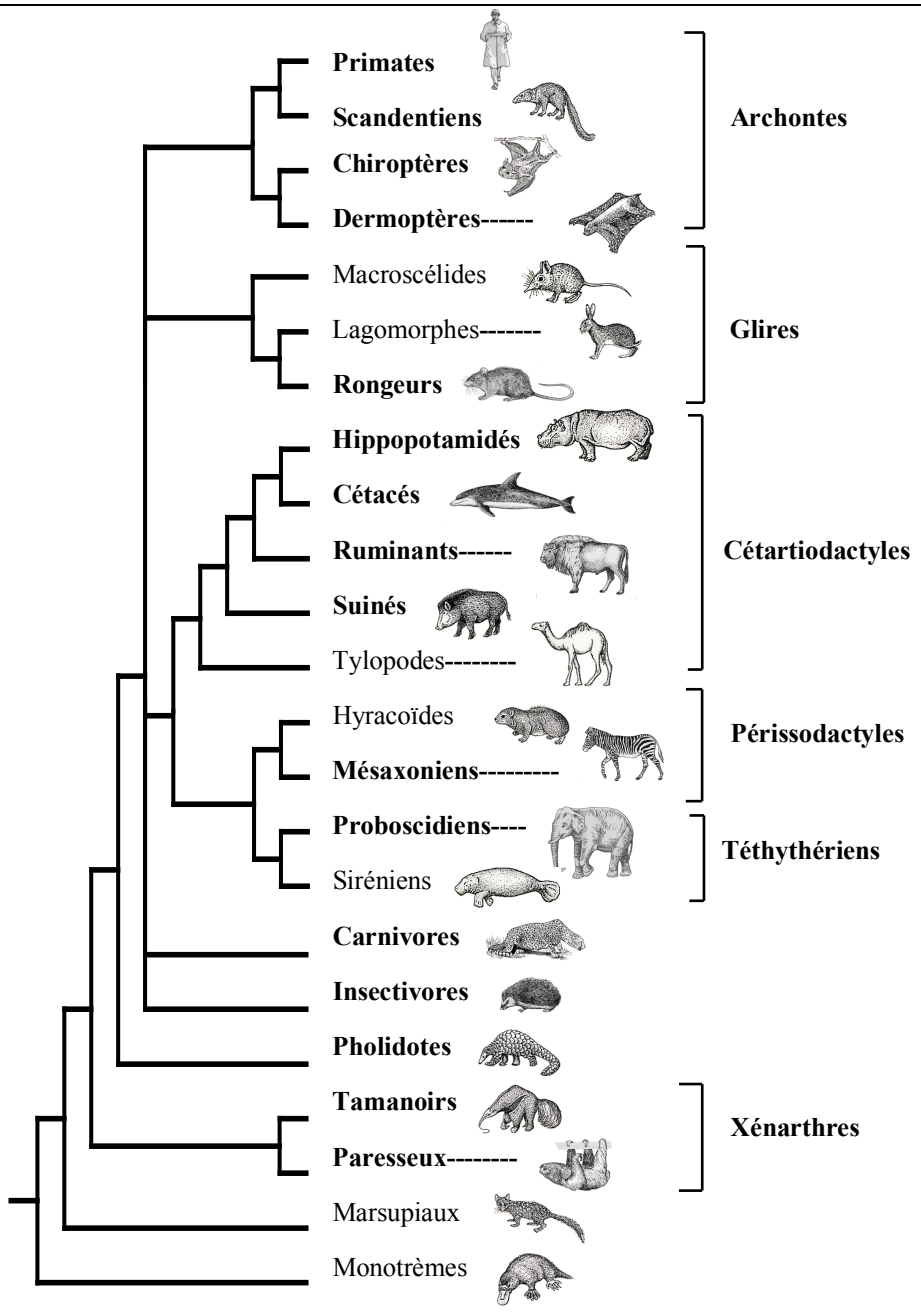


Figure 1. Relations de parenté des principales lignées de mammifères d'après Lecointre et Le Guyader (2001), un consensus des phylogénies des mammifères existantes. Les taxons en gras ont été étudiés dans le présent travail.

Squamates

Voir chapitre "Evolution de l'amélogénine chez les reptiles, Tableau 1" pour la liste des espèces de squamates étudiées et la Figure 2, ci-après, pour les liens de parentés de ces espèces.

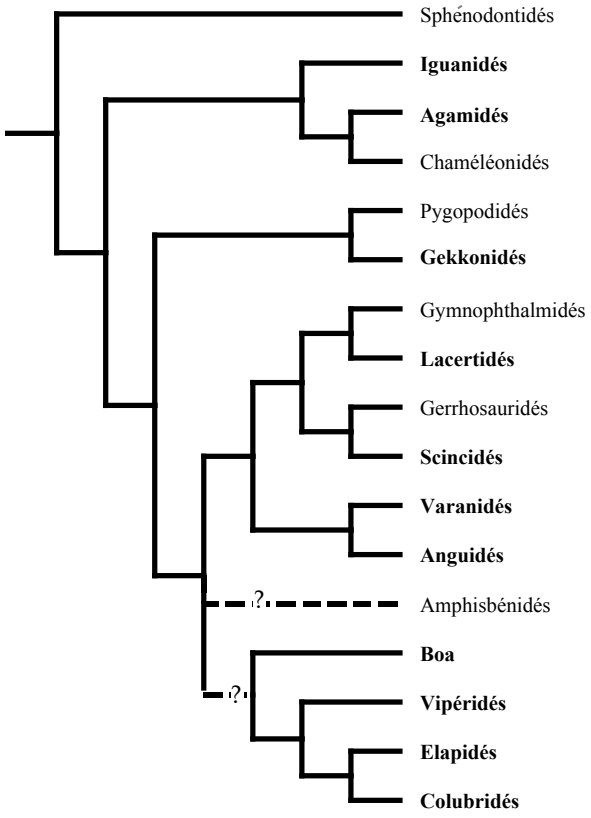


Figure 2. Relations de parenté des principales lignées de squamates d'après Estes *et al.* (1988). Les taxons en gras ont été étudiés dans le présent travail. La position des serpents et des amphisbaenidés n'est pas bien déterminée.

Analyses phylogénétiques

Mammifères

Les séquences d'amélogénine des chromosomes Y ont été exclues de l'analyse phylogénétique car elles évoluent indépendamment des séquences d'amélogénine des chromosomes X (voir chapitre "Evolution moléculaire de l'amélogénine chez les mammifères").

La région du gène de l'amélogénine utilisée pour l'analyse phylogénétique comprend 341 pb (après alignement); elle est localisée dans l'exon 6, le plus grand exon du gène et le seul accessible par PCR sur ADNg. Cette région commence à la 41ème base de l'exon 6 et se termine 53 bases avant la fin de cet exon. Cela représente la seule partie de l'exon 6 connue chez tous les taxons utilisés. Les insertions localisées au centre de l'exon 6 chez certaines espèces ont été exclues de l'analyse sachant quelles n'ont pas de signification évolutive puisqu'elles ont été acquises indépendamment (voir chapitre "Evolution moléculaire de l'amélogénine chez les mammifères").

Les alignements de séquences ont été réalisés avec CLUSTAL X (1.8) (Thompson *et al.*, 1997), puis corrigés à la main à l'aide de Se-Al (2.0).

Les analyses phylogénétiques ont été réalisées avec PAUP 4.0 (Swofford, 1993) en utilisant les méthodes de phénétique, de parcimonie et de maximum de vraisemblance.

Les analyses phénétiques ont été réalisées à l'aide de l'algorithme BioNJ (Gascuel, 1997), une version améliorée du neighbor-joining, avec les mesures de distance suivantes : Tajima-Nei (Tajima et Nei, 1984), Kimura 2 paramètres (Kimura, 1980), Kimura 3 paramètres (Kimura, 1981), Tamura-Nei (Tamura et Nei, 1993), F84 (Felsenstein, 1984), HKY85 (Hasegawa *et al.*, 1985) et maximum de vraisemblance. Les caractéristiques de ces différents modèles sont décrites dans PAUP 4.0 (Swofford, 1993).

Les analyses en parcimonie ont été réalisées avec l'algorithme "Heuristique" (ou "algorithme exact"), la méthode "Exhaustive" ne pouvait pas être utilisée pour cette analyse en raison du nombre trop élevé de taxons étudiés et la méthode du "Branch-and-bound" s'avérait trop longue pour réaliser les calculs.

Le maximum de vraisemblance a été utilisé avec l'algorithme de "Quartet puzzling", dans un modèle de substitution qui suppose que les transitions sont deux fois plus nombreuses que les transversions et des fréquences de bases utilisant le modèle à 2 paramètres de Hasegawa, Kishino et Yano (1985).

Les arbres ont été enracinés avec les monotrèmes (ornithorynque et échidné) considérés comme les mammifères actuels les plus basaux (Rowe, 1999).

Squamates

Seules les séquences d'ADN ont été utilisées dans ce travail car une étude préliminaire utilisant les séquences protéiques a montré de nombreuses polytomies et une très faible solidité des branches.

La région du gène de l'amélogénine utilisée pour l'analyse phylogénétique comprend 393 pb (après alignement) localisées dans l'exon 6, depuis la 39ème base jusqu'à 24 bases avant la fin de cet exon. Cette partie est la seule qui est connue chez tous les taxons étudiés.

Les alignements de séquences ont été réalisés de la même façon que chez les mammifères et les analyses phylogénétiques (distance, parcimonie et maximum de vraisemblance) ont utilisé les mêmes méthodes.

Les arbres ont été enracinés avec la méthode des points moyens car je ne disposais pas du groupe-frère des squamates, les rhynchocéphales (sphénodon) et l'amélogénine des crocodiliens était trop divergente ce qui rendait l'alignement difficile à utiliser en raison d'un trop grand nombre d'insertions/délétions.

Résultats

Mammifères

Etude par la méthode de distance (ou méthode phénétique)

Les analyses des séquences de l'exon 6 du gène de l'amélogénine par la méthode de distance révèlent plusieurs groupes monophylétiques qui sont toujours conservés quels que soient les modèles évolutifs utilisés (Figure 2).

Les dermoptères (lémur volant), les scandentiens (toupaye) et les primates (homme) forment un groupe monophylétique dont les rongeurs sont toujours le groupe-frère.

Cet ensemble (primates/dermoptères/scandentiens/rongeurs) est lui-même le groupe-frère d'un grand clade qui comprend : les carnivores (guépard, fennec, phoque), les artiodactyles (vache, chèvre, cochon, hippopotame), les périssodactyles (cheval, zèbre), les cétacés (dauphin), les pholidotes (pangolin), les euinsectivores (hérisson), les chiroptères (chauve-souris) et les xénarthres (paresseux et tamanoir).

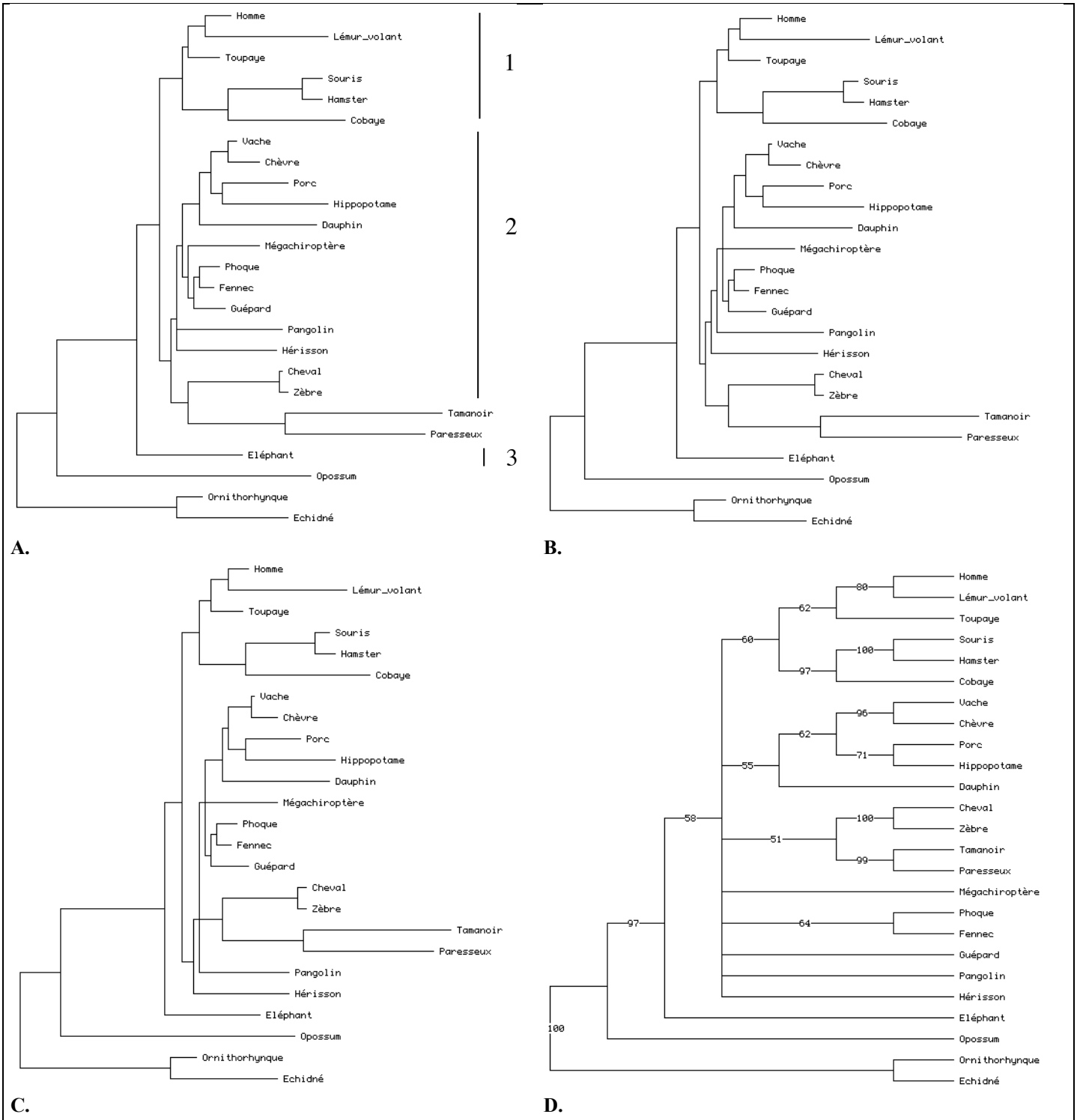


Figure 3. Résultats des études phylogénétiques avec la méthode de distance. **A.** Arbre de distance obtenu par la méthode de BioNJ à partir du modèle de Tamura-Nei (Tamura et Nei, 1993) On obtient un arbre identique avec le modèle Juke-Cantor (Juke et Cantor, 1969) et Tajima-Nei (Tajima et Nei, 1984). **B.** Arbre de distance obtenu par la méthode de BioNJ à partir du modèle de Kimura 3 paramètres (Kimura, 1981); **C.** Arbre de distance obtenu par l'algorithme de BioNJ utilisant les distances du maximum de vraisemblance; **D.** Arbre bootstrap obtenu par la méthode de BioNJ à partir de 1000 tests de bootstrap. Seules les valeurs supérieures à 50% sont indiquées. 1 = clade III; 2 =Laurasiathériens; 3 = Afrothérien (groupes définis par Madsen *et al.* (2001); Murphy *et al.* (2001)).

Dans ce vaste ensemble on retrouve la monophylie des carnivores, ainsi que celle des artiodactyles dont les cétacés est le groupe frère. Enfin, on observe un clade qui inclut les périssodactyles avec les xénarthres.

Les proboscidiens (éléphant) constituent un ensemble à part, groupe-frère des autres euthériens.

Trois taxons ont une position qui varie quelque peu selon l'algorithme utilisé. Il s'agit des chiroptères, des pholidotes et des euinsectivores. Les chiroptères sont soit le groupe-frère des carnivores (Figure 3A) soit en trichotomie avec les carnivores et les cétartiodactyles (Figure 3B). Les pholidotes sont soit le groupe-frère de l'ensemble cétartiodactyles /carnivores/chiroptères (Figure 3B), soit en trichotomie avec les euinsectivores et le groupe cétartiodactyles/carnivores/chiroptères (Figure 3A). La méthode de BioNJ avec les distances du maximum de vraisemblance donne un résultat légèrement différent puisque la position des pholidotes est plus basale.

En effet, ils forment une trichotomie avec les périssodactyles/xénarthres et les carnivores/cétartiodactyles/chiroptères (Figure 3C). Enfin, les euinsectivores sont soit le groupe-frère de l'ensemble cétartiodactyles/carnivores/pangolin (Figure 3B), soit en trichotomie avec ces derniers et le groupe périssodactyles/xénarthres (Figure 3A). La méthode de BioNJ avec les distances du maximum de vraisemblance indique pour les euinsectivores une position plus basale comme groupe frère de tous les taxons cités ci-dessus (périssodactyles, cétartiodactyles, carnivores, pholidotes et chiroptères) (Figure 3C).

Etude par la méthode de parcimonie

La méthode de parcimonie a permis d'obtenir 34 arbres minimum qui ont été résumés en deux arbres consensus (Figure 4). L'un représente le consensus strict (Figure 4A) et l'autre le consensus obtenu par la règle de 50% de majorité (Figure 4B).

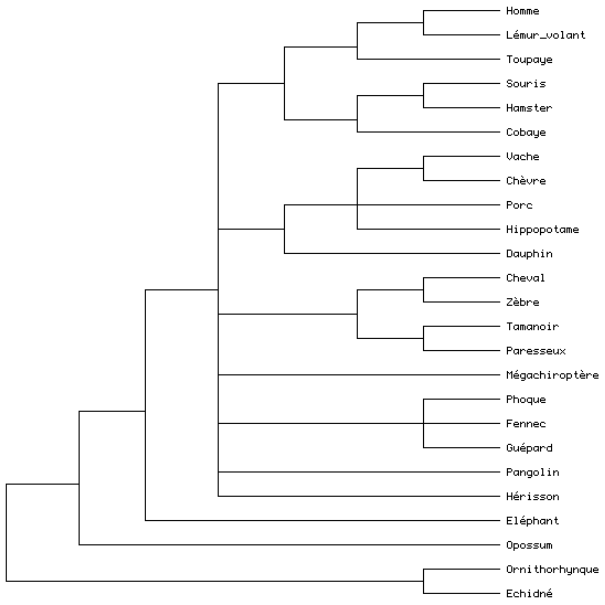


Figure 4.A. Consensus strict des 34 arbres obtenus par la méthode de parcimonie à l'aide de l'algorithme heuristique.

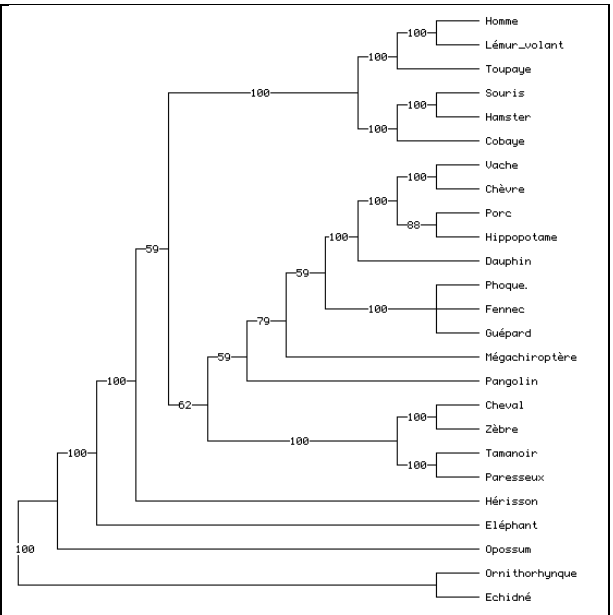


Figure 4.B. Consensus (règle de 50% de majorité) des 34 arbres obtenus par la méthode de parcimonie à l'aide de l'algorithme heuristique.

Dans l'arbre de consensus strict, on retrouve les mêmes groupes monophylétiques que dans l'analyse précédente : (primates/dermoptères/scandentiens)/rongeurs, cétartiodactyles, périssodactyles/xénarthres, carnivores. Comme précédemment, tous ces groupes, avec

également les pholidotes, les euinsectivores et les chiroptères, forment un ensemble d'où l'éléphant est exclu. Le consensus strict ne permet pas de résoudre les branchements à l'intérieur de ce grand groupe qui exclut l'éléphant. Ils forment une polytomie.

L'arbre de consensus obtenu par la règle de 50% de majorité permet de regrouper les taxons qui sont présents dans plus de 50% des 34 arbres en indiquant la fréquence des branchements. Les carnivores deviennent le groupe-frère des cétartiodactyles, les chiroptères sont le groupe frère de cet ensemble, puis les pholidotes, les périssodactyles/xénarthres et enfin les rongeurs/(primates/dermoptères/scandentiens). Les euinsectivores deviennent le groupe-frère de l'ensemble, puis les proboscidiens.

Maximum de vraisemblance

La méthode des quartets (Figure 5) permet d'obtenir un arbre identique à l'arbre de distance (Figure 3A) sauf pour les pholidotes qui deviennent le groupe-frère du clade périssodactyles/xénarthres (voir précédemment).

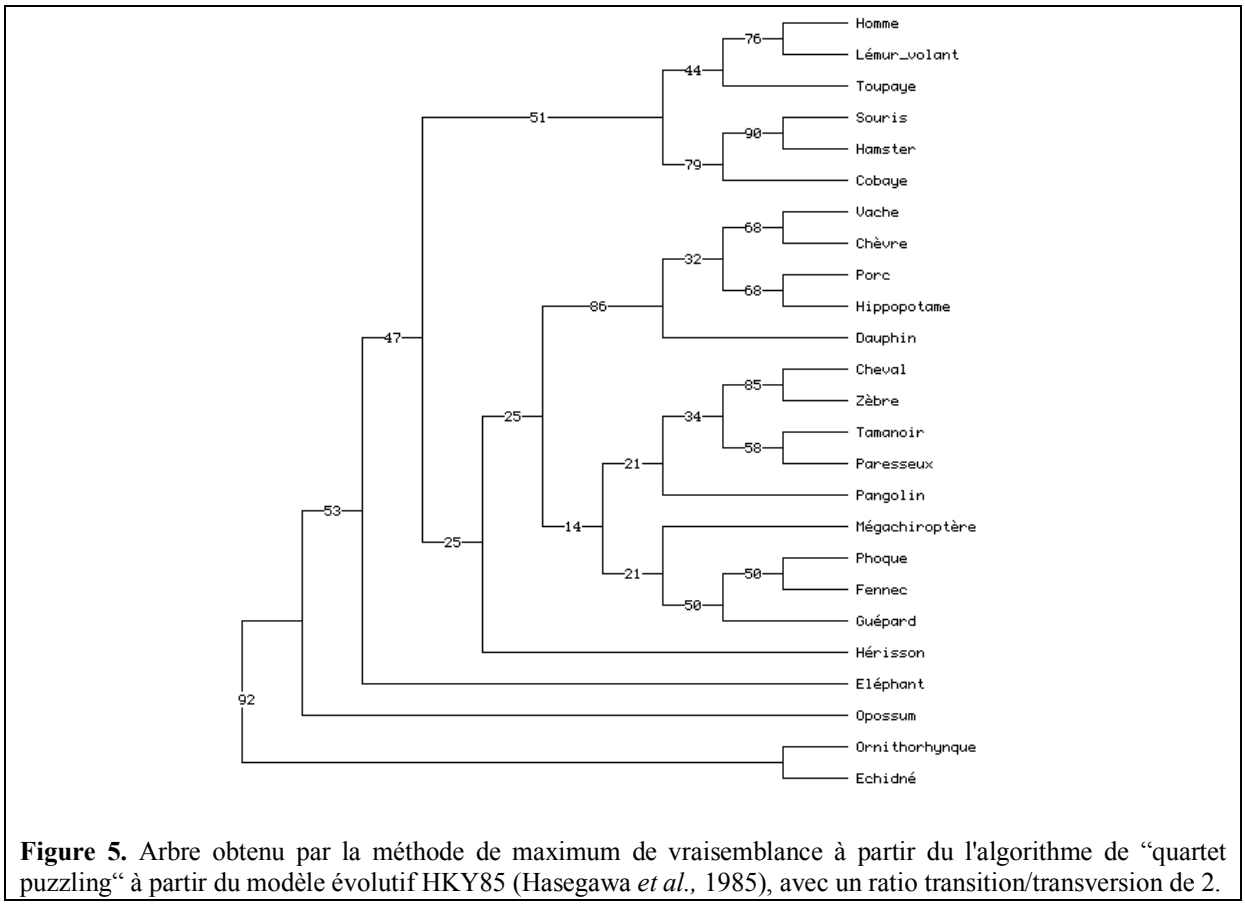


Figure 5. Arbre obtenu par la méthode de maximum de vraisemblance à partir du l'algorithme de "quartet puzzling" à partir du modèle évolutif HKY85 (Hasegawa *et al.,* 1985), avec un ratio transition/transversion de 2.

Squamates

Etude par la méthode de distance (ou méthode phénétique)

Tous les modèles évolutifs utilisés (Figure 6) aboutissent au même arbre dans lequel les gekkonidés constituent le groupe le plus basal des squamates et les serpents (boa, vipère, serpent corail et élaphe) forment un groupe monophylétique avec les boidés comme groupe basal. Les varanidés sont le groupe-frère des serpents (Figure 6A). Les lacertidés (*Gallotia*,

lézard des murailles et psammodrome) forment également un groupe monophylétique dont les anguidés (orvet) sont le groupe-frère. Ces deux clades forment un ensemble dont les agamidés (agame) sont le groupe-frère. En position plus basale, on trouve les iguanidés (iguane) puis les scincidés (*Chalcides*). Le test de bootstrap montre que certaines branches ont des valeurs très fortes alors que d'autres sont très faibles (en dessous de 50%) et aboutissent à une polytomie (Figure 6B).

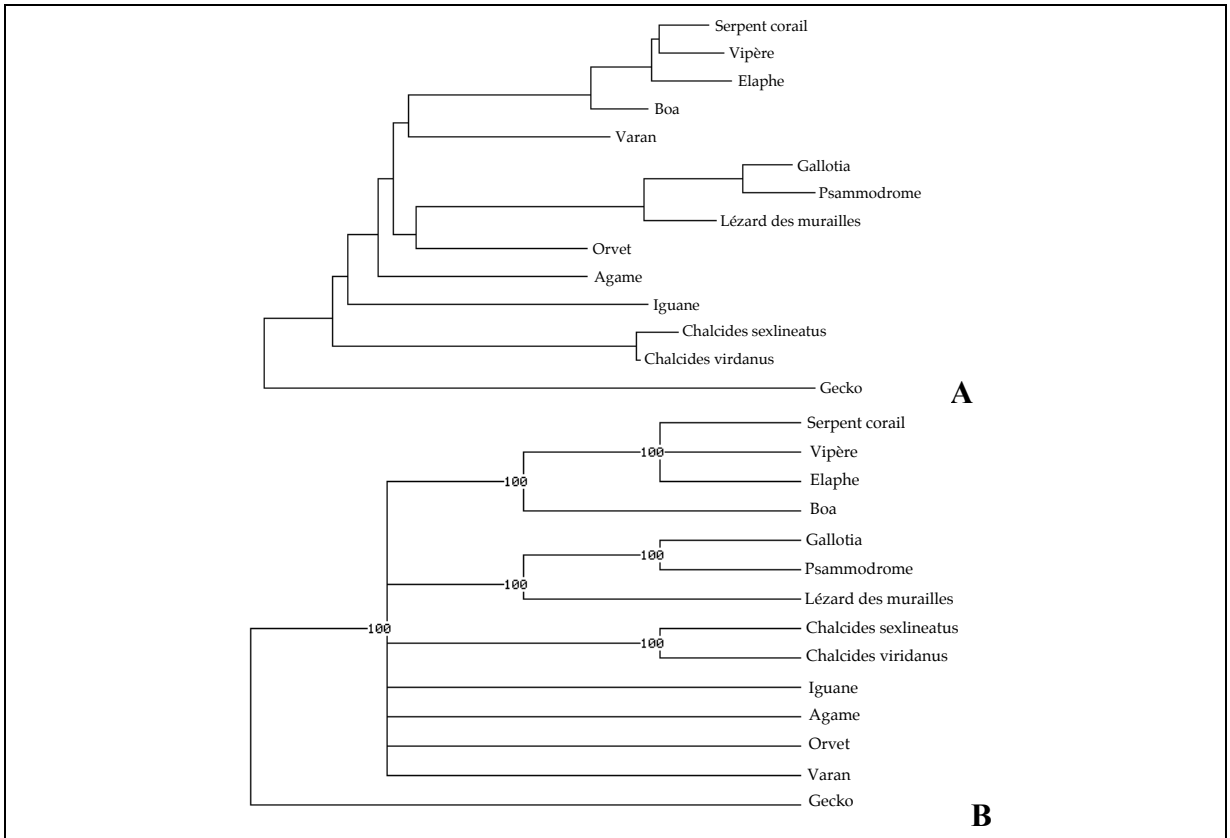


Figure 6. A. Arbre de distance obtenu par la méthode de BioNJ à partir du modèle de Tamura-Nei (Tamura et Nei, 1993). On obtient un arbre identique avec les autres modèles évolutifs existant dans PAUP 4.0. **B.** Arbre bootstrap obtenu par la méthode de BioNJ à partir de 1000 tests de bootstrap. Seules les valeurs supérieures à 50% sont indiquées.

Parcimonie

La recherche d'arbres en parcimonie permet d'obtenir un seul arbre de longueur minimum dans lequel on trouve les mêmes regroupements (serpents, lacertidés) qu'avec la méthode de distance. Les varanidés sont apparentés aux serpents et les anguidés sont proches des lacertidés. Seule la place des agamidés change puisqu'ils sont placés comme groupe-frère des serpents/varanidés (Figure 7).

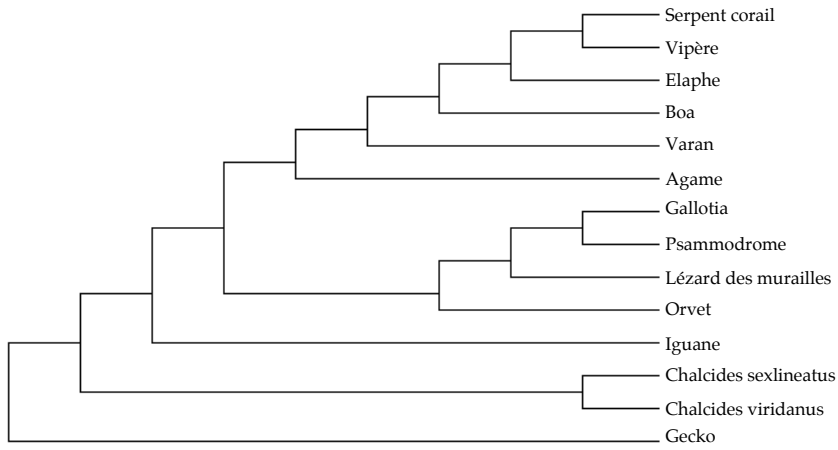


Figure 7. Arbre le plus court obtenu par la méthode de parcimonie à l'aide de l'algorithme heuristique.

Maximum de vraisemblance

La méthode de maximum de vraisemblance permet d'obtenir un arbre ayant la même topologie que les arbres précédents, mais à une exception près, les iguanidés, qui sont regroupés avec les scincidés (Figure 8). Cependant, le pourcentage de regroupement de ces branches est assez faible ce qui indique un branchement peu solide.

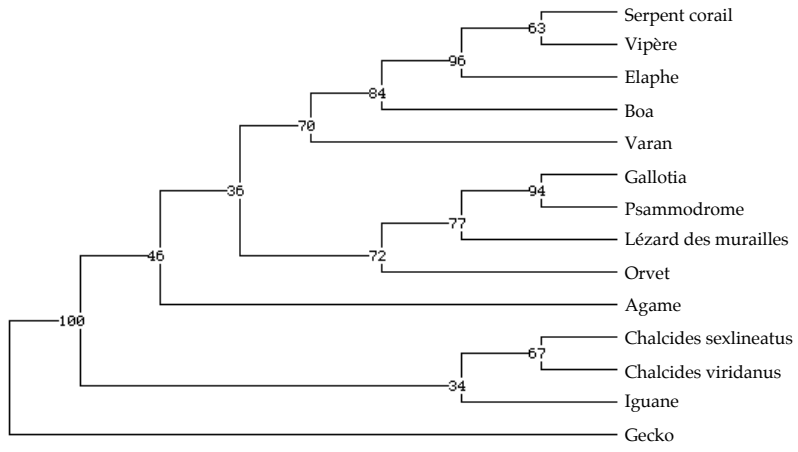


Figure 8. Arbre obtenu par la méthode de maximum de vraisemblance à partir du l'algorithme de "quartet puzzling" à partir du modèle évolutif HKY85 (Hasegawa *et al.,* 1985), avec un ratio transition/transversion de 2.

Discussion

Mammifères

Bilan des analyses phylogénétiques

Toutes les méthodes phylogénétiques utilisées dans cette étude s'accordent sur plusieurs points. Premièrement, il existe deux clades d'euthériens, l'un constitué par les proboscidiens, seul clade d'afrothériens représenté ici, et l'autre incluant tous les autres clades étudiés. Deuxièmement, ce dernier grand groupe est subdivisé en deux ensembles. Le premier

(clade 1) contient les primates, les scandentiens, les dermoptères et les rongeurs, et le second (clade 2), contient les cétartiodactyles, les carnivores, les pholidotes, les euinsectivores et les chiroptères. Au sein du clade 1, on observe deux groupes monophylétiques. Dans le premier, qui constitue les rongeurs, la souris est plus proche du hamster que du cochon d'inde ; dans le second, les primates sont plus proches des dermoptères que des scandentiens.

Dans le clade 2, on observe trois groupes monophylétiques. Les carnivores (guépard, fennec et phoque), les cétartiodactyles (dauphin, vache, chèvre, porc et hippopotame) et les xénarthres/périssodactyles (paresseux + tamanoir et cheval + zèbre). Dans tous les arbres, les cétartiodactyles sont toujours plus proches des carnivores que des xénarthres/périssodactyles.

En revanche, les différents algorithmes utilisés ne s'accordent pas tout à fait sur la place de trois taxons. Il s'agit des euinsectivores (hérisson), des chiroptères (chauve-souris géante) et des pholidotes (pangolin). Ils sont tous les trois situés dans le clade 2. En revanche, il reste des incertitudes sur leur place dans ce clade. Les chiroptères sont proposés comme groupe-frère des carnivores dans plusieurs de nos analyses, mais ils pourraient être également le groupe-frère de l'ensemble carnivores/cétartiodactyles. Cette branche n'est pas très solide, ce qui est démontré par l'arbre de bootstrap (Figure 3C). Les pholidotes sont aussi mal définis, mais ils apparaissent souvent comme le groupe-frère de l'ensemble carnivores /cétartiodactyles/chiroptères. Enfin, les euinsectivores apparaissent comme le groupe le plus basal des laurasiathériens, mais il reste à déterminer s'ils sont plus basaux que le clade périssodactyles/xénarthres.

Comparaisons avec les récentes phylogénies

Les données récentes de la phylogénie moléculaire des mammifères (Stanhope *et al.*, 1998; Madsen *et al.*, 2001; Murphy *et al.*, 2001), basées sur l'analyse de 2600 à 10000 pb provenant de gènes nucléaires (jusqu'à 15 gènes selon les études) et mitochondriaux, suggèrent l'existence de 4 grands clades d'euthériens (Figure 9):

- le clade III (appelé ainsi par Murphy *et al.*, (2001)) regroupant les rongeurs, les primates, les dermoptères et les scandentiens, sans toutefois s'accorder sur les relations de parenté entre ces divers groupes (Figure 9, n°1).

- les laurasiathériens, comprenant les cétartiodactyles, les périssodactyles, les euinsectivores, les chiroptères et les carnivores, sans toutefois s'accorder sur les relations de parenté entre ces divers clades (excepté pour les cétartiodactyles qui sont toujours regroupés avec les périssodactyles) (Figure 9, n°2).

- les xénarthres (Figure 9, n°3).

- les afrothériens (Figure 9, n°4).

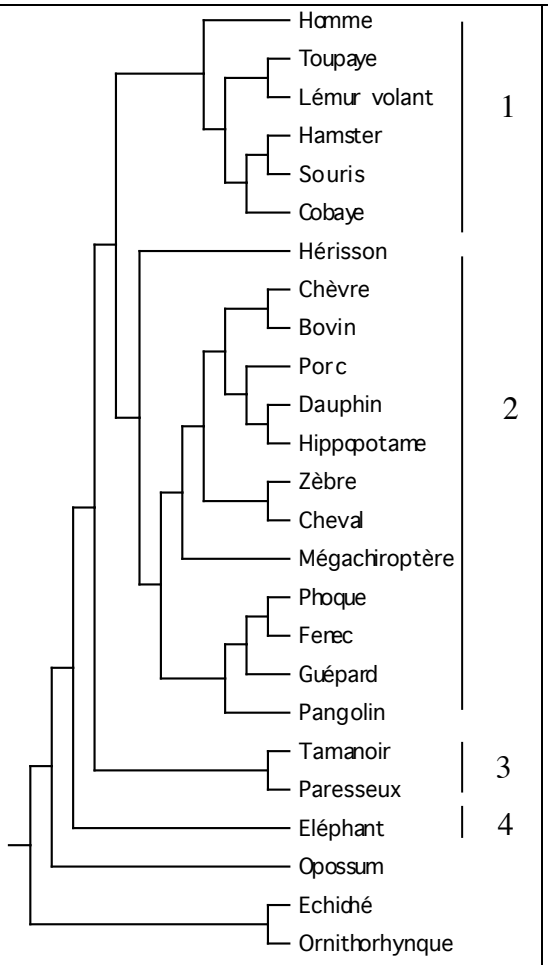


Figure 9.A. Arbre phylogénétique des mammifères d'après Madsen *et al.* (2001). Seules les espèces que nous avons étudiées y ont été placées.

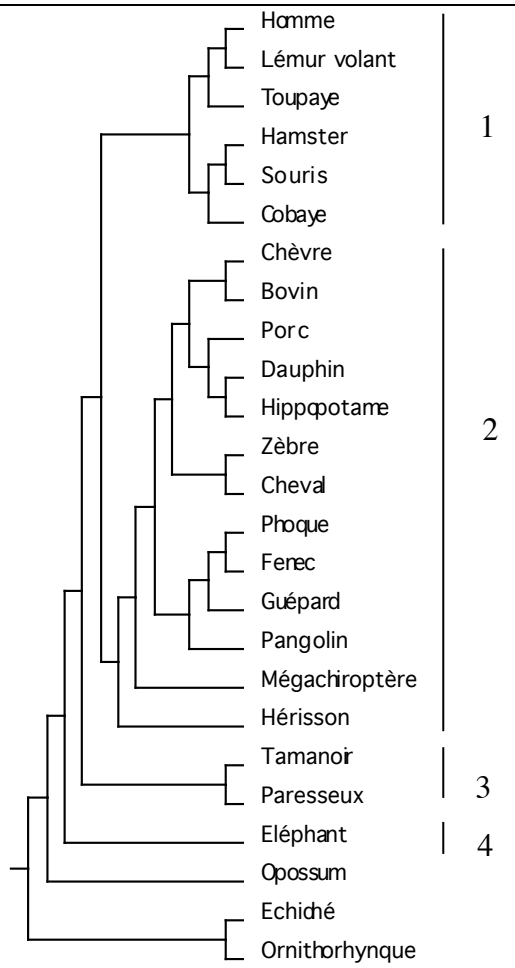


Figure 9.B. Arbre phylogénétique des mammifères d'après Murphy *et al.* (2001). Seules les espèces que nous avons étudiées y ont été placées.

Ces données récentes apportées par la phylogénie moléculaire ont pour conséquence de voir disparaître certains clades comme les ongulés, les archontes et de morceler les insectivores. Dans les phylogénies plus anciennes, les taxons que l'on retrouve maintenant regroupés au sein les afrothériens étaient dispersés. Par exemple, il était admis depuis longtemps que les proboscidiens, les siréniens et les macroscélidés étaient proches parents des périssodactyles. Ils faisaient partie d'un groupe appelé altongulés (voir dans Lecointre et Le Guyader, 2001; et dans "The Tree of Life ": http://tolweb.org/tree/phylogeny.html (consensus de 82 articles)). A présent, les altongulés n'existeraient plus et les périssodactyles seraient très éloignés des proboscidiens. De la même manière, les tubulidentés, qui étaient auparavant placés à la base des ongulés, et les tenrécidés, qui étaient classés dans les insectivores, font à présent partie des afrothériens.

La création du grand clade des laurasiathériens entraîne le rapprochement des carnivores, des euinsectivores, des pholidotes, des chiroptères et des périssodactyles /artiodactyles, alors que ces taxons appartenaient tous à des clades différents. Il s'agit d'un changement radical des liens de parenté entre ces taxons par rapport à leur position précédente. La nouvelle position des chiroptères supprime, par exemple, l'ancien groupe des archontes dans lequel se trouvaient les chiroptères, les primates, les scandentiens et les dermoptères.

Les résultats obtenus avec l'amélogénine s'accordent parfaitement avec l'existence des laurasiathériens, des afrothériens et du clade III (voir Figure 3A) confirmant ainsi les études moléculaires (Stanhope *et al.*, 1998; Madsen *et al.*, 2001; Murphy *et al.*, 2001). Les proboscidiens (seuls représentants des afrothériens dans cette étude) sont séparés du reste des euthériens et les laurasiathériens rassemblent les carnivores, les cétartiodactyles, les périssodactyles, les pholidotes, les euinsectivores et les chiroptères (Figures 9 et 10). De même, le regroupement des rongeurs, des primates, des scandentiens et des dermoptères correspond au clade III (Figures 9A et B). Les laurasiathériens sont plus proches du clade III que des afrothériens ce qui est parfaitement en accord avec les résultats de Madsen *et al.* (2001) et de Murphy *et al.* (2001).

En revanche, la position des xénarthres est différente de celle qui est proposée dans les études précédentes. Quelle que soit la méthode que nous avons utilisée, les xénarthres sont toujours les plus proches parents des périssodactyles (cheval et zèbre) (Figures 3, 4 et 5). Par ailleurs, le test de bootstrap montre une valeur importante pour ce regroupement (100%) qui est donc solide dans cette étude. Jusqu'à présent, les xénarthres étaient très difficiles à placer au sein des mammifères. Selon les phylogénies, ils sont soit des euthériens basaux (Lecointre et Le Guyader, 2001) soit le groupe-frère des afrothériens (Stanhope *et al.*, 1998; Madsen *et al.*, 2001) soit le groupe-frère de tous les non-afrothériens (Murphy *et al.,* 2001). Il ne faut cependant pas négliger le fait que les deux espèces étudiées ici ont un émail très réduit (même absent chez l'adulte tamanoir). Cette réduction de l'émail correspond sans doute à une dégénération de ce tissu dans ce groupe, ce qui pourrait expliquer l'accumulation de mutations dans le gène codant pour l'amélogénine. Cette évolution plus rapide pourrait être alors à l'origine de la position de ce groupe non conforme à la littérature.

Dans toutes nos analyses, le dauphin est proposé comme groupe-frère des artiodactyles (Figures 3, 4 et 5). Bien que cette position est restée longtemps admise sur des bases morphologiques (voir dans Rowe, 1999), elle est assez controversée aujourd'hui dans les phylogénies moléculaires récentes (Madsen *et al.*, 2001; Murphy *et al.,* 2001) et d'autres, plus anciennes (voir dans Tree of Life), dans lesquelles il est proposé que les cétacés sont proches des hippopotamidés. L'amélogénine de dauphin présente, cependant certaines divergences (insertions/délétions) qui peuvent expliquer son exclusion du reste des artiodactyles; cependant, malgré ces importantes délétions, il est intéressant de noter que le gène de l'amélogénine parvient à conserver un signal phylogénétique important qui permet aux cétacés de rester apparentés aux artiodactyles. Il serait donc intéressant de compléter l'échantillon des cétacés par d'autres espèces (cachalot, etc.)

Entre les primates, les scandentiens et les dermoptères, l'étude phylogénétique montre un rapprochement des primates et des dermoptères, les scandentiens étant groupe frère des précédents. Ce résultat est en accord avec les travaux de Murphy *et al.* (2001), mais pas avec ceux Madsen *et al.* (2001), pour lesquels les scandentiens et les dermoptères sont plus proches des rongeurs que des primates, ni avec les études précédentes (voir dans Lecointre et Le Guyader, 2001) pour qui les scandentiens sont plus proches des primates que les dermoptères. La topologie que nous avons obtenue est cependant assez solide comme en témoignent les valeurs de bootstrap.

Conclusion

Des études moléculaires récentes ont bouleversé la vision que l'on avait des relations de parenté des mammifères. Cependant, ces études nécessitaient une confirmation avec d'autres phylogénies moléculaires. Notre étude confirme ces résultats et offre un nouvel outil pour l'étude des liens de parenté entre les grands groupes de mammifères : le gène de l'amélogénine. Par ailleurs, ces résultats, dont beaucoup confirment les relations de parenté

entre certains groupes de mammifères, ont été obtenus avec 380 pb seulement alors que les précédents utilisaient de 2600 à 10 000 pb ce qui indique que le gène de l'amélogénine contient beaucoup d'informations phylogénétiques. Cependant, cette étude nécessite d'être complétée par de nouvelles données afin de pour confirmer certains résultats. Le séquençage de l'amélogénine chez d'autres afrothériens que les proboscidiens est en cours (siréniens, tenrécidés, tubulidentés, chrysochloridés) afin de confirmer les relations de parenté au sein de ce clade. D'autres données sont aussi essentielles afin de conforter certaines relations de parenté. Il serait par exemple souhaitable d'obtenir des données chez d'autres insectivores (taupe, musaraigne), d'autres cétacés (cachalot, …), d'autres périssodactyles (rhinocéros). Ces données nouvelles permettront sans doute de confirmer le regroupement xénarthres/périssodactyles, qui n'a jamais été décrit auparavant. Elles permettrons aussi, de déterminer avec plus de précision la position des chiroptères, des pholidotes et des euinsectivores.

Squamates

Bilan des analyses phylogénétiques

Les différentes méthodes phylogénétiques qui ont été utilisées dans cette étude s'accordent pour définir des relations de parenté au sein des squamates. Les gekkonidés sont le groupe le plus basal et les scincidés forment le groupe-frère des autres familles de squamates (lacertidés, anguidés, serpents et varanidés). Les varanidés sont proches parents des serpents et les anguidés sont apparentés aux lacertidés. Au sein des serpents, dont le groupe le plus basal est les boïdés, les vipéridés sont plus proches des élapidés que des colubridés. Parmi les lacertidés, le genre *Gallotia* est plus proche du genre *Psammodromus* que de *Lacerta*.

En revanche, la position des agamidés est mal définie. Selon les phylogénies, ce clade est soit proche parent des serpents/varanidés (Figures 7 et 8) soit groupe-frère de tous les squamates en dehors des scincidés, des iguanidés et des gekkonidés (Figure 6A). Il existe également une incertitude sur la position des iguanidés. Ils seraient apparentés soit aux scincidés (Figure 8) soit ils occuperaient une position moins basale (Figures 7 et 6A).

La méthode de bootstrap révèle que les branches des arbres obtenus ne sont pas très solides.

Comparaisons avec les récentes phylogénies

Les phylogénies des squamates obtenues à partir de données morphologiques (Estes *et al.,* 1988) ont été peu discutées dans la littérature et, malgré les données récentes d'Evans et Barbadillo (1998), les relations de parenté n'ont pas évolué (Figure 10A). Trois phylogénies moléculaires ont été réalisées en ce qui concerne les squamates. Harris *et al.* (1999) ont utilisé le gène *C-mos* (375 pb) de 30 squamates dont 3 serpents de la même famille (boïdés) (Figure 10B). Auparavant, en 1995, Heise *et al.* avaient publié la phylogénie des serpents (36 espèces) fondée sur le gène *C-mos* et Fu (2000) celle des lacertidés basée sur l'ADN mitochondrial (plus de 50 espèces).

D'après les caractères morphologiques, le groupe le plus basal des squamates est celui qui rassemble les iguanidés et les agamidés, et la place des serpents n'est pas résolue (Figure 10A). Les scincidés sont les plus proches parents des lacertidés et les anguidés sont proches des varanidés.

Les relations de parenté déduites des données moléculaires diffèrent de celles qui ont été obtenues par les données morphologiques par la position basale des gekkonidés (Harris *et*

al., 1999). En revanche, les relations de parenté entre agamidés et iguanidés, varanidés et anguidés, et scincidés et lacertidés sont conservées. Les serpents formeraient une trichotomie avec les deux ensembles précédents.

Les relations phylogénétiques obtenues à l'aide du gène de l'amélogénine sont en accord avec les travaux de Harris *et al.* (1999) pour ce qui concerne la position du taxon le plus basal, les gekkonidés. De plus, les relations de parenté au sein des lacertidés sont en accord avec la phylogénie de Fu (2000). En revanche, les relations de parenté au sein des serpents sont en désaccord avec la phylogénie des serpents communément admise (par exemple Heise *et al.,* 1995), excepté pour la position basale des boïdés. En effet, les travaux de Heise *et al.* (1995) indiquent que les élapidés (serpent corail) sont le groupe-frère des colubridés (*Elaphe*) et non celui des vipéridés (vipère) comme le montrent nos résultats.

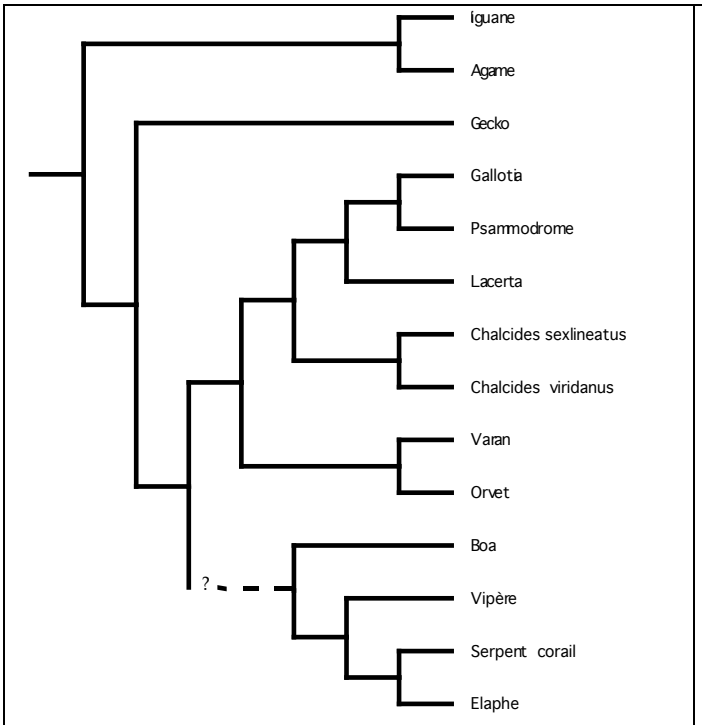


Figure 10.A. Arbre phylogénétique des squamates d'après Estes *et al.* (1988) (Heise *et al.* (1995) pour la phylogénie des serpents et Fu (2000) pour celle des lacertidés). Seules les espèces que nous avons étudiées y ont été placées.

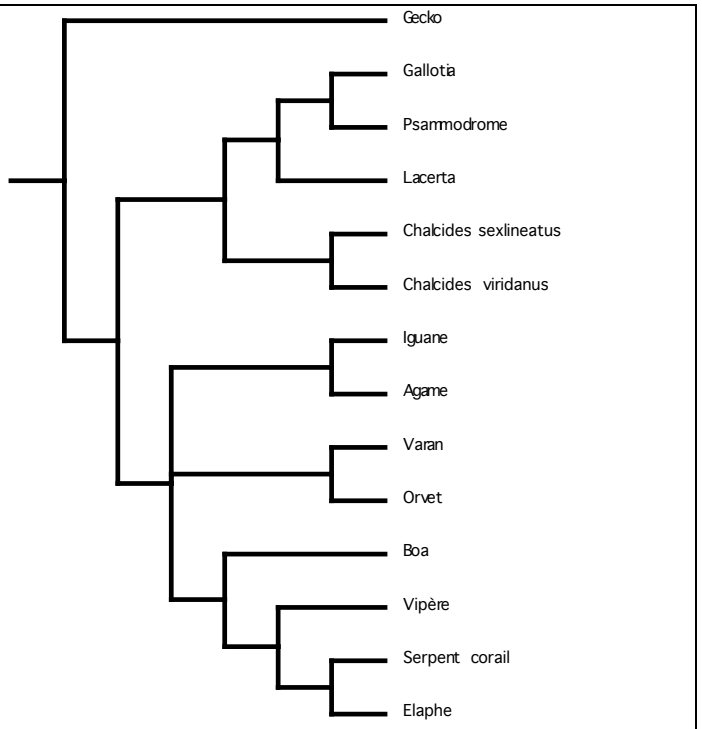


Figure 10.B. Arbre phylogénétique des squamates, d'après Harris *et al.*, (1999) (Heise *et al.*, 1995 et Fu, 2000). Seules les espèces que nous avons étudiées y ont été placées.

Le résultat le plus inattendu est la séparation entre les scincidés et les lacertidés alors que les autres phylogénies les placent ensemble dans les Scincomorpha (Estes *et al.,* 1988; Evans et Barbadillo, 1998; Harris *et al.,*1999). En supposant que les scincidés soient effectivement proches des lacertidés, ce qui semble assez probable au vu de la littérature, deux hypothèses peuvent être proposées pour expliquer les résultats. Soit le genre *Chalcides* a été classé à tort dans les scincidés et il appartiendrait en réalité à un groupe plus basal dont la position est alors déterminée par notre analyse, soit *Chalcides* appartient effectivement aux scincidés, mais pour expliquer la "mauvaise" position de ce groupe, il faut supposer que le gène de l'amélogénine de *Chalcides* a dû accumuler plus de mutations que celui des lacertidés. Cette dernière hypothèse est soutenue par l'étude du gène de l'amélogénine chez les squamates (voir chapitre "Evolution du gène de l'amélogénine"). En effet, l'amélogénine de

Chalcides présente des mutations (insertions/délétions) que l'on ne retrouve pas chez les autres espèces. Est-ce suffisant pour que les scincidés soient séparés des lacertidés ? Pour en savoir plus sur cette question, il faudrait inclure dans cette étude des séquences d'autres genres de scincidés afin d'augmenter la robustesse de l'analyse.

Conclusion

Quelques incertitudes demeurent sur les relations de parenté obtenues avec l'amélogénine chez les squamates. Notre analyse montre en effet que certaines branches sont peu soutenues. Pour résoudre ces incertitudes, une solution serait d'augmenter l'échantillonnage (il existe plus de 20 familles de squamates dont on ne connaît pas le gène de l'amélogénine). En effet, du fait que les familles de squamates se sont séparées sur un temps beaucoup plus long que celles des mammifères, les différents genres et espèces ont pu accumuler plus de mutations. Cela peut conduire dans certains cas au phénomène bien connu "d'attraction des longues branches". Pour éviter ce phénomène, il est alors possible d'ajouter des taxons afin de "casser" ces longues branches. Ce phénomène peut expliquer la position basale des scincidés ou l'incertitude sur la position des iguanidés.

De plus, on sait que les squamates accumulent plus de mutations que les autres groupes de vertébrés (crocodiliens et mammifères) (Hughes et Mouchiroud, 2001) et ce phénomène pourrait brouiller le signal phylogénétique. L'élimination des troisièmes bases des codons, trop variables, pourrait améliorer les résultats ainsi que l'utilisation de séquences de plus grande taille (mais cela nécessite d'obtenir la séquence des autres exons).

Quoi qu'il en soit, si le signal phylogénétique du gène de l'amélogénine se révèle trop peu discriminant pour séparer les différentes familles de squamates, il pourrait cependant devenir un bon outil pour résoudre les relations de parenté à l'intérieur de ces familles comme il a été possible de le constater avec les lacertidés. Pour le vérifier il faudrait obtenir de nombreuses séquences de lézards dans différentes familles (lacertidés, scincidés, etc.). Enfin, pour bien enraciner l'arbre des squamates, la séquence du sphénodon (tuatara) est absolument nécessaire. Une collaboration est en cours avec la Nouvelle-Zélande (seule région où il existe) pour obtenir de l'ADN de ce groupe basal des lépidosauriens.

CHAPITRE 3
-
AMÉLOGÉNINE ET DÉVELOPEMENT DENTAIRE CHEZ *CHALCIDES VIRIDANUS* (SQUAMATE, SCINCIDÉ)

Le chapitre précédent était entièrement consacré aux études moléculaires et il était nécessaire, pour que ce travail soit complet, de faire le lien entre les molécules et les niveaux d'intégration supérieurs : les cellules et les tissus dentaires.

La deuxième partie de ce chapitre a donc pour but d'étudier de l'expression de l'amélogénine dans les tissus dentaires d'un squamate: *Chalcides viridanus*. Suite au séquençage du gène de l'amélogénine chez ce lézard, il était important dans un contexte évolutif, de comparer l'expression du gène avec celle, déjà bien étudiée, des mammifères. L'originalité de ces travaux est d'avoir utilisé pour la première fois (1) une sonde d'ADN spécifique de l'amélogénine de *Chalcides* pour détecter les ARNm qui s'expriment dans les dents en formation et (2) un anticorps spécifique de l'exon 3 de l'amélogénine de ce lézard pour étudier l'expression de la protéine. A ce jour, ce type de technique n'avait jamais été utilisé chez un reptile et la seule étude d'expression de l'amélogénine existant pour cette lignée utilisait des anticorps anti-amélogénine mammalienne (Ishiyama *et al.*, 1998). Une étude histologique du développement des dents de *Chalcides* a également été réalisée afin d'identifier précisément les stades au cours desquels l'émail était formé.

En préambule à ces travaux, il était nécessaire d'étudier la dentition et le mode de remplacement dentaire chez cet animal pour mieux définir les régions et les niveaux des coupes histologiques à réaliser pour étudier des bourgeons dentaires avec de l'émail en formation. Ce travail constitue donc la première partie de ce chapitre.

Première Partie

LE PATRON DE REMPLACEMENT DENTAIRE CHEZ *CHALCIDES* (SQUAMATA; SCINCIDAE)

Mandibule gauche d'un Chalcides viridanus observé au microscope électronique à balayage

LA DENTITION ET LE PATRON DE REMPLACEMENT DENTAIRE CHEZ *CHALCIDES* (SQUAMATA; SCINCIDAE)

Traduit de l'anglais à partir de l'article à paraître dans

Journal of Morphology,

The Dentition and Tooth Replacement Pattern in *Chalcides* (Squamata; Scincidae),

par **Sidney Delgado**, Tiphaine Davit-Béal & Jean-Yves Sire

UMR 8570, Équipe "Evolution et Développement du Squelette Dermique", Paris.

Remerciements: Nous sommes reconnaissant envers le Prof. A. Huysseune (Université de Gand, Belgique) et le Dr M. Laurin (Paris) pour leurs commentaires et leurs suggestions sur le manuscrit. Nous remercions le Prof. J. Castanet pour nous avoir fourni les spécimens de *Chalcides* et pour son expertise dans l'estimation de l'âge. M.M. Loth est remerciée pour son assistance dans la réalisation des coupes.

Résumé

Cette étude a été entreprise comme une étape préalable aux investigations sur la différenciation des dents chez un squamate, les scincidés des Canaries du genre *Chalcides*. Notre but principal était de déterminer si le patron de remplacement des dents, connu pour être régulier chez les lézards, pouvait être utile pour prévoir exactement chaque étape du développement des dents. Une série de croissance de 20 spécimens élevés en laboratoire, âgés de 15 jours à environ 6 ans, a été utilisée. La dentition (dents fonctionnelles et dents de remplacement) a été étudiée à l'aide de radiographies de quadrants de mâchoire. Le nombre de positions, le nombre de dents en fonction de l'âge et des saisons et la taille des dents de remplacement ont été enregistrés. Chez les *Chalcides*, une rangée simple de dents fonctionnelles pleurodontes se trouve à la marge labiale du dentaire**,** du prémaxillaire et du maxillaire. Quel que soit l'âge des spécimens, 16 positions ont été enregistrées, en moyenne, dans chaque quadrant de mâchoire suggérant ainsi que les positions sont maintenues tout au long de la vie. Les dents de remplacement sont nombreuses, indépendamment de l'âge et de la saison, alors que le nombre de dents fonctionnelles est variable. La symétrie dans le développement des dents a été évaluée en comparant les dents deux par deux de chaque côté opposé dans les quatre quadrants des mâchoires de plusieurs spécimens. Bien que la taille relative de quelques dents de remplacement corresponde parfaitement, le critère de symétrie n'est pas fiable pour prévoir le stade de développement de la dent opposée, que la paire de dents comparées soit gauche-droite ou supérieure-inférieure. La meilleure correspondance a été trouvée en comparant la taille des dents de remplacement successives antéro-postérieurement sur la mâchoire. Chaque dent de remplacement qui a atteint 40 à 80% de sa taille définitive est suivie, sur l'arcade, par une dent de remplacement qui est en moyenne 20% moins développée. En considérant les dents en position alternée (séries paires et impaires), chaque dent de remplacement était un peu plus développée que la précédente (de 0,5 à 20% quand les dents ont atteint 10 à 40% de leur taille finale). Le remplacement des dents a lieu en position alternée de l'arrière vers l'avant formant ainsi des rangées plus ou moins régulières (c'est-à-dire selon la théorie du "Zahnreihen"). Chez les *Chalcides*, le stade de développement d'une dent de remplacement à une position p sur la mâchoire peut être exactement prévu à partir du stade de développement de la dent de remplacement à la position p-1 ou, à un degré moindre, à la position p-2. Cette observation sera particulièrement utile pour mener les études structurales et ultrastructurales de la différenciation des dents chez ce lézard.

Introduction

Cette étude fait partie d'un vaste programme de recherche consacré au développement et à l'évolution du squelette dermique et particulièrement des tissus dentaires, qui montrent une grande diversité chez les vertébrés (voir revue dans Huysseune et Sire, 1998). Actuellement, notre intérêt est de retracer l'histoire évolutive d'un tissu superficiel fortement minéralisé, l'émail et particulièrement de sa protéine majoritaire, l'amélogénine (Girondot et Sire, 1998 ; Delgado *et al*., 2001). Si chez les mammifères, l'amélogénèse est bien connue, l'information manque dans d'autres taxa (*e. g.*, Carlson, 1990 ; Sander, 1999). Nous avons voulu commencer notre programme en étudiant l'expression de l'amélogénine pendant la différenciation des dents dans la lignée la plus proche des mammifères, c'est-à-dire chez les reptiles. Notre but principal est de chercher les différences temporelles et/ou spatiales qui pourraient expliquer les différences structurales de l'émail (émail prismatique chez les mammifères *versus* émail sans prismes chez la plupart des reptiles).

Cependant, une étape préalable à un tel travail est de connaître le modèle de dentition dans l'espèce reptilienne prise pour modèle. En effet, il est particulièrement recommandé de savoir où et quand les stades appropriés de la formation de l'émail peuvent être trouvés le long des mâchoires chez des spécimens récemment nés, des juvéniles et des adultes, puisque l'expression de l'amélogénine peut seulement être étudiée dans des dents en différenciation (les protéines de l'émail étant entièrement dégradées par les protéases pendant le processus de maturation de l'émail). Parmi les reptiles, seuls les crocodiles et les lépidosaures (sphénodon, serpents et lézards) possèdent des dents. Il est difficile d'obtenir des séries de croissance d'espèces reptiliennes dans la nature ou en laboratoire. C'est pourquoi nous avons profité du succès de notre élevage de deux espèces de scincidés du genre *Chalcides* (un lézard), *C. sexlineatus* et *C. viridanus*.

Par opposition avec les mammifères qui ont seulement deux générations de dents (diphyodontie), la plupart des reptiles remplacent leurs dents tout au long de leur vie (polyphyodontie). Par conséquent, toutes les étapes de développement des dents (et donc de la formation de l'émail) sont disponibles chez un *Chalcides* juvénile ou adulte, il n'est donc pas nécessaire de chercher des germes de dent chez l'embryon. De plus, les dents de scincidés, comme celles de la plupart des squamates, sont régulièrement remplacées par vagues successives en position alternée (*e. g.*, Edmund, 1960, 1969). Un patron de remplacement si régulier implique qu'il pourrait être possible, en théorie, de prévoir le stade de développement de n'importe quelle dent dans la mâchoire d'un spécimen de *Chalcides*, pourvu que le patron de remplacement soit bien connu pour cette espèce. Dans la famille Scincidae, le patron de remplacement des dents a été brièvement décrit par Edmond dans plusieurs genres, mais pas chez *Chalcides*, pour qui la dentition (c'est-à-dire, le nombre de dents fonctionnelles et de remplacement) tout comme le patron de remplacement sont inconnus.

La présente étude a été entreprise pour répondre à la question suivante : le patron de remplacement des dents chez les *Chalcides* peut-il nous aider à prévoir chaque étape de différenciation des dents ? Pour s'assurer d'une prédiction précise nous devions savoir i) si le remplacement des dents varie avec l'âge et/ou avec les saisons, ii) s'il y a une symétrie dans le remplacement des dents (côtés gauches/droits et/ou supérieurs/inférieurs des mâchoires) et iii) si les dents sont remplacées régulièrement par vagues successives comme cela a été décrit dans la littérature (*e. g.*, Edmund, 1960, 1969 ; Cooper, 1965, 1966 ; Osborn, 1970, 1972, 1973, 1974, 1978). Avec ces questions à l'esprit, nous avons étudié la dentition et le patron de remplacement des dents dans une série de croissance de 20 *Chalcides* au moyen de la radiographie.

Matériels et méthodes

Matériel biologique

Deux espèces de scinques des Canaries, *Chalcides sexlineatus* Steindachner, de l'île Gran Canaria et *C. viridanus* (Gravenhorst), de Tenerife, sont élevées dans notre laboratoire depuis le printemps 1996, quand 30 adultes ont été capturés dans la nature et apportés au laboratoire. Les deux espèces de *Chalcides* ont été élevées dans des conditions contrôlées, c'est-à-dire, une période "d'hiver" de 2,5 mois, de la fin novembre à la mi-février (mais avec une température ambiante ne tombant jamais en-dessous de 10°C grâce à des sources de chaleur qui se déclenchent à 10°C) et une période "tempérée" (température ambiante) pour le reste de l'année. Les mâles et les femmes de chaque espèce ont été distribués dans plusieurs aquariums (80cm de long / 40cm de large, avec un fond sablonneux) et nourris (sauf pendant la période "d'hiver") deux fois par semaine avec des larves d'insectes (surtout des asticots).

Les scinques des Canaries sont vivipares. Deux à trois jeunes scinques par femelle ont été obtenus chaque année depuis l'été 1996, à partir de la deuxième moitié de juin jusqu'à la fin juillet. Peu après leur naissance les jeunes scinques ont été placés dans de petits aquariums (40/20 cm) et nourris avec de petits asticots.

Un total de 20 spécimens (11 *Chalcides viridanus* et 9 *C. sexlineatus*), âgés de 15 jours après la naissance à 6 ans ont été utilisés. Tous les lézards ont été mesurés (du museau au cloaque), décapités et leur mâchoire disséquée. Les mâchoires inférieures ont été sectionnées au niveau de la symphyse et les mâchoires supérieures au niveau de la zone de fusion des prémaxillaires. Pour chaque spécimen, les demi-mâchoires ont été fixées et conservées dans de l'éthanol à 70%.

Méthodes

Estimation de l'âge

Parce que la plupart des spécimens sont nés en laboratoire, l'âge de leur mort était connu. Cependant, cinq adultes appartenant à l'échantillon initial capturé dans la nature étaient d'âge inconnu. L'âge de ces individus a été évalué par squelettochronologie (Castanet *et al.*, 1997). Les fémurs gauches ont été disséqués et les tissus mous enlevés. Les fémurs ont été décalcifiés pendant 8 h dans l'acide nitrique dilué à 5%. Des coupes de 10 à 15 µm d'épaisseur ont été réalisées au cryostat (Zeiss Cryomat), au milieu de la diaphyse du fémur, et ont été colorées avec de l'hématoxyline d'Ehrlich. Les lignes d'arrêt de croissance (LAC), qui apparaissent colorées plus sombre que le reste de la matrice de l'os, ont été comptées. Un

LAC reflète l'arrêt du dépôt de la matrice de l'os pendant la période d'hiver (c'est-à-dire, 1 LAC = 1 hiver). Dans les îles Gran Canaria et Tenerife, les *Chalcides* naissent en juin. La date de la mort des cinq spécimens étant connue, l'âge a été évalué au mois près après leur naissance et inclus dans les données (Tableau I).

Radiographies

La radiographie a été choisie parce qu' (i) elle révèle toutes les dents en formation à condition que leur matrice soit minéralisée (la minéralisation de la dent commence tôt, quand la taille des germes est inférieure à 10% de leur hauteur définitive) et ii) les positions des dents ne sont pas modifiées. Dans du matériel sec observé à la loupe binoculaire, de petits germes sont souvent masqués par les tissus mous environnants. La position des dents par rapport aux autres peut aussi varier selon l'angle d'observation. De plus, si l'on nettoie les mâchoires avec de l'eau de Javel pour enlever les tissus mous (par exemple, pour une étude au microscope à balayage) toutes les dents de remplacement sont perdues (elles ne sont pas encore attachées à l'os de support) et, parfois, des dents fonctionnelles lorsqu'elles sont fortement résorbées.

Pour chaque individu utilisé dans cette étude, les quatre demi-mâchoires ont été déshydratées dans une série d'éthanol de plus en plus concentrée et séchées. Les demi-mâchoires (quadrants) ont été placées, le côté lingual au-dessus, sur un film Kodak SO343 à haute résolution, puis exposées aux rayons X (10 mA, 25 kV, 30 cm de la source, temps d'exposition de 5 à 40 minutes selon la taille de la mâchoire). Les radiographies ont été directement observées dans une loupe binoculaire Olympus en lumière transmise. Les dessins ont été faits grâce à une "chambre claire".

Microscopie électronique à balayage (MEB)

Les observations au MEB étaient nécessaires dans la présente étude pour obtenir des informations sur la forme et l'orientation des dents fonctionnelles. Certaines des demi-mâchoires inférieures, qui avaient été précédemment radiographiées, ont été réhydratées et immergées dans de l'eau de Javel à 10% pendant 20 à 40 minutes à température ambiante (selon la taille de la mâchoire). Les tissus mous ont été délicatement enlevés à l'aide de pinces fines et de petits pinceaux. Les demi-mâchoires ont été alors déshydratées dans une série croissante d'éthanol, séchées, collées sur un support en cuivre, couvertes avec une couche mince d'or/palladium et observées dans un JEOL SEM 35 fonctionnant à 25 kV.

Comptages et mesures

Tous les comptages et les mesures ont été réalisés à partir des radiographies et vérifiés à l'aide du matériel original quand cela était nécessaire. Les longueurs et les hauteurs ont été mesurées et les valeurs arrondies au plus proche 50 µm. Le nombre de positions dentaires a été déterminé dans chaque quadrant, en commençant antérieurement et en incluant les positions où ils manquaient des dents. Toutes les dents fonctionnelles et de remplacement ont été comptées. Les mesures suivantes ont été aussi prises : la longueur de la rangée de dents dans chaque demi-mâchoire (depuis la dent la plus antérieure jusqu'à la plus postérieure), l'intervalle moyen entre les dents (entre les deux pointes de dents successives le long de la

rangée) et la hauteur de chaque dent fonctionnelle et de remplacement. La hauteur d'une dent fonctionnelle a été mesurée le long d'une ligne perpendiculaire allant de la pointe de la dent à la base de la dent du côté lingual ; la hauteur d'une dent de remplacement a été mesurée de la même façon, c'est-à-dire de la pointe de la dent à la partie minéralisée visible la plus basale.

Patron de remplacement

Taille de la dent. - La taille de chaque dent de remplacement a été exprimée comme un pourcentage de sa hauteur définitive probable, c'est-à-dire, la hauteur de la dent fonctionnelle qu'elle allait remplacer ou, quand cette dernière manquait, de la dent fonctionnelle la plus proche. Des valeurs relatives ont été choisies pour éviter des erreurs en raison de la taille différente des spécimens et de la variation de la hauteur des dents le long de la rangée. Pour chaque individu on peut raisonnablement supposer que les différences de la taille relative des dents de remplacement correspondent aux diverses étapes de différenciation des dents. Ces valeurs relatives ont été groupées dans 19 indices de développement dentaire variant de 0 (= dent de remplacement non visible sur la radiographie) à 1 (= dent fonctionnelle). Un diagramme a été obtenu pour chaque quadrant de mâchoire inférieure pour les 20 spécimens (c'est-à-dire, 40 demi-mâchoires). La taille relative a été aussi exprimée en fonction de l'âge et de la saison pour évaluer les effets possibles de ces facteurs sur le nombre et la taille des dents de remplacement. Dans chaque cas, le nombre de dents de remplacement a été recalculé pour obtenir un nombre égal d'individus (5) par saison et dans chacune des quatre classes d'âge (1, 2, 3 et > 3 ans) choisie.

Le patron de dentition. - Trois étapes de différenciation dentaire ont été arbitrairement définies pour étudier la symétrie de dentition et la position des dents de remplacement par rapport à leurs prédécesseurs fonctionnels. À cette fin, le nombre total de dents de remplacement dans les 40 quadrants inférieurs, c'est-à-dire, 560 dents, a été divisé par 3 (= 187). Trois classes de taille (petit, moyen et grand) totalisant un nombre de dents proche de 187 (c'est-à-dire respectivement, 188, 173 et 199) ont été alors définies à partir de la taille relative de toutes les dents de remplacement comme suit : petit = les dents de remplacement dont la hauteur n'excède pas 25% de la hauteur des dents fonctionnelles; moyen = les dents de remplacement mesurant de 26 à 40% des dents fonctionnelles; grand = les dents de remplacement plus grandes de 40% mais non fonctionnelles. Une quatrième classe comprenait toutes les dents fonctionnelles, y compris celles qui étaient soumises à la résorption. La position de toutes les dents fonctionnelles et des dents de remplacement a été reportée sur des dessins schématiques des quadrants supérieurs et inférieurs des 20 spécimens. Les radiographies et le matériel original ont été utilisés pour définir exactement les positions.

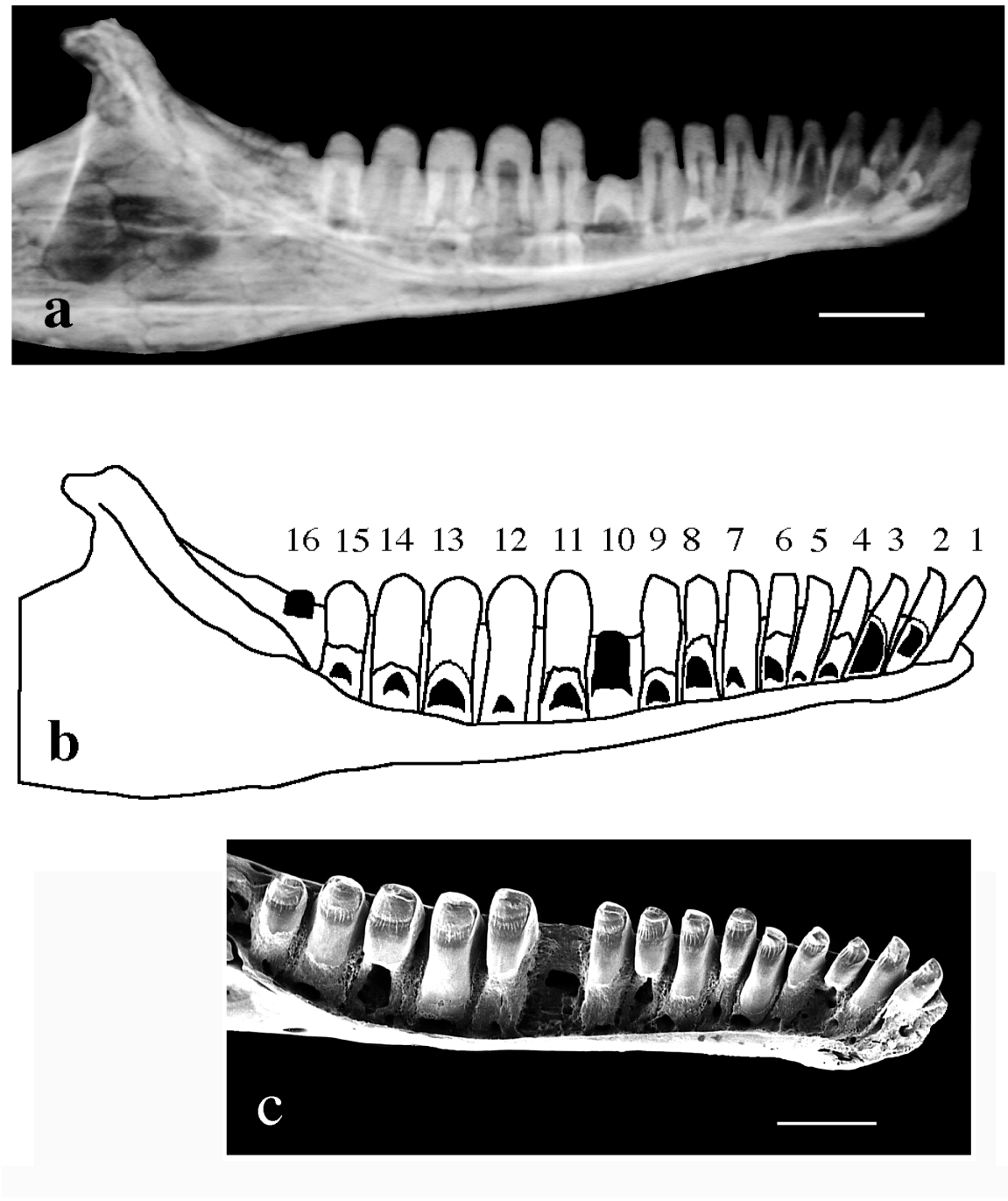
Résultats

Chez les deux espèces de *Chalcides*, la dentition, la forme, le nombre de dents et le patron de remplacement sont semblables. Les observations au microscope électronique à balayage et le patron de remplacement des dents suggèrent que les variations interspécifiques ne diffèrent pas des variations intraspécifiques. Pour ces raisons les résultats obtenus sur les deux espèces ont été utilisés ensemble et, par commodité, nous nous référerons aux deux espèce en employant seulement le terme de *Chalcides*.

La dentition

Chez *Chalcides*, la dentition consiste en une rangée de dents fonctionnelles placées sur la marge labiale du dentaire, prémaxillaire et maxillaire et d'une rangée de dents de remplacement sur le côté lingual de ces os (Fig. 1a, b). Il n'y a aucune autre dent à un autre emplacement.

Les dents fonctionnelles sont localisées dans un sillon dentaire, une dépression située entre la paroi labiale et la paroi linguale de l'os (Figure 1.c.). Les dents ont un mode d'attache pleurodonte et sont ankylosées sur le côté intérieur de la paroi labiale de la mâchoire, qui est plus haute que la paroi linguale, laissant le côté lingual de la base de la dent exposé. Dans chaque arcade dentaire, une augmentation de la taille des dents est observée le long de la rangée. Les dents placées à l'arrière (sauf la plus postérieure) sont plus grandes (plus larges et un peu plus hautes) que les dents antérieures, mais la forme de toutes les dents est grossièrement semblable (c'est-à-dire, que les *chalcides* sont homodontes) (Figure 1.c.). Les dents antérieures des positions 1 à 5 sont inclinées antérieurement tandis que celles des positions suivantes sont grossièrement verticales. Toutes les pointes sont recourbées du côté lingual.



(Echelle = 1 mm)

Figure 1. Vue linguale de la mâchoire inférieure gauche d'un *Chalcides* âgé de 70 mois. **a** : Radiographie; **b** : Dessin de la radiographie pour mettre en évidence les dents de remplacement (en noir); **c** : Image en "microscopie électronique à balayage" de la même mâchoire montrant la taille et l'orientation des dents fonctionnelles (toutes les dents de remplacement ont été perdues après le traitement à l'eau de Javel).

La dentition (dents fonctionnelles et de remplacement) se trouve être semblable dans les mâchoires inférieures et supérieures de chaque spécimen étudié. Par commodité, nous avons choisi de travailler sur les quadrants inférieurs. Les quadrants supérieurs ont été utilisés comme matériel comparatif pour étudier seulement la symétrie de la dentition.

Nombre de dents fonctionnelles et de remplacement

Quinze jours après la naissance, 15 positions ont été comptées dans les deux quadrants inférieurs (Tableau I). Chez les 19 autres spécimens âgés de 1 à 70 mois, le nombre de positions de dent variait de 14 à 18 (moyenne : 16,0 à gauche, 16,5 à droite). Dans les mâchoires supérieures (10 quadrants examinés seulement), il y avait 9 dents sur le prémaxillaire et 13 ou 14 positions sur chaque maxillaire gauche et droit. Le nombre impair de dents prémaxillaires est dû à la présence d'une dent simple dans la région médiane. Cette dent est placée dans la position qui était occupée à la naissance par la dent de l'œuf (cette dent étant elle même dérivée d'un des deux premiers germes en formation à cette position dans l'embryon). Donc, à part cette dent, il y avait 17 ou 18 dents sur les deux quadrants supérieurs (moyenne : 17,8 à gauche, 17,6 à droite).

Chez le spécimen âgé de 15 jours, la longueur de la rangée de dents sur la mâchoire inférieure gauche était de 2,76 mm. Chez les 19 autres individus, la longueur de la rangée de dents était de 3,4 mm pour celui âgé d'un mois et 7,0 mm chez le spécimen âgé de 70 mois. Dans la même série, l'intervalle moyen entre les dents a augmenté respectivement de 0,20 à 0,46 mm (Tableau I).

Les dents ont été comptées sur les radiographies des mâchoires inférieures (40 quadrants). Un total de 560 dents de remplacement et 430 dents fonctionnelles a été enregistré (Figure 2).

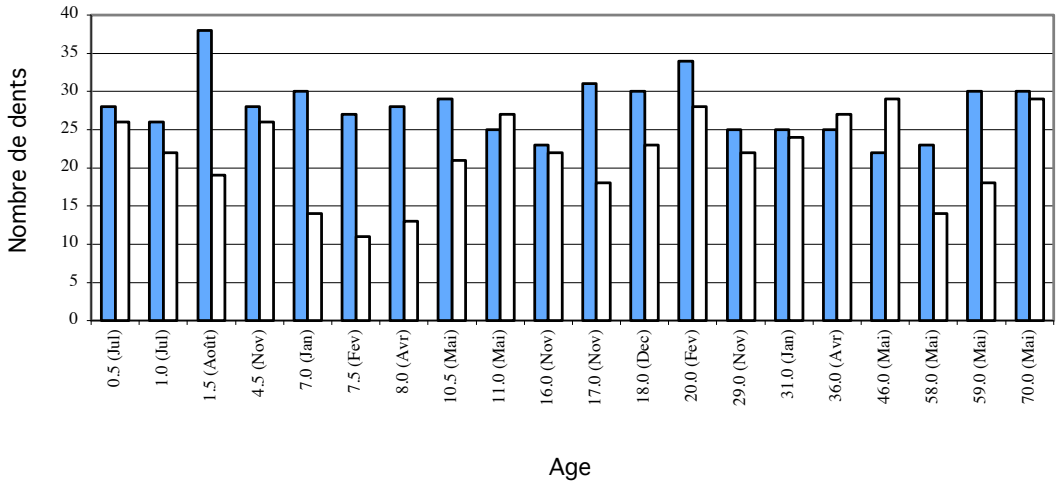


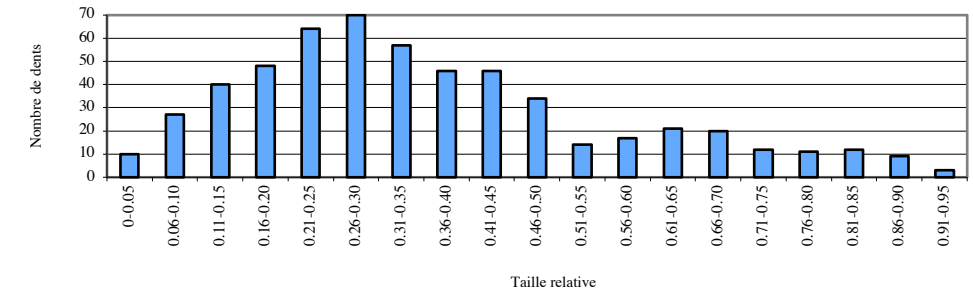
Figure 2. Nombre de dents de remplacement (■) et de dents fonctionnelles (□) à la mâchoire inférieure (quadrants gauches et droits) des 20 *Chalcides* étudiés en fonction de leur âge (de 15 jours à 70 mois). Le mois de la mort est indiqué entre parenthèses.

No- et espèces	Age après la naissance (en mois)	Jour de mort	Taille museau / cloaque (mm)	Nombre de dents à la mâchoire inférieure *gauche/droite*	Taille de la rangée de dents gauche (mm)	Moyenne des intervalles dentaires (en mm)
1- C.v.	15 jours	13 Août	38	15/15	2,76	0,20 (0,17-0,20)
2- C.s.	1	13 Jul,	42	16/16	3,40	0,24 (0,17-0,30)
3- C.v.	1,5	28 Août	40	17/17	5,14	0,24 (0,20-0,40)
4- C.v.	4,5	7 Nov,	50	14/15	3,53	0,24 (0,20-0,30)
5- C.s.	7	10 Jan,	42	16/15	3,79	0,25 (0,17-0,27)
6- C.v.	7,5	14 Fev,	42	15/16	2,95	0,23 (0,17-0,23)
7- C.s.	8	20 Apr,	31	16/15	2,83	0,21 (0,17-0,23)
8- C.v.	10,5	26 May	60	17/16	3,98	0,24 (0,20-0,33)
9- C.v.	11	22 May	55	16/18	3,79	0,22 (0,20-0,37)
10- C.v.	16	7 Nov,	70	16/17	4,50	0,29 (0,23-0,37)
11- C.v.	17	28 Nov,	51	16/17	3,73	0,28 (0,23-0,33)
12- C.s.	18	6 Dec,	51	16/17	3,79	0,27 (0,20-0,37)
13- C.v.	20	14 Fev,	78	17/18	4,90	0,29 (0,23-0,37)
14- C.v.	29	7 Nov,	70	17/17	4,30	0,28 (0,23-0,47)
15- C.s.	31	25 Jan,	49	16/16	3,53	0,27 (0,23-0,37)
16- C.s.	36*	20 Avr,	70	16/17	5,07	0,31 (0,27-0,37)
17- C.s.	46*	7 Mai	80	16/17	4,80	0,30 (0,28-0,37)
18- C.s.	58*	4 Mai	85	15/16	5,72	0,31 (0,27-0,47)
19- C.v.	59*	28 Mai	82	18/18	4,75	0,30 (0,27-0,43)
20- C.s.	70*	10 Mai	90	16/17	7,00	0,46 (0,33-0,53)

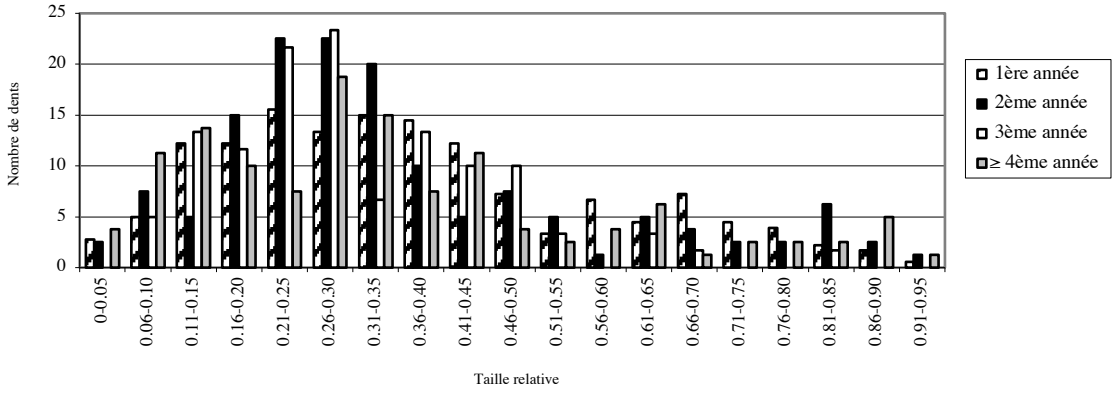
Tableau I. Age, jour de la mort et mesures diverses des 20 individus étudiés. C.s. = *Chalcides sexlineatus*; C.v. = *Chalcides viridanus*. * : spécimens pour lesquels l'âge a été évalué par la méthode de squelettochronologie (voir Matériel et Méthodes).

Chez 17 des 20 spécimens étudiés, on a observé plus de dents de remplacement que de dents fonctionnelles. Cette différence est plus à rapprocher des variations dans le nombre de dents fonctionnelles que celles des dents de remplacement. En effet, dans la série de croissance, le nombre de dents de remplacement était toujours supérieur à 22 pour les deux quadrants inférieurs (variation : de 23 à 38; moyenne : 28) indépendamment de l'âge des spécimens et des saisons, même durant la période d'hiver. Ce nombre élevé signifie qu'il y a probablement une dent de remplacement pour chaque position en tenant compte que les petits germes de dents ne peuvent être vus sur les radiographies. Par opposition, le nombre de dents fonctionnelles était plus petit (variation : de 11 à 29; moyenne : 21,5) et cela variait beaucoup

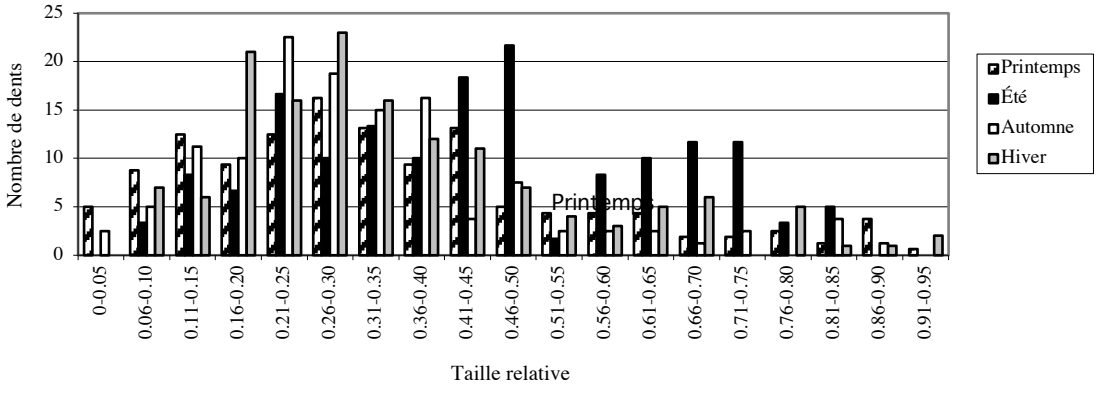
entre les individus. La valeur la plus petite, 11 dents fonctionnelles (c'est-à-dire, 5 ou 6 dents fonctionnelles sur les quadrants gauches et droits), a été observée chez le spécimen âgé de 7,5 mois, qui est mort à la fin de son premier hiver (contre plus de 20 dents fonctionnelles dans les lézards qui sont morts l'été d'après) (Figure 2). Cependant, le nombre d'individus était insuffisant pour pouvoir évaluer si cette différence était significative.



a



b



c

Figure 3. a : Distribution de la taille relative des 560 dents de remplacement à la mâchoire inférieure (quadrants gauches et droits) des 20 *Chalcides* étudiés. **b** : Distribution de la taille des dents de remplacement en fonction de l'âge. **c** : Distribution de la taille des dents de remplacement en fonction des saisons.

La taille relative des 560 dents de remplacement comptées dans les 20 mâchoires inférieures n'est pas régulièrement distribuée dans les 19 classes de taille choisies (Figure 3A). En effet, 78,8% des dents de remplacement ont une taille relative inférieure à 0,51 et, parmi celles-ci, 34,0% ont une taille comprise entre 0,21 et 0,35. Une distribution semblable est obtenue quand la taille relative est tracée en fonction de l'âge (Figure 3B). Dans les quatre classes d'âge (1, 2, 3 et > 4 ans) les dents de remplacement montrent la même distribution de taille (respectivement 77,3, 80,4, 93,5 et 80,9% des dents de remplacement inférieures à 0,51) avec un grand nombre de dents dans les classes de taille 0,21-0,35 (respectivement 31,2, 44,6, 42,3 et 31,7%). Les grands bourgeons dentaires > 0,51 sont plus nombreux chez les spécimens âgés de 1 à 2 ans (65%) que chez les spécimens âgés de 3 ans et plus (35%). Quand la taille de la dent est tracée en fonction des saisons, une distribution différente est observée en raison du plus grand nombre de dents de remplacement en été que dans les autres saisons. Au-dessus d'une taille relative de 0,41, l'été représente 45,5% du nombre de dents de remplacement contre 13,4, 18,8 et 22,3% pour l'automne, le printemps et l'hiver, respectivement. Au contraire, les petits bourgeons dentaires (de taille relative inférieure à 0,41) sont plus nombreux en hiver (28,0%) et en automne (27,4%) qu'au printemps (23,0%) et en été (21,6%).

Symétrie de la dentition

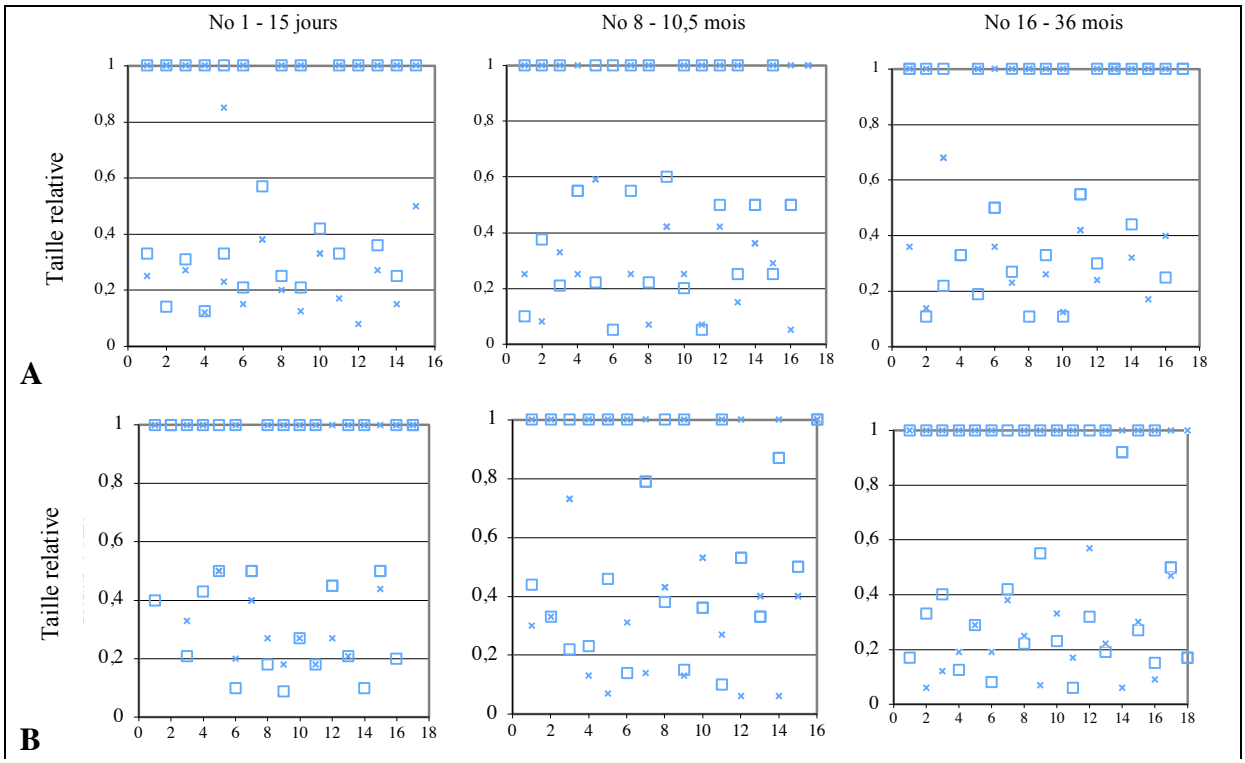
Pour savoir si le remplacement des dents est symétrique, les quatre quadrants ont été étudiés en les comparant deux par deux (gauche/droit, supérieur gauche/inférieur gauche, etc.) chez plusieurs spécimens. Trois spécimens (âgés de 0,5, 10,5 et de 36 mois) ont été choisis comme représentatifs de la série de croissance et la symétrie de la dentition est présentée sur les Figures 4 et 5.

Dans un premier temps, nous avons comparé la taille relative des dents à chaque position en les rapportant sur des diagrammes (Figure 4). Les trois spécimens présentés ici ont 16 dents sur chaque quadrant ; 64 comparaisons de deux dents opposées dans la même position sont possibles (mâchoire supérieure gauche/droite; mâchoire inférieure gauche/droite; mâchoire inférieure gauche/supérieure gauche; mâchoire supérieure gauche/supérieure droite). Ainsi, 192 comparaisons peuvent être faites pour les trois spécimens étudiés. Une dent de remplacement symétrique était absente dans 39 cas (20,3%).

Parmi les 153 autres comparaisons possibles, les deux dents ont montré un stade presque identique de développement (c'est-à-dire, une symétrie parfaite) dans 25 cas (13,0%). Pour les 128 comparaisons restantes, aucune symétrie stricte n'a été observée quels que soient les côtés des mâchoires comparées. Cependant, dans un certain nombre de positions, la différence de taille relative entre les deux dents opposées était inférieure à 10% (une différence pouvant être positive ou négative). Par exemple, chez le spécimen âgé de 15 jours,

14 sur 62 comparaisons (22,6%) n'avaient pas de symétrie entre les dents de remplacement opposées, 10 ont montré une bonne symétrie (16,1%), 27 (43,5%) ont montré une différence de taille n'excédant pas 10% et 11 (17,8%) ont montré une différence de taille supérieure à 10% (Figure 4). Chez les trois spécimens, la meilleure correspondance (12 des 49 comparaisons (24,5%) avec une bonne symétrie) a été observée en comparant des mâchoires supérieures gauches et droites (Figure 4B). Au contraire des dents de remplacement, la présence ou l'absence de dents fonctionnelles a montré une bonne symétrie : de 61,0 à 87,0% selon le spécimen et le quadrant étudié. Cependant, le niveau de résorption n'a pas été pris en compte pour ces dents.

Chez ces mêmes spécimens nous avons comparé la taille des dents de remplacement opposées dans chacune des positions en employant des représentations schématiques des patrons de dentition dans les quadrants (Figure 5). Dans les 25 cas étudiés, bien que les dents de remplacement aient été groupées dans seulement trois catégories de taille (petit, moyen et grand, voir Matériels et Méthodes), aucune symétrie dans le développement des dents n'est apparue quel que soit le côté de la mâchoire, la position de la dent et l'âge des individus.



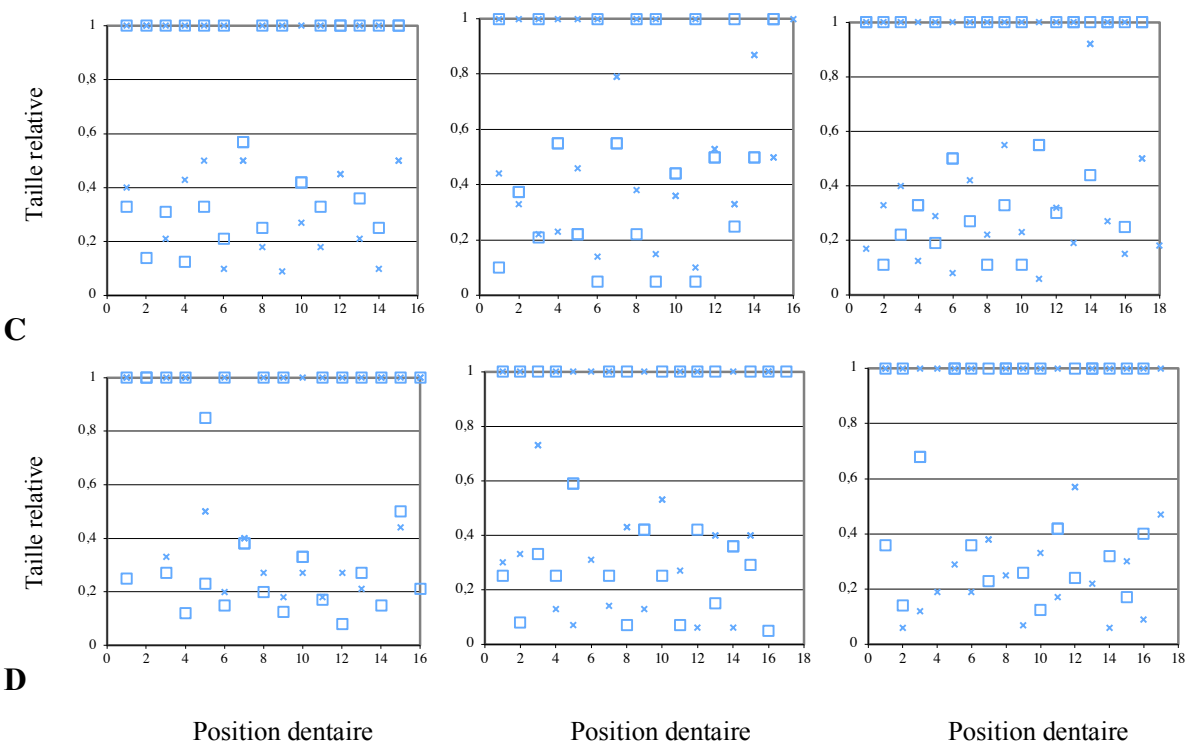


Figure 4. Diagrammes de la dentition chez trois spécimens de *Chalcides* (âgé de 0,5, 10,5 et 36 mois) montrant la taille relative des dents de remplacement dans chaque position (1 à 18). Relations entre le stade de développement des dents de remplacement dans les quadrants gauches (□) et droits (x) des mâchoires inférieures (A) et supérieures (B) et dans les quadrants inférieurs (□) et supérieurs (x) des côtés gauches (C) et droits (D). Les dents fonctionnelles sont indiquées par la ligne de valeur 1.

Processus de remplacement

La connaissance de l'emplacement précis des dents de remplacement par rapport aux dents fonctionnelles est utile pour interpréter de futures coupes sériées. Le patron de dentition a été schématiquement dessiné en 40 quadrants inférieurs, comme pour les trois spécimens présentés sur la Figure 5. Chez *Chalcides*, le remplacement des dents s'effectue dans une position disto-linguale chez tous les spécimens étudiés. Cependant, la position des bourgeons dentaires par rapport aux dents qu'ils remplacent et, en conséquence, la résorption qu'ils provoquent, diffère en fonction de l'age (Figure 5).

Chez le spécimen âgé de 15 jours, les bourgeons dentaires apparaissent légèrement postérieurs et à quelque distance des dents fonctionnelles qu'ils vont remplacer ; mais ils sont placés dans le même plan transversal (Figure 5A).

En grandissant, les dents de remplacement restent latérales et viennent progressivement contre les dents fonctionnelles. Chez les spécimens âgés de 10,5 mois et 36

mois, les bourgeons dentaires apparaissent également postérieurement aux dents fonctionnelles, mais au cours de la croissance, particulièrement chez les vieux spécimens, ils sont situés plus profondément dans sillon dentaire et ils sont plus près des dents fonctionnelles par rapport à ce qui se passe chez les jeunes individus (Figure 5C). Lorsque la taille des dents de remplacement augmente, elles viennent rapidement au contact de la base de la dent fonctionnelle.

Chez les jeunes spécimens, la résorption commence du côté disto-lingual de la base de la dent et se développe latéralement et vers le haut. En grandissant, le côté mésio-labial de la dent de remplacement entre partiellement dans la cavité pulpaire de la dent fonctionnelle tandis que la résorption continue vers le sommet. Entre temps, la surface mésio-labiale de la dent fonctionnelle reste inchangée et, bien que sévèrement résorbée, elle reste toujours fonctionnelle. Plus tard, la résorption de la paroi mésio-labiale de la dent fonctionnelle s'effectue à partir de l'intérieur de la cavité pulpaire. Finalement, la plus grande partie de la dent fonctionnelle est résorbée et seulement un petit sommet (l'émail) de la dent persiste avant d'être perdu. Dans quelques positions (seulement chez de jeunes spécimens), une dent fortement résorbée mais fonctionnelle peut toujours être présente, pendant que la dent de remplacement provoquant la résorption est bien développée et qu'un nouveau bourgeon est visible à sa base.

Chez les *Chalcides* adultes, les dents de remplacement induisent la résorption du côté disto-lingual des dents fonctionnelles comme chez les jeunes spécimens. Mais, au contraire de ce qui est observé chez les jeunes, l'érosion des dents se produit par en-dessous et plus près de la base de la dent fonctionnelle. Quand le germe grandit, la résorption affecte la région basale entière de la dent fonctionnelle, pénètre dans la cavité pulpaire par en-dessous et une grande partie de la dent fonctionnelle est perdue une fois que le côté mésio-labial est résorbé.

Le patron de remplacement

La taille relative de la dent (dents de remplacement et dents fonctionnelles) à chaque position a été transférée sur un diagramme pour chacune des 40 quadrants inférieurs radiographiés. L'analyse des patrons de remplacement a concerné 560 dents de remplacement et 430 dents fonctionnelles. Les diagrammes obtenus pour 9 spécimens, choisis comme représentatifs de la série de croissance (c'est-à-dire, de 0,5 à 70 mois), ont été illustrés (Figure 6). Les résultats étaient semblables chez d'autres individus. Deux résultats différents

ont été obtenus selon que l'on considère la taille des dents de remplacement en position adjacente ou alternée.

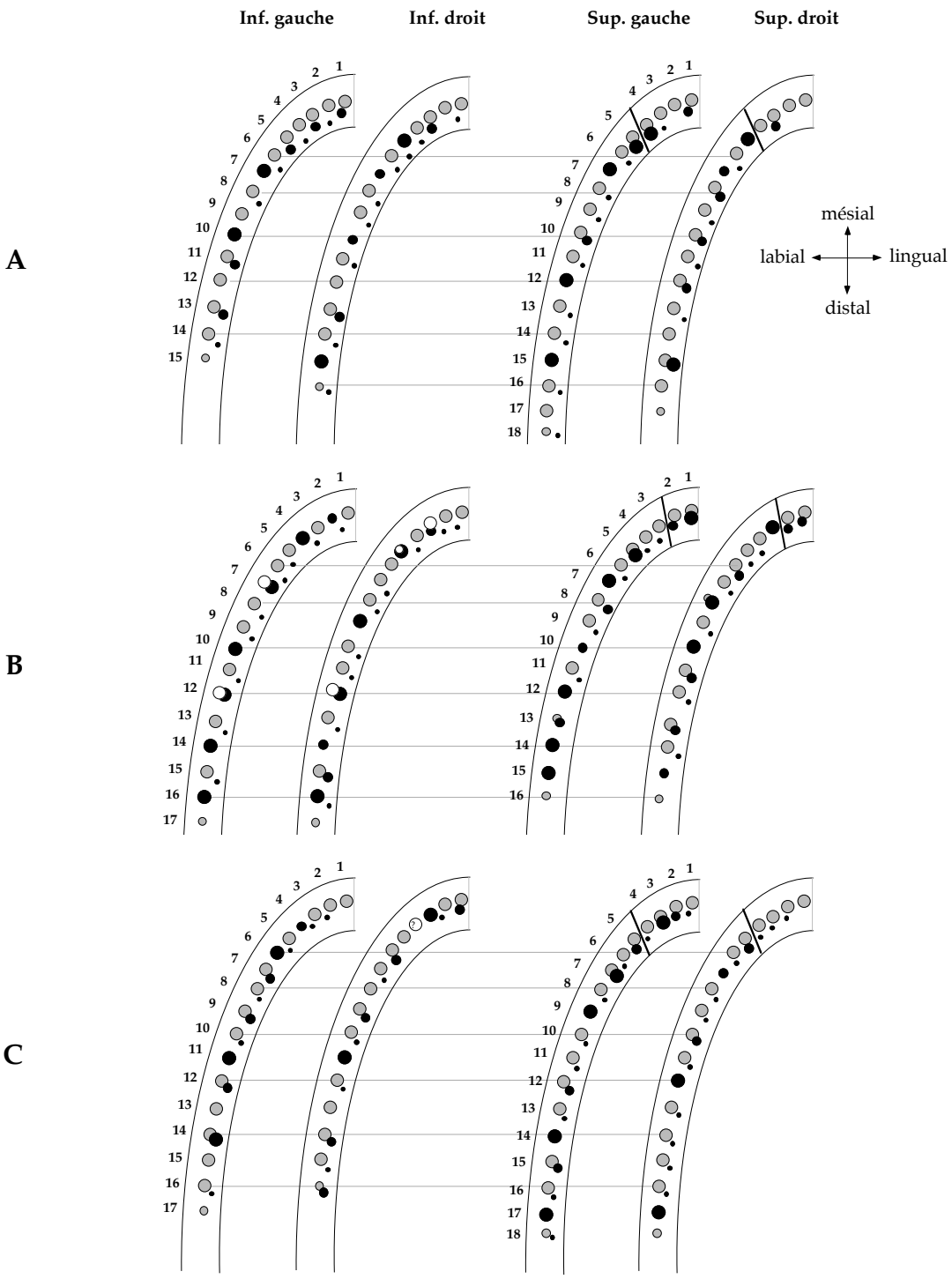

Inf. gauche
Inf. droit
Sup. gauche
Sup. droit
A
B
C
mésial
labial
lingual
distal

Figure 5. Représentation schématique de la dentition dans les quatre quadrants des mêmes individus que sur la Figure 4 (A = spécimen âgé de 15 jours; B = spécimen âgé de 10,5 mois; C = spécimen âgé de 36 mois). Dans les quadrants supérieurs, le prémaxillaire est séparé du maxillaire par une ligne. 1 à 18 = positions dentaires; cercles gris = dents fonctionnelles; cercles noirs = dents de remplacement petites, moyennes et grandes (voir le texte pour explication). Les dents de remplacement ont été précisément placées par rapport aux dents fonctionnelles qu'elles remplaceront.

Les dents de remplacement dans les positions adjacentes

Pour chacun des spécimens étudiés, chaque dent de remplacement successive est séparée de la suivante par un intervalle grossièrement régulier. L'analyse détaillée des 40 diagrammes révèle les tendances générales suivantes : premièrement, indépendamment de la position des dents et de l'âge de l'individu, une dent de remplacement dans un position p est suivie à la position p+1 par une dent de remplacement qui est en moyenne 20% moins développée (0,2 en la taille relative). Par exemple, quand la dent de remplacement dans la position p est à 60% de sa taille définitive, la dent de remplacement en p+1 est développée à environ 40% et la suivante en p+2 l'est environ à 20%. Deuxièmement, quand la taille d'une dent de remplacement à une position p est inférieure à 40% de sa taille définitive, la position p+1 contient soit une petite dent de remplacement, approximativement 20% moins développée, soit une dent de remplacement bien développée. Dans ce dernier cas, nous supposons qu'un bourgeon dentaire non minéralisé commence probablement son processus morphogénétique en position p+1, mais il n'est pas visible sur la radiographie. Quelques exemples sont présentés ci-dessous pour illustrer cette règle (Fig. 6).

Dans le diagramme correspondant au *Chalcides* âgé de 15 jours, six séries de dents de remplacement (positions 1-2, 3-4, 5-6, 7-8, 10-11, 13-14) ont été enregistrées. La première série commence dans la position 1 avec une dent de remplacement ayant une taille relative de 0,32. La position suivante, la position 2, montre un bourgeon avec une taille approximativement moins développée de 20%, c'est-à-dire une taille relative de 0,15. La position 3 contient probablement un bourgeon qui a commencé son programme morphogénétique, mais n'est pas visible sur la radiographie. Cependant, il y a une dent de remplacement bien développée dans cette position. En effet, la position 3 contient une dent avec la taille relative de 0,32, suivi d'une dent avec une taille relative de 0,15, etc.

Chez le spécimen âgé de 11 mois, les séries successives de dents de remplacement sont 1-2, 3-4, 5-6-7, 8-9-10, 11-12-13, 14-15. La dent dans la position 1 a une taille relative de 0,45, en position 2 de 0,18 et il n'y a aucune dent plus petite visible (bourgeon non minéralisé ?) en position 3; une dent de remplacement bien développée est présente dans la position 3 (d'une taille relative de 0,50), la suivante en position 4 a une taille relative de 0,20, etc. Chez ce spécimen, quelques positions de dents successives montrent une différence de taille plus grande que 20%, mais la tendance générale s'applique toujours.

Finalement, chez le spécimen âgé de 70 mois les 14 dents de remplacement prennent les positions successives 1-2, 3-4-5, 6-7, 8-9, 10-11-12 et 13-14. Dans chaque groupe la deuxième dent est 20% moins développée que la précédente.

Ces tendances générales se trouvent être observées pour 95% des dents de remplacement chez *Chalcides*, c'est-à-dire que seulement quelques exceptions peuvent être trouvées.

Les dents de remplacement en position alternée

En considérant les 40 quadrants inférieurs (560 dents de remplacement), la taille relative de la dent de remplacement dans la position alternée suivante, plus postérieure, est plus grande pour 80% des dents de remplacement, égale pour 15% et plus petite pour 5% (Figure 6). Cependant, lorsque l'on considère l'augmentation de la taille, ce n'est pas aussi régulier que ce qui a été décrit précédemment pour les intervalles entre les dents successives de remplacement. Aucune tendance générale ne peut être déduite du patron de remplacement que nous avons étudié. Cependant, en général (65%) quand une dent de remplacement est petite, c'est-à-dire, qu'elle a une taille relative entre 0,1 et 0,4, la dent dans la position p+2 montre seulement une petite augmentation (en moyenne, la dent est 10% plus grande).

Par exemple, chez le spécimen âgé de 4,5 mois la taille relative de la dent de remplacement dans même la série (1-3-5-7) est : 0,28 en position 1, 0,32 en position 3, 0,39 en position 5, mais de 0,6 en position 7 et une dent fonctionnelle dans la position 9. Chez le spécimen âgé de 18 mois, la même série (2-4-6-8) est : 0,28 en position 2, 0,42 en position 4, 0,52 en position 6, 0,59 en position 8 et une dent fonctionnelle en position 10.

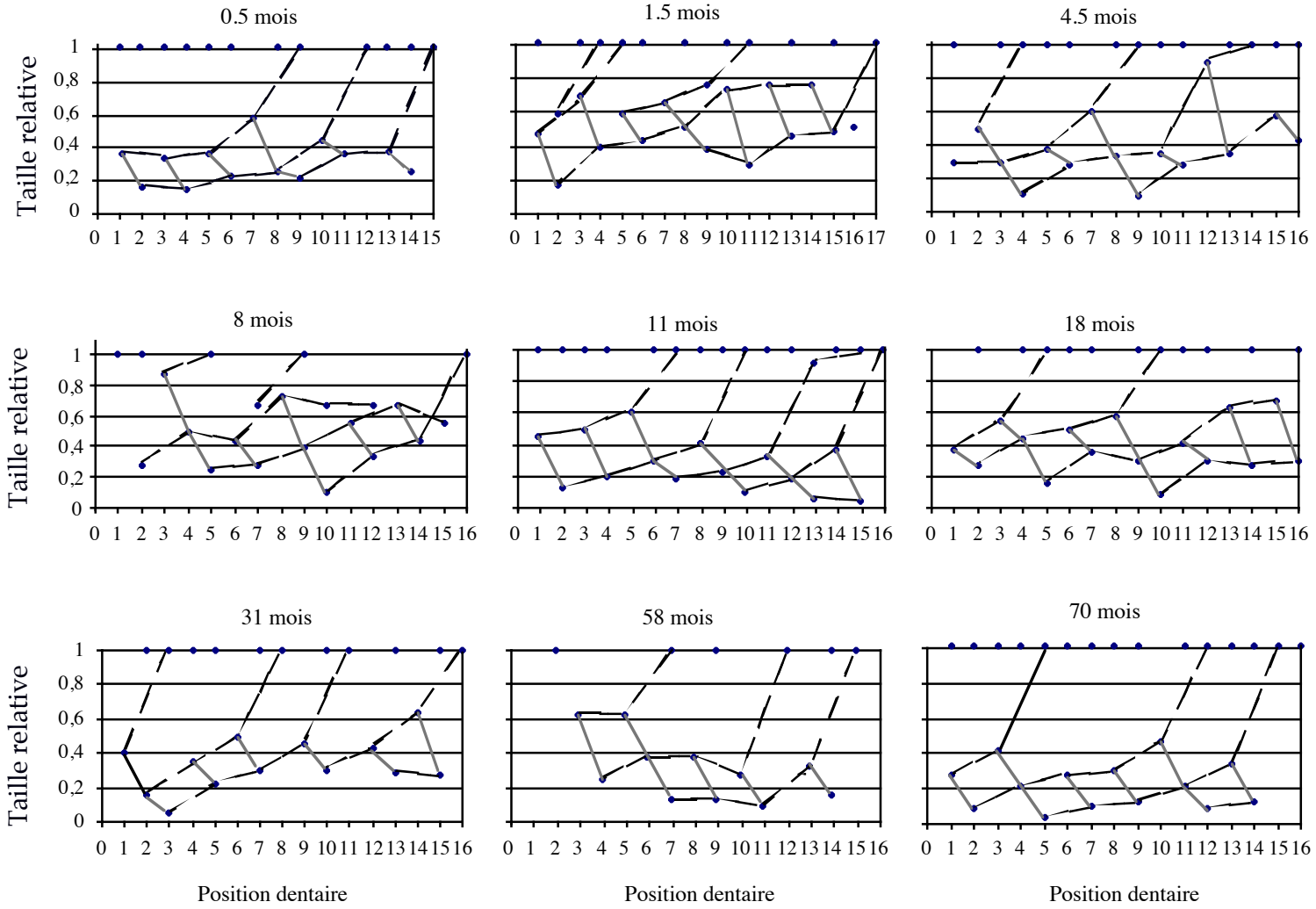


Figure 6 : Diagrammes de la taille relative des dents (dents de remplacement et dents fonctionnelles) dans toutes les positions de la mâchoire inférieure gauche de 9 spécimens représentatifs de la série de croissance. Des lignes pleines lient les dents de remplacement de taille décroissante dans les positions successives; les lignes pointillées lient les dents dans des positions alternées (séries impaires et paires).

Discussion

Parmi toutes les études qui ont traité de la dentition et du patron de remplacement des dents chez les squamates, cette étude chez *Chalcides* est à notre connaissance, la première qui soit (a) basée sur une série de croissance de spécimens d'âge connu ayant vécu dans des conditions semblables et (b) qui tienne compte de la taille précise (taille relative) des dents de remplacement. Les études précédentes chez les reptiles ont concerné soit (a) la comparaison du patron de dentition dans un grand nombre de familles (par exemple, Edmund 1960, 1969), soit (b) ont traité une espèce en particulier, mais avec une série des spécimens d'âge inconnu (Cooper, 1966 ; *Anguis* fragilis) ou (c) avec des spécimens matures seulement (Rocek, 1980 ; *Lacerta viridis*). Dans les travaux d'Edmund, seuls quelques spécimens (en général un ou deux adultes), mais dans un grand nombre d'espèces, ont été utilisés. Le patron de remplacement des dents chez ces espèces a servi à établir la théorie bien connue sous le nom de "Zahnreihen". Les spécimens étaient d'âge inconnu et provenaient de collections de Musée. De plus, dans des études précédentes la taille des dents de remplacement avait été seulement enregistrée comme petite, moyenne ou grande.

Dans la nature il est difficile d'obtenir un nombre de spécimens d'une d'espèce représentatifs d'une série de croissance et il est même plus difficile de capturer des individus dans toutes les saisons. Cependant, il est possible, comme cela est montré dans la présente étude, d'estimer l'âge des spécimens sauvages en utilisant la squelettochronologie, une méthode qui est couramment utilisée dans l'étude des populations de lézards (Castanet *et al*., 1993). En obtenant la reproduction de *Chalcides* en captivité, nous avons obtenu une série de croissance. Cela a l'avantage (1) de fournir des spécimens d'âge connu régulièrement distribués dans une série de croissance et (2) de fournir des individus qui sont morts à différentes saisons. Nous nous rendons compte que les conditions de reproduction des *Chalcides* dans notre laboratoire diffèrent des conditions rencontrées dans la nature, mais les spécimens étudiés ont toujours été en bonne santé.

Cette étude a été entreprise pour déterminer si la connaissance du patron de remplacement des dents chez *Chalcides* pouvait nous aider à prévoir des stades de différenciation des dents. Nous pouvons maintenant répondre à cette question de manière affirmative. En supplément, un certain nombre de données relatives à la dentition et au remplacement des dents chez ce scincidé doivent être traitées.

La dentition et le remplacement des dents chez Chalcides

Le nombre de position de dents durant l'ontogénie chez Chalcides

Le nombre de positions dentaires chez *Chalcides* (16-17) est grossièrement semblable chez les 20 individus étudiés, c'est-à-dire, chez les spécimens âgés de 0,5 à 70 mois. La conservation d'un nombre constant de positions dentaires sur la rangée, depuis la naissance jusqu'à la mort a deux explications possibles qui s'excluent mutuellement. Premièrement, le nombre de familles de dents reste constant au cours de l'ontogénie (c'est-à-dire qu'une dent à la position p à la naissance appartient à la même famille que la dent dans la position p, 6 ans

plus tard). Alternativement, des familles de dents ont été perdues dans la rangée durant l'ontogénie mais de nouvelles positions ont été acquises, postérieurement ou entre les deux. Le nombre de positions est donc resté constant mais les positions ne sont pas nécessairement homologues chez les juvéniles et les adultes. Chez *Chalcides*, d'une part, la longueur de la rangée dentaire, la taille des dents et l'augmentation des intervalles entre les dents augmentent pendant l'ontogénie et, d'autre part, le remplacement dentaire s'effectue de manière légèrement postérieure (disto-lingualement) aux dents fonctionnelles. Cela suggère que l'augmentation des intervalles entre les dents intervient dans l'augmentation de la longueur de la rangée dentaire et permet d'accommoder de plus grandes dents. Cela pourrait signifier que le nombre de positions est conservé durant l'ontogénie avec une légère dérive des positions vers l'arrière, dérive facilitée par le remplacement disto-lingual. Une observation similaire d'un nombre inchangé de positions dentaires a été faite par Gans (1957) chez un Amphisbaenidae et par Cooper (1966) chez l'orvet, *Anguis fragilis*.

Cependant, le maintien d'un nombre constant de positions dentaires n'est pas la condition commune chez la majorité des reptiles pour lesquels le nombre de dents augmente généralement avec l'âge par l'addition de nouvelles dents postérieurement sur chaque rangée (*e. g.*, Edmund, 1960). Par exemple, l'agame *Agama agama* a neuf dents sur la rangée au moment de l'éclosion tandis qu'il y en a 19 chez l'adulte ; la rangée dentaire s'étend postérieurement par le complément de nouvelles positions alors que la mâchoire grandit (Cooper *et al*., 1970). De manière similaire, un jeune iguane ajoute plusieurs positions à la région postérieure de la rangée en un an (Edmund, 1969).

Chez les squamates, la réduction du nombre de dents (c'est-à-dire, la perte de positions dentaires) avec l'âge a été observée chez *Lacerta* et chez l'orvet (Cooper, 1963, 1966) et chez un amphisbaenidé (Schmidt, 1960). Cette réduction est apparemment réalisée par le remplacement de deux dents adjacentes par un seul successeur. Dans les mâchoires qui ont cessé de grandir, il est probable qu'une des dents n'arrivent pas à se développer à cause de la présence d'une plus grande dent voisine. Il est aussi possible que le remplacement cesse chez les vieux spécimens comme cela a été suggéré par Gans (1957) chez un amphisbaenidé et par Bellairs et Milles (1960, 1961) chez un varan. Chez quelques *Chalcides* adultes, le nombre de positions a été trouvé plus petit que chez les juvéniles, ce qui suggère qu'une ou deux positions pourraient avoir été perdues. Cependant, le *Chalcides* le plus âgé étudié dans le présent travail n'a pas montré d'arrêt du remplacement et il n'y a aucune preuve que le remplacement cesse chez la plupart des squamates examinés, même chez les très vieux individus (Edmund, 1960). Ces résultats, quelquefois contradictoires, peuvent être mis en relation avec les différentes méthodes d'études utilisées.

Nombre de dents en fonction de l'âge et de la saison

Le nombre élevé de dents de remplacement dans la série de croissance des *Chalcides* implique que (a) le processus de renouvellement des dents est actif quel que soit l'âge (l'individu le plus âgé étudié avait 6 ans), et (b) que les dents de remplacement sont présentes pendant des périodes d'automne et d'hiver. Il y a au moins une dent en développement à chaque position sur la rangée, mais deux dents de remplacement (une grande et une petite) ont été fréquemment observées chez les jeunes spécimens ; également, quelques petits bourgeons n'étaient certainement pas visibles sur les radiographies parce qu'ils n'étaient pas suffisamment minéralisés. Ainsi le nombre de bourgeons dentaires a sans doute été sous-estimé. La présence de deux dents de remplacement à la même position chez les jeunes spécimens suggère qu'il n'y a aucune liaison entre l'âge de la dent fonctionnelle et l'initiation d'une nouvelle dent. Cela indique plutôt que les dents situées à une position sont initiées à un rythme donné, même si la dent précédente n'est pas encore fonctionnelle. Une horloge

génétique pourrait être responsable de l'initiation périodique de la dent à chaque position. Est-ce que cette horloge est la même pour chaque position dentaire chez *Chalcides* ? Probablement, si nous considérons « l'intervalle de développement » régulier séparant chaque dent de remplacement successive dans la rangée. Cette horloge est-elle synchronisée pour toutes les dents de remplacement dans la rangée ? Probablement pas si nous considérons « l'intervalle de développement » irrégulier séparant les dents de remplacement en position alternée. Une horloge ainsi synchronisée aurait pu souligner le mécanisme génétique de ce qu'Edmund (1960, 1969) a appelé « la vague de dents de remplacement » sur laquelle sa théorie "Zahnreihen" est fondée.

Il est connu que le temps nécessaire pour former une dent augmente avec l'âge. Par exemple, chez les amphibiens urodèles, le développement de la dent prend seulement quelques jours chez les larves, mais il exige plusieurs mois chez les adultes (Chibon, 1977). Chez les lézards *Iguana iguana* et *Varanus bengalensis* le développement de la dent prend environ 3 mois (Edmund, 1960) et il a besoin de 12 mois chez le crocodile âgé de deux ans (Edmund, 1962). La période fonctionnelle d'une dent augmente aussi avec l'âge : de 20 jours pour les larves à plusieurs mois pour les adultes amphibiens (Chibon, 1977); au moins 12 mois pour les crocodiles âgés de 2 ans (Edmund, 1962); de 1,5 à 4 mois chez les orvets âgés de 3 à 10 ans (Cooper, 1966); 3 mois chez les *I. iguana* et 3,5 mois chez les *V. bengalensis* (Edmund, 1960). Cela implique que le temps séparant les initiations des dents à une position donnée augmente avec l'âge et cela explique pourquoi seulement une dent de remplacement est trouvée chez le juvénile et l'adulte *Chalcides*, au lieu de deux chez les *Chalcides* les plus jeunes. Toutes ces données suggèrent que l'horloge génétique contrôlant l'initiation des dents est sous le contrôle de facteurs épigénétiques comme le métabolisme général. Cela pourrait expliquer pourquoi le nombre de dents de remplacement est plus élevé en été que durant les autres saisons chez *Chalcides*.

Par contraste avec les dents de remplacement, le nombre de dents fonctionnelles chez *Chalcides* varie beaucoup avec l'âge et la saison, particulièrement pendant la période d'hiver de la première année de la vie. Le nombre d'individus à chaque classe d'âge et à chaque saison était trop petit pour le traitement statistique, mais les résultats méritent néanmoins un bref commentaire. Les *Chalcides* sont nés en juillet avec des dents fonctionnelles et avec des dents de remplacement en formation. En effet, chez les reptiles, les dents sont d'abord formées chez l'embryon et on sait que le remplacement même des dents est initié avant la naissance (*e. g.*, Röse, 1894 ; Woerdeman, 1919 ; Osborn, 1971 ; Westergaard, 1986 ; Westergaard et Ferguson, 1986). Chez les jeunes lézards, les dents sont remplacées dans une période de quelques semaines (Osborn, 1971 ; Westergaard, 1986). Donc, chez un jeune *Chalcides* quelques générations de dents se développent pendant l'été et l'automne, puis le renouvellement des dents s'arrête quand la diète arrive en hiver, comme illustré par le faible nombre de dents fonctionnelles. De nombreuses dents fonctionnelles sont perdues pendant le premier hiver, suggérant que le processus de résorption a continué pendant cette période. Un tel phénomène n'a pas été observé chez les spécimens âgés de plus d'un an qui sont morts durant l'hiver suivant. Cela pourrait être dû soit à la présence de plus grandes dents fonctionnelles soit à un renouvellement de dents plus lent chez les juvéniles et les adultes comparés aux individus les plus jeunes. Bien que ces observations doivent être confirmées avec un plus grand nombre de spécimens morts pendant la période d'hiver, elles pourraient indiquer une différence dans le remplacement des dents entre les spécimens âgés d'un an et ceux plus âgés. De la même façon, les *Anguis fragilis* âgé de 3 à 10 ans ne perdent pas leurs dents pendant l'hiver (Cooper, 1966).

Vitesse de différenciation des dents

Les tailles des 560 dents de remplacement étudiées ne sont pas régulièrement distribuées dans les différentes classes de taille. Cela signifie que le taux de développement n'est pas régulier durant la formation de chaque dent. La distribution de taille suggère fortement que les dents grandissent rapidement jusqu'à ce qu'elles atteignent 20% de leur taille définitive. Alors, la vitesse de développement se réduit pendant les 2 à 3 cinquièmes de leur acquisition de taille (jusqu'à ce que les dents acquièrent approximativement 50% de leur hauteur définitive) et s'accélère finalement pendant les deux derniers cinquièmes, jusqu'a ce que les dents s'attachent et percent. Aucune différence majeure n'a été observée dans la distribution de taille entre des individus jeunes et vieux et cela est confirmé par les diagrammes de taille des dents.

Ces observations diffèrent de la vue traditionnelle (par exemple, Edmund, 1960 ; Cooper, 1966) selon laquelle les dents augmentent de taille à un taux régulier. En effet, nous nous attendions à un taux constant de développement plutôt qu'à deux étapes de développement rapide, séparées par une période de croissance plus lente. Cependant, ces variations des taux de croissance des dents chez *Chalcides* ont été déduites des mesures prises sur un grand nombre de dents de remplacement dans une série de croissance. La période de croissance lente correspond à la période où la jeune dent de remplacement s'approche de la base de la dent fonctionnelle. En effet, celle-ci doit être entièrement résorbée (ou tombée) avant que la dent de remplacement puisse prendre sa place. Le processus de résorption ralentit probablement le développement des dents de remplacement pendant la période de 20% à 40-50% de leur hauteur définitive. Quand la dent fonctionnelle est perdue, cela ne limite plus la croissance de la nouvelle dent et cela pourrait expliquer la croissance finale accélérée. Chez les amphibiens, une telle accélération arrive quand les dents de remplacement sont à 45% de leur taille finale (Chibon, 1977). Cela pourrait être dû au même phénomène que celui que nous avons observé chez *Chalcides*.

Le remplacement des dents chez Chalcides

Chez *Chalcides*, selon que l'on considère les juvéniles ou les adultes, les dents de remplacement sont localisées dans une position différente de celle des dents fonctionnelles qu'elles remplacent. Ce point n'a pas été clairement exposé dans la littérature en raison de l'absence de série de croissance et/ou de la connaissance de l'âge des spécimens. Chez les *Chalcides* de 15 jours, les dents de remplacement se développent dans une position disto-linguale de la dent fonctionnelle, mais latérale, tandis que chez les adultes de 2 ans et plus elles se développent dans une position linguale mais basale. Ces différences sont certainement liées à la croissance des mâchoires. Chez les jeunes spécimens, chez lesquels les mâchoires grandissent rapidement et les dents augmentent de taille après plusieurs générations, le remplacement disto-lingual se produit. Chez les adultes, chez lesquels les mâchoires grandissent lentement, si cela est le cas, le remplacement est seulement lingual. La position

linguale protège des chocs les bourgeons en croissance. La position latérale des dents de remplacement chez les jeunes spécimens *versus* la position basale chez les adultes peut aussi être discutée en termes de protection (les forces de préhension et d'écrasement sont plus grandes chez les adultes) mais cela pourrait aussi être dû au logement de plus grandes dents dans les sillons dentaires qui s'accroissent plus en profondeur qu'en largeur.

Edmund (1960) a reconnu deux méthodes de remplacement des dents chez les squamates : les méthodes varanidé et iguanidé. Dans la méthode du varan, qui est rencontrée chez la plupart des lézards anguinomorphes, les dents de remplacement poussent en position interdentaire, c'est-à-dire, postérieurement aux dents fonctionnelles. La méthode iguanidé concerne quant à elle des lézards non-anguinomorphes; la dent de remplacement aborde directement le côté lingual de la dent fonctionnelle. Rieppel (1978) a trouvé un troisième type pour les lézards anguinoïdes, la méthode intermédiaire, dans laquelle la dent de remplacement se développe disto-lingualement à la dent fonctionnelle. Rieppel a trouvé que cette méthode est facilement dérivable de la méthode iguanidé, que l'on considère comme un type primitif, par un déplacement des dents de remplacement en position disto-linguale. Malheureusement, ces études concernent surtout des spécimens adultes. L'utilisation d'une série de croissance chez *Chalcides* nous a permis d'analyser la variation de remplacement des dents pendant l'ontogénie et de prendre en considération la croissance de la mâchoire. Nous avons constaté que le remplacement des dents chez les *Chalcides* jeunes et juvéniles, suivait la méthode intermédiaire (comme celle décrite par Rieppel, 1978) et que chez les adultes, cette méthode est progressivement remplacée par la méthode iguanidé (comme celle décrite par Edmund, 1960). Nous pouvons spéculer que la méthode intermédiaire sera trouvée chez les jeunes et les juvéniles des espèces de squamates pour lesquelles la méthode iguanidé a été décrite chez les adultes. De plus, cela pourrait impliquer qu'il est difficile de considérer plus longtemps le mode de remplacement iguanidé comme primitif. Cependant, cela pourrait être vérifié dans des séries de croissance d'iguanes.

Remplacement dentaire et prédiction des étapes précises du développement des dents

Cette étude a été entreprise parce que deux particularités dans le patron de développement des dents chez les reptiles étaient, a *priori*, favorables pour une prédiction précise des étapes précises du développement des dents chez les squamates : la symétrie dans la dentition (*e. g.*, Cooper, 1966) et la théorie "Zahnreihen" (Edmund, 1960, 1969). Nos résultats fondés sur une série de croissance de *Chalcides* indiquent clairement que la symétrie n'est pas un critère fiable et que l'initiation successive des dents de remplacement dans des loci alternés n'est pas aussi régulière que supposée dans la théorie de "Zahnreihen". En fait, une troisième particularité, qui n'avait pas été envisagée jusqu'alors, c'est-à-dire, les stades de développement des dents de remplacement dans des positions successives, est apparue être fortement fiable pour prévoir les stades de développement de dents chez ce lézard.

Symétrie dans le remplacement

Dans la série de croissance de *Chalcides*, aucune symétrie stricte (synchronie) n'a été observée dans le remplacement des dents, soit bilatéralement soit entre les mâchoires supérieures et inférieures. En effet, bien que la symétrie ait été observée pour quelques positions, il y avait trop de variation d'une position à l'autre pour considérer que les dents opposées (droite/gauche haut/bas) montraient le même stade de développement. Cette absence de symétrie précise chez les *Chalcides* contraste avec les études précédentes sur la symétrie bilatérale chez les squamates, avec des stades de remplacement semblables dans les mâchoires supérieures et inférieures et sur les côtés gauches et droits (Cooper, 1965, 1966). Chez *Chalcides*, une telle symétrie semble ne pas être nécessaire pour maintenir une dentition fonctionnelle à tout moment.

La théorie "Zahnreihen" et la situation chez Chalcides

Les dents de la plupart des reptiles (avec quelques exceptions pour les sphénodons et les agamidés, Cooper et *al.*, 1970 ; Cooper et Poole, 1973 ; Robinson, 1976 ; Throckmorton, 1979) sont remplacées durant toute la vie et les dents de remplacement se développent par vagues successives qui se déplacent le long des positions alternées (Edmund, 1960). Ces observations n'ont jamais été mises en doute par les auteurs qui ont par la suite travaillé sur le sujet (*e. g.*, Cooper, Osborn, DeMar). Edmond (1960, 1962, 1969) a été le premier à essayer d'expliquer l'organisation des dentitions de reptiles. Sa théorie de "Zahnreihen" est basée sur l'initiation successive de dents à des loci alternés par la propagation d'un "stimulus" le long de la lame dentaire. Chaque stimulus produit une rangée dentaire, dans laquelle les dents de remplacement sont dans des stades de développement successifs. Edmund a appelé ces stimuli "vagues de remplacement". Cependant, la question sur la façon dont une vague de remplacement est générée et/ou contrôlée a été très débattue en termes d'implications évolutives (Osborn, 1970, 1984 ; DeMar, 1972, 1974 ; Osborn et Crompton, 1973) et de réalité embryologique (*e. g.*, Osborn, 1971, 1972). Quelques auteurs ont utilisé les données d'Edmund (c'est-à-dire, seulement la condition adulte de la dentition) pour établir des modèles mathématiques du patron de remplacement des dents (Osborn, 1970, 1972 ; DeMar, 1972, 1973, 1974). L'étude de l'ontogénie de la succession des dents chez *Lacerta vivipara* a mené Osborn (1971), à proposer la théorie de famille dentaire (c'est-à-dire les générations successives de dents à une même position) comme une explication possible des vagues de remplacement : ces vagues pourraient être la suite automatique du patron de développement des dents chez les embryons. Cependant, la théorie d'Osborn a été mise en doute par Westergaard (1986) et par Westergaard et Ferguson (1986), qui pensent que la théorie de la famille dentaire ne peut pas expliquer le développement de la dentition précoce chez les reptiles.

Notre étude sur la dentition de *Chalcides* semble confirmer que, pendant l'ontogénie, le remplacement des dents s'effectue par vagues, qui avancent rapidement à des positions dentaires alternées. Cependant, ce modèle de remplacement est difficile à établir dans des spécimens jeunes et juvéniles parce que ce n'est pas aussi régulier que ce que suggère la théorie de "Zahnreihen" proposée par Edmund (1960). Cependant, quand on regarde les spécimens les plus vieux, ce modèle est plus régulier. Donc, il semble que la théorie de "Zahnreihen" s'applique plus aux dentitions adultes qu'aux dentitions juvéniles. Autrement dit, lorsque la fréquence de remplacement de dent est rapide le modèle est moins régulier (bien qu'il semble exister) que lorsque la fréquence de remplacement est lente. Cette différence avec les résultats d'Edmund pourrait être aussi due à la façon dont nous avons mesuré la hauteur des dents comparée à Edmund et aux auteurs suivants. En effet, tous nos patrons sont déduits de la taille précise des dents de remplacement (à 50 µm prés), contrairement aux quatre conditions seulement (petit, moyen, grand, très grand) utilisées par

Edmund. En effet, grouper les tailles de dent dans des classes pourrait réduire la gamme de variations et pourrait modifier un peu le modèle.

En résumé, l'étude de développement des dents chez les *Chalcides* pourrait profiter à la connaissance du patron de remplacement dans les positions alternées, mais les prédictions seront plus précises chez les adultes que chez les juvéniles.

Prédiction des étapes successives de développement des dents chez Chalcides

Nos résultats chez le squamate *Chalcides* indiquent que, de l'avant à l'arrière, nous pouvons prévoir la taille d'une dent de remplacement (et ainsi son stade de développement) pourvu que nous ayons obtenu l'information précise sur le stade de développement de la dent de remplacement précédente dans la rangée. Par exemple, en utilisant les coupes sériées de n'importe quel quadrant de *Chalcides*, la prédiction d'un stade de développement d'un germe de dent pourrait être obtenue comme suit :

En regardant les positions successives: Si une dent de remplacement dans la position p est à 40% ou plus de sa hauteur finale, la dent dans la position p+1 sera 20% moins développée. Si la dent dans la position p est inférieure à 40% de sa hauteur finale, p+1 contient une dent 20% moins développée, c'est-à-dire, soit une petite dent soit une dent récemment initiée. Dans le dernier cas, une dent de remplacement bien développée sera trouvée dans p+1.

En regardant les positions alternées: Si la taille d'une dent de remplacement dans la position p est inférieure à 40% de sa hauteur finale, la dent dans la position p+2 sera légèrement plus grande et à une étape légèrement plus avancée de développement. Il est clair qu'une comparaison entre des positions adjacentes apporte la meilleure prédiction possible sur l'état de développement d'un germe particulier.

Deuxième Partie

EXPRESSION DE L'AMÉLOGÉNINE LORS DU DÉVELOPPEMENT DES DENTS DE *CHALCIDES*

ÉTUDE DU DÉVELOPPEMENT DENTAIRE CHEZ UN LÉZARD SCINCIDÉ, *CHALCIDES*, AVEC UNE ATTENTION PARTICULIÈRE POUR L'AMÉLOGÉNÈSE: CLONAGE, SÉQUENCAGE ET EXPRESSION DE L'AMÉLOGÉNINE, LA PROTÉINE MAJEURE DE L'ÉMAIL EN FORMATION

Ce travail a été réalisé par :

Sidney Delgado*, Marie-Lise Couble[§] et Jean-Yves Sire*.

* UMR 8570, Équipe "Evolution et Développement du Squelette Dermique", Paris.
[§] Faculté d'Odontologie, Laboratoire"Développement des Tissus Dentaires", Lyon.

Remerciements : Nous sommes particulièrement reconnaissants envers F. Allizard pour son excellente assistance technique dans la préparation de coupes. Nous remercions le Prof. J. Castanet (Université Paris 6) pour les embryons, les juvéniles et les adultes de *Chalcides viridanus*, qu'il nous a fournis. Le travail en MET et MEB a été réalisé au « Service de Microscopie électronique - Université Paris 6 et CNRS ».

Résumé

L'analyse comparative du développement dans les principales lignées de vertébrés est nécessaire si nous voulons disposer de bases de connaissances solides pour comprendre les différentes voies évolutives qui ont mené à la dentition des vertébrés actuels. Jusqu'à présent, ce sujet n'a pas été étudié en profondeur chez les squamates (lézards et serpents), pour lesquels les études précédentes sont dispersées et limitées à certains aspects spécifiques. Dans la présente étude, nous avons obtenu de nombreuses données sur les principaux stades de développement chez une espèce de scincidé, *Chalcides viridanus* élevé au laboratoire depuis plusieurs années. En utilisant la microscopie optique, électronique à balayage et à transmission, le développement dentaire a été décrit, du stade d'initiation à la différenciation complète, à partir d'une série de croissance depuis des embryons tardifs jusqu'à des adultes âgés de 6 ans. Nous avons concentré principalement notre attention sur l'amélogénèse et complété notre description en étudiant l'amélogénine, la protéine principale de l'émail en formation chez les mammifères. Nous avons cloné le gène de l'amélogénine chez ce lézard et étudié son expression durant la formation de l'émail en utilisant l'hybridation *in situ* sur coupes.

Durant l'ontogénèse, la forme des dents et leur orientation changent, mais le nombre de positions dentaires n'augmente pas. A partir d'une forme conique, d'un type très général chez l'embryon, la morphologie dentaire se modifie progressivement pour atteindre un aspect typique, caractérisé par la formation de cuspides et de crêtes, résultant en une ornementation propre à l'espèce. Ces modifications n'ont été possibles qu'après des remplacements dentaires successifs à chaque position qui représentent une famille dentaire. Le mode d'attachement dentaire est pleurodonte et le remplacement a lieu à partir de la face linguale. Le cône de dentine de la dent fonctionnelle est résorbé sur cette face, la résorption commençant à partir d'une région plus profonde chez les adultes que chez les juvéniles. Le développement dentaire commence par l'initiation d'une première dent à chaque position après une interaction entre la lame dentaire et une région profonde du mésenchyme. La lame dentaire est directement issue de la couche basale de l'épithélium oral. Quand une dent de remplacement est initiée, la lame dentaire s'étend à partir de l'organe de l'émail de la dent précédente dans la même famille. Chez *C. viridanus*, la morphogénèse et la différenciation dentaire sont similaires à celles qui ont été décrites chez les mammifères. Les cellules épithéliales localisées à l'extrémité de la lame dentaire prolifèrent et se différencient pour donner les diverses couches de l'organe de l'émail. L'épithélium dentaire interne se différencie en améloblastes qui synthétiseront la matrice de l'émail. La population de cellules du mésenchyme située en face de l'organe de l'émail forme une papille dentaire qui donnera naissance aux odontoblastes qui produiront la matrice de la prédentine dans la région située en face de l'épithélium dentaire interne. Les odontoblastes se différencient en premier et une fine couche de prédentine est déposée avant que les premiers éléments de la matrice de l'émail soient synthétisés par les améloblastes. La minéralisation des matrices dentaires commence dans la partie supérieure de la dent quand les améloblastes arrêtent de déposer la matrice de l'émail, alors qu'ils continuent leur travail le long du cône de dentine. Puis, la maturation de la matrice de l'émail progresse du sommet vers la base, alors que la minéralisation de la dentine se fait de manière centripète, depuis la jonction énamélo-dentinaire vers la cavité pulpaire.

L'hybridation *in situ* révèle l'expression du gène de l'amélogénine dans les améloblastes différenciés à différents stades du développement dentaire. La sonde d'amélogénine montre la forte présence des transcrits dans les améloblastes en activité et une diminution de l'expression avec lorsque la minéralisation de l'émail débute. Les améloblastes post-sécréteurs ne sont pas marqués et les transcrits d'amélogénine n'ont jamais été observés au niveau des odontoblastes.

Introduction

Grâce à de nombreuses investigations, consacrées au développement du squelette dermique dans de nombreuses lignées de vertébrés, nous sommes maintenant capables de réaliser l'analyse développementale comparée de ses différents éléments (Huysseune et Sire, 1998; Sire et Huysseune, 2002; Sire *et al.*, 2002). Ces travaux démontrent clairement combien peuvent être cruciales de telles comparaisons développementales pour la compréhension des relations évolutives qui lient les différents éléments du squelette dermique (*e. g.,* écailles, os dermique, rayons des nageoires, denticules, dents, ...) et, particulièrement, en ce qui concerne l'histoire des tissus dentaires, dentine et émail, et nombre de tissus apparentés aux dents (Sire et Huysseune, 2002). Cette dernière revue a bénéficié principalement des études développementales conduites sur le squelette dermique de plusieurs actinoptérygiens. En effet, cette lignée possède une grande diversité, non seulement par leur grand nombre d'espèces, mais aussi par le nombre de types différents d'éléments du squelette dermique.

Une étude comparée récente, consacrée au développement dentaire chez les vertébrés (morphogenèse et différenciation, au niveau tissulaire et de la cellulaire), a mis en évidence une absence d'information sur un tel sujet dans plusieurs lignées (Sire *et al*., 2002). En effet, chez les tétrapodes non-mammaliens, les travaux antérieurs se sont limités à certains aspects spécifiques du développement dentaire. C'est, par exemple, le cas des études sur le développement dentaire des squamates (lézards et serpents). A l'exemple des travaux d'Edmund (1960, 1969) sur le patron de remplacement dentaire, les auteurs qui se sont intéressés au développement des dents chez les reptiles ont principalement concentré leur attention sur le développement de la dentition chez les embryons et les juvéniles avec l'objectif de trouver l'origine du patron de remplacement (*e. g.*, Osborn, 1971; Westergaard, 1986; Westergaard et Ferguson, 1986). Cependant, une description qui couvrirait les principaux stades de développement, de l'initiation à l'attachement, ainsi que les processus de résorption, chez les squamates, est nécessaire pour comparer le développement dans toutes les lignées de vertébrés dans une perspective évolutive.

Le principal objectif de la présente étude était de combler ce manque d'informations en étudiant le développement dentaire d'un lézard scincidé, *Chalcides viridanus*. Ce travail bénéficie des résultats récents sur la dentition et le patron de remplacement dentaire de cette espèce (Voir chapitre 3, 1ère partie). Le développement dentaire de *C. viridanus* a montré qu'il suivait un patron régulier et reproductible permettant de prédire le stade de développement d'une dent à une position donnée.

Le second objectif qui nous a conduit à réaliser cette étude concernait plus spécifiquement l'amélogenèse. En effet, nous avons récemment entrepris une analyse évolutive de l'amélogenèse chez les vertébrés et il manquait des données chez les squamates pour pouvoir les comparer à la formation de l'émail chez les mammifères. Pour combler ce manque, il est apparu crucial de savoir si oui ou non l'amélogénine, la protéine majeure de l'émail en formation chez les mammifères, était exprimée durant le développement des dents chez *Chalcides viridanus* de la même manière que ce qui avait été décrit chez les mammifères. Cela nous a conduit à cloner le gène de l'amélogénine dans cette espèce et à analyser son expression durant le développement des dents.

Cette étude présente (1) la morphologie dentaire au cours l'ontogénèse à l'aide d'observations au microscope électronique à balayage, (2) le développement dentaire en utilisant la microscopie photonique et la microscopie électronique à transmission, avec une attention particulière pour l'amélogenèse et (3) l'expression du gène de l'amélogénine par hybridation *in situ* sur coupes à différents stades de développement de la dent.

Matériel et méthode

Matériel biologique

Depuis 1996, le scinque des Canaries, *Chalcides viridanus* (Gravenhorst) est élevé dans notre laboratoire où il se reproduit en conditions contrôlées. Les scinques sont vivipares et atteignent leur maturité sexuelle à l'âge de trois ans (J. Castanet, communication personnelle). Avant cet âge, les individus sont considérés comme juvéniles. Dans notre laboratoire, les naissances ont lieu chaque année, de la seconde moitié de juin à la fin juillet.

De nombreux spécimens, du stade embryonnaire 39 (approximativement 2 semaines avant la naissance) à l'adulte de 70 mois, ont été utilisés. Le stade 39 a été estimé d'après la table de développement de Dufaune et Hubert (1961) pour *Lacerta vivipara*. Tous les lézards ont été mesurés, tués par overdose de chloroforme, décapités puis les mâchoires disséquées.

Méthodes

Microscopie électronique à balayage (MEB)

Le quadrant inférieur droit de la mâchoire de 5 spécimens, âgés de 1; 4,5; 17; 20 et 70 mois (35, 39, 60, 70, 75 mm de longueur museau-cloaque, respectivement), ont été immergés dans de l'eau de javel à 10% pendant 20 à 40 min (en fonction de la taille de la mâchoire). Les tissus mous ont été délicatement retirés avec des pinces et un pinceau. chaque échantillon a été ensuite déshydraté dans des séries croissantes d'éthanol, séché à l'air, collé sur un support en cuivre, couvert par une fine couche d'or/palladium, et observé dans un SEM JEOL 35 opérant à 25 kV.

Microscopie photonique (MP) et électronique à transmission (MET)

La région antérieure de la mâchoire inférieure de 5 spécimens (embryon au stade 39, 30 mm de longueur museau-cloaque; 1 jour après la naissance, 41 mm; 1 mois, 43 mm; 5 mois, 42 mm; et 60 mois, femelle adulte, 72 mm) a été sectionnée en petits morceaux et fixé pendant 2 h, à température ambiante, dans une mixture contenant 1,5% de glutaraldéhyde et 1,5% de paraformaldéhyde dans 0,1 M de tampon cacodylate (pH 7,4). Certains échantillons ont été décalcifiés durant 7 jours (petits spécimens) ou 14 jours, à 4°C dans le même fixateur auquel de l'EDTA 5% a été ajouté. Après rinçage dans le même tampon pour la nuit à 4°C, les échantillons ont été post-fixés dans du tampon cacodylate avec de l'osmium tétroxide à 1%, puis rapidement rincés. Ils ont ensuite été déshydratés en utilisant des séries croissantes d'éthanol et inclus dans l'épon pour des coupes transversales et longitudinales.

Des bourgeons dentaires ont été sélectionnés à des stades bien précis pour illustrer différentes étapes du développement dentaire ou, plus particulièrement, de l'amélogenèse. Dans ce but, cette sélection s'est appuyée sur le résultat de l'étude précédente qui montre que l'on peut prédire le stade de développement d'un bourgeon dentaire à une position donnée en connaissant le stade de développement du bourgeon situé à la position précédente (Voir chapitre 3, 1ère partie). De nombreux échantillons ont été sectionnés en coupes semi-fines transversales et longitudinales (1 μm d'épaisseur) qui ont été colorées au bleu de toluidine; les séries transversales ont été interrompues au niveau du premier bourgeon dentaire rencontré afin d'en déterminer le stade de développement. En fonction de ce stade, le choix s'est porté sur l'un ou l'autre des bourgeons suivant qui sont toujours moins développés. Par endroits, des coupes ultra-fines ont été confectionnées, contrastées avec de l'acétate uranyle et du citrate de plomb et visualisées dans un microscope électronique à transmission Philips 201 opérant à 60 kV.

ADNc d'amélogénine

Immédiatement après dissection, les mâchoires inférieures (contenant des dents en développement) de deux spécimens (45 et 47 mm de longueur museau-cloaque) ont été plongées dans l'azote liquide et réduites en fine poudre. Les ARN totaux ont été purifiés (Rneasy Midi, Qiagen S.A.), puis les ARNm ont été isolés (Oligotex, Qiagen S.A.), et transformés en ADNc par RT-PCR en utilisant une amorce oligo$(dT)_{18}$ (First Strand cDNA, MBI fermentas).

Amplification par PCR

Les amorces suivantes ont été définies après comparaison avec les séquences nucléotidiques des gènes de l'amélogénine déjà connus chez les mammifères et les reptiles:

Ex2 (sens; 5'-CTGGACTTTGGTTATGTGCC-3') a été construite pour s'hybrider à l'exon 2 et *Amel 529* (antisens; 5'-CACTTCTTCTTGCTTGGTCTT-3') a été construite pour s'hybrider à la région 5' de l'exon 6.

L'ADNc (1 µl) a été amplifié par PCR dans une mixture composée de 5 µl de tampon de Taq (10X) pH 8,8, 3µl de $MgCl_2$ 25 mM, et 1 µl de dNTP 10 mM, en présence des amorces *Ex2* et *Amel529*, et de 0,3 µl de polymérase Red Hot (Advanced Biotechnologies Ltd.). L'amplification a été réalisée dans un thermocycleur (Genius Techne) durant 38 cycles, dont chacun a consisté en : 1 min de dénaturation à 94°C, 1 min d'hybridation des amorces à 59°C et 1 min de synthèse de l'ADN à 72°C. La synthèse finale a duré 20 min à 72°C. Un fragment d'amélogénine de 600 pb a été amplifié.

Clonage

Un microlitre de produit de PCR a été isolé, inséré dans un plasmide pCR 2.1-TOPO (InVitrogen S.A.) utilisant la méthode du TA-cloning, puis ce plasmide a été utilisé pour transformer des bactéries compétentes *E. coli* TOP10F (InVitrogen S.A.).Les bactéries transformées ont été incubées à 37°C pendant la nuit, sur boîte de pétri, dans un milieu Luria-broth agar contenant de l'ampicilline. Les clones positifs ont été incubés à nouveau à 37°C pendant la nuit dans 3 ml de milieu Luria-broth liquide contenant de l'ampicilline, puis les bactéries ont été lysées dans 200 µl de NaOH 0.2 M – SDS 1%, à 0°C pendant 5 min.

Ensuite, 150 µl d'acétate de potassium 3M ont été ajoutés à 0°C pendant 5 min pour précipiter les protéines. Les plasmides ont été purifiés en utilisant un mélange de phénol/chloroforme. Le séquençage a été réalisé par Génome Express S.A.

Préparation de la sonde et marquage

Les plasmides ont été réamplifiés pour obtenir un fragment d'amélogénine et l'utiliser pour fabriquer la sonde. Une PCR asymétrique, amplification linéaire de l'ADN, a été réalisée comme suit : le fragment d'amélogénine de 600 pb (*Ex2-Amel529*) (25 ng) a été ajouté à une solution (50 µl) contenant du Tris HCl 10 mM (pH 8.3), du KCl 50mM, du $MgCl_2$ 15 mM, une mixture de dATP, dGTP et dTTP, 5 µl de [α-^{33}P] dCTP (3000 Ci/mmol), de la Taq polymérase (2,5 U), et l'amorce antisens (20 pmol). L'amplification a été obtenue après 10 cycles à 55°C et la réaction a été arrêtée en ajoutant de l'EDTA 0.2 M (5 µl). L'amorce sens a été utilisée pour synthétiser la sonde contrôle.

Une sonde d'ADN simple brin avec une activité spécifique de $2{,}8.10^6$ dpm/pmol a été obtenue. La sonde radio-marquée a été précipitée en présence d'ARNt avec de l'éthanol froid.

Hybridation in situ sur coupes

Obtention des coupes

Deux spécimens de 12 mois (55 et 60 mm de longueur museau-cloaque) ont été tués en juillet, les mâchoires inférieures ont été disséquées et les quadrants gauche et droit séparés au niveau de la symphyse. Les mâchoires ont été immédiatement plongées dans une solution de Formoy (10% d'acide acétique, 30 % d'éthanol absolu et 60% de paraformaldéhyde à 37%) pendant 24 h à température ambiante, puis placées dans une solution de paraformaldéhyde 37% et d'acide acétique 1 M pendant 24 h à 4°C. La décalcification a été complétée dans une solution d'acide acétique 1 M pendant 24 h à 4°C, puis les mâchoires ont été déshydratées dans une série croissante d'éthanol, incluses dans la paraffine (pour des coupes transversales ou longitudinales) et les blocs ont été conservés à 4°C. Les coupes sériées de 10 µm d'épaisseur ont été transférées sur des lames recouvertes de 3-aminopropyl-triehoxysilane et les stades d'amélogenèse ont été sélectionnés comme cela a été décrit précédemment (page 168, *MET*). Les lames ont été séchées à l'air, déparaffinées dans du xylène, réhydratées dans une série décroissante d'éthanol, fixées dans le paraformaldéhyde à

4% dans du PBS pendant 20 min, lavées 2 fois dans le PBS, puis déshydratées dans une série croissante d'éthanol et séchées à l'air.

Hybridation

Les coupes ont été préhybridées pendant 2 h à 37°C dans une chambre humide avec une solution contenant du formamide déionisé à 50%, du $NaPO_4$ 20 mM (pH 7,4), du SSC 2x (SSC 1x = NaCl 150 mM, Trisodium Citrate, pH 7), de l'EDTA 5mM, de la solution de Denhart, de l'ADN de sperme de hareng natif (50 µg/ml) et des ARNt (500 µg/ml). Les sections ont été ensuite rapidement rincées dans du SSC (2x), déshydratées dans l'éthanol et séchées à l'air. La sonde marquée a été ajoutée à la solution d'hybridation (solution de préhybridation dans laquelle le dextran sulfate remplace l'ARNt). Après dénaturation pendant 10 min à 100°C, 50 µg/ml d'ADN de sperme de Hareng a été ajouté. L'hybridation a eu lieu pendant la nuit, à 37°C dans une chambre humide, après que 30 µl de solution d'hybridation ont été déposés sur chaque coupe. Les coupes ont été rincées avec 50% de SSC 2x pendant 30 min, puis dans une mixture de SSC 1x / formamide 50% pendant 30 min et dans du SSC 0,5x pendant 30 min. Les coupes ont été lavées dans du SSC 0,25x, déshydratées dans l'éthanol, puis séchées à l'air.

Les autoradiographies ont été obtenues après immersion des lames dans une émulsion LM-1 (Amersham, Buckinghamshire, England). Après une durée d'exposition de 1 à 6 jours, l'émulsion a été révélée dans du D-19 (Kodak) et fixée dans du sodium thiosulfate. Les sections ont été colorées avec de l'Hemalun de Masson et séchées à l'air.

Résultats

Forme des dents, usure et résorption

Les observations au MEB des mâchoires de *Chalcides viridanus* apportent des informations (1) sur la forme, la taille et l'orientation dans dents (chez les juvéniles *versus* les adultes, et les dents antérieures *versus* postérieures), (2) sur les relations des dents par rapport à l'os de support, le dentaire, (3) sur différents aspects de l'usure dentaire, et (4) sur les caractéristiques de la résorption des dents fonctionnelles en relation avec la présence d'une dent de remplacement en développement. Ces données, en complément de celles récemment acquises sur la description du patron de remplacement (voir ce chapitre, 1ère partie), aident à mieux localiser les bourgeons dentaires par rapport à la position de la dent fonctionnelle et à interpréter la position et l'orientation des coupes. Chez *Chalcides viridanus*, les dents sont localisées dans un sillon dentaire, une dépression de l'os de la mâchoire bordée par une paroi labial et lingual, la première étant plus haute que l'autre (Figures 1a-d). Les dents sont pleurodontes dans leur mode d'attache sur l'os, c'est-à-dire qu'elles sont ankylosées sur la face interne de la paroi labiale de l'os de la mâchoire. La forme du sillon dentaire et des dents changent durant l'ontogénèse. En conséquence, les relations des dents avec l'os de support changent également. Chez les adultes, le sillon dentaire est plus profond et étroit, et la paroi labiale plus haute que chez les juvéniles. Plus de la moitié de la hauteur de la dent dépasse du sommet de la paroi labiale chez les juvéniles contre seulement le tiers ou le quart chez les adultes (comparer, *e. g.*, Figures 1a et d). Dans chaque arcade, il y a une gradation des dimensions des dents le long de la rangée dentaire. Les dents localisées à l'arrière sont plus larges que les dents antérieures. Chez les spécimens âgés de 1 mois, les dents ont, en moyenne, une hauteur de 450 μm et la couronne un diamètre de 150 μm, alors que chez les adultes de 70 mois, la dent atteint une hauteur de 1 mm et la couronne un diamètre de 400 μm. La taille de la dent est globalement similaire sur chaque arcade dentaire (homodontie) mais varie durant l'ontogénèse alors que la rangée dentaire devient plus régulière. Les dents des juvéniles sont coniques et séparées, en moyenne, par un espace de 200 μm. La couronne est soit droite, soit légèrement incurvée en arrière et la base de dentine est entourée par du tissu osseux (Figures 1a, b). Les dents adultes sont cylindriques et séparées, en moyenne, par un espace de 100 μm seulement. La largeur est similaire du sommet à la base et la courbure de la couronne est plus prononcée que chez les juvéniles (Figures 1c, d). Les couronnes sont incurvées vers l'intérieur, perpendiculairement au contour de l'arcade et, durant l'ontogénie, les dents acquièrent leur forme caractéristique (grâce remplacement permanent des dents) (Figures 1 à g). La surface de la couronne (émail) est plutôt lisse chez les lézards de 1 mois et une unique cuspide est observable sur la face linguale, soulignée par une arête. Chez les

juvéniles de 17 mois, cette arête est renforcée et la surface linguale de la couronne est ornée de deux ou trois crêtes (Figure 1f). Chez les lézards âgés de 20 mois, l'arête principale se différencie progressivement pour former une cuspide secondaire alors que les crêtes verticales sont devenues proéminentes et que leur nombre a augmenté sur la surface linguale (Figure 1g). Chez les juvéniles de 1 et 4,5 mois, la base du cône de dentine a un diamètre de 300 µm qui est le double de celui de la couronne (Figures 1a, b, h, i). A ce stade, la base de la dent est entièrement entourée par de l'os qui s'étend dans l'espace séparant les dents successives. Chez les adultes, la racine et la couronne ont un diamètre similaire et l'os est moins visible autour de la base des dents qui sont plus proches les unes des autres que chez les juvéniles (Figure 1j). Un simple foramen est observé à la base de la dent du côté lingual. Ces foramens mènent à la cavité pulpaire et permettent probablement le passage du réseau vasculaire et nerveux.

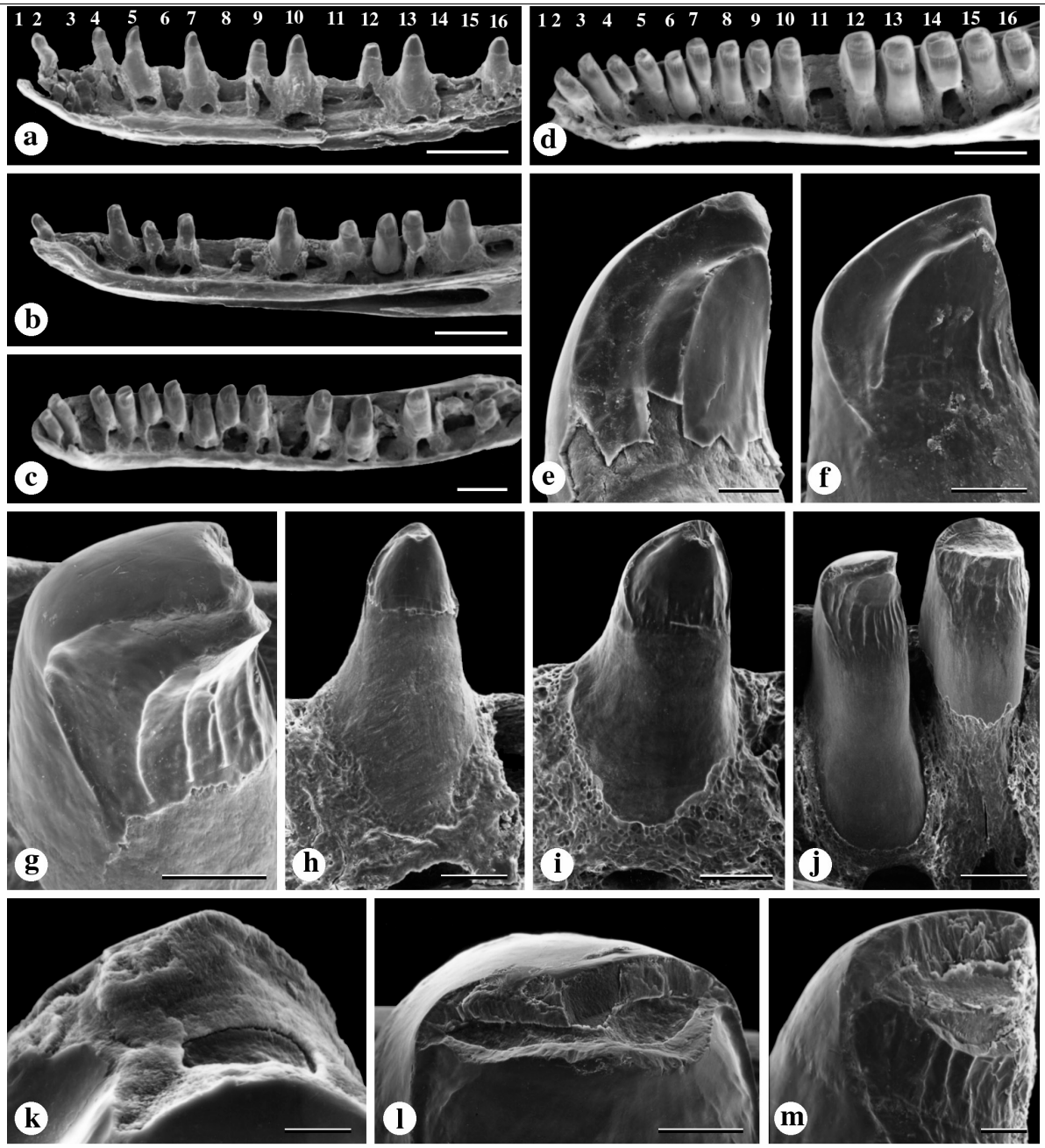


Figure 1. Photographies prises au microscope électronique à balayage de quadrants inférieurs droits de spécimens sélectionnés à partir d'une série de croissance du lézard scincidé, *Chalcides viridanus* [âgés de 1 mois (**a, e, h, k**) 4,5 mois (**b, i**), 17 mois (**f**), 20 mois (**c, g**) et 70 mois (**d, j, l, m**)]. La partie antérieure est sur la gauche des images et toutes les photos présentent la vue linguale. Les dents de remplacement en développement et les dents fonctionnelles qui étaient en phase avancée de résorption ont été perdues durant le nettoyage des mâchoires, laissant libres quelques positions dentaires. Les positions dentaires sont numérotées de 1 à 16 en (**a**) et (**b**). **a-d**: Faible grandissement des quadrants de mâchoires. Noter le changement de forme et l'orientation des dents durant l'ontogénèse. Plus de positions sont perdues chez les spécimens jeunes (**a, b**) que chez les vieux (**c, d**), indiquant que plus de dents étaient en cours de remplacement. **e-g**: Fort grandissement de la couronne d'une dent antérieure nouvellement fonctionnelle (positions 4 et 5). Noter la différenciation progressive des cuspides et des ornementations de la couverture d'émail sur la face linguale, durant l'ontogénèse. **h-j**: Aspect de la base d'une dent récemment fonctionnelle. Noter la modification progressive du cône de dentine, de conique à cylindrique, durant l'ontogénèse. **k-m**: Aspects variés de l'usure dentaire. Les usures sont localisées au sommet de la dent, dès le premier mois de la vie (**k**). Une usure plus importante est observée chez les spécimens plus vieux, chez lesquels une grande partie de l'émail est érodée, laissant apparaître la surface de la dentine qui est aussi érodée (**l, m**). Echelle : a, b, c = 500 μm; d = 1 mm; e = 25 μm; f, g = 50 μm; h, i = 100 μm; j = 200 μm; k = 10 μm; l = 100 μm; m = 50 μm.

L'observation détaillée de la couronne dentaire révèle que cette région est souvent soumise à l'usure (Figures 1 k-m), une caractéristique qui est à peine visible à faible grandissement. Chez les juvéniles, aussi bien chez les individus âgés de 1 mois que chez les autres, de nombreuses dents usées sont observées sur chaque arcade, bien avant leur remplacement.

Cependant, je n'ai pas pu observer la surface de la couronne de dents très résorbées (probablement avec une usure plus prononcée) car elles sont tombées durant la préparation des mâchoires pour la MEB. La couche d'émail est usée (ou partiellement cassée) au niveau du sommet des cuspides sur la face linguale (Figure 1k), puis l'aire d'érosion s'étend lingualement à toute la surface du sommet de la couronne (Figures 1l, m). Dans ces dents, l'émail est suffisamment usé pour exposer la surface de la dentine qui est aussi érodée. Ces usures permettent de déterminer que la couche d'émail est plus épaisse chez les adultes que chez les juvéniles (65 μm contre 20 μm, respectivement, au sommet de la dent) et plus épaisse du côté labial que du côté lingual (65 μm contre 35 μm, respectivement, au sommet de la dent, 65 μm contre 20 μm au milieu de la couronne) (Figures 1k-m). Dans toutes les dents examinées, l'usure concerne toujours le sommet de la couronne, et elle est seulement localisée du côté lingual. Le côté labial et les parois de la couverture d'émail ne sont pas touchés. L'usure ne s'étend pas sur le cône de dentine, même chez les spécimens les plus vieux (approximativement 6 ans), mais la partie supérieure de la dentine est érodée dans le cas de larges usures. Il est probable que le remplacement dentaire empêche une telle usure de s'étendre plus profondément.

Chez *Chalcides viridanus*, quel que soit le stade ontogénétique, chaque position dentaire possède au moins une dent de remplacement plus ou moins développée (voir ce chapitre, 1ère partie). La position du bourgeon dentaire peut varier en fonction du stade de développement du bourgeon et de l'âge du lézard, et cela se voit à la façon dont la base de la dent fonctionnelle est résorbée (Figure 2). Tous les bourgeons débutent leur morphogenèse et leurs premiers stades de différenciation sur la face linguale. A cet endroit, ils sont protégés des chocs par la dent fonctionnelle et le sillon dentaire. Chez les juvéniles, les bourgeons sont localisés latéralement, c'est-à-dire, qu'ils se développent en position adjacente à la base de la dent et de manière légèrement postérieure à la dent fonctionnelle. En conséquence, la résorption commence à la base de la dentine et concerne la surface de dentine de la face linguale (Figures 2a, c). La paroi de dentine située, en face du bourgeon qui se développe, est résorbée et la cavité pulpaire est alors ouverte. Le bourgeon grandissant, la surface linguale

du cône de dentine est entièrement résorbée et la dent en développement pénètre progressivement dans la cavité pulpaire. La plus grande partie de la région supérieure et labiale de la dent fonctionnelle reste non-résorbée durant ce processus (Figure 2b). Chez le lézard pré-mature et adulte, le bourgeon est également localisé latéralement, mais il est moins profondément enfoncé dans le sillon dentaire. En conséquence, la résorption touche directement toute la base de la dent et le bourgeon pénètre directement dans la cavité pulpaire, très tôt au cours du développement (Figures 2d, e). Le bourgeon dentaire grandissant, la paroi de dentine est résorbée par la base (Figure 2e) et la partie supérieure de la dent fonctionnelle peut rester intacte alors que le reste est entièrement résorbé. Comme on l'observe dans la figure 2e, la couronne de la dent en développement est complètement formée longtemps avant que le cône de dentine commence à se mettre en place.

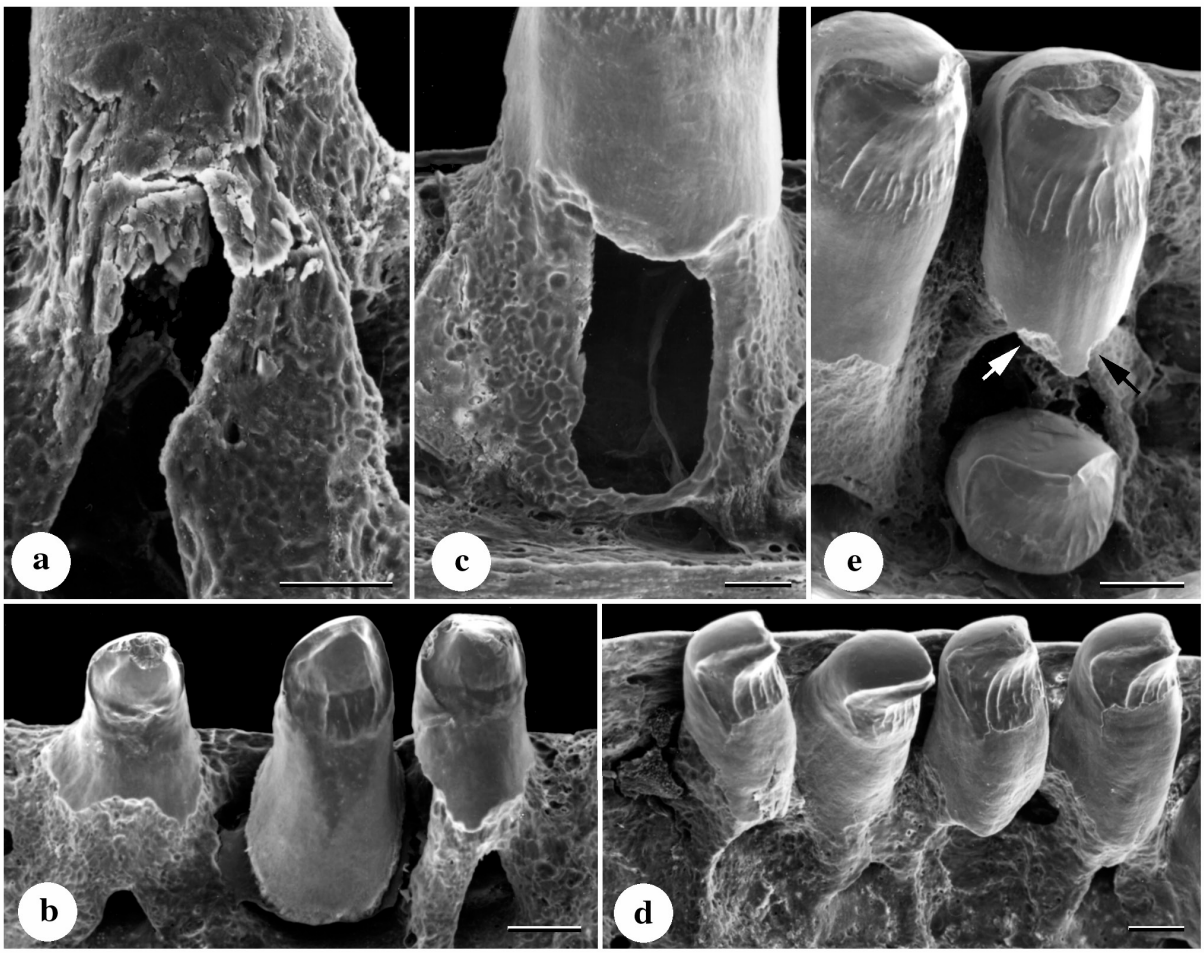


Figure 2. Différents aspects de la résorption en relation avec la croissance de la dent de remplacement chez *Chalcides viridanus* [spécimens âgés de 1 mois (**a**), 4,5 mois (**b**), 17 mois (**c**), 20 mois (**d**) et 70 mois (**e**)]. Les dents de remplacement ont été perdues durant le nettoyage des mâchoires excepté en (**e**). L'emplacement du cône de dentine en résorption change durant l'ontogénèse. La dentine est résorbée latéralement (sur la face linguale) chez les jeunes lézards (**a, b, c**) et par dessous chez les vieux spécimens (**d, e**) chez lesquels le bourgeon pénètre dans la cavité pulpaire. Noter que la couronne de la dent de remplacement est complètement formée avant que le cône de dentine ne se développe et que l'apex de la dent qu'il va remplacer est usé (**e**). L'activité ostéoclastique est clairement visible sur la surface de la base de la dent (flèches).
Echelle : a, c = 50 μm; b-d = 100 μm; e = 200 μm.

Développement dentaire chez Chalcides viridanus

Les mâchoires débitée en coupes sériées dans les séries de croissance de *Chalcides viridanus* nous ont fourni de nombreux stades de développement. Par commodité, nous avons sélectionné les stades de développement de divers spécimens pour illustrer les différentes étapes de développement avec un intérêt particulier pour la formation de l'émail. Les données ont été obtenues par microscopie photonique (Figures 3, 4, 5, 8) et les détails de

l'amélogenèse ont été observés au microscope électronique à transmission (TEM) (Figures 6, 7).

Initiation de la dent

L'étude précédente de la dentition et du patron de remplacement chez *Chalcides viridanus* a révélé que le nombre de positions dentaires est déjà complet lorsque le lézard vient au monde et aucune nouvelle position n'est acquise durant la croissance (voir ce chapitre, 1ère partie). A la naissance, le jeune scincidé est équipé de dents fonctionnelles occupant chacune des 16 positions dentaires (voir 1ère partie). De plus, à chaque position, l'initiation des dents de première génération a lieu durant le développement embryonnaire. Chez l'embryon, l'initiation primaire est définie comme celle de la première dent de chaque position sur la rangée. Cette initiation commence par un épaississement local de la couche basale de l'épithélium buccal (non illustré). Puis, une extension lamellaire de la couche basale de l'épithélium buccal (appelée lame dentaire) s'enfonce dans le mésenchyme. L'extrémité de la lame dentaire interagit avec une population de cellules du mésenchyme localisée près de la surface de l'os en développement qui deviendra l'os de support de la future dent (Figure 3a). Contrairement à l'initiation primaire, l'initiation secondaire est définie comme l'initiation de chaque dent de remplacement ayant lieu à chaque position. De plus, ce type d'initiation secondaire est le principal processus d'initiation dentaire quand on considère que chaque position est remplacée plusieurs fois durant la vie du lézard. La lame dentaire se développe à partir de la région impliquée dans l'initiation de la dent précédente de la même famille (Figure 3b). Les deux types de processus d'initiation sont similaires. L'extension de la lame dentaire a une épaisseur de 10 μm et elle est composée de deux couches de cellules indifférenciées issues de la couche basale de l'épithélium buccal (Figures 3a, b, c).

L'interaction épithélio-mésenchymateuse qui précède le développement dentaire se présente comme (1) un épaississement de l'extrémité de la lame dentaire résultant d'une augmentation du volume des cellules et (2) une accumulation des cellules du mésenchyme entourant cette extrémité (voir insert sur la Figure 3a et Figure 3c). A ce stade, la lame dentaire est profondément enfoncée, à 75 μm de la surface de l'os en développement (Figure 3a). La région distale de la lame dentaire et, en particulier, sa face labiale (interne) est impliquée dans la première étape de la morphogenèse dentaire et de la différenciation, alors que la couche linguale (externe) ne l'est apparemment pas (Figures 3b, d). Cette dernière semble impliquée dans la conservation de la continuité de la lame dentaire alors que les cellules à partir desquelles les trois couches composant l'organe de l'émail se différencieront (épithélium dentaire externe, réticulum étoilé et épithélium dentaire interne), semblent dériver de la première. L'initiation dentaire de type secondaire commence après que la dent précédente a démarré sa différenciation. La lame dentaire se développe depuis la face linguale de la dent précédente jusqu'à la région profonde du mésenchyme où une nouvelle dent de remplacement sera initiée (Figure 3b). Chez l'embryon, pour lequel la dent n'est pas encore fonctionnelle (c'est-à-dire, attachée à l'os de support et perçant l'épithélium buccal), il y a seulement deux dents en formation à la même position : une dent déjà bien différenciée avec de l'émail en maturation, une couche de dentine en extension, et un bourgeon dentaire récemment initié. Au contraire, les juvéniles et les adultes, chez qui les dents sont fonctionnelles, possèdent trois dents à la même position (elles constituent une famille dentaire) : la dent fonctionnelle (attachée à l'os de support et perçant l'épithélium), une dent de remplacement déjà bien différenciée (qui provoque la résorption de la base de la dent fonctionnelle) et un bourgeon dentaire chez lequel les tissus dentaires sont en train de se

déposer (Figure 3e). Ces trois dents sont liées, l'une à l'autre, à la lame dentaire, qui est déjà en extension dans une région plus profonde du mésenchyme pour initier une quatrième dent.

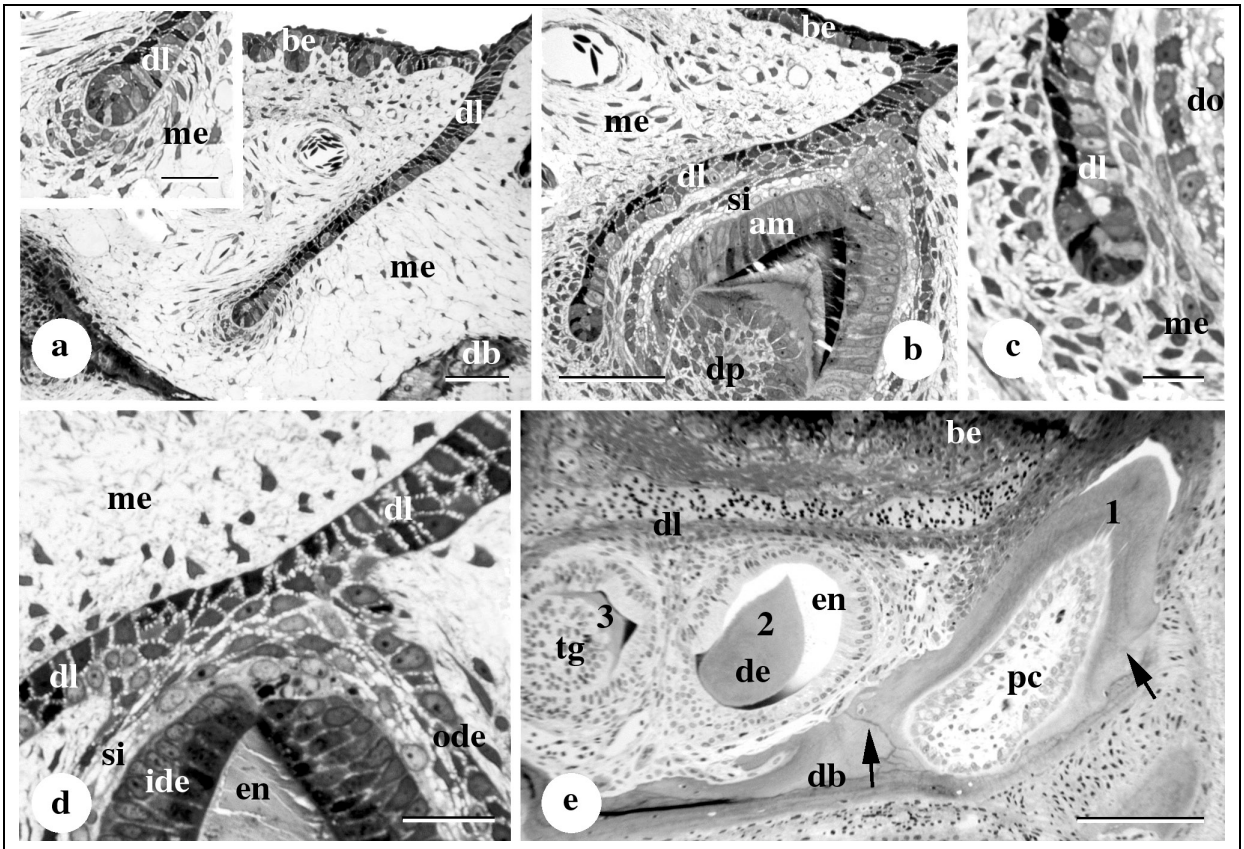


Figure 3. *Chalcides viridanus*. Coupe transversale de 1 μm d'épaisseur colorée au bleu de toluidine. Différents types d'initiation dentaire chez l'embryon (**a-d**) et chez l'adulte (**e**). Mâchoire inférieure droite, côté labial sur la droite. **a:** Initiation de type primaire. La lame dentaire s'enfonce dans le mésenchyme dans lequel de forme l'os dentaire. Insert = détail de la lame dentaire interagissant avec une population de cellules du mésenchyme. **b:** Initiation de type secondaire. La lame dentaire se développe depuis la face linguale de l'épithélium dentaire entourant une dent en formation. **c:** Détail de l'extrémité de la lame dentaire en (**b**) montrant les interactions avec la population de cellules mésenchymateuses. **d:** Détail de la connexion de la dent avec la lame dentaire. **e:** Trois dents constituent une famille dentaire, chez un adulte; une dent fonctionnelle (1), qui s'est attachée récemment à l'os de support (flèches); une dent de remplacement (2) dont l'émail est en maturation, et un bourgeon dentaire (3) en développement alors que l'extrémité de la lame s'enfonce dans le mésenchyme. Ces dents sont liées les unes aux autres par la lame dentaire issue de l'épithélium buccal. Les flèches indiquent les limites (lignes de cément) entre la matrice de l'os d'attache et celle du dentaire. Echelle : a = 50 μm; insert = 20 μm; b = 50 μm; c = 10 μm; d = 20 μm; e = 100 μm.

Morphogenèse dentaire et cytodifférenciation

Embryon (stade 39). L'interaction épithélio-(région distale de la lame dentaire) mésenchymateuse (région proche de l'os de support), résulte dans l'accumulation d'une population dense de cellules du mésenchyme entourant l'extrémité de la lame dentaire (Figure 4a). Cette région distale de la lame dentaire est maintenant composée par de nombreuses cellules qui ont proliféré et commencé un processus de différenciation (Figure 4b). L'extrémité de la lame dentaire s'est repliée dans une direction labiale et forme une petite cloche (30 μm de diamètre) qui entoure un petit groupe de cellules du mésenchyme constituant la papille dentaire.

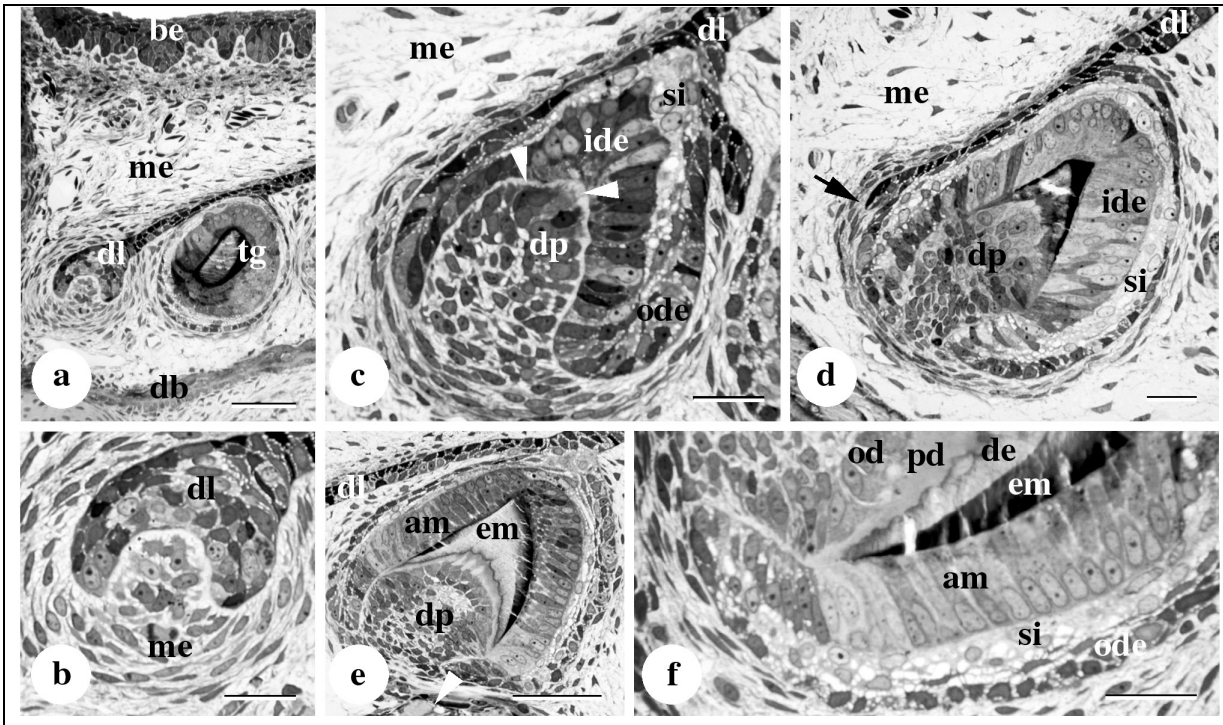


Figure 4. *Chalcides viridanus*, embryon, stade 39. Coupe transversale de 1 μm d'épaisseur colorée au bleu de toluidine. Divers stades de développement dentaire. Mâchoire inférieure droite, côté labial sur la droite. **a:** Les cellules localisées à l'extrémité de la lame dentaire forment une cloche qui entoure la papille dentaire. **b:** Détail de (**a**) montrant la région d'interaction entre la papille dentaire et les cellules de la lame dentaire. **c:** La cloche s'est développée et les cellules de l'organe dentaire et de la papille dentaire sont différenciées. Les premiers éléments de la matrice dentaire (prédentine) sont déposés (flèches). **d:** La dent est maintenant bien différenciée. Les matrices de la dentine et de l'émail ont commencé à se minéraliser. La flèche indique la surface du dentaire. **f:** Détail de la base de la dent montrant des améloblastes bien différenciés déposant la matrice de l'émail, et les odontoblastes déposant la prédentine. Echelle : a, e = 50 μm; b-d, f = 20 μm.

Les cellules de la papille sont condensées et deviennent plus arrondies alors que la majorité des cellules du mésenchyme entourant cette papille restent indifférenciées et organisées en couches concentriques autour de la cloche. Puis l'organe dentaire grandit (100 μm de hauteur, 80 μm d'épaisseur à sa base), les bordures de la cloche s'étendent à l'intérieur du mésenchyme, et les différentes couches dentaires se différencient. Trois couches peuvent être distinguées : l'épithélium dentaire interne, le réticulum étoilé et l'épithélium dentaire externe (Figure 4c). Les cellules de l'épithélium dentaire interne (= pré-améloblastes) se sont différenciées dans la région faisant face à la papille dentaire alors que l'épithélium dentaire externe est situé en face des cellules indifférenciées du mésenchyme. Le réticulum étoilé est localisé entre les deux couches précédentes et il est composé de cellules indifférenciées arrondies et séparées par le larges espaces intercellulaires. Le côté labial de la cloche est plus épais et les pré-améloblastes sont plus différenciés que sur le côté lingual (Figure 4c). La papille dentaire a maintenant un diamètre de 40 à 50 μm et le nombre de cellules a considérablement augmenté. Les cellules situées en face des améloblastes sont différenciées en odontoblastes et la première matrice de collagène non minéralisée a été déposée (prédentine). L'étape suivante du développement dentaire est caractérisée par une meilleure organisation de la cloche épithéliale (maintenant 160 μm de hauteur). Les matrices de la dent, l'émail et la dentine, sont plus épaisses, la première couche déposée est minéralisée et la minéralisation a lieu de façon centrifuge (Figure 4d). La matrice de l'émail est déposée par

des améloblastes allongés de 20 μm de haut et de 3 μm de large, dont le noyau est localisé distalement. La cuspide dentaire est formée par les améloblastes dans la région supérieure de l'organe de l'émail (Figure 4b). La prédentine est déposée par des odontoblastes arrondis localisés dans la région supérieure de la papille dentaire dense. Dans la région basale de l'organe dentaire, et à partir de la surface linguale de l'épithélium dentaire externe, la lame dentaire a commencé à se développer à nouveau vers le mésenchyme (Figure 4d). Cette lame dentaire s'allonge ensuite et se compose de deux couches (Figure 4e). La dent en formation continue de grandir. La maturation de l'émail augmente dans la région supérieure de la dent, alors que la matrice de l'émail est toujours déposée à la base de la dent par les améloblastes bien différenciés (Figure 4f). Les odontoblastes sont organisés en épithélium et ils produisent la matrice de la prédentine dans la région supérieure de la dent (Figure 4e). Le dépôt de la prédentine s'étend vers la base de la dent, qui est maintenant proche de la surface de l'os de support, mais la dent n'est pas encore attachée (Figures 4e, f). La cavité pulpaire n'est pas réellement distincte et ni les nerfs, ni les vaisseaux sanguins y sont visibles à ce stade. Toutes les dents bien formées examinées chez cet embryon ont atteint ce stade de développement. Aucune dent n'a été trouvée attachée à l'os de support, alors que toutes les dents fonctionnelles étaient attachées chez le spécimen âgé de 1 jour (non illustré).

Juvéniles et adultes. Les descriptions ci-dessus sont aussi valables pour le développement dentaire chez l'adulte, le subadulte et les juvéniles. Les seules variations concernent la taille et la forme de la dent en formation (mais pas les cellules qui sont seulement plus nombreuses), et l'os d'attache et de support. La première partie présente une description générale au microscope photonique (Figure 5) et elle est suivie par des descriptions détaillées de certains stades au niveau du MET (Figures 6 h, 7).

Les améloblastes commencent à déposer la matrice de l'émail dès que les premiers éléments de la matrice de prédentine sont déposés par les odontoblastes (Figure 5a). Les premières protéines de la matrice de l'émail, fortement métachromatiques avec le bleu de toluidine, sont déposées par les améloblastes à l'apex de la dent et la couche s'étend sur une petite surface de chaque côté. Les améloblastes sécréteurs sont allongés, étroits et juxtaposés les uns aux autres, perpendiculairement à la surface de la dent, et leur noyau est localisé dans leur région distale. Puis, la matrice de l'émail s'épaissit et elle est déposée sur toute la surface de la dent, alors que la prédentine déposée dans cette région est déjà imprégnée par des protéines non-collagéniques (Figure 5b). La forme de l'organe de l'émail change quand la dent passe d'une forme conique à une forme bicuspidée. Au sommet de la dent, l'aspect des cellules change. De nombreux espaces inter-cellulaires apparaissent entre ces améloblastes par rapport aux cellules voisines qui restent juxtaposées et proches les unes des autres. Ces changements sont concomitants au processus d'initiation de minéralisation de la matrice de l'émail (Figure 5c). La matrice de l'émail s'épaissit uniquement grâce à l'activité des améloblastes. La membrane des améloblastes qui entoure la matrice de l'émail est régulière et les processus de Tomes, décrits lors de l'amélogénèse des mammifères, n'existent pas (Figure 5d). La minéralisation de la matrice de l'émail commence dans la partie supérieure de la dent quand la matrice de la dentine est minéralisée. Puis la minéralisation se poursuit vers la base de la dent (Figure 5e). La maturation de l'émail commence puis s'étend, depuis le sommet vers la base et elle s'intensifie de telle sorte qu'aucune matrice organique n'est visible après décalcification (Figures 5f, g). Au sommet de la dent, seuls quelques granules noirs et éparses restent en place dans la matrice de l'émail et il disparaissent finalement quand la maturation de l'émail est achevée. Pendant ce temps, la matrice de l'émail est toujours déposée à la base de la dent où elle deviendra mature plus tard, quand la dent sera presque attachée (Figures 5f, g).

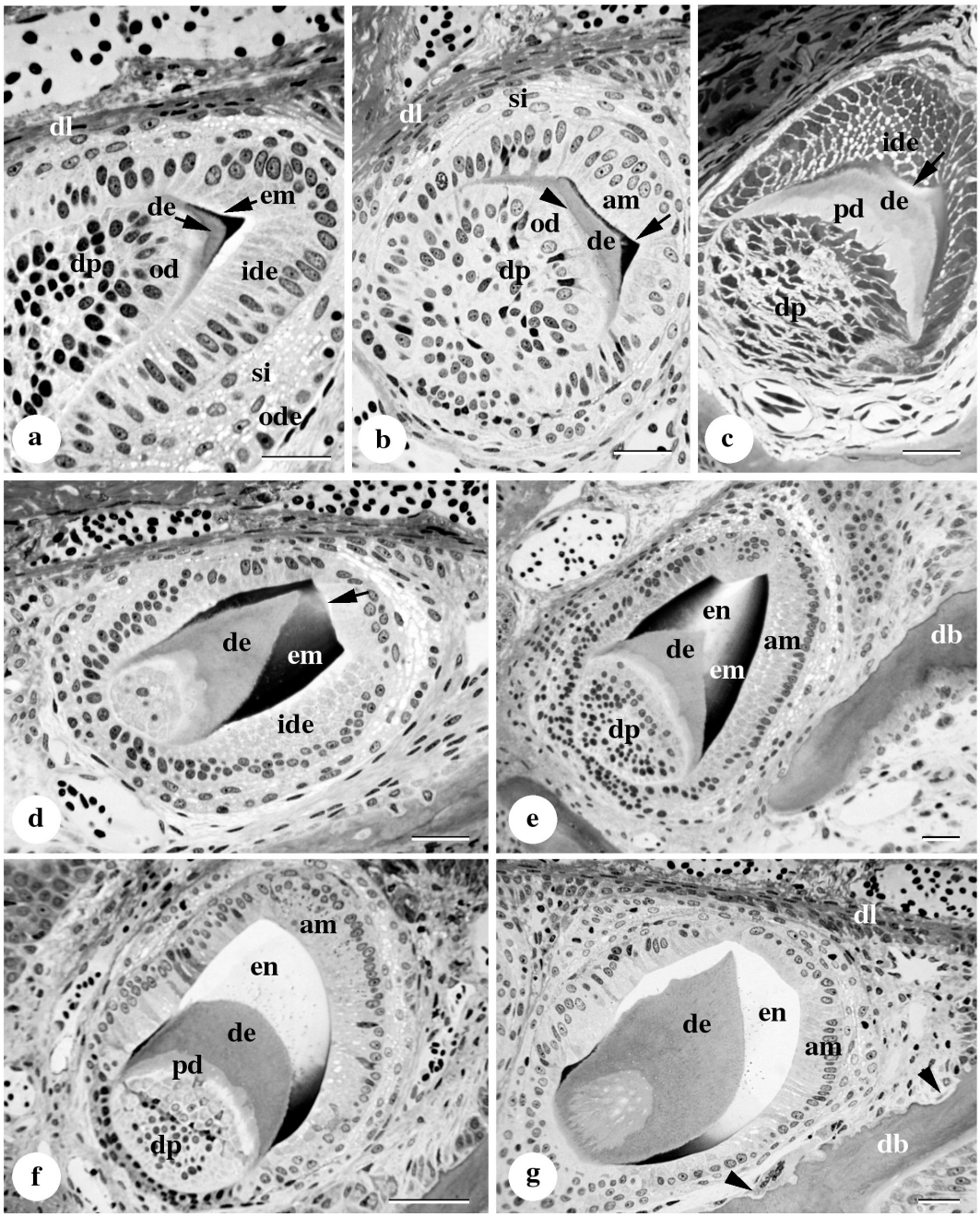


Figure 5. *Chalcides viridanus*, juvénile de 5 mois et adulte de 70 mois. Coupes transversales de 1 µm d'épaisseur colorées au bleu de toluidine. Mâchoire inférieure droite, côté labial sur la droite. Échantillons décalcifiés à l'EDTA. Divers stades de différenciation de la dent de remplacement avec une attention particulière pour la formation de l'émail. **a:** Stade précoce de dépôt des tissus dentaires. La matrice de l'émail est fortement métachromatique. La limite entre dentine et prédentine est visible grâce à une ligne dense. **b:** Extension de la surface couverte par la matrice de l'émail (flèche). La flèche indique la limite entre dentine et prédentine. **c:** La matrice de l'émail est déposée à la surface de la dentine, le long des côtés de la dent. Les cuspides sont visibles à l'apex de la dent où le processus de minéralisation a commencé (flèche). **d:** Une fine couche d'émail est déposée. Les cuspides sont visibles à la moitié de la partie supérieure de l'émail (flèche). Noter l'absence des processus de Tomes. **e:** La minéralisation concerne l'apex de la dent où la matrice est progressivement éliminée de la région située en dessous de la dentine. **f:** Coupe de la région supérieure d'une dent. La maturation de l'émail est bien avancée dans cette région de la dent et a eu lieu de l'apex vers la base où la matrice est toujours en dépôt. Dans la région de maturation, la matrice de l'émail restante est seulement visible par de petits granules noirs. **g:** Coupe à travers la région moyenne d'une dent. La maturation de l'émail a atteint la base de la dent où la matrice est, cependant, toujours en dépôt. Noter la lacune de résorption du dentaire qui fait face à la dent en développement (tête de flèche). Echelle = 50 µm.

Au niveau ultrastructural, avant le dépôt de la matrice de l'émail, les deux populations de cellules épithéliales et mésenchymateuses, les améloblastes et les odontoblastes, se font face et sont séparées par une membrane basale ininterrompue (Figure 6a). Le long de l'interface épithélio-mésenchymateux, la surface des améloblastes est régulière et lisse, alors que de nombreuses prolongations cytoplasmiques issues des odontoblastes sont au contact de la membrane basale. A ce moment, les améloblastes sont juxtaposés, allongés et ne montrent aucune indication d'une synthèse protéique active. Ils possèdent un grand noyau localisé dans la région moyenne de la cellule et le cytoplasme contient une quantité relativement faible de mitochondries, de citernes du réticulum endoplasmique granuleux (REG) et de vésicules (Figure 6a). Au contraire, les odontoblastes qui font face sont polarisés et leur cytoplasme contient de nombreuses citernes du REG et des mitochondries, indiquant une synthèse protéique active. Une petite quantité de matrice collagénique a été déposée dans l'espace extracellulaire séparant les odontoblastes des améloblastes. Un peu plus tard, l'espace entre les deux populations de cellules s'est agrandi permettant le dépôt d'une fine couche de matrice collagénique par les odontoblastes (Figure 6b). Cette matrice de prédentine non-minéralisée contient de nombreux prolongements cellulaires appartenant aux odonloblastes qui montrent une intense activité, comme l'indique la quantité importante de mitochondries et de citernes du REG. Face à la couche de prédentine, les améloblastes n'ont pas beaucoup changé. Une membrane basale est encore visible, soulignant la surface des améloblastes.

La seule différence avec le stade précédent est un accroissement du nombre d'organites cytoplasmiques dans la région faisant face à la prédentine (Figure 6b). La matrice de la prédentine localisée près de l'interface épithélio-mésenchymateux, c'est-à-dire le premier dépôt de matrice collagénique, est un peu différente de celle des autres régions. De nombreuses fibrilles de collagène sont perpendiculaires à la surface de la prédentine et atteignent la membrane basale. Cette particularité est clairement visible dans le stade suivant, où la matrice de prédentine a été imprégnée de protéines non-collagéniques et électroniquement denses, et la membrane basale a disparu (Figure 6c). Le cytoplasme des améloblastes qui font face à la surface de la prédentine montre des signes d'une grande activité de synthèse protéique : grandes quantité de REG, mitochondries et vésicules avec un contenu dense aux électrons (Figure 6c). de plus, la membrane cellulaire n'est plus lisse et régulière. En effet, des espaces extracellulaires sont apparus le long de la matrice de prédentine. Ces espaces sont remplis d'une matrice très dense aux électrons et dans laquelle il est difficile de reconnaître les fibrilles de collagène (Figure 6d).

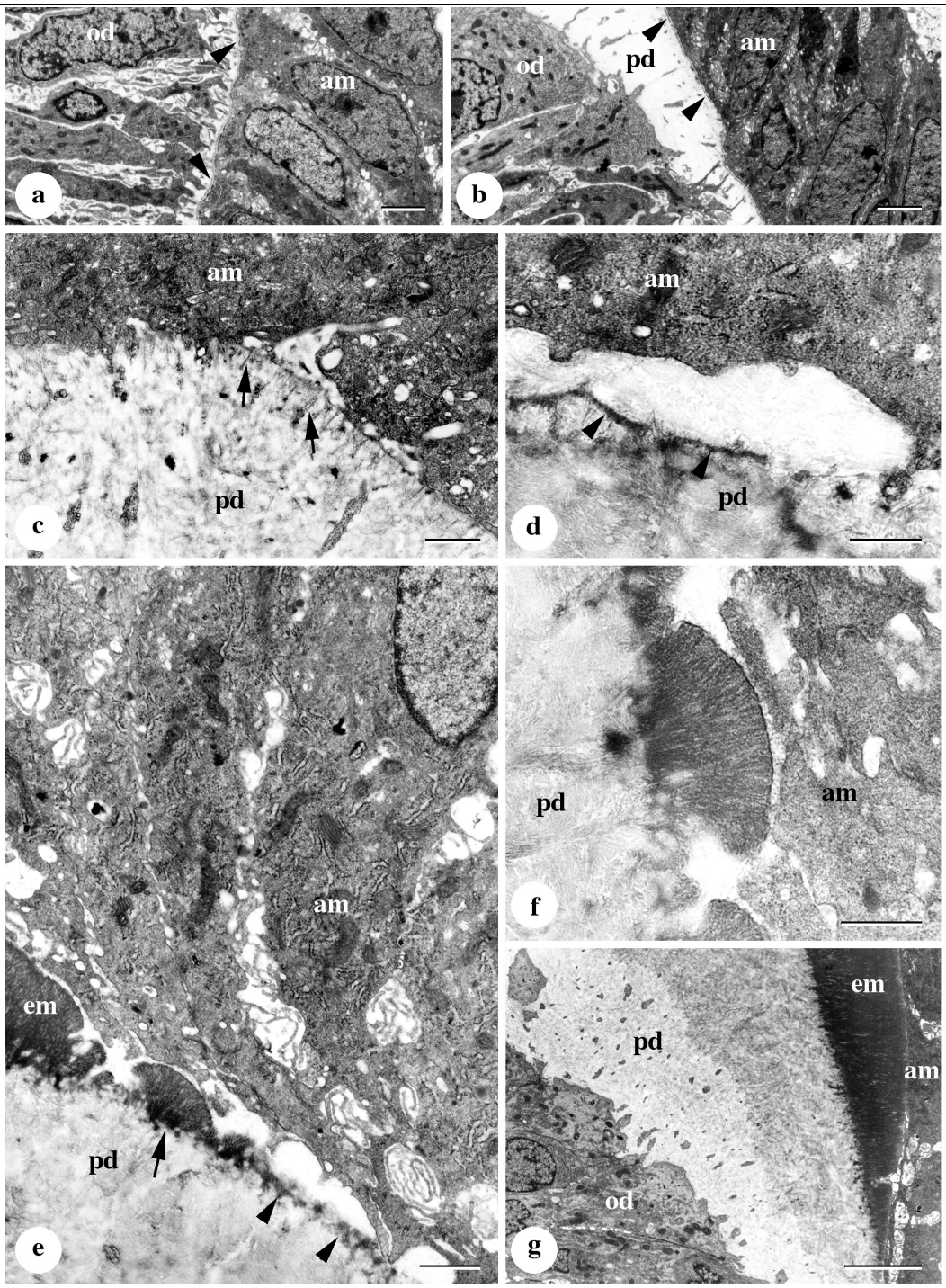
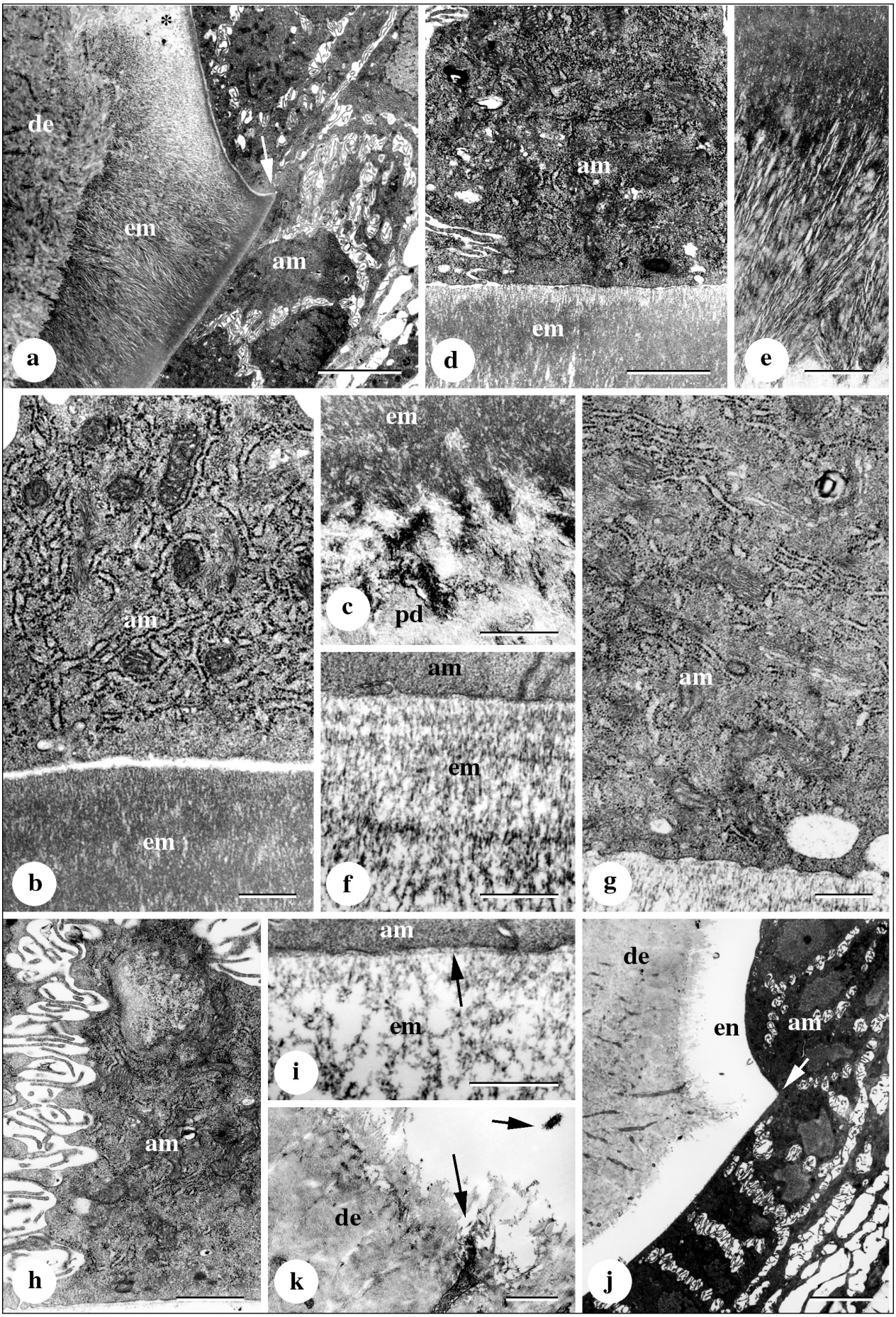


Figure 6. *Chalcides viridanus*, spécimen âgé de 1 mois. Mâchoire inférieure gauche décalcifiée à l'EDTA. Certains aspects du dépôt de la matrice de l'émail en formation observés au microscope électronique à transmission. **a:** Interaction entre odontoblastes (sur la gauche) et améloblastes (sur la droite). Noter les nombreuses extensions cytoplasmiques des odontoblastes. Elles atteignent la membrane basale qui limite les améloblastes (tête de flèche). **b:** Dépôt de la prédentine par les odontoblastes. La membrane basale est toujours présente (flèches). **c:** La membrane basale a disparu et les améloblastes limitent la surface de prédentine. **d:** Un espace extracellulaire contenant une matrice fortement visible a été créé entre les améloblastes et la surface de prédentine. Les flèches indiquent les restes de la membrane basale. **e:** Premier dépôt de la matrice de l'émail. Des petits amas de matrice d'émail sont déposés à la surface de la prédentine (têtes de flèche) puis elles s'épaississent (flèche). **f:** Détail d'un grand amas de matrice d'émail récemment déposé. **g:** Faible grandissement montrant l'épaisse couche de matrice d'émail récemment déposée, la prédentine imprégnée par des protéines non collagéniques et la matrice de prédentine récemment déposée. Noter l'absence des processus de Tomes.
Echelle : a-b = 2 nm; c, e = 1 nm; d, f = 500 nm; g = 5 nm.

Certains espaces intercellulaires sont visibles entre des améloblastes adjacents, alors que leur noyau est localisé dans une région distale de la cellule. Dans la région proximale qui borde la surface de la prédentine, le cytoplasme est riche en organites et de petites vésicules sont abondantes. La membrane cellulaire est irrégulière et nombre de courtes digitations cellulaires sont visibles (Figure 6e). Ces prolongations cytoplasmiques pénètrent l'espace entre les améloblastes et la surface de la prédentine, où sont délimitées de petites régions dans lesquelles des paquets de matrice denses aux électrons apparaissent. Ces paquets représentent les premiers éléments de la matrice de l'émail (Figure 6e). Ces éléments pionniers de la matrice de l'émail sont déposés tout d'abord à la surface de la prédentine où ils sont localisés au milieu du réseau collagénique. Cette région constituera la jonction énamélo-dentinaire. La matrice de l'émail est facilement distinguable de la matrice de la prédentine par son aspect dense aux électrons. Puis, les paquets de matrice d'émail s'agrandissent à l'intérieur de l'espace extracellulaire qui fait face à la surface des améloblastes. Ils forment de grands amas, denses et fibrillaires dans lesquels les fibrilles sont parallèles les unes aux autres et orientées perpendiculairement à la surface des améloblastes. Ces fibrilles sont incluses à l'intérieur d'une substance fine et granuleuse (Figure 6f). Puis les amas s'agrandissent et fusionnent les uns avec les autres pour constituer une couche homogène de matrice de l'émail (Figure 6g). Cette couche épaisse et dense aux électrons s'épaissit grâce à l'activité des améloblastes, dont la surface devient de plus en plus régulière à mesure que la matrice de l'émail devient plus épaisse. Des lignes sombres, parallèles à la surface de l'émail apparaissent à l'intérieur de la matrice et suggèrent une croissance périodique de cette couche (Figure 6g). Pendant ce temps, les odontoblastes très actifs ont déposé une fine couche de prédentine, dont la couche supérieure a été imprégnée par des protéines non collagéniques denses aux électrons, un phénomène qui indique un processus de minéralisation imminent.

Au niveau des cuspides, l'organe de l'émail est organisé différemment par rapport aux autres régions. En particulier, l'orientation des améloblastes diffère des deux côtés de la surface de la cuspide où les deux couches sont perpendiculaires (Figure 7a). Les améloblastes qui déposent la matrice de l'émail sont riches en citernes de REG, en mitochondries, en appareil de Golgi, desquelles de petites vésicules sont issues (Figure 7b). Certaines de ces vésicules sont fusionnées avec la membrane des cellules faisant face à la matrice de l'émail. La totalité de l'épaisseur de l'émail est atteinte avant que la minéralisation démarre. En effet, quand la plus grande partie de la couche d'émail est déposée, la minéralisation n'a pas encore débuté au niveau de la jonction énamélo-dentinaire, qui est irrégulière et montre de petits amas de matrice d'émail déposés au hasard à l'intérieur de la matrice de collagène (Figure 7c). Quand le processus de minéralisation commence, il concerne la totalité de l'épaisseur de la couche d'émail mais seulement la région du dépôt le plus ancien. La maturation de l'émail progresse ensuite à partir de ces régions localisées dans la partie supérieure de la dent vers la matrice déposée plus récemment, vers la base de la dent. Il y a donc un gradient dans la minéralisation de l'émail, du sommet vers la base. Sur les échantillons décalcifiés que nous avons utilisés, le processus de minéralisation de la matrice peut être indirectement évalué par l'aspect de dégradation (par protéolyse) de la matrice durant la maturation. Plus la matrice est éliminée et plus la maturation de l'émail est avancée. Quand la minéralisation commence, le cytoplasme des améloblastes qui font face à la zone en minéralisation ne semble pas très différent de celui des étapes précédentes, excepté la présence de petites vacuoles (Figures 7a, d). Certains espaces vides apparaissent le long des fibrilles de la matrice de l'émail, sur toute l'épaisseur de la couche, à la surface aussi bien qu'à la jonction énamélo-dentinaire (Figure 7e).


a
de
em
am
*
d
am
em
e
b
am
em
c
em
pd
f
am
em
g
am
h
am
i
am
em
k
de
j
de
en
am

Quand la maturation de l'émail a lieu, la matrice organique est progressivement éliminée, mais avec un petit retard au niveau des lignes de croissance (Figure 7f). Les améloblastes qui bordent la région de maturation de l'émail n'ont pas changé, pas plus que leur contenu cytoplasmique, excepté la présence de grandes vacuoles le long de la surface de l'émail et l'aspect irrégulier de la membrane cellulaire (Figure 7g). Dans une étape plus avancée de la maturation, les améloblastes ont réduit leur quantité d'organites cellulaires et augmenté les espaces extracellulaires, dans lesquels de nombreuses interdigitations sont visibles (Figure 7h). La matrice de l'émail a été éliminée en grande partie et les protéines restantes ont perdu leur aspect fibrillaire. Il y a maintenant plus de granules et une fine membrane apparaît à la surface de l'émail près de la membrane cellulaire des améloblastes (Figure 7i). A la fin de la maturation de l'émail, la matrice organique a complètement disparu. La bordure des améloblastes est très interdigitée, pauvre en organites, et apparaît dense aux électrons en raison de sa richesse en micro-filaments (Figure 7j). Après décalcification, la jonction énamélo-dentinaire a conservé les traces des interactions entre les protéines de l'émail et de la dentine sous la forme d'une ligne de petits granules denses aux électrons (Figure 7k). La présence de prolongements cytoplasmiques des odontoblastes à l'intérieur de la matrice de l'émail indique que, dans certaines régions, les protéines de l'émail ont imprégné la région supérieure de la dentine sur une assez grande épaisseur et cette partie de la jonction énamélo-dentinaire a été éliminée par protéolyse (Figure 7j, k).

De tous les stades de développement et sections examinées, il est clair que (1) les processus de Tomes, et (2) les prismes d'émail similaires à ceux décrits durant la formation de l'émail chez les mammifères, ne se forment à aucun stade de dépôt de l'émail chez ce lézard.

Résorption

Chez *Chalcides viridanus*, le remplacement périodique des dents tout le long de la vie implique que les dents fonctionnelles sont résorbées et sont perdues avant l'attache et l'éruption d'une nouvelle dent. En effet, au sein d'une famille dentaire, la dent de remplacement occupera la même position que la dent fonctionnelle. La croissance des bourgeons dentaires ayant lieu près de la base de la dent fonctionnelle, elle provoque évidemment la résorption de cette dernière. La résorption commence par la face linguale (c'est-à-dire par le côté faisant face à la dent en développement) et près de la base de la dent (Figure 8a).

A ce stade, le bourgeon est déjà bien formé, la couche d'émail étant en maturation, et la surface de l'organe dentaire borde la base du cône de dentine de la dent fonctionnelle. La croissance de la dent de remplacement exerce une pression sur la surface de la dent fonctionnelle ce qui induit le recrutement de cellules clastiques. Nous appelons odontoclastes les cellules de résorption multinucléées qui érodent le cône de dentine, quelle que soit leur localisation (à l'extérieur ou à l'intérieur de la cavité pulpaire), et ostéoclastes, les cellules ayant la même morphologie, mais qui résorbent la surface de l'os. Les odontoclastes sont observés le long de la surface extérieure, linguale, du cône de dentine de la dent fonctionnelle. Ils éliminent la matrice de la dentine à partir de l'extérieur jusqu'à ce que le cône de dentine soit percé (Figure 8b). Quand la cavité pulpaire est ouverte, les odontoclastes pénètrent dans la cavité pulpaire et résorbent la paroi de dentine de l'intérieur alors que d'autres odontoclastes continuent d'éliminer la matrice de la dentine sur la face linguale (Figure 8c). A ce moment, la dent de remplacement a terminé son processus de maturation et elle est maintenant localisée, pour une large part, dans la cavité pulpaire de la dent

fonctionnelle, la paroi de dentine étant extrêmement réduite. Peu de temps avant l'éruption de la dent de remplacement, la paroi labiale de la dentine est attaquée (Figure 8d). Nous n'avons jamais vu d'odontoclastes résorbant la couche d'émail et nous pouvons conclure que c'est la seule partie de la dent fonctionnelle qui tombe. Durant la croissance de la dent de remplacement, les ostéoclastes sont fréquemment observés le long de la surface de l'os correspondant à la région où la nouvelle dent sera attachée.

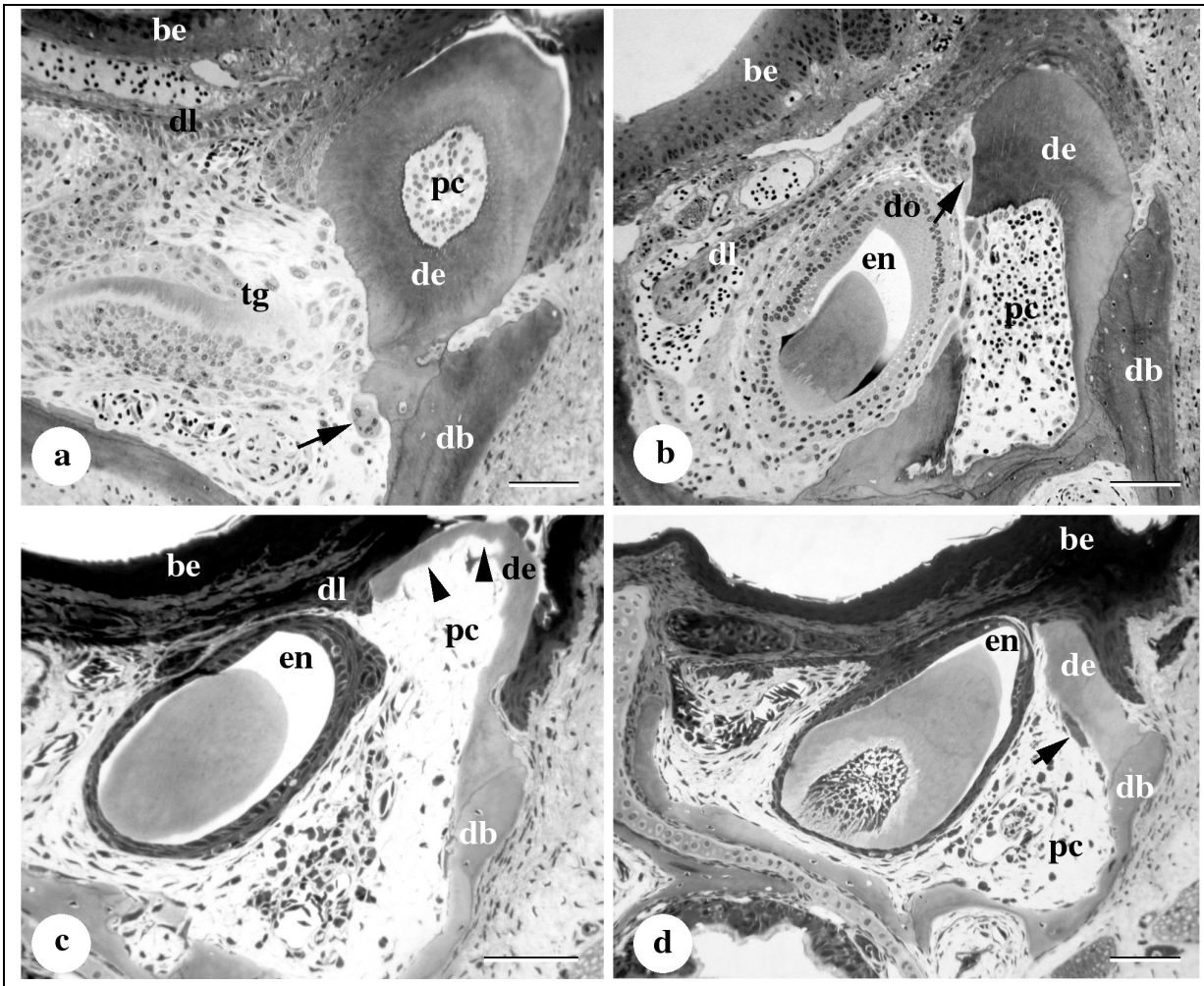


Figure 8. *Chalcides viridanus*, adulte de 70 mois. Échantillons décalcifiés à l'EDTA. Coupes transversales de 1 μm d'épaisseur colorées au bleu de toluidine. Résorption de la dent fonctionnelle en relation avec le stade de développement de la dent de remplacement. Mâchoire inférieure gauche, côté labial sur la droite. **a:** La matrice de l'émail de cette dent de remplacement est partiellement minéralisée mais la région n'est pas visible à ce niveau de la coupe. La région basale du cône de dentine est résorbée latéralement sur sa face linguale, qui est au contact avec l'organe de l'émail de la dent en développement. **b:** La dent de remplacement a grandi et la minéralisation de la matrice de l'émail est presque terminée. **c:** La dent de remplacement (son émail étant maintenant mature) a provoqué la résorption de la plus grande partie de la face linguale de la dent fonctionnelle et d'une partie de la dentine depuis l'intérieur. Les pointes de flèche indiquent les lacunes de résorption. La cavité pulpaire est largement ouverte. **d:** Dent de remplacement dont le cycle est presque achevé. La dent fonctionnelle est entièrement résorbée, à l'exception de la face labiale du cône de dentine qui reste fixée au dentaire. La flèche indique un odontoclaste. Echelle : a = 20 μm; b-d = 50 μm.

Expression du gène de l'amélogénine

L'hybridation *in situ* sur coupes a prouvé qu'elle est une technique efficace pour démontrer l'expression du gène de l'amélogénine dans les améloblastes durant la formation de l'émail chez *Chalcides viridanus* (Figure 9). Utilisant une sonde marquée au P^{33}, nous avons obtenu un très fort signal à chaque étape de l'amélogenèse. Les coupes longitudinales de la mâchoire inférieure, parallèles à la racine de la dent, nous fournissent des photos de plusieurs dents de remplacement sur la même coupe (Figure 9a). Le fort signal observé est le résultat de la détection des transcrits par la sonde d'amélogénine et ce n'est pas une réaction artéfactuelle comme le montre le témoin négatif. Quand les premiers éléments de la matrice de l'émail sont déposés à la surface de la prédentine, les ARNm de l'amélogénine sont détectés dans les améloblastes qui font face à la matrice déposée (Figures 9c, d). Puis, à l'apex de la dent, dans la région où les premiers signes de la minéralisation de la matrice de l'émail sont visibles, le signal s'affaiblit alors qu'il reste très fort dans les régions latérales où la matrice de l'émail est encore en cours de dépôt (Figures 9e, f). Puis le signal disparaît complètement des améloblastes situés en face de l'émail mature dans la partie supérieure de la dent alors qu'il reste fortement visible le long de la base de la dent où la matrice de l'émail est toujours déposée (Figure 9g, h). Les transcrits d'amélogénine sont encore détectés à la base de la dent alors qu'elle est bien formée mais pas encore attachée (Figure 9i, j). L'expression de l'amélogénine n'a jamais été observée dans les dents fonctionnelles (Figure 9a), et les ARNm de l'amélogénine n'ont jamais été détectés dans les odontoblastes, ni dans les autres cellules de l'organe de l'émail, ni à la surface de l'os de support.

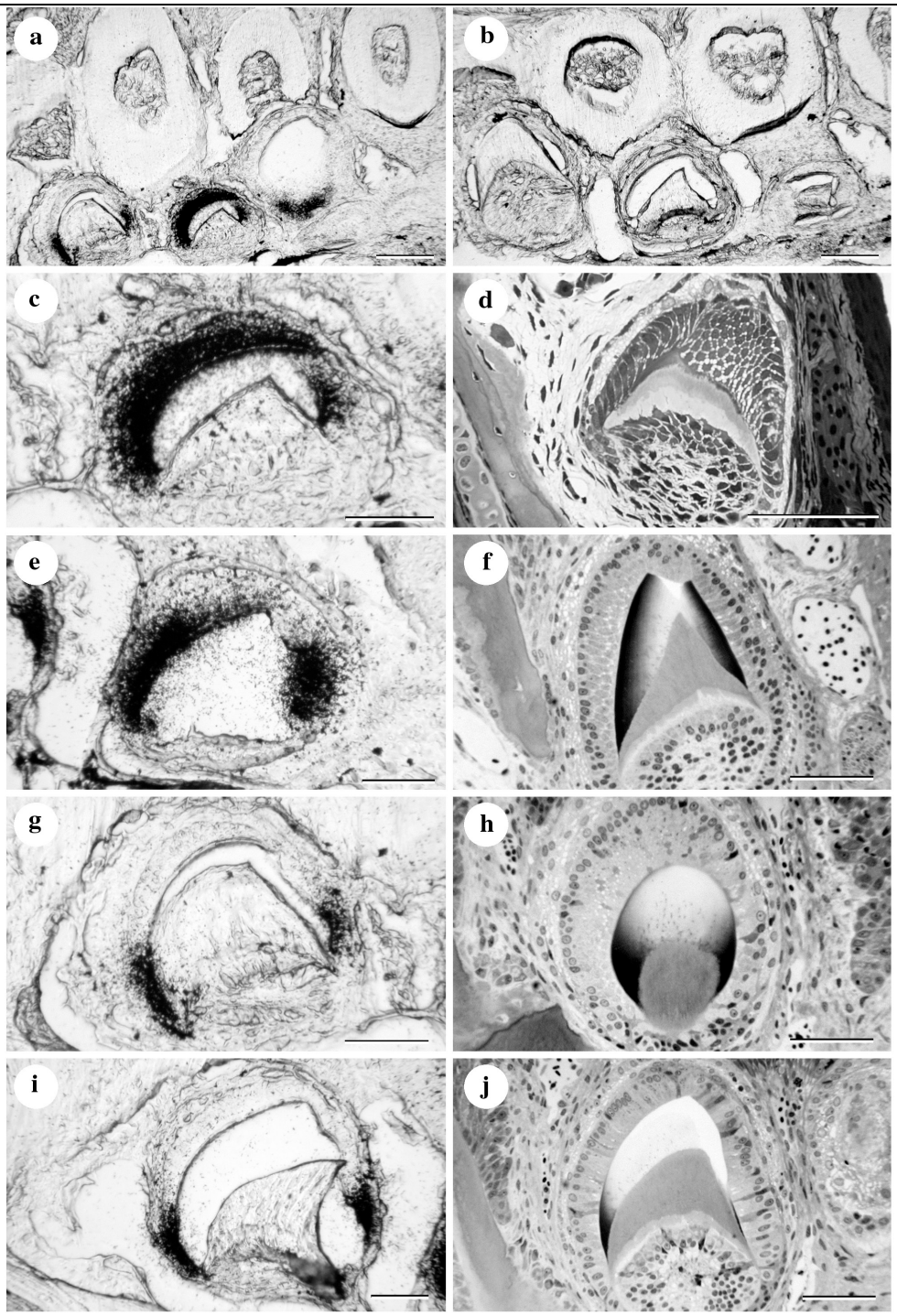


Figure 9. Expression du gène de l'amélogénine durant le développement dentaire chez *Chalcides viridanus.* Hybridation *in situ* sur coupes utilisant une sonde d'amélogénine marquée au P^{33} (grains noirs). Mâchoire inférieure, côté labial sur la droite. **a:** Faible grandissement montrant trois positions dentaires. L'amélogénine n'est pas exprimée dans les dents fonctionnelles alors que les transcrits sont détectés dans l'organe de l'émail de la dent de remplacement. **b:** Contrôle négatif pour (**a**). **c-j:** Matrice de l'émail en dépôt dans des dents de remplacement sélectionnées pour illustrer les étapes successives de l'amélogenèse, du stade le plus précoce (**c, d**) à la fin de la maturation de l'émail (**i, j**). Sur la gauche (**c, e, g, i**): hybridation *in situ* montrant la localisation des ARNm de l'amélogénine dans l'organe de l'émail (améloblastes); sur la droite (**d, f, h, j**): sélection de coupes de 1 μm montrant l'aspect de la matrice de l'émail à des étapes semblables de l'amélogénèse. Les ARNm de l'amélogénine sont fortement détectés dans les améloblastes qui sont en face de la matrice de l'émail récemment déposée, alors que les transcrits sont moins nombreux dans les améloblastes faisant face à l'émail en maturation et sont absents des cellules faisant face à l'émail mature. Noter les nombreux ostéoclastes (flèches) agissant contre la surface du dentaire proche des dents en développement. Echelle : a, b = 100 μm; c-j = 50 μm.

Discussion

Jusqu'à maintenant, la plupart des études sur les dents des reptiles (squamates et crocodiles) n'ont concerné que l'étude du patron de la dentition et plus spécialement, la question de la génération des vagues de remplacement dentaire (Edmund, 1960, 1969; Cooper, 1966; Cooper *et al.*, 1970; Rieppel, 1978; Rocek, 1980). La connaissance du patron du remplacement dentaire a généré un large débat en terme d'implications évolutives (Osborn, 1970, 1984; DeMar, 1972, 1974; Osborn et Crompton, 1973) et de réalité embryologique (*e. g.*, Osborn, 1970, 1972; Westergaard, 1986; Westergaard et Ferguson, 1986). En revanche, les études histologiques du développement dentaire chez les reptiles sont soit anciennes (*e. g.*, Harrison, 1901; Woederman, 1919; Poole, 1957) ou ont plutôt servi d'outil pour comprendre le patron du développement dentaire (*e. g.*, Osborn, 1971). A notre connaissance, le présent travail est la seule étude qui fournit autant de détails sur l'amélogenèse chez un squamate à l'aide de la microscopie photonique et de la MET. Cette étude est également la première description chez un reptile de l'expression du gène de l'amélogénine par hybridation *in situ.*

Morphologie dentaire

La forme, la taille et l'orientation des dents changent au cours de l'ontogénèse: à partir d'une forme très généraliste présente chez les nouveaux-nés (c'est-à-dire une dent conique) avec une surface lisse, une petite taille, de larges espaces entre les dents et une orientation droite, on aboutit à la forme spécifique de l'espèce (c'est-à-dire, dent rectangulaire) avec des dents cuspidés, de grande taille, proches les unes aux autres et incurvées postérieurement chez les juvéniles prématures et les adultes. Toutes ces caractéristiques ont progressivement été acquises durant l'ontogénèse grâce au remplacement successif des dents. Chez ce lézard, les dents de première génération (c'est-à-dire les premières dents à être fonctionnelles) ont une morphologie généraliste qui est différente de celle de l'adulte mais leur structure est déjà similaire, puisque la première dent apparaît comme une miniature de la dent adulte (Sire et al., 2002). Elle diffère des dents de première génération des actinoptérygiens et de la larve des amphibiens urodèles, dans lesquelles la structure est minimaliste en relation avec la petite taille de la dent (Sire et al., 2002). Chez Chalcides viridanus, il y a seulement une dent fonctionnelle à une position donnée, et la dent de remplacement ne s'attache pas tant que la dent fonctionnelle n'est pas tombée. De plus, Chez C. viridanus, le nombre de dents ne change pas durant l'ontogénèse (voir 1ère partie), la taille plus grande de la dent adulte étant adaptée par un changement de forme de conique à rectangulaire, et par réduction de l'espace entre deux positions adjacentes. A l'exception d'une légère variation de la taille (dents un peu plus grandes depuis l'extrémité antérieure vers l'arrière), la morphologie dentaire est la même pour toutes les dents chez chaque individu (ils sont homodontes). L'homodontie est la règle chez la majorité des lézards (Edmund, 1960).

De nombreuses dents usées ont été observées. Dans nos conditions d'élevage, ces usures ne peuvent pas être provoquées par la mastication de la nourriture (tomates et asticots) mais elles pourraient s'expliquer par les grains de sable qui sont collés à la nourriture (le fond des terrariums est couvert de sable).

Chez *Chalcides viridanus*, l'attache de la dent est pleurodonte (c'est-à-dire, ankylosée sur la face interne de la paroi labiale) comme c'est généralement le cas chez les reptiles (*e. g.,* Edmund, 1960), avec cependant une exception chez quelques espèces acrodontes comme les agamidés et les crocodiles. Durant l'ontogénèse, le mode d'attache de la dent ne change pas, mais les relations avec le sillon dentaire et la dent de remplacement varient un peu ; le

bourgeon dentaire se forme plus profondément chez les adultes que chez les juvéniles, provoquant la résorption à partir de la base chez l'adulte plutôt que latéralement chez le juvénile. Des ostéoclastes et des odontoclastes résorbant l'os et la surface de la dentine près de la dent en croissance ont été observés. Le nom de ces cellules a été attribué en fonction du tissu qu'ils résorbent, les ostéoclastes étant impliqués dans la préparation de la surface de l'os pour la future attache de la dent et les odontoclastes favorisant la croissance de la dent au dépend de la dent fonctionnelle, mais pas par rapport à leur origine qui semble être la même. Il apparaît que le processus de remplacement est indépendant de l'état d'usure ou de la durée de la vie de la dent fonctionnelle. La seule existence d'un tel remplacement dentaire "automatique" pourrait expliquer pourquoi une dent est remplacée alors qu'elle n'est pas usée ou cassée.

Le Nettoyage des mâchoires avec de l'eau de javel a aidé à retirer les tissus mous, mais cela a provoqué la perte de la plupart des bourgeons dentaires. Heureusement, une jeune dent de remplacement est restée en place (probablement parce qu'elle était protégée par la dent fonctionnelle chez cet adulte). Nous avons observé que la couronne de la dent était finie alors que le cône de dentine et la base de la dent n'étaient pas encore formés. Dans ce bourgeon, la couronne de la dent (couverture d'émail) était la même (forme, taille et ornementation) que pour la dent fonctionnelle qu'il remplacera. Ceci indique clairement que, chez ce spécimen adulte, la dent ne change pas de morphologie quand elle est remplacée et que la couronne de la dent est terminée longtemps avant que la dent de remplacement ne soit attachée. Nous pouvons alors supposer que, durant l'ontogénèse, le principal changement morphologique qui concerne la couronne d'émail (cuspides et ornementation) pourrait être étudié en comparant l'aspect de la surface de la dent en développement à celui de la dent fonctionnelle, et ses successeurs, dans un grand nombre de positions dentaires. Cela pourrait également apporter une information sur le moment où la couronne d'émail est terminée et mener, par les moyens de l'histologie et de l'expression des divers gènes, à identifier les facteurs qui déterminent la formation des cuspides.

Développement dentaire

Le développement histologique de la dentition a été étudié chez plusieurs reptiles : chez *Lacerta vivipara* par Osborn (1971), chez *Sphenodon punctatus*, *Cnemaspis kandiana* (gekkonidés), *Lacerta agilis* et *Anguis fragilis* (orvet) par Westergaard (1986) et chez *Alligator mississipiensis* par Westergaard et Fergusson (1986). Cependant ces études ne fournissent que peu de données comparatives et seuls d'anciens travaux peuvent être utilisés (*e. g.*, Röse, 1894; Woerdeman, 1919; Poole, 1957). Chez les scincidés, le développement dentaire n'a été étudié que très récemment. Les dents de première génération ont été étudiée chez *Chalcides viridanus* et *C. sexlineatus* (Sire *et al.*, sous presse). Chez *C. viridanus*, la première dent en développement se forme précocement, chez l'embryon, mais la plupart d'entre elles sont dégénératives et ne deviennent jamais fonctionnelles (Sire *et al.*, 2002). Ces dents dégénératives (rudimentaires) sont remplacées par des dents fonctionnelles avant la naissance ou peu de temps après, comme cela a déjà été décrit chez quelques reptiles incluant les lézards et les crocodiles (*e. g.*, Röse, 1894; Woerdeman, 1919; Edmund, 1969; Osborn, 1971; Westergaard, 1986; Westergaard et Fergusson, 1986). Le patron de remplacement dentaire chez *C. viridanus* est prédictible (voir ce chapitre, 1ère partie) et cette caractéristique a été utilisée avec précision dans ce travail pour rechercher les stades représentatifs du développement dentaire dans les mâchoires d'une série de croissance de *C. viridanus* et, en particulier, pour étudier l'amélogenèse.

Initiation

Chez les espèces de crocodiliens et de lépidosauriens étudiées dans la littérature, la toute première dent (dégénérative) a été décrite comme étant initiée directement à partir de l'épithélium oral avant que la lame dentaire soit formée (Westergaard 1986; Westergaard et Fergusson, 1986). Chez *Alligator mississipiensis,* Westergaard et Fergusson (1986) ont montré que la dentition commence par l'initiation d'une dent longtemps avant que d'autres dents apparaissent. Le stimulus d'initiation précoce peut provenir d'une région particulière de l'épithélium (appelée placode par Westergaard et Fergusson, 1986), suivi par une induction locale de l'épithélium (lame dentaire) vers le mésenchyme. Malheureusement, le stade embryonnaire disponible dans la présente étude était trop âgé (une à deux semaines avant la naissance) pour observer le stade placode. Les dents dégénératives n'ont pas été observées car elles avaient déjà disparu (c'est-à-dire, été résorbées) à ce stade embryonnaire. La première dent fonctionnelle qui se développe, s'initie à l'extrémité d'une lame dentaire étroite et rectiligne qui s'enfonce dans le mésenchyme qui se condense. Ceci est en accord avec les observations précédentes indiquant que toutes les dents fonctionnelles chez les reptiles se développent à partir d'une lame dentaire bien formée (Westergaard, 1986; Westergaard et Fergusson, 1987, 1990).

Aussitôt que la dent est différenciée, la lame basale s'étend et une nouvelle dent est initiée. Ceci a été observé aussi bien chez l'embryon que chez les juvéniles et les adultes, et signifie que lorsqu'une nouvelle dent de remplacement commence à s'initier, la dent qu'elle remplace n'est pas encore fonctionnelle (c'est-à-dire qu'elle na pas encore percé et n'est pas encore attachée à l'os de support). Ceci signifie clairement que les nouvelles dents sont initiées longtemps avant que la dent précédente de la même famille ne devienne fonctionnelle et soulève la question de l'identité des facteurs qui contrôlent l'initiation de la dent de remplacement. Ceci pourrait expliquer l'existence des vagues de remplacement déjà bien documentées dans la littérature (Edmund, 1960, 1969; Osborn, 1970, 1984). Une hypothèse possible est que cette nouvelle extension de la lame basale, à partir de l'épithélium dentaire externe, pourrait être liée au stade de différenciation de la dent précédente dans la même famille. Cette hypothèse doit être confirmée en comparant différents stades sur de nombreuse coupes sériées de dents en développement à différentes positions et dans une série de croissance. Est-ce que le stade de différenciation est le même pour toutes les dents de remplacement qui se développent et à quel moment la lame dentaire commence-t-elle à se développer à partir de l'épithélium dentaire externe pour initier une nouvelle dent ? Chez l'homme, pour lequel la dentition est diphyodonte, le remplacement des dents temporaires a lieu de manière semblable. La dent permanente est initiée et se différencie durant l'ontogenèse des dents temporaires. La lame dentaire de la dent de remplacement s'étend, sur la face linguale, à partir de la surface de l'organe de l'émail bien différencié. Par contre, l'initiation des molaires, qui n'ont pas de dents temporaires, a lieu directement à partir de la lame dentaire issue de l'épithélium oral (Ruch, 2001).

Contrairement aux embryons, chez lesquels une ou deux dents sont déjà présentes à une position donnée (les dents successives dans une même position constituent une famille), chez les spécimens plus vieux, trois dents d'une même famille étaient déjà présentes, la troisième dent étant au stade de la cloche avec la matrice de l'émail et la prédentine en train de se déposer. Toutes les dents d'une même famille sont liées à la lame dentaire par son épithélium dentaire externe. De plus, une nouvelle extension de la lame basale dans le mésenchyme était déjà observée depuis l'épithélium dentaire externe de la troisième dent, suggérant que l'initiation d'une quatrième dent a commencé.

Ce résultat confirme l'hypothèse proposée dans la première partie de ce chapitre sur la probable existence d'une troisième dent en développement, non minéralisée, en plus des deux dents qui sont visibles sur les radiographies.

Différenciation

Invariablement, chez *Chalcides viridanus*, la dent de première génération et la dent en de remplacement se développent comme cela a été décrit chez d'autres espèces de squamates : les cellules de l'épithélium de la lame dentaire et une population de cellules du mésenchyme sont entrées en contact et se sont différenciées en un organe dentaire et une papille dentaire, respectivement, à une courte distance de la surface de l'os de support (Osborn, 1971; Ogawa, 1977). Nous avons observé effectivement que la région de la lame dentaire, faisant face à la région labiale est impliquée dans la différenciation dentaire alors que la région labiale est seulement concernée par l'épithélium dentaire externe. Ce sont les mêmes observations que celles décrites lors du développement des dents chez les mammifères. Chez *Chalcides,* une couche de prédentine est déposée en premier par les odontoblastes, avant le dépôt de la matrice de l'émail par les améloblastes. La minéralisation des matrices de prédentine et d'émail commence seulement dans les régions où l'épaisseur entière de la couche de l'émail est atteinte, signifiant que les améloblastes ont arrêté leur rôle actif dans le dépôt de la matrice de l'émail. Puis, les deux couches s'épaississent et se minéralisent à partir du premier dépôt du sommet et vers la base.

La seule différence notable avec le développement des dents de mammifères concerne l'absence des fibres de Tomes à la surface des améloblastes durant le dépôt de la matrice de l'émail. La différence la plus importante avec l'émail des mammifères est l'absence des prismes, chez *C. viridanus,* comme cela est décrit chez toutes les espèces de reptiles étudiées (voir Sander, 2001, pour une revue détaillée de l'organisation des cristaux d'émail chez les reptiles).

L'organe de l'émail se différencie à partir de la zone de prolifération à l'extrémité de la lame dentaire comme cela est décrit chez les mammifères. Cet organe de l'émail est composé de trois couches : un épithélium dentaire interne (couche d'améloblastes) et externe qui sont séparées par une fine couche que nous avons interprété comme du réticulum étoilé (Ogawa, 1977). Ceci diffère un peu de la condition observée dans les dents des mammifères où une autre couche a été décrite, le stratum intermedium, mais nos observations doivent être complétées sur des stades de développement plus avancés.

Amélogénine

La présence de l'amélogénine dans des coupes consécutives à différents stades de développement dentaire a été révélée par hybridation *in situ*. Nous avons utilisé une procédure rapide et facile pour préparer une sonde d'ADN simple-brin en utilisant une PCR asymétrique comme cela a déjà été décrit (Bleicher *et al.*, 1999). A notre connaissance, c'est la première étude qui utilise une sonde spécifique pour démontrer que les améloblastes reptiliens expriment l'amélogénine durant le dépôt de la matrice de l'émail. Jusqu'à 1998, avant que le clonage du gène de l'amélogénine ait été obtenu chez un serpent par Ishiyama *et al.* et, simultanément, chez un crocodile et un xénope par Toyosawa *et al.*, la détection de l'amélogénine chez les espèces non-mammaliennes par anticorps a longtemps été controversée (voir Zylberberg *et al.*, 1997). Cependant, en plus des réactions positives des anticorps mammaliens avec la matrice de l'émail de squamates (Herold *et al.,* 1989), le fait que l'émail des squamates, bien que dépourvu de prismes, soit très similaire à l'émail des

mammifères (*e. g.*, Sander, 2001) était un argument fort pour soutenir que l'amélogénine était certainement présente chez *Chalcides viridanus.*

L'étude présente, en montrant l'expression du gène de l'amélogénine dans les améloblastes durant la formation de la matrice de l'émail, confirme donc la présence d'amélogénine chez les squamates. La description de l'expression du gène de l'amélogénine dans les dents en développement de *C. viridanus* est en accord avec les descriptions disponibles dans la littérature mammalienne et avec amélogenèse (Snead *et al.*, 1988; Inage *et al.*, 1996; Bleicher *et al.*, 1999). L'expression spécifique des améloblastes a été observée dans les cellules qui sécrètent la matrice de l'émail. La sonde d'amélogénine révèle une forte présence de transcrits dans les améloblastes à un stade de différenciation avancé, puis l'expression diminue avec le début de la minéralisation et disparaît complètement dans les améloblastes en activité post-sécrétoire faisant face à l'émail mature. Aucun transcrit d'amélogénine n'a été détecté dans les odontoblastes à un quelconque stade de leur développement. L'expression de l'amélogénine a été signalée dans les odontoblastes du porc à un stade précoce de la différenciation dentaire (Oida *et al.,* 2002). Cependant, cette période est caractérisée par une forte interaction entre les améloblastes et les odontoblastes. Les deux populations de cellules sont séparées par une membrane basale, mais les cellules sont si proches les unes des autres (par l'intermédiaire des prolongations des odontoblastes qui entrent en contact avec la membrane basale) que les molécules de signalisation peuvent migrer à travers la membrane basale d'un type cellulaire à l'autre. Une telle interaction pourrait expliquer l'expression (1000 fois plus faible) de l'amélogénine dans les odontoblastes, comme elle pourrait expliquer l'expression de la sialo-protéine dentinaire (DSP) dans les améloblastes (Goldberg, 2001). Le début de la minéralisation de la dentine semble être un processus clef qui interrompt le processus d'interaction, agissant comme une barrière physique.

Ces résultats ont besoin d'être complété par des études plus précises de l'expression du gène de l'amélogénine. En particulier, il pourrait être intéressant (1) de vérifier à quel moment les pré-améloblastes de *Chalcides viridanus* expriment le gène de l'amélogénine et à quel moment les premières protéines sont exprimées, en utilisant des anticorps spécifiques, et (2) de comparer le profil d'expression du gène de l'amélogénine avec d'autres gènes (ou d'autres protéines) de l'émail ou de la dentine, comme l'améloblastine, l'énaméline, DSP, ...

CHAPITRE 4
-
DISCUSSION - CONCLUSION GÉNÉRALES

I. Étude d'une séquence d'amélogénine de myxine

Lorsque Slavkin et Diekwish publient, en 1996, une séquence protéique très partielle d'amélogénine de myxine, un crâniate sans mâchoire, ils apportent un témoignage important sur plus de 500 millions d'années d'évolution de l'amélogénine. Les conséquences de l'existence de cette séquence sont multiples. Cette séquence étant très conservée par rapport aux mammifères, cela signifiait que l'amélogénine a subi très peu de changements en 500 millions d'années et qu'elle est présente chez tous les vertébrés, y compris les plus basaux comme les actinoptérygiens et les chondrichtyens (requins, raies) puisque la myxine appartient au groupe-frère des vertébrés.

A cette époque, les résultats contradictoires de l'utilisation d'anticorps anti-amélogénine de mammifères chez les vertébrés font s'opposer deux théories. Les partisans de l'une supposent l'apparition de l'amélogénine assez récente et la situent chez les tétrapodes. Pour cela ils se fondent sur des résultats négatifs de l'utilisation d'anticorps chez les vertébrés non-tétrapodes aquatiques (voir introduction - I.5. “Utilisation d'anticorps...“). Au contraire, leurs adversaires soutiennent l'idée d'une apparition ancienne de l'amélogénine justifiée par des résultats positifs d'anticorps anti-amélogénine chez des actinoptérygiens et des chondrichtyens (voir introduction - I.5.). La séquence d'amélogénine d'une myxine était donc un encouragement pour les seconds, d'autant que la conservation de la séquence chez un taxon aussi éloigné des mammifères, signifiait, à cette époque, l'espoir d'obtenir facilement la séquence d'ADN de l'amélogénine dans toutes les lignées de vertébrés et, ainsi, apporter la preuve définitive par rapport aux résultats des anticorps, et peut être aussi mieux comprendre le fonctionnement de l'amélogénine.

Pour Slavkin, cette séquence protéique était aussi une preuve que ses travaux précédents (Slavkin *et al.,* 1983b), qui montraient (toujours avec des anticorps) la présence d'amélogénine dans les dents cornées (kératinisées) de la myxine, étaient fondés. La conséquence est importante : le rôle de l'amélogénine ne serait plus exclusivement voué à la formation de l'émail dentaire puisque les dents de la myxine, kératinisées, ne possèdent pas d'émail (ou d'émailloïde). Si le gène de l'amélogénine a pu se maintenir pendant 500 Ma dans une lignée sans servir à la minéralisation de l'émail, alors il est possible d'envisager sa présence chez d'autres taxons dépourvus d'émail et de dents comme les oiseaux et les tortues.

Les résultats de notre étude montrent clairement que la séquence d'amélogénine de myxine est une erreur et qu'il s'agit en réalité de l'ADN d'un rongeur. C'est un espoir déçu pour trouver l'amélogénine chez les vertébrés non-tétrapodes aquatiques, mais cela aurait pu relancer le débat sur l'apparition de l'amélogénine. Ce n'est plus vraiment le cas aujourd'hui, puisqu'à la suite des travaux présentés dans le chapitre II, il est maintenant prouvé que l'amélogénine existe depuis au moins 630 millions d'années. Donc, on doit logiquement pouvoir trouver le gène de l'amélogénine dans toutes les lignées de vertébrés qui possèdent de l'émail ou de l'émailloïde (à condition de réussir à concevoir des amorces capables d'amplifier le gène chez tous ces taxons !).

Qu'en est-il de l'amélogénine de myxine ? Rien n'empêche, *a priori*, que le gène de l'amélogénine ait été conservé chez un crâniate sans mâchoires, comme la myxine, puisque la lignée est séparée des vertébrés depuis moins de 630 millions d'années (Kumar & Hedges, 1998 ; Lecointre et Le Guyader, 2001). Cependant, si les ancêtres des myxines possédaient des odontodes (homologues des dents) sur le corps, les myxines actuelles n'en possèdent plus. Il faut alors se demander ce qu'ont pu détecter au niveau des dents kératinisées, les anticorps d'amélogénine de mammifère utilisés par Slavkin. Il est difficile de croire à la présence d'amélogénine chez la myxine d'autant qu'il est maintenant démontré que des taxons ayant perdu les dents (et donc l'émail), comme les oiseaux et les tortues, n'ont plus le gène de

l'amélogénine (Girondot et Sire, 1998). Est-ce que l'amélogénine ancestrale (il y a plus de 500 millions d'années) avait une fonction différente que celle qu'on lui connaît aujourd'hui ? On pourrait en effet admettre que la molécule ait pu avoir une fonction localisée dans l'épithélium buccal d'un ancêtre vertébré et que ce ne soit que secondairement que l'amélogénine ait été recrutée comme protéine spécifique de l'émail. Dans ce cas, rien n'empêche de penser que c'est bien de l'amélogénine qui a été reconnue par les anticorps au niveau des dents kératinisées de myxine. Quelle serait alors sa fonction ?

II. Preuve moléculaire de l'origine précambrienne de l'amélogénine, la protéine majeure de l'émail des vertébrés

Cette étude a permis d'estimer, grâce à des méthodes moléculaires, la date d'apparition de l'amélogénine au cours de l'évolution et, par voie de conséquence, de pouvoir dater en même temps l'origine de l'émail. En effet, chez les vertébrés, l'amélogénine est connues pour être exclusivement associée à la formation et à la minéralisation de l'émail (pas effet pléiotropique), un tissu uniquement connu chez les vertébrés et dont les caractéristiques sont entièrement différentes de celles des autres tissus minéralisés.

Pour estimer une origine évolutive, il faut disposer d'un gène homologue à partir duquel le gène de l'amélogénine aurait été crée. Or, aucun gène homologue n'était connu et ce n'est qu'un petit détail relevé dans un article de Liao *et al.* (1998) (les auteurs signalaient l'existence de 10 bases identiques entre le gène de l'amélogénine et celui de la SPARC (ostéonectine)) qui nous a conduit à montrer que l'exon 2 de l'amélogénine et celui de la SPARC avait une origine commune. En fait, l'exon 2 de l'amélogénine provient d'une duplication de l'exon 2 du gène de la SPARC. Cet exon 2 contient le peptide signal, un élément capital dans la voie de synthèse de ces deux protéines.

La datation de l'origine de cette duplication a été estimée à partir de la date de séparation entre le gène de la SPARC et un gène orthologue de SPARC, SC1, qui est postérieure à la séparation SPARC/amélogénine. En utilisant la méthode de maximum de vraisemblance (Rambaut et Bromham, 1998) la date de séparation entre SPARC et SC1 a été estimée à environ 630 millions d'années et l'amélogénine est apparue avant. Cette date ou une date légèrement postérieure correspondrait également à l'origine de l'émail.

La datation de l'amélogénine a permis d'apporter un élément de discussion à la controverse sur "l'explosion Cambrienne" qui a longtemps animé le débat entre la vision gradualiste de l'évolution et la théorie des équilibres ponctués (Eldredge et Gould, 1972) qui propose des phases de stabilité ponctuées par des phases d'évolution rapide. Cette "explosion", correspondant à l'apparition soudaine de nombreux clades de Bilateralia dans plusieurs gisements fossiles datés du Cambrien Ancien (il y a environ 543 millions d'années), a longtemps été interprétée (par exemple, Valentine *et al.*, 1999) comme le résultat d'une grande radiation évolutive se déroulant sur un temps très court (10 millions d'années). Cependant, cette interprétation a été contredite par certains auteurs qui ont proposé que cette abondance soudaine de fossiles de métazoaires dans le Cambrien pourrait être un artefact taphonomique lié à une meilleure fossilisation de taxons déjà présents depuis longtemps. Cette dernière hypothèse est soutenue à la fois par des données moléculaires et par des données paléontologiques. La datation moléculaire de la séparation entre les protostomiens et les deutérostomiens, estimée entre 1000 et 830 millions d'années (par exemple, Bromham et Hendy, 2000), indique que de nombreuses lignées de Bilateralia devaient être présentes avant "l'explosion Cambrienne". L'observation d'une acquisition simultanée de la minéralisation dans la plupart des phyla animaux pendant le Cambrien (Bengston *et al.*, 1992), alors que les quelques fossiles de Bilateralia du Protérozoïque n'ont pas de tissus minéralisés, soutient l'hypothèse d'une absence de minéralisation avant le Cambrien.

Si l'amélogénine et, par conséquent, l'émail sont apparus dans la lignée des métazoaires il y a plus de 630 millions d'années, soit 100 millions d'années avant l'explosion Cambrienne, cela apporte un élément important en faveur de l'hypothèse d'une radiation évolutive ancienne et lente des métazoaires contre celle, proposée au départ, d'une diversification rapide des plans d'organisation au Cambrien. Puisque l'amélogénine est spécifique de l'émail, que l'émail est spécifique des dents et que les dents ne se trouvent que chez les vertébrés, cela signifierait que les premiers vertébrés étaient déjà présents il y a 630 millions d'années mais qu'aucun d'entre eux n'a encore été découvert dans les gisements fossiles datant du Précambrien (peu d'individus ou période de fossilisation difficile). Il est alors difficile d'envisager la présence d'amélogénine chez la myxine, un crâniate qui ne possède pas de dents ni de tissus apparenté à l'émail.

Cependant, cela n'exclut pas la présence de l'amélogénine dans d'autres groupes de métazoaires à condition (1) qu'ils se soient séparés des autres métazoaires après l'apparition de l'amélogénine, donc il y a moins de 630 millions d'années, et (2) qu'ils aient conservé des tissus apparentés à l'émail. Parmi les candidats possibles se trouvent les conodontes qui s'enracineraient avant les gnathostomes et possédaient des appareils pharyngiens ressemblant à dents (Smith et Coates, 1998). Dans des groupes plus basaux de métazoaires, comme les céphalocordés ou les urocordés, la présence possible du gène de l'amélogénine est sujette au doute. En effet, ces organismes filtreurs ne semblent pas posséder d'odontodes ou d'organe homologue des dents. De plus, même si les premiers fossiles appartenant à ces deux groupes apparaissent au moment de l'explosion Cambrienne, rien n'interdit de supposer que ces lignées ne soient pas plus anciennes et qu'elles aient divergé avant l'apparition de l'amélogénine. Si l'on envisage maintenant que l'amélogénine ancestrale avait une autre fonction que celle de former l'émail, cela pourrait expliquer (1) son origine ancienne en absence de vertébrés possédant de l'émail et (2) sa présence au niveau des denticules cornés de myxine nous venons de le voir dans la partie consacrée à la séquence de myxine. Toutefois, la détection d'amélogénine par des anticorps chez cette dernière suggère une bonne conservation de certaines régions de la molécule, ce qui n'est pas confirmé par la recherche très difficile du gène dans les groupes non-tétrapodes. Les études en cours apporteront peut-être une réponse à cette apparente contradiction.

III. Evolution moléculaire de l'amélogénine chez les mammifères

L'originalité de ce chapitre tient avant tout dans l'apport de 14 nouvelles séquences d'amélogénine dans des lignées mammaliennes représentatives de l'ensemble du clade mammifère. Ce grand éventail de séquences a permis de développer une approche évolutive fondée sur l'étude comparative de toutes les séquences disponibles. Ce type de démarche n'était pas possible auparavant en raison du peu de séquences disponibles et de leur mauvaise répartition dans le cladogramme des mammifères. Les recherches menées depuis 20 ans sur l'amélogénine ont mis en évidence de nombreuses caractéristiques biochimiques de la protéine, de même que la présence d'insertions de triplets d'acides aminés dans l'exon 6. Les études en génétique humaine ont révélé que certains résidus étaient essentiels au bon fonctionnement de la protéine et, lorsqu'ils étaient substitués, l'amélogénèse était imparfaite. Cependant, ces observations étant limitées à quelques groupes seulement, elles ne pouvaient pas être exploitées à l'échelle évolutive.

La comparaison d'un grand nombre de séquences (26) a permis d'apporter des résultats originaux comme, par exemple, la mise en évidence de "points chauds" de mutation en relation avec une structure particulière de l'ADN. Certaines régions du gène de l'amélogénine accumulent plus de mutations que d'autres alors que les pressions sélectives sont les mêmes, uniquement parce qu'elles sont riches en couples CpG. Ce qui est remarquable, c'est certaines

régions riches en CpG possèdent très peu de substitutions, ce qui indique la présence de fortes pressions de sélection à ces endroits. Ces régions sont impliquées 1) dans l'épissage alternatif de l'exon 6 (LRAP) et 2) dans la coupure protéolytique de la protéine (TRAP). Elles semblent donc jouer un rôle majeur dans la formation de l'émail chez les mammifères.

Le fait qu'une grande partie de l'exon 6 soit soumise à plus de variation que le reste du gène a été aussi relevé mais la caractéristique principale de cet exon est d'être riche en prolines et en glutamines. Le grand nombre de séquences a permis également de mieux caractériser le rôle de cette partie de la protéine. Si les pourcentages de proline et de glutamine sont bien conservés chez tous les mammifères, et cela malgré la variabilité de cette région, c'est qu'il existe des pressions de sélection sur ces deux résidus, sans doute importants pour cette région hydrophobe de la molécule.

La présence dans l'exon 6 de répétitions de triplets d'acides aminés contenant presque toujours une proline et souvent une glutamine, a permis, dans un premier temps, d'analyser la genèse de ce phénomène. Chez certains mammifères (opossum, vache, etc.) certaines répétitions, au centre de l'exon 6, se sont autogénérées indépendamment dans chaque taxon, probablement par un système de glissement de l'ADN polymérase. En fait, la presque totalité de l'exon 6 semble avoir été générée de la même façon, par insertion de triplets d'acides aminés à différentes étapes de l'évolution de l'amélogénine. Les plus anciens événements sont peu reconnaissables car ils ont accumulé trop de mutations, mais plus l'événement s'est produit récemment plus il est possible de d'identifier ces triplets. Un fait remarquable, observé dans cette région de l'amélogénine, est la conservation de son caractère hydrophobe, quel que soit le nombre de substitutions, délétions ou insertions, ce qui faciliterait les interactions protéine-protéine (Williamson, 1994). En s'appuyant sur les données comparatives, il semble que cette région hydrophobe pourrait contribuer à augmenter la solidité de l'émail nécessaire au mode de dentition particulier des mammifères chez lesquels les dents ne sont pas remplacées continuellement au cours de la vie.

L''étude évolutive a aussi montré la conservation de certains acides aminés, chez toutes les espèces étudiées et à des emplacements dont la fonction n'est pas encore connue; il est certain que ces acides aminés jouent un rôle crucial dans la fonction de l'amélogénine. C'est grâce à l'existence de sites liés à l'amélogénèse imparfaite qu'une erreur produite dans une séquence d'amélogénine de cochon d'inde (GenBank n° AJ-012200) a pu être corrigée. En effet, une mutation était décrite chez l'homme au niveau Pro40, un acide aminé qui lorsqu'il est substitué, est responsable d'une amélogenèse imparfaite (Ravindranath *et al.*, 1999). Chez le cochon d'inde, la présence d'une mutation à cet emplacement était suspecte car elle aurait dû provoquer une amélogenèse imparfaite. Le séquençage de l'amélogénine chez un autre individu a permis de constater qu'il n'y avait pas de mutation en Pro41, et que la mutation décrite était probablement une erreur de séquençage.

Ce chapitre n'aurait pas été complet sans évoquer la présence de l'amélogénine sur les chromosomes sexuels X et Y et d'en étudier l'évolution par rapport à l'évolution des chromosomes sexuels. Jusqu'à présent, les comparaisons entre les gènes X et Y chez l'homme et la vache avaient conduit à supposer que des événements de duplication indépendants avaient créé le gène du chromosome Y à partir de celui du X dans les deux lignées (Girondot et Sire, 1998). L'absence du gène de l'amélogénine sur le chromosome Y chez d'autres mammifères avait conforté cette hypothèse. Cependant, de nouvelles séquences d'amélogénine sur les chromosomes X et Y ont été publiées et l'interprétation de ces nouvelles données à la lumière de cette hypothèse était très difficile. Les dernières connaissances de l'évolution du chromosome Y (par exemple, Iwase *et al.*, 2001; Lahn *et al.*, 2001) ont permis d'apporter une nouvelle interprétation. A partir des données dont nous disposons actuellement, le scénario suivant peut être proposé : un ancêtre des euthériens possédait le gène de l'amélogénine sur les chromosomes X et Y, deux copies maintenues

identiques grâce au phénomène de recombinaison chromosomique. Cependant, le chromosome Y a été soumis, au cours de l'évolution des mammifères à plusieurs inversions chromosomiques. La dernière inversion sur le chromosome Y a eu lieu chez un ancêtre des euthériens peu de temps avant la divergence des différentes lignées et elle était très proche du gène de l'amélogénine. Les recombinaisons ont progressivement cessé de se produire entre les deux copies X et Y du gène de l'amélogénine qui ont alors commencé à diverger. La recombinaison X/Y a cessé indépendamment chez les artiodactyles, les périssodactyles et les primates, dont les amélogénines X et Y ont commencé à diverger différemment. Dans d'autres lignées, comme celle des rongeurs, la divergence a été telle (événement plus ancien ou évolution accélérée) que le gène de l'amélogénine a totalement disparu sur le chromosome Y. Chez les artiodactyles les recombinaisons ont également cessé indépendamment dans les différentes lignées. Les deux copies X et Y de l'amélogénine de la chèvre étant identiques on peut penser que les recombinaisons X et Y continuent à avoir lieu.

Bien que les séquences de l'amélogénine des chromosomes Y étaient toutes disponibles au moment où a débuté cette étude, ces interprétations n'ont été possibles que grâce aux études comparatives des différentes séquences dans un cadre évolutif.

Ce chapitre de discussion générale permet de revenir sur de deux séquences particulières d'amélogénine qui ont été très rapidement évoquées dans le chapitre II (“Évolution de l'amélogénine chez les mammifères“). En effet, j'ai pu obtenir les séquences d'amélogénine chez deux espèces appartenant à un groupe très particulier de mammifères, les édentés. Le paresseux et le tamanoir sont connus dans les traités de zoologie ou d'anatomie dentaire pour ne pas posséder d'émail (ni même de dents chez le tamanoir). Comme on l'a vu dans les chapitres précédents, l'amélogénine est exclusivement utilisée dans la matrice de l'émail et les taxa qui ont perdu les dents n'ont pas d'amélogénine (Girondot et Sire, 1998). Comment expliquer alors la présence d'amélogénine chez ces deux édentés ? Les données concernant les dents de cette famille de mammifères sont peu nombreuses, anciennes et souvent fragmentaires et contradictoires. Cependant, on sait que chez certains édentés, les Dasypodidae, l'émail est très fin et visible uniquement sur coupe histologique à fort grossissement (Martin, 1916). Les données concernant le développement embryonnaire des édentés sont quasi inexistantes, il est donc impossible de vérifier s'ils possèdent des dents à l'état embryonnaire. L'amélogénine de ces deux édentés est bien conservée si on la compare à celle des autres mammifères, dans les régions hydrophiles. En revanche, la richesse en proline et glutamine dans l'exon 6, que l'on soupçonne être en relation avec la solidité de l'émail est bien inférieure à celle des autres groupes. Leur émail serait donc peu solide. De plus, l'amélogénine de ces deux espèces présente des mutations dans l'exon 6 qui empêchent 1) l'épissage alternatif à l'origine de la protéine LRAP et 2) la coupure protéolytique de l'amélogénine à l'origine des peptides TRAP impliquée dans le contrôle de la minéralisation. Nous pouvons supposer que l'émail qui en résulte doit être peu solide, peut-être transitoire et donc difficile à observer. Cette étude cependant permet d'affirmer que ces deux espèces possèdent de d'émail à un moment de leur existence (embryon ? juvénile ?) alors qu'il n'a jamais été observé jusqu'à aujourd'hui. Les caractéristiques de cet émail sont vraisemblablement différentes de celles des autres mammifères.

IV. Évolution moléculaire de l'amélogénine chez les reptiles

L'étude de l'évolution de l'amélogénine chez les reptiles s'inscrit dans la continuité de la partie précédente consacrée aux mammifères et pour laquelle 14 nouvelles séquences d'amélogénine avaient été obtenues. L'étude de l'évolution de l'amélogénine chez les reptiles apporte plus de nouveauté encore, puisque seules deux séquences d'amélogénine étaient

connues chez les reptiles et que 16 nouvelles séquences, dont 14 sont réparties dans les principales lignées de squamates et 2 appartiennent aux crocodiliens, ont été obtenues.

Les amélogénines de ces deux lignées de reptiles (crocodiliens et lépidosauriens) différent fortement ce qui s'explique sachant qu'elles sont séparées depuis environ 276 Ma (Kumar et Hedges, 1998). De plus, les amélogénines des lépidosauriens diffèrent plus les unes des autres que celles des mammifères entre eux. Ceci est la conséquence d'un temps de divergence plus long entre les différentes familles de squamates qu'entre les différentes familles de mammifères dans toutes les régions de l'amélogénine, même dans les régions les plus conservées. Mais, cette étude a également permis d'observer une très grande divergence de l'exon 6 chez les squamates. Cet exon est déjà connu chez les mammifères pour avoir, dans sa partie centrale, un taux d'évolution plus élevé que le reste de la molécule, mais les différences observées ici sont telles que de nombreuses insertions et délétions ont été introduites pour pouvoir les aligner. Cette divergence pourrait être due principalement aux faibles contraintes sélectives dans cette région (voir partie précédente), mais elle confirme aussi les théories actuelles indiquant qu'il existe un taux d'évolution plus important dans le génome des squamates que dans les autres lignées de reptiles (Hughes et Mouchiroud, 2001).

L'étude de l'évolution de l'amélogénine chez les squamates a été particulièrement informative en raison de ce taux d'évolution élevé. En effet, les résidus qui sont conservés chez tous les squamates sont supposés être très importants pour le bon fonctionnement de la protéine, ce qui pourrait ne pas être le cas chez les mammifères. Evolutivement parlant, les mammifères n'apportent pas une grande aide pour étudier l'évolution des régions de l'amélogénine qui sont soumises à de fortes pressions sélectives car les principales lignées mammaliennes se sont séparées dans une période de temps assez court, et il en résulte une grande similitude des différents gènes de l'amélogénine. Il était alors difficile chez les mammifères de distinguer les résidus importants de ceux qui avaient été hérités uniquement d'un ancêtre commun. Les résidus connus pour entraîner, lorsqu'ils sont substitués, une amélogenèse imparfaite chez l'homme sont conservés chez tous les squamates, et chez tous les crocodiliens, ainsi que les résidus impliqués dans la coupure protéolytique, TRAP. Cette étude montre que les coupures protéolytiques ne sont pas une particularité de l'amélogénine des mammifères mais qu'elles jouent aussi un rôle important dans la minéralisation de l'émail chez tous les amniotes. De plus, les nombreux résidus conservés chez les reptiles et dont on ne connaît pas le rôle, devraient faire l'objet d'études approfondies dans le futur.

L'épissage alternatif intra-exonique de l'exon 6 de l'amélogénine, observé chez les mammifères, existe également chez les reptiles, mais peut-être pas chez toutes les espèces. En effet, le site de l'exon 6 lié à cet épissage n'est pas présent chez tous les reptiles. Cependant, le manque de connaissance des sites consensus d'épissage alternatif n'exclu pas la présence d'autres sites ailleurs dans l'exon 6.

Dans la partie consacrée aux mammifères, il avait été proposé que la partie centrale de l'exon 6 des mammifères aurait été créée par l'insertion de nombreuses répétitions. La présence des répétitions Pxx est également observée chez les reptiles, un peu moins chez les squamates à cause de leur taux d'évolution plus rapide permet de constater que la création de l'exon 6 par insertion de triplets de répétitions est peut-être un phénomène ancien. La conséquence de ces insertions est que l'exon 6 de l'amélogénine est très hydrophobe et riche en prolines et en glutamines. Cependant, le taux de glutamine n'est pas aussi important chez les reptiles que chez les mammifères, et cela a amené à une hypothèse intéressante sur le rôle de ces insertions. Nous avons supposé que chez les mammifères, l'observation d'insertions récentes de triplets PxQ entraînant une richesse en glutamine, interviendrait dans la minéralisation pour permettre d'obtenir un émail plus solide capable de résister à l'abrasion. Chez les reptiles où les dents se remplacent continuellement, l'émail n'a pas besoin d'être aussi solide, d'où les caractéristiques différentes de l'exon 6 dans les deux lignées.

Enfin, il est étrange de constater que les séquences des crocodiliens sont très semblables, en particulier entre Alligatorinae et Crocodilinae, alors que les deux sous-familles seraient séparées depuis environ 66 Ma (Brochu, 1997). Cette conservation permet de se demander si les temps de divergences n'ont pas été surestimés ou s'il y a une forte conservation de l'amélogénine chez les crocodiles ou si le génome des crocodiliens est très stable. Dans tous les cas, la réponse serait très intéressante.

V. Étude phylogénétique utilisant l'amélogénine

L'apport principal de cette étude est de fournir un nouvel outil pour la réalisation de phylogénies moléculaires chez les mammifères et les squamates.

Pour cela, ce chapitre s'est appuyé sur plusieurs études moléculaires récentes (Stanhope *et al.*, 1998; Madsen *et al.*, 2001; Murphy *et al.*, 2001) présentant une phylogénie des mammifères très différente de celle qui était couramment admise. La création de deux nouveaux groupes d'euthériens, les laurasiathériens et les afrothériens, est le résultat le plus marquant de ces études. Des familles de mammifères qui étaient autrefois séparées sont maintenant regroupées chez les laurasiathériens ou les afrothériens: tenrécidés, proboscidiens (éléphants), tubulidentés, macroscélidés et chrysochloridés font partie des afrothériens et carnivores, cétartiodactyles, périssodactyles, chiroptères et euinsectivores font partie des laurasiathériens. Nos résultats confirment ces analyses malgré un nombre de bases analysées très faible (341 pb), alors que les autres études ont fait appel à plus de 10 000 pb dans certains cas. De plus, mes résultats sont très solides malgré l'absence de taxons dans certains groupes comme les afrothériens. Cette étude a permis de résoudre des points sur lesquels les autres études moléculaires n'étaient pas en accord comme, par exemple, la parenté des laurasiathériens avec le groupe primates/rongeurs, la parenté des xénarthres avec les périssodactyles et celle des chiroptères avec les carnivores. Ces résultats, en particulier la position des xénarthres qui n'avaient jamais été décrits jusqu'ici comme parents des périssodactyles, demandent une confirmation. L'ajout de nouveaux taxons permettra de confirmer ces liens de parenté en apportant de nouvelles données pour les analyses. Le signal phylogénétique présent dans cette région du gène de l'amélogénine est fort, puisqu'il permet d'obtenir des phylogénies valables avec un petit nombre de paires de bases.

sur ce qui concerne les squamates, il existe peu de travaux moléculaires récents à ce sujet (Harris *et al.,* 1999). Nos résultats confirment la présence des gekkonidés en position basale et montrent des liens de parenté entre varanidés et serpents, lacertidés et anguidés. La présence des scincidés en position assez basale n'est pas conforme à l'existence du groupe des Scincomorpha qui rassemble lacertidés et scincidés. Il est possible que les *Chalcides* choisis pour représenter les scincidés soient placés par erreur dans ce groupe, ce qui pourra être confirmé par des séquences d'amélogénine d'autres scincidés. Cependant, la solidité de la phylogénie des squamates réalisée avec le gène de l'amélogénine est un peu faible. Les familles de squamates se sont séparées il y a longtemps, ce qui a permis aux mutations de s'accumuler, brouillant un peu le signal phylogénétique. De plus, les gènes des squamates sont connus pour évoluer plus rapidement que ceux des autres reptiles ou ceux des mammifères (Hughes et Mouchiroud, 2001), brouillant encore plus l'information évolutive des gènes. Pour tenter d'améliorer les résultats, un des moyens envisagés est l'utilisation d'un plus grand nombre d'espèces avec des représentants d'autres familles de squamates qui n'ont pas été étudiées dans ce chapitre.

VI. Dentition chez *Chalcides*

Dans la perspective d'une étude de l'expression du gène de l'amélogénine dans des bourgeons dentaires en formation, il était nécessaire d'étudier la dentition chez *Chalcides* afin de déterminer s'il était possible, à partir de l'observation d'une coupe de transversale de dent, de prévoir les stades développement des dents suivantes dans une série histologique. Cette idée de pouvoir "prédire" le patron de dentition est venu est venue d'une particularité des dents chez les reptiles : la théorie "de Zahnreihen" (Edmund, 1960, 1969) qui est basée sur l'initiation successive des dents à des loci alternés. Une autre possibilité était d'étudier la symétrie de la dentition (e. g., Cooper, 1966) qui pourrait permettre de connaître l'emplacement de bourgeons dentaires dans une série de coupes à partir de l'observation de la mâchoire opposée (gauche/droite ou supérieure/inférieure).

Ayant bénéficié d'un élevage en laboratoire d'un scincidé, *chalcides*, c'est la première étude fondée sur de nombreux individus d'âge connu élevés dans les mêmes conditions et morts à des périodes différentes de l'année; de plus, les bourgeons dentaires ont été mesurés précisément afin de définir des stades de développement précis.

La théorie "Zahnreihen" et symétrie dentaire chez Chalcides

Pendant l'ontogénie, le remplacement des dents s'effectue par vagues en position dentaire alternée, comme le prévoit la théorie de "Zahnreihen" (Edmund, 1960). Cependant, ce modèle de remplacement est difficile à établir chez des jeunes et juvéniles. Il semble en effet que la théorie de "Zahnreihen" s'applique plus aux dentitions adultes qu'aux dentitions juvéniles plus irrégulières. Autrement dit, lorsque la fréquence de remplacement de dent est rapide le modèle est moins régulier (bien qu'il semble exister) que lorsque la fréquence de remplacement est lente.

Bien qu'une symétrie bilatérale ou entre mâchoires supérieure et inférieure ait souvent été observée pour quelques positions dentaires, les résultats obtenus montrent clairement qu'il y a trop de variation pour parler réellement de symétrie dans la dentition chez *Chalcides*.

Ainsi, les deux particularités connues du patron de dentition chez les reptiles, qui auraient pu permettre de prévoir des stades de développement dentaires dans une série histologique, se sont avéré être inutilisables. En revanche, contrairement a ce qui était attendu au départ, les résultats indiquent que, de l'avant à l'arrière, il est possible de prévoir la taille d'une dent de remplacement (et donc son stade de développement) en connaissant le stade de développement de la dent de remplacement précédente dans la rangée.

Autres résultats de cette étude

Au cours de la croissance de *Chalcides*, le nombre de positions dentaires sur chaque quadrant est relativement constant alors qu'il augmente dans la majorité des reptiles par l'addition de nouvelles dents postérieurement. L'hypothèse la plus simple est de considérer que les positions dentaires sont conservées au cours de l'ontogénie, mais il est également possible de supposer que la perte de certaines positions dentaires pourrait être compensé par

l'apparition de nouvelles positions en arrière de la mâchoire. L'augmentation de l'intervalle entre les dents avec l'âge et le mode de remplacement disto-lingual sont des arguments en faveurs de la première hypothèse.

Chez quelques adultes, le nombre de positions dentaires est plus petit que chez les juvéniles. Il semble que cette réduction puisse être la conséquence du remplacement de deux dents adjacentes par une seule.

> *A partir de la répartition irrégulière de la taille de toutes les dents mesurées, nous avons déterminé chez* Chalcides *que la vitesse de développement des bourgeons dentaires n'est pas régulière. A une première étape de croissance rapide succède une période de ralentissement correspondant au temps nécessaire pour la résorption de la dent fonctionnelle. Quand cette dernière est tombée la croissance n'est plus freinée et s'accélère à nouveau. Aucune étude de ce type n'a été entreprise jusqu'ici. Il se peut alors qu'une croissance régulière des dents ne soit pas la condition générale chez les squamates, ni chez les reptiles, contrairement à ce qui est généralement admis.*

L'étude du remplacement des dents et, en particulier, la position des bourgeons dentaires par rapport à la dent qu'ils vont remplacer, a un intérêt évolutif puisque trois méthodes de remplacement ont été décrites jusqu'ici : la méthode de type iguanidé, la méthode de type varanidé (Edmund, 1960) et une méthode intermédiaire qui serait dérivée du type iguanidé considéré comme primitif (Rieppel, 1978). Chez les *Chalcides* jeunes et les juvéniles, c'est la méthode intermédiaire qui est observée alors que chez les adultes c'est la méthode iguanidé. Comme les descriptions précédentes ont été réalisées uniquement chez des adultes, la plupart des lézards qui présentent le type iguanidé pourrait aussi remplacer leurs dents selon la méthode intermédiaire chez les jeunes et les juvéniles. L'hypothèse présentant la méthode de type intermédiaire comme dérivée de celle, primitive, de type iguanidé est difficile à soutenir parce qu'il semble que l'une succède à l'autre dans la série ontogénétique.

Deux dents de remplacement de taille différente peuvent être observées à une même position chez les jeunes *Chalcides*. Le fait qu'une dent soit initiée alors que la précédente n'est pas encore fonctionnelle, signifie que l'initiation ne dépend pas de l'âge de la dent fonctionnelle qui va être remplacée. Il semble donc qu'il existe un phénomène d'initiation à intervalle régulier, comme une horloge, mais dont la période augmente avec l'âge et probablement en fonction des saisons puisque le nombre de dents de remplacement est plus élevé en été. Ces deux observations montrent qu'en réalité des facteurs épigénétiques pourraient aussi intervenir dans le rythme de l'initiation des dents de remplacement.

VI. Développement dentaire et expression du gène de l'amélogénine

Ce chapitre présente une étude du développement dentaire et l'expression de l'amélogénine au cours de la formation de l'émail chez un lézard, *Chalcides viridanus*. D'un point de vue évolutif, une telle étude est essentielle pour compléter les études comparatives du développement dentaire des vertébrés, d'autant que ce thème a été peu étudié chez les squamates.

Ce travail décrit la morphogénèse et la différenciation dentaire depuis le stade de l'initiation jusqu'à la formation complète de la dent et sa résorption lors du remplacement

dentaire. Cette étude, montre que la morphogénèse et la différenciation dentaire sont semblables à celles qui ont été décrites chez les mammifères. Pour résumé le phénomène, l'initiation commence par une interaction épithélio-mésenchymateuse, puis se poursuit par la différenciation des types cellulaires (améloblastes et odontoblastes). La matrice de la prédentine est ensuite déposée par les odontoblastes avant celle de la matrice de l'émail par les améloblastes. Le dépôt de l'émail a été suivi plus attentivement afin de comparer les résultats histologiques avec l'expression du gène de l'amélogénine. Cela a permis de constater que la minéralisation commence dans la partie supérieure de la dent et se poursuit vers la base, ce qui est similaire phénomène décrit chez les mammifères. En revanche, les processus de Tomes et les prismes d'émail qui sont caractéristiques de l'amélogénèse dentaire chez les mammifères n'ont jamais été observés chez *Chalcides*.

En comparant les stades de développement dentaire avec l'expression de l'amélogénine, on constate que les transcrits d'amélogénine sont fortement exprimés dans les améloblastes au moment du dépôt de la matrice de l'émail, puis ils disparaissent peu après le début de la minéralisation. Au cours de la croissance, les d'ARNm d'amélogénine disparaissent donc progressivement du sommet de la dent pour continuer à être exprimé à la base.

Les images des dents de *Chalcides* obtenues par la microscopie électronique a balayage, ont permis d'observer les traces d'usures au niveau des dents et de constater qu'elles ne sont pas en relation avec le remplacement dentaire, ce qui confirme les résultats du chapitre précédent au sujet des vagues de remplacement dentaire chez ce scincidé. De plus, les dents de *Chalcides* changent au cours de l'ontogénèse dentaire. A partir d'un type “embryonnaire“ elles acquièrent, grâce au remplacement permanent, une forme typique à l'espèce “adulte“. La présence d'ornementations (crêtes et cuspides) a été observée chez ce lézard et celles-ci sont progressivement acquises au cours de l'ontogénèse.

CHAPITRE 5
-
BIBLIOGRAPHIE

Akaike H (1974). A new look at the statistical model identification. *IEEE Trans Automatic Control* 19:716-23.

Akita H, Fukae M, Shimoda S, Aoba T (1992). Localization of glycosylated matrix proteins in secretory porcine enamel and their possible functional roles in enamel mineralization. *Arch Oral Biol* 37(11)*:*953-62.

Aldred MJ, Crawford PJ, Roberts E, Thomas NS (1992). Identification of a nonsense mutation in the amelogenin gene (AMELX) in a family with X-linked amelogenesis imperfecta (AIH1). *Hum Genet* 90(4)*:*413-6.

Aldridge RJ, Donoghue PCJ (1998). Conodonts: a sister group to hagfishes? In: The biology of hagfishes. JM Jørgensen, JP Lomholt, RE Weber and H Malte editors. London: Chapman and Hall, pp. 15-31.

Aoba T, Moreno EC, Kresak M, Tanabe T (1989). Possible roles of partial sequences at N- and C-termini of amelogenin in protein-enamel mineral interaction. *J Dent Res* 68(9)*:*1331-6.

Aoba T (1996). Recent observations on enamel crystal formation during mammalian amelogenesis. *Anat Rec* 245(2)*:*208-18.

Ayala FJ, Rzhetsky A (1998). Origin of the metazoan phyla: molecular clocks confirm paleontological estimates. *Proc Natl Acad Sci USA* 95(2)*:*606-11.

Bartlett JD, Simmer JP, Xue J, Margolis HC, Moreno EC (1996). Molecular cloning and mRNA tissue distribution of a novel matrix metalloproteinase isolated from porcine enamel organ. *Gene* 183(1-2)*:*123-8.

Bashir MM, Abrams WR, Tucker T, Sellinger B, Budarf M, Emanuel B, Rosenbloom J (1998). Molecular cloning and characterization of the bovine and human tuftelin genes. *Connect Tissue Res* 39(1-3)*:*13-24.

Bellairs AdA, Miles AEW (1960). Apparent failure of tooth replacement in monitor lizards. *Brit J Herpetol* 2:189-194.

Bellairs AdA, Miles AEW (1965). Apparent failure of tooth replacement in monitor lizards. Addendum. *Brit J Herpetol* 3:14-5.

Bengtson S, Conway Morris S (1992). Early radiation of biomineralizing phyla. In: Origin and evolution of the Metazoa. JH Lipps and PWS Ignor editors. New York: Plenum Press, pp. 447-81.

Bengtson S (1998). Animal embryos in deep time. *Nature* 391:529-30.

Bleicher F, Couble ML, Farges JC, Couble P, Magloire H (1999). Sequential expression of matrix protein genes in developing rat teeth. *Matrix Biol* 18(2)*:*133-43.

Bonass WA, Robinson PA, Kirkham J, Shore RC, Robinson C (1994). Molecular cloning and DNA sequence of rat amelogenin and a comparative analysis of mammalian amelogenin protein sequence divergence. *Bioch Biophys Res Com* 198(2)*:*755-63.

Bonass WA, Kirkham J, Brookes SJ, Shore RC, Robinson C (1994). Isolation and characterisation of an alternatively-spliced rat amelogenin cDNA: LRAP--a highly conserved, functional alternatively-spliced amelogenin? *Biochim Biophys Acta* 1219(3)*:*690-2.

Brekken RA, Sage EH (2000). SPARC, a matricellular protein: at the crossroads of cell-matrix. *Matrix Biol* 19(7)*:*569-80.

Brochu CA (1997). Morphology, fossils, divergence timing, and the phylogenetic relationships of *Gavialis*. *Syst Biol* 46(3)*:*479-522.

Bromham L, Rambaut A, Fortey R, Cooper A, Penny D (1998). Testing the Cambrian explosion hypothesis by using a molecular dating technique. *Proc Natl Acad Sci USA* 95(21)*:*12386-9.

Bromham LD, Hendy MD (2000). Can fast early rates reconcile molecular dates with the Cambrian explosion? *Proc Roy Soc Lond B Biol Sci* 267(1447)*:*1041-7.

Bulmer M, Wolfe KH, Sharp PM (1991). Synonymous nucleotide substitution rates in mammalian genes: implications for the molecular clock and the relationship of mammalian orders. *Proc Natl Acad Sci USA* 88(14)*:*5974-8.

Burgess RC, MacLaren C (1965). Proteins in developing bovine enamel. In: Tooth Enamel, its Composition, Properties and Fundamental Structure. R Fearnhead and MV Stack editors. Bristol: Wright Bros., pp. 74-82.

Carlson S (1990). Vertebral dental structures. In: Skeletal Biomineralizations: Patterns, Processes and Evolutionary Trends. JG Carter editor. New York: Van Nostrand Reinhold, pp. 531-56.

Castanet J, Francillon-Vieillot H, Meunier FJ, de Ricqlès A (1993). In: Bone and individual Aging. BK Hall editor: CRC Press, pp. 245-83.

Catalano-Sherman J, Palmon A, Burstein Y, Deutsch D (1993). Amino acid sequence of a major human amelogenin protein employing Edman degradation and cDNA sequencing. *J Dent Res* 72(12)*:*1566-72.

Cerny R, Slaby I, Hammarstrom L, Wurtz T (1996). A novel gene expressed in rat ameloblasts codes for proteins with cell binding domains. *J Bone Miner Res* 11(7)*:*883-91.

Cerny R, Hammarstrom L (1998). Cloning, cDNA sequence, and alternative slicing of guinea-pig amelogenin mRNA. Sixth International Conference on the Chemistry and Biology of Mineralized Tissues, Vittel. (Résumé).

Cerny R, Hammarstrom L (1999). Guinea-pig amelogenin sequence: NCBI Protein Database.

Charlesworth B (1991). The evolution of sex chromosomes. *Science* 251(4997)*:*1030-3.

Charlesworth B, Charlesworth D (2000). The degeneration of Y chromosomes. *Philos Trans R Soc Lond B Biol Sci* 355(1403)*:*1563-72.

Chen J-Y, Huang DY, Li CW (1999). An early Cambrian craniate-like chordate. *Nature* 402:518-22.

Chibon P (1977). Vitesse de croissance et renouvellement des dents chez les Amphibiens. *J Embryol exp Morphol* 42:43-63.

Cina MS, Collins KA, Fitts M, Pettenati MJ (2000). Isolation and identification of male and female DNA on a postcoital condom. *Arch Pathol Lab Med* 124(7)*:*1083-6.

Cipollaro M, Di Bernado G, Forte A, Galano G, De Masi L, Galderisi U, Guarino FM, Angelini F, Cascino A (1999). Histological analysis and ancient DNA amplification of human bone remains found in caius iulius polybius house in pompeii. *Croat Med J* 40(3)*:*392-7.

Collier PM, Sauk JJ, Rosenbloom SJ, Yuan ZA, Gibson CW (1997). An amelogenin gene defect associated with human X-linked amelogenesis imperfecta. *Arch Oral Biol* 42(3)*:*235-42.

Collura RV, Stewart C-B (1995). Insertions and duplications of mtDNA in the nuclear genomes of Old World monkeys and hominoids. *Nature* 378:485-9.

Cooper JS (1965). Tooth replacement in amphibians and reptiles. *Brit J Herpetol* 3:214-7.

Cooper JS (1966). Tooth replacement in the Slow worm (*Anguis fragilis*). *J Zool Lond* 150*:*235-48.

Cooper JS, Poole DFG (1970). The dentition of agamid lizards with special reference to tooth replacement. *J Zool Lond* 162:85-98.

Cooper JS, Poole DFG (1973). The dentition and dental tissues of the agamid lizard, *Uromastyx*. *J Zool Lond* 162:85-98.

Couwenhoven RI, Snead ML (1994). Early determination and permissive expression of amelogenin transcription during mouse mandibular first molar development. *Dev Biol* 164(1)*:*290-9.

Darlu P, Tassy P (1993). Reconstruction Phylogénétique. Masson ed.

Dayhoff MO, Schwartz R, Orcutt BC (1978). Atlas of Protein Sequence and Structure.

Deakins M (1942). Changes in ash, water and organic content of pig enamel during calcification. *J Dent Res* 21:429.

Deldado S, Sire J-Y, Girondot M (1998). Evolutionnary analysis of non-mammalian amelogenin Genes. Sixth International Conference on the Chemistry and Biology of Mineralized Tissues, Vittel. (Résumé).

Delgado S, Casane D, Bonnaud L, Laurin M, Sire JY, Girondot M (2001). Molecular evidence for precambrian origin of amelogenin, the major protein of vertebrate enamel. *Mol Biol Evol* 18(12)*:*2146-53.

Delgado S, Davit-Béal T, Sire J-Y (2002, sous presse). The dentition and tooth replacement pattern in *Chalcides* (Squamata; Scincidae). *J Morphol*.

DeMar RE (1972). Evolutionary implications of Zahnreihen. *Evolution* 26:435-50.

DeMar RE (1973). The functional implications of the geometrical organization of dentitions. *J Paleontol* 47:452-61.

DeMar RE (1974). On the reality of Zahnreihen and the nature of reality in morphological studies. *Evolution* 28:328-30.

Deutsch D, Palmon A, Catalano-Sherman J, Laskov R (1987). Production of monoclonal antibodies against enamelin and against amelogenin proteins of developing enamel matrix. *Adv Dent Res* 1(2)*:*282-8.

Deutsch D (1989). Structure and function of enamel gene products. *Anat Rec* 224(2)*:*189-210.

Deutsch D, Palmon A, Fisher LW, Kolodny N, Termine JD, Young MF (1991). Sequencing of bovine enamelin ("tuftelin") a novel acidic enamel protein. *J Biol Chem* 266(24)*:*16021-8.

Deutsch D, Catalano-Sherman J, Dafni L, David S, Palmon A (1995a). Enamel matrix proteins and ameloblast biology. *Connect Tissue Res* 32(1-4)*:*97-107.

Deutsch D, Palmon A, Dafni L, Catalano-Sherman J, Young MF, Fisher LW (1995b). The enamelin (tuftelin) gene. *Int J Dev Biol* 39(1)*:*135-43.

Duffaune JP, Hubert J (1961). table de développement du lézard vivipare: *Lacerta* (Zootoca)*vivipara*. *Arch Anat Microscop* 50(3)*:*309-27.

Easteal S (1999). Molecular evidence for the early divergence of placental mammals. *Bioessays* 21(12)*:*1052-8; discussion 1059.

Eastoe JE (1963). The amino acid composition of proteins from the oral tissue. II. The matrix proteins in dentine and enamel from developing deciduous human teeth. *Arch Oral Biol* 8:633-52.

Eastoe JE (1965). The chemical composition of bone and tooth. *Adv Fluorine Res Caries Prevent* 3:5-16.

Eastoe JE (1979). Enamel protein chemistry--past, present and future. *J Dent Res* 58(Spec Issue B)*:*753-64.

Edman P (1970). Sequence determination. *Mol Biol Biochem Biophys* 8:211-55.

Edmund AG (1960). Tooth replacement phenomena in the lower vertebrates. *Contrib R Ont Mus Life Sci Div* 52:1-190.

Edmund AG (1962). Sequence and rate of tooth replacement in the Crocodilia. *Contrib R Ont Mus Life Sci Div* 56:1-42.

Edmund AG (1969). Dentition. In: Biology of Reptilia. C Gans, AdA Bellairs and TS Parsons editors. London: Academic Press.

Eggert FM, Allen GA, Burgess RC (1973). Amelogenins. Purification and partial characterization of proteins from developing bovine dental enamel. *Biochem J* 131(3)*:*471-84.

Elder Jr. JFJ, Turner BJ (1995). Concerted evolution of repetitive DNA sequences in eukaryotes. *Q Rev Biol* 70(3)*:*297-320.

Eldredge N, Gould SJ (1972). Punctuated equilibria: an alternative to phyletic gradualism. In: Models in Paleobiology. TJM Schopf editor. San Francisco: Freeman, Cooper, & Co., pp. 82-115.

Estes R, de Queiroz K, Gauthier J (1988). Phylogenetic relationships within Squamata. In: Phylogenetic Relationships of the Lizard Families. R Estes and G Pregill editors. Stanford, CA: Stanford University Press.

Evans SE, Barbadillo LJ (1998). An unusual lizard (Reptilia: Squamata) from the Early Cretaceous of Las Hoyas, Spain. *Zool J Linnean Soc* 124:235-65.

Falconi M, Pelotti S, Pappalardo G (2001). A method for sex assignment in mixed samples. *Hum Genet* 108(3)*:*267-8.

Fedonkin MA, Waggoner BM (1998). The Late Pre-cambrian fossil Kimberella is a mollusc-like bilateralian organism. *Nature* 388:868-71.

Felsenstein J (1978). Cases in which parsimony or compatibility methods will be positively misleading. *Syst Zool* 27:401-10.

Felsenstein J (1981). Evolutionary trees from DNA sequences: a maximum likelihood approach. *J Mol Evol* 17:368-76.

Felsenstein J (1984). Distance methods for inferring phylogenies: A justification. *Evolution* 38:16-24.

Fincham AG (1979). The amelogenin problem: a comparison of purified enamel matrix proteins. *Calcif Tissue Int* 27(1)*:*65-73.

Fincham AG, Belcourt AB, Termine JD, Butler WT, Cothran WC (1981). Dental enamel matrix: sequences of two amelogenin polypeptides. *Bioscience Reports* 1(10)*:*771-8.

Fincham AG, Belcourt AB, Lyaruu DM, Termine JD (1982). Comparative protein biochemistry of developing dental enamel matrix from five mammalian species. *Calcif Tissue Int* 34(2)*:*182-9.

Fincham AG, Belcourt AB, Termine JD, Butler WT, Cothran WC (1983). Amelogenins. Sequence homologies in enamel-matrix proteins from three mammalian species. *Biochem J* 211(1)*:*149-54.

Fincham AG, Hu YY, Lau E, Pavlova Z, Slavkin HC, Snead ML (1990). Isolation and partial characterization of a human amelogenin from a single foetal dentition using HPLC techniques. *Calcif Tissue Int* 47(2)*:*105-11.

Fincham AG, Hu Y, Lau EC, Slavkin HC, Snead ML (1991). Amelogenin post-secretory processing during biomineralization in the postnatal mouse molar tooth. *Arch Oral Biol* 36(4)*:*305-17.

Fincham AG, Moradian_Oldak J (1993). Amelogenin post-translational modifications: carboxy-terminal processing and the phosphorylation of bovine and porcine "TRAP" and "LRAP" amelogenins. *Biochem Biophys Res Com* 197(1)*:*248-55.

Fincham AG, Moradian-Oldak J (1995). Recent advances in amelogenin biochemistry. *Connect Tissue Res* 32(1-4)*:*119-24.

Fincham AG, Moradian-Oldak J (1996). Comparative mass spectrometric analyses of enamel matrix proteins from five species suggest a common pathway of post-secretory proteolytic processing. *Connect Tissue Res* 35(1-4)*:*151-6.

Fincham AG, Moradian-Oldak J, Simmer JP (1999). The structural biology of the developing dental enamel matrix. *J Struct Biol* 126(3)*:*270-99.

Fong CD, Slaby I, Hammarstrom L (1996). Amelin: an enamel-related protein, transcribed in the cells of epithelial root sheath. *J Bone Miner Res* 11(7)*:*892-8.

Forsman K, Lind L, Backman B, Westermark E, Holmgren G (1994). Localization of a gene for autosomal dominant amelogenesis imperfecta (ADAI) to chromosome 4q. *Hum Mol Gen* 3(9)*:*1621-5.

Fortey RA, Briggs DE, Wills MA (1997). The Cambrian evolutionary 'explosion' recalibrated. *BioEssays* 19:429-434.

Fu J (2000). Toward the phylogeny of the family Lacertidae - Why 4708 base pairs of mtDNA sequences cannot draw the picture. *Biol J Linnean Soc* 71(2)*:*203-17.

Fukae M, Ijiri H, Tanabe T, Shimizu M (1979). Partial amino acid sequences of two proteins in developing porcine enamel. *J Dent Res* 58(Spec Issue B)*:*1000-1.

Fukae M, Tanabe T, Ijiri H, Shimizu M (1980). Studies on porcine enamel proteins: a possible original enamel protein. *Tsurumi Shigaku. Tsurumi Univ Dent J* 6(2)*:*87-94.

Fukae M, Shimizu M (1983). Amino Acid sequence of main component of porcine enamel proteins. *Jpn J Oral Biol* 25 (suppl.):29.

Fukae M, Tanabe T (1985). Separation of non-amelogenin component from purified amelogenin preparation of immature porcine enamel. *Jpn J Oral Biol* 27:1249-51.

Fukae M, Tanabe T (1987). Nonamelogenin components of porcine enamel in the protein fraction free from the enamel crystals. *Calcif Tissue Int* 40(5)*:*286-93.

Fukae M, Tanabe T, Uchida T, Yamakoshi Y, Shimizu M (1993). Enamelins in the newly formed bovine enamel. *Calcif Tissue Int* 53(4)*:*257-61.

Fukae M, Tanabe T, Murakami C, Dohi N, Uchida T, Shimizu M (1996). Primary structure of the porcine 89-kDa enamelin. *Adv Dent Res* 10(2)*:*111-8.

Gans C (1957). "Anguimorph" tooth replacement in Amphisbaena alba Linnaeus, 1758, and A. fuliginosa Linnaeus, 1758 (Reptilia: Amphisbaenidae). *Breviora* 70:1-12.

Gascuel O (1997). BIONJ: An improved version of the NJ algorithm based on a simple model of sequence data. *Mol Biol Evol* 14(7)*:*685-95.

Gibson C, Golub E, Herold R, Risser M, Ding W, Shimokawa H, Young M, Termine J, Rosenbloom J (1991a). Structure and expression of the bovine amelogenin gene. *Biochemistry* 30:1075-9.

Gibson CW, Golub E, Ding W, Shimokawa H, Young M, Termine J, Rosenbloom J (1991b). Identification of the leucine-rich amelogenin peptide (LRAP) as the translation product of an alternatively splice transcript. *Biochem Biophys Res Commun* 174:1306-12.

Gibson CW, Golub EE, Abrams WR, Shen G, Ding W, Rosenbloom J (1992). Bovine amelogenin message heterogeneity: alternative splicing and Y-chromosomal gene transcription. *Biochemistry* 31:8384-8.

Gibson CW, Kucich U, Collier P, Shen G, Decker S, Bashir M, Rosenbloom J (1995). Analysis of amelogenin proteins using monospecific antibodies to defined sequences. *Connect Tissue Res* 32(1-4)*:*109-14.

Gibson CW (1999). Regulation of amelogenin gene expression. *Crit Rev Eukaryot Gene Expr* 9(1)*:*45-57.

Gibson CW, Yuan ZA, Hall B, Longenecker G, Chen E, Thyagarajan T, Sreenath T, Wright JT, Decker S, Piddington R, Harrison G, Kulkarni AB (2001). Amelogenin-deficient mice display an amelogenesis imperfecta phenotype. *J Biol Chem* 276(34)*:*31871-5.

Gill P, Ivanov PL, Kimpton C, Piercy R, Benson N, Tully G, Evett I, Hagelberg E, Sullivan K (1994). Identification of the remains of the Romanov family by DNA analysis. *Nat Genet* 6(2)*:*130-5.

Girondot M, Casane D (1998). Détection d'une évolution darwininenne positive sur les gènes de l'amélogénine humaine. Premières Journées Françaises de Biologie des Tissus Minéralisés, 22-23 janvier 1998, Lyon.

Girondot M, Delgado S, Laurin M (1998). Evolutionary analysis of "hagfish amelogenin". *Anat Rec* 252(4)*:*608-11.

Girondot M, Sire J-Y (1998). Evolution of the amelogenin gene in toothed and tooth-less vertebrates. *Europ J Oral Sci* 106 (suppl. 1):501-8.

Glimcher MJ (1964). The amino acid composition of the organic matrix and the neutral soluble and acid soluble components of embryonic bovine enamel. *biochem J* 93:198-202.

Glimcher MJ (1979). Phosphopeptides of enamel matrix. *J Dent Res* 58(Spec Issue B)*:*790-809.

Goldberg M (2001). La dent normale et pathologique.

Gonnet GH, Cohen MA, Benner SA (1992). Exhaustive matching of the entire protein sequence database. *Science* 256(5062)*:*1443-5.

Greene SR, Yuan ZA, Wright JT, Amjad H, Abrams WR, Buchanan JA, Trachtenberg DI, Gibson CW (2002). A new frameshift mutation encoding a truncated amelogenin leads to X-linked amelogenesis imperfecta. *Arch Oral Biol* 47(3)*:*211-7.

Gu X (1998). Early metazoan divergence was about 830 million years ago. *J Mol Evol* 47(3)*:*369-71.

Hänni C (2001). Fresh news on ancient DNA from bones and teeth. 7th International Conference on Tooth Morphogenesis and Differenciation, June 16-21, 2001, La Londe-les-Maures, France. (Résumé).

Harris DJ, Sinclair EA, Mercader NL, Marshall JC, Crandall KA (1999). Squamate relationships based on *C-mos* nuclear DNA sequences. *Herpetol J* 9:147-51.

Harrison HS (1901). The development and succession of teeth in *Hatteria punctata*. *J Microsc Sci* 44:161-219.

Hart PS, Vlaservich AC, Hart TC, Wright JT (2000a). Polymorphism (g2035C>T) in the amelogenin gene. *Hum Mutat (Online)* 15(3)*:*298.

Hart PS, Aldred MJ, Crawford PJ, Wright NJ, Hart TC, Wright JT (2002). Amelogenesis imperfecta phenotype-genotype correlations with two amelogenin gene mutations. *Arch Oral Biol* 47(4)*:*261-5.

Hart S, Hart T, Gibson C, Wright JT (2000b). Mutational analysis of X-linked amelogenesis imperfecta in multiple families. *Arch Oral Biol* 45(1)*:*79-86.

Hasegawa M, Kishino H, Yano T (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J Mol Evol* 22(2)*:*160-74.

Hasegawa M, Kishino H (1994). Accuracies of the simple methods for estimating the bootstrap probability of a maximum-likelihood tree. *Mol Biol Evol* 11:142-5.

Hasegawa T, Sato F, Ishida N, Fukushima Y, Mukoyama H (2000). Sex determination by simultaneous amplification of equine SRY and amelogenin genes. *J Vet Med Sci* 62(10)*:*1109-10.

Heise PJ, Maxson LR, Dowling HG, Hedges SB (1995). Higher-level snake phylogeny inferred from mitochondrial DNA sequences of 12S rRNA and 16S rRNA genes. *Molecular Biology and Evolution* 12(2)*:*259-65.

Herold R, Rosenbloom J, Granovsky M (1989). Phylogenetic distribution of enamel proteins: immunohistochemical localization with monoclonal antibodies indicates the evolutionary appearance of enamelins prior to amelogenins. *Calcif Tissue Int* 45:88-94.

Herold RC, Graver HT, Christner P (1980). Immunohistochemical localization of amelogenins in enameloid of lower vertebrate teeth. *Science* 207:1357-8.

Higgins DG, Thompson JD, Gibson TJ (1996). Using CLUSTAL for multiple sequence alignments. *Methods Enzymol* 266:383-402.

Hu CC, Bartlett JD, Zhang CH, Qian Q, Ryu OH, Simmer JP (1996a). Cloning, cDNA sequence, and alternative splicing of porcine amelogenin mRNAs. *J Dent Res* 75(10)*:*1735-41.

Hu CC, Zhang C, Qian Q, Ryu OH, Moradian-Oldak J, Fincham AG, Simmer JP (1996b). Cloning, DNA sequence, and alternative splicing of opossum amelogenin mRNAs. *J Dent Res* 75(10)*:*1728-1734.

Hu CC, Fukae M, Uchida T, Qian Q, Zhang CH, Ryu OH, Tanabe T, Yamakoshi Y, Murakami C, Dohi N, Shimizu M, Simmer JP (1997a). Sheathlin: cloning, cDNA/polypeptide sequences, and immunolocalization of porcine enamel sheath proteins. *J Dent Res* 76(2)*:*648-57.

Hu CC, Fukae M, Uchida T, Qian Q, Zhang CH, Ryu OH, Tanabe T, Yamakoshi Y, Murakami C, Dohi N, Shimizu M, Simmer JP (1997b). Cloning and characterization of porcine enamelin mRNAs. *J Dent Res* 76(11)*:*1720-9.

Hu CC, Qian Q, Zhang C, Simmer JP (1998a). Cloning of human enamelin. Sixth International Conference on the Chemistry and Biology of Mineralized Tissues, Vittel, France. (Résumé).

Hu CC, Simmer JP, Bartlett JD, Qian Q, Zhang C, Ryu OH, Xue J, Fukae M, Uchida T, MacDougall M (1998b). Murine enamelin: cDNA and derived protein sequences. *Connect Tissue Res* 39(1-3)*:*47-61.

Hu CC, Hart TC, Dupont BR, Chen JJ, Sun X, Qian Q, Zhang CH, Jiang H, Mattern VL, Wright JT, Simmer JP (2000). Cloning human enamelin cDNA, chromosomal localization, and analysis of expression during tooth development. *J Dent Res* 79(4)*:*912-9.

Hu CC, Zhang C, Simmer JP (2001). The pig amelogenin gene expresses a unique exon 4. 7th International Conference on Tooth Morphogenesis and Differenciation, June 16-21, 2001, La Londe-les-Maures, France.

Hu JC, Zhang CH, Yang Y, Karrman-Mardh C, Forsman-Semb K, Simmer JP (2001b). Cloning and characterization of the mouse and human enamelin genes. *J Dent Res* 80(3)*:*898-902.

Huang W, Chang BH-J, Hewett-Emmett D, Li W-H (1997). Sex differences in mutation rate in higher primates estimated from AMG intron sequences. *J Mol Evol* 44:463-5.

Huysseune A, Sire JY (1998). Evolution of patterns and processes in teeth and tooth-related tissues in non-mammalian vertebrates. *Eur J Oral Sci* 106 Suppl 1:437-81.

Inage T, Shimokawa H, Wakao K, Sasaki S (1996). Gene expression and localization of amelogenin in the rat incisor. *Adv Dent Res* 10(2)*:*201-7.

Ishiyama M, Mikami M, Shimokawa H, Oida S (1998). Amelogenin protein in tooth germs of the snake Elaphe quadrivirgata, immunohistochemistry, cloning and cDNA sequence. *Arch Histol Cytol* 61(5)*:*467-74.

Ishiyama M, Inage T, Shimokawa H (1999). An immunocytochemical study of amelogenin proteins in the developing tooth enamel of the gar-pike, Lepisosteus oculatus (Holostei, Actinopterygii). *Arch Histol Cytol* 62(2)*:*191-7.

Iwase M, Satta Y, Takahata N (2001). Sex-chromosomal differentiation and amelogenin genes in mammals. *Mol Biol Evol* 18(8)*:*1601-3.

Jamain S, Girondot M, Leroy P, Clergue M, Quach H, Fellous M, Bourgeron T (2001). Transduction of the human gene FAM8A1 by endogenous retrovirus during primate evolution. *Genomics* 78(1-2)*:*38-45.

Janvier P (1993). Patterns of diversity in the skull of jawless fishes. In: The skull. J Hanken and BK Hall editors. Chicago: The University of Chicago Press, pp. 131-188.

Janvier P (1996a). Early vertebrates Oxford, UK.: Clarendon Press.

Janvier P (1996b). The dawn of the vertebrates: characters versus common ascent in the rise of current vertebrate phylogenies. *Paleontology* 39:259-87.

Jegalian K, Page DC (1998). A proposed path by which genes common to mammalian X and Y chromosomes evolve to become X inactivated. *Nature* 394(6695)*:*776-80.

Jukes TH, Cantor CR (1969). Evolution of protein molecules. In: Mammalian protein mztabolism. HN Munro editor. New-York: Academic Press, pp. 21-132.

Katz EP, Mechanic GL, Glimcher MJ (1965). The ultracentrifugal and free zone electrophoretic characterization of neutral soluble proteins of embryonic bovine enamel. *Biochim Biophys Acta* 107:471-84.

Kimura M (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* 16:111-20.

Kimura M (1981). estimation of evolutionary distances between homologous nucleotide sequences. *Proc Natl Acad Sci U S A* 78:454-8.

Kindelan SA, Brook AH, Gangemi L, Lench NJ, Wong FS, Fearne J, Jackson Z, Foster G, Stringer BM (2000). Detection of a novel mutation in X-linked amelogenesis imperfecta. *J Dent Res* 79(12)*:*1978-82.

Knoll AH, Carroll SB (1999). Early animal evolution: emerging views from comparative biology and geology. *Science* 284(5423)*:*2129-37.

Krebsbach PH, Lee SK, Matsuki Y, Kozak CA, Yamada KM, Yamada Y (1996). Full-length sequence, localization, and chromosomal mapping of ameloblastin. A novel tooth-specific gene. *J Biol Chem* 271(8)*:*4431-5.

Kumar S, Hedges SB (1998). A molecular timescale for vertebrate evolution. *Nature* 392:917-20.

Lagerstrom M, Dahl N, Nakahori Y, Nakagome Y, Backman B, Landegren U, Pettersson U (1991). A deletion in the amelogenin gene (AMG) causes X-linked amelogenesis imperfecta (AIH1). *Genomics* 10(4)*:*971-5.

Lagerstrom-Fermer M, Pettersson U, Landegren U (1993). Molecular basis and consequences of a deletion in the amelogenin gene, analyzed by capture PCR. *Genomics* 17:89-92.

Lahn BT, Page DC (1999). Four evolutionary strata on the human X chromosome. *Science* 286(5441)*:*964-7.

Lahn BT, Pearson NM, Jegalian K (2001). The human Y chromosome, in the light of evolution. *Nat Rev Genet* 2(3)*:*207-16.

Lau EC, Mohandas TK, Shapiro LJ, Slavkin HC, Snead ML (1989). Human and mouse amelogenin gene loci are on the sex chromosomes. *Genomics* 4(2)*:*162-8.

Lecointre G (1993). Etude de l'impact de l'échantillonage des espèces et de la longueur des séquences sur la robustesse des phylogénies moléculaires. Implications sur la phylogénie des Téléostéens (Thèse de Doctorat), Université Paris 7.

Lecointre G, Le Guyader H (2001). Classification Phylogénétique du Vivant: Belin.

Lench NJ, Brook AH, Winter GB (1994). SSCP detection of a nonsense mutation in exon 5 of the amelogenin gene (AMGX) causing X-linked amelogenesis imperfecta (AIH1). *Hum Mol Genet* 3(5)*:*827-8.

Lench NJ, Winter GB (1995). Characterisation of molecular defects in X-linked amelogenesis imperfecta (AIH1). *Hum Mutat* 5(3)*:*251-9.

Li R, Li W, DenBesten PK (1995). Alternative splicing of amelogenin mRNA from rat incisor ameloblasts. *J Dent Res* 74(12)*:*1880-5.

Li W, Mathews C, Gao C, DenBesten PK (1998). Identification of two additional exons at the 3' end of the amelogenin gene. *Arch Oral Biol* 43(6)*:*497-504.

Liao D (1999). Concerted evolution: molecular mechanism and biological implications. *Am J Hum Genet* 64(1)*:*24-30.

Liao H, Brandsten C, Lundmark C, Christersson C, Wurtz T (1998). Osteonectin RNA and collagen alpha1(I) RNA in the developing rat maxilla. *Eur J Oral Sci* 106 Suppl 1:418-23.

Llano E, Pendas AM, Knauper V, Sorsa T, Salo T, Salido E, Murphy G, Simmer JP, Bartlett JD, Lopez-Otin C (1997). Identification and structural and functional characterization of human enamelysin (MMP-20). *Biochemistry* 36(49)*:*15101-8.

Lopez P, Forterre P, Philippe H (1999). The root of the tree of life in the light of the covarion model. *J Mol Evol* 49(4)*:*496-508.

Lyaruu DM, Hu CC, Zhang C, Qian Q, Ryu OH, Moradian-Oldak J, Woltgens JH, Fincham AG, Simmer JP (1998). Derived protein and cDNA sequences of hamster amelogenin. *Eur J Oral Sci* 106 Suppl 1:299-307.

Lyngstadaas SP, Risnes S, Nordbo H, Flones AG (1990). Amelogenin gene similarity in vertebrates : DNA sequences encoding amelogenin seem to be conserved during evolution. *J Comp Physiol B* 160:469-72.

Lyngstadaas SP, Risnes S, Sproat BS, Thrane PS, Prydz HP (1995). A synthetic, chemically modified ribozyme eliminates amelogenin, the major translation product in developing mouse enamel in vivo. *Embo J* 14(21)*:*5224-9.

MacDougall M, DuPont BR, Simmons D, Reus B, Krebsbach P, Karrman C, Holmgren G, Leach RJ, Forsman K (1997). Ameloblastin gene (AMBN) maps within the critical region for autosomal dominant amelogenesis imperfecta at chromosome 4q21. *Genomics* 41(1)*:*115-8.

MacDougall M, Simmons D, Dodds A, Knight C, Luan X, Zeichner-David M, Zhang C, Ryu OH, Qian Q, Simmer JP, Hu CC (1998). Cloning, characterization, and tissue expression pattern of mouse tuftelin cDNA. *J Dent Res* 77(12)*:*1970-8.

Maddison WP, Maddison DR (1992). MacClade: Analysis of Phylogeny and Character Evolution Massachusetts: Sunderland.

Madsen O, Scally M, Douady CJ, Kao DJ, DeBry RW, Adkins R, Amrine HM, Stanhope MJ, de Jong WW, Springer MS (2001). Parallel adaptive radiations in two major clades of placental mammals. *Nature* 409(6820)*:*610-4.

Marshall CR, Raff EC, Raff RA (1994). Dollo's law and the death and resurrection of genes. *Proc Natl Acad Sci USA* 91(25)*:*12283-7.

Martin BE (1916). Tooth development in *Dasypus novemcinctus*. *J Morphol* 27:647-91.

Mechanic GL (1971). The multicomponent re-equilibrating protein system of bovine embryonic enamelin (dental enamel protein): Chromatography in deaggregating solvents. In: Tooth Enamel II. R Fearnhead and MV Stack editors. Bristol: Wright Bros., pp. 88-92.

Meunier FJ (1983). Les tissus osseux des Ostéichthyens. Structure, genèse, croissance et évolution (Thèse de Doctorat ès Sciences), Université Paris 7.

Min GS, Kim SH, Kim W (1998). Molecular phylogeny of arthropods and their relatives: polyphyletic origin of arthropodization. *Mol Cell* 8(1)*:*75-83.

Moradian-Oldak J, Simmer JP, Lau EC, Sarte PE, Slavkin HC, Fincham AG (1994a). Detection of monodisperse aggregates of a recombinant amelogenin by dynamic light scattering. *Biopolymers* 34(10)*:*1339-47.

Moradian-Oldak J, Simmer JP, Sarte PE, Zeichner-David M, Fincham AG (1994b). Specific cleavage of a recombinant murine amelogenin at the carboxy- terminal region by a proteinase fraction isolated from developing bovine tooth enamel. *Arch Oral Biol* 39(8)*:*647-56.

Moradian-Oldak J, Simmer JP, Lau EC, Diekwisch T, Slavkin HC, Fincham AG (1995). A review of the aggregation properties of a recombinant amelogenin. *Connect Tissue Res* 32(1-4)*:*125-30.

Moradian-Oldak J, Tan J, Fincham AG (1998). Interaction of amelogenin with hydroxyapatite crystals: an adherence effect through amelogenin molecular self-association. *Biopolymers* 46(4)*:*225-38.

Moradian-Oldak J, Bouropoulos N, Wang L, Gharakhanian N (2002). Analysis of self-assembly and apatite binding properties of amelogenin proteins lacking the hydrophilic C-terminal. *Matrix Biol* 21(2)*:*197-205.

Morris SC (1997). Defusing the Cambrian 'explosion'? *Curr Biol* 7(2)*:*R71-4.

Murphy WJ, Eizirik E, Johnson WE, Zhang YP, Ryder OA, O'Brien SJ (2001). Molecular phylogenetics and the origins of placental mammals. *Nature* 409(6820)*:*614-8.

Nikiforuk GH, Simmons NS (1965). Purification and properties of protein from embryonic bovine enamel. *J Dent Res* 44:1119-1122.

Ogawa T (1977). A histological study of the gekko tooth. *Shigaku Odontol* 64:1377-1388.

Ohta T (2000). Evolution of gene families. *Gene* 259(1-2)*:*45-52.

Oida S, Nagano T, Yamakoshi Y, Ando H, Yamada M, Fukae M (2002). Amelogenin gene expression in porcine odontoblasts. *J Dent Res* 81(2)*:*103-8.

Ørvig T (1989). Histologic studies of ostracoderms, placod-erms and fossil elasmobranchs. 6. Hard tissues of Ordovician vertebrates. *Zool Scripta* 18:427-446.

Osborn JW (1970). New approach to Zahnreihen. *Nature* 225(230)*:*343-6.

Osborn JW (1971). The ontogeny of tooth succession in Lacerta vivipara Jacquin (1787). *Proc Roy Soc Lond B Biol Sci* 179(56)*:*261-89.

Osborn JW (1973). The evolution of dentitions. The study of evolution suggests how the development of mammalian dentitions may be controlled. *Am Sci* 61(5)*:*548-59.

Osborn JW, Crompton AW (1973). The evolution of mammalian from reptilian dentitions. *Breviora* 399:1-18.

Osborn JW (1974). On the control of tooth replacement in reptiles and its relationship to growth. *J Theor Biol* 46(2)*:*509-27.

Osborn JW (1978). Morphogenetic gradients: Fields versus clones. In: Development, Function and Evolution of Teeth. PM Butler and KA Joysey editors. London: Academic Press, pp. 171-201.

Osborn JW (1984). From reptile to mammals: Evolutionary consisderations of the dentition with emphasis on tooth attachment. *Symp Zool Soc Lond* 52:549-574.

Page RDM (1996). TreeView: An application to display phylogenetic trees on personal computers. *CABIOS* 12:357-8.

Pagel M (1999). Inferring the historical patterns of biological evolution. *Nature* 401:877-884.

Paine ML, Zhu DH, Luo W, Bringas P, Jr., Goldberg M, White SN, Lei YP, Sarikaya M, Fong HK, Snead ML (2000). Enamel biomineralization defects result from alterations to amelogenin self-assembly. *J Struct Biol* 132(3)*:*191-200.

Papas A, Seyer JM, Glimcher MJ (1977). Isolation from embryonic bovine dental enamel of a polypeptide (E3) containing as its only phosphorylated sequence, Glu-O-phosphoserine-Leu. *Febs Letters* 79(2)*:*276-80.

Perrière G, Gouy M (1996). WWW-Query: an on-line retrieval system for biological sequence banks. *Biochimie* 78:364-69.

Piez KA (1960). The nature of the protein matrix of human enamel. *J Dent Res* 39:712.

Prakash SK, Van den Veyver IB, Franco B, Volta M, Ballabio A, Zoghbi HY (1999). Characterization of a novel chromo domain gene in xp22.3 with homology to Drosophila msl-3. *Genomics* 59(1)*:*77-84.

Prakash SK, Paylor R, Jenna S, Lamarche-Vane N, Armstrong DL, Xu B, Mancini MA, Zoghbi HY (2000). Functional analysis of ARHGAP6, a novel GTPase-activating protein for RhoA. *Hum Mol Genet* 9(4)*:*477-88.

Quintana-Murci L, Jamain S, Fellous M (2001). Origine et évolution des chromosomes sexuels des mammifères. *C R Acad Sci III* 324(1)*:*1-11.

Rajpar MH, Harley K, Laing C, Davies RM, Dixon MJ (2001). Mutation of the gene encoding the enamel-specific protein, enamelin, causes autosomal-dominant amelogenesis imperfecta. *Hum Mol Genet* 10(16)*:*1673-7.

Rambaut A, Bromham L (1998). Estimating divergence dates from molecular sequences. *Mol Biol Evol* 15(4)*:*442-8.

Ravassipour DB, Hart PS, Hart TC, Ritter AV, Yamauchi M, Gibson C, Wright JT (2000). Unique enamel phenotype associated with amelogenin gene (AMELX) codon 41 point mutation. *J Dent Res* 79(7)*:*1476-81.

Ravindranath RM, Moradian-Oldak J, Fincham AG (1999). Tyrosyl motif in amelogenins binds N-acetyl-D-glucosamine. *J Biol Chem* 274(4)*:*2464-71.

Ravindranath RM, Tam WY, Nguyen P, Fincham AG (2000). The enamel protein amelogenin binds to the N-acetyl-D-glucosamine-mimicking peptide motif of cytokeratins. *J Biol Chem* 275(50)*:*39654-61.

Renugopalakrishnan V, Strawich ES, Horowitz PM, Glimcher MJ (1986). Studies of the secondary structures of amelogenin from bovine tooth enamel. *Biochemistry* 25(17)*:*4879-87.

Renugopalakrishnan V, Pattabiraman N, Prabhakaran M, Strawich E, Glimcher MJ (1989a). Tooth enamel protein, amelogenin, has a probable beta-spiral internal channel, Gln112-Leu138, within a single polypeptide chain: preliminary molecular mechanics and dynamics studies. *Biopolymers* 28(1)*:*297-303.

Renugopalakrishnan V, Prabhakaran M, Huang SG, Balasubramaniam A, Strawich E, Glimcher MJ (1989b). Secondary structure and limited three-dimensional structure of bovine amelogenin. *Connect Tissue Res* 22(1-4)*:*131-8.

Renugopalakrishnan V (2002). A 27-mer tandem repeat polypeptide in bovine amelogenin: synthesis and CD spectra. *J Pept Sci* 8(4)*:*139-43.

Repetski JE (1978). A fish from the Upper Cambrian of North America. *Science* 200:529-31.

Rieppel O (1978). Tooth replacement in Anguinomorph lizards. *Zoomorphology* 91:77-90.

Robinson C, Lowe NR, Weatherell JA (1975). Amino acid composition, distribution and origin of "tuft" protein in human and bovine dental enamel. *Arch Oral Biol* 20(1)*:*29-42.

Robinson C, Brookes SJ, Shore RC, Kirkham J (1998). The developing enamel matrix: nature and function. *Eur J Oral Sci* 106 Suppl 1:282-91.

Robinson PL (1976). How *Sphenodon* and *Uromastyx* grow their teeth and use them. In: Morphology and Biology of Reptiles. AdA Bellairs and CB Cox editors: Linn Soc Symp Ser n° 3, Academic Press, pp. 43-64.

Rocek Z (1980). Intraspecific and ontogenetic variation of the dentition in the green lizard Lacerta viridis (Reptilia, Squamata). *Vest Spolec Zool* 44:272-8.

Rogan PK, Schneider TD (1995). Using information content and base frequencies to distinguish mutations from genetic polymorphisms in splice junction recognition sites. *Hum Mut* 6(1)*:*74-6.

Röse C (1894). Ueber die Zahnentwicklung der Crocodile. *Morphol Arbeit* 3:195-228.

Rowe T (1999). At the roots of the mammalian family tree. *Nature* 398(6725)*:*283-4.

Ruch JV (2001). In: La dent normale et pathologique. M Goldberg and M Piette editors.

Saitou N, Nei M (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4:406-425.

Salido EC, Yen PH, Koprivnikar K, Yu LC, Shapiro LJ (1992). The human enamel protein gene amelogenin is expressed from both the X and the Y chromosomes [see comments]. *Am J Hum Genet* 50(2)*:*303-16.

Sambrook J, Fritsch EF, Maniatis T (1989). Molecular cloning: a laboratory manual. 2nd. ed. ed. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Sander PM (2001). Primless enamel in amniotes: terminology, function and evolution. In: Development, Function and Evolution of Teeth. M Teaford, MWJ Ferguson and MM Smith editors. New York: Cambridge University Press, pp. 92-106.

Sansom IJ, Smith MP, Armstrong HA, Smith MM (1992). Presence of the earliest vertebrate hard tissue in conodonts. *Science* 256(5061)*:*1308-11.

Sasaki S, Shimokawa H (1995). The amelogenin gene. *Int J Dev Biol* 39(1)*:*127-33.

Satchell PG, Shuler CF, Diekwisch TG (2000). True enamel covering in teeth of the Australian lungfish *Neoceratodus forsteri*. *Cell Tissue Res* 299(1)*:*27-37.

Schmidt WJ (1960). Uber einen Zwillingzahn bein einer Eidechse (Gerrhonotus multicarinatus webbii). *Dt zahnärztl Z* 15:1149-51.

Scholz M, Hengst S, Broghammer M, Pusch CM (2001). Intrapopulational relationships in ancient societies: a multidisciplinary study. *Z Morphol Anthropol* 83(1)*:*5-21.

Schultze HP (1996a). The scales of Mesozoic actinopterygians. In: Mesozoic fishes - Systematics and paleoecology. G Arratia and G Viohl editors. München, Germany: Verlag Dr Freidrich Pfeil, pp. 83-93.

Schultze HP (1996b). Conodont histology: an indicator of vertebrate relationship? *Mod Geol* 20:275-85.

Schwarzbauer JE, Spencer CS (1993). The *Caenorhabditis elegans* homologue of the extracellular calcium binding protein SPARC/osteonectin affects nematode body morphology and mobility. *Mol Biol Cell* 4(9)*:*941-52.

Seyer JM (1972). Bovine enamel proteins. In: The comparative molecular Biology of Extracellular Matrices. HC Slavkin editor. New-york: Academic Press, pp. 273-95.

Seyer JM, Glimcher MJ (1977). Evidence for the presence of numerous protein components in immature bovine dental enamel. *Calcif Tissue Res* 24(3)*:*253-7.

Shintani S, Kobata M, Toyosawa S, Fujiwara T, Sato A, Ooshima T (2002). Identification and characterization of ameloblastin gene in a reptile. *Gene* 283(1-2)*:*245-54.

Shoshani J (1986). Mammalian phylogeny: Comparison of morphological and molecular Results. *Mol Biol Evol* 3:222-42.

Shu D-G, Luo H-L, Conway Morris S, Zhang X-L, Hu S-X, Chen L, Han J, Zhu M, Li Y, Chen L-Z (1999). Lower Cambrian vertebrates from south China. *Nature* 402(6757)*:*21-2.

Simmer JP, Hu CC, Lau EC, Sarte P, Slavkin HC, Fincham AG (1994). Alternative splicing of the mouse amelogenin primary RNA transcript. *Calcif Tissue Int* 55(4)*:*302-10.

Simmer JP (1995). Alternative splicing of amelogenins. *Connect Tissue Res* 32(1-4)*:*131-6.

Simmer JP, Fukae M, Tanabe T, Yamakoshi Y, Uchida T, Xue J, Margolis HC, Shimizu M, DeHart BC, Hu CC, Bartlett JD (1998). Purification, characterization, and cloning of enamel matrix serine proteinase 1. *J Dent Res* 77(2)*:*377-86.

Simpson GG (1945). The principles of classification and a classification of mammals. *Bull Am Museum Nat Hist* 85:1-135.

Sire J-Y (1988). Evidence that mineralized spherules are involved in the formation of the super-ficial layer of the elasmoid scale in the cichlids *Hemichromis bimaculatus* and *Cichlasoma octo-fasciatum* (Pisces, Teleostei): an epidermal active participation? *Cell Tissue Res* 253:165-72.

Sire J-Y (1990). From ganoid to elasmoid scales in the actinopterygian fishes. *Neth J Zool* 40:75-92.

Sire J-Y, Quilhac A, Bourguignon J, Allizard F (1997). Evidence for participation of the epidermis in the deposition of superficial layer of scales in zebrafish (*Danio rerio*): a SEM and TEM study. *J Morphol* 231:161-74.

Sire J-Y, Davit-Béal T, Deldado S, Van Der Heyden C, Huysseune A (2002). First-generation teeth in nonmamalian lineages: Evidence for a conserved ancestral character? *Micros Res Techn* 59 (5): 408-34.

Sire J-Y, Huysseune A (2002, sous presse). Formation of skeletal and dental tissues in fish: A comparative and evolutionary approach. *Biol Rev*.

Sire JY, Géraudie J, Meunier FJ, Zylberberg L (1987). On the origin of ganoine: histological and ultrastructural data on the experimental regeneration of the scales of *Calamoichthys calabaricu*s (Osteichthyes, Brachyopterygii, Polypteridae). *Am. J Anat* 180:391-402.

Sire JY (1989). Scales in young *Polypterus senegalus* are elasmoid: new phylogenetic implications. *Am J Anat* 186(3)*:*315-23.

Slavkin HC, Zeichner-David M, Ferguson MWJ, Termine JD, Graham E, MacDougall M, Bringas Jr. P, Bessem C, Grodin M (1982). Phylogenetic and immunogenetic aspects of enamel proteins. In: Oral immunogenetic aspects of enamel proteins. GR Riviere and WH Hildemann editors. New York: Elsevier, pp. 241-51.

Slavkin HC, Graham EE, Zeichner-David M, Hildemann W (1983a). Enamel-like antigenes in hagfish; possible evolutionary significance. *Evolution* 37:404-12.

Slavkin HC, Samuel N, Bringas PJ, Nanci A, Santos V (1983b). Selachian tooth development: II. Immunolocalization of amelogenin polypeptides in epithelium during secretory amelogenesis in *Squalus acanthias*. *J Craniofac Genet Dev Biol* 3(1)*:*43-52.

Slavkin HC, Snead ML, Zeichner-David M, Jaskoll TF, Smith BT (1984). Concepts of epithelial-mesenchymal interactions during development: tooth and lung organogenesis. *J Cell Biochem* 26(2)*:*117-25.

Slavkin HC, Krejsa RJ, Fincham AG, Bringas P Jr., Santos V, Sasano Y, Snead ML, Zeichner-David M (1991). Evolution of enamel proteins: a paradigm for mechanisms of biomineralization. In: Mechanisms and phylogeny of mineralisation in biological systems. S Suga editor. Tokyo: Springer-Verlarg, pp. 383-9.

Slavkin HC, Diekwish T (1996). Evolution in tooth developmental biology: of morphology and molecules. *Anat Rec* 245:131-50.

Smith MM, Hall B (1993). A developmental model for evolution of the vertebrate exoskeleton and teeth. *Evol Biol* 27:387-448.

Smith MM, Coates MI (1998). Evolutionary origins of the vertebrate dentition: phylogenetic patterns and developmental evolution. *Eur J Oral Sci* 106 Suppl 1:482-500.

Smith MM, Coates MI (2000). Evolutionary origins of teeth and jaws: Developmental models and phylogenetic patterns. In: Development, Function and Evolution of Teeth. MF Teaford, MM Smith and MWJ Ferguson editors. Cambridge: Cambridge University Press, pp. 133-51.

Smith MP, Sansom IJ, Repetski JE (1996). Histology of the first fish. *Nature* 380:702-4.

Snead ML, Zeichner-David M, Chandra T, Robson KJ, Woo SL, Slavkin HC (1983). Construction and identification of mouse amelogenin cDNA clones. *Proc Natl Acad Sci USA* 80(23)*:*7254-8.

Snead ML, Lau EC, Zeichner-David M, Fincham AG, Woo SL, Slavkin HC (1985). DNA sequence for cloned cDNA for murine amelogenin reveal the amino acid sequence for enamel-specific protein. *Biochem Biophys Res Commun* 129(3)*:*812-8.

Snead ML, Luo W, Lau EC, Slavkin HC (1988). Spatial- and temporal-restricted pattern for amelogenin gene expression during mouse molar tooth organogenesis. *Development* 104:77-85.

Soares-Vieira JA, Billerbeck AE, Iwamura ES, Cardoso L, Romero Munoz D (2000a). Post-mortem forensic identity testing: application of PCR to the identification of fire victim. *Sao Paulo Med J* 118(3)*:*75-7.

Soares-Vieira JA, Billerbeck AE, Iwamura ES, Cardoso LA, Munoz DR (2000b). Parentage testing on blood crusts from firearms projectiles by DNA typing settles and insurance fraud case. *J Forensic Sci* 45(5)*:*1142-3.

Stanhope MJ, Waddell VG, Madsen O, de Jong W, Hedges SB, Cleven GC, Kao D, Springer MS (1998). Molecular evidence for multiple origins of Insectivora and for a new order of endemic African insectivore mammals. *Proc Natl Acad Sci U S A* 95(17)*:*9967-72.

Stephan W (1989). Tandem-repetitive noncoding DNA: Forms and forces. *Mol Biol Evol* 6(2)*:*198-212.

Stone AC, Stoneking M (1999). Analysis of ancient DNA from a prehistoric Amerindian cemetery. *Philos Trans R Soc Lond B Biol Sci* 354(1379)*:*153-9.

Strimmer K, von Haeseler A (1996). Quartet puzzling: a quartet maximum-likelihood method for reconstruction tree topologies. *Mol Biol Evol* 13*:*964-9.

Sullivan KM, Mannucci A, Kimpton CP, Gill P (1993). A rapid and quantitative DNA sex test: fluorescence-based PCR analysis of X-Y homologous gene amelogenin. *Biotechniques* 15(4)*:*636-8, 640-1.

Swofford DL (1993). PAUP: Phylogenetic Analysis Using Parsimony, Vers. 3.1.1. Washington, DC: Smithsonian Institution.

Tajima F, Nei M (1984). Estimation of evolutionary distance between nucleotide sequences. *Mol Biol Evol* 1(3)*:*269-85.

Tajima F, Nei M (1993). Unbiased estimation of evolutionary distance between nucleotide sequences. *Mol Biol Evol* 10:677-88.

Takagi T, Suzuki M, Baba T, Minegishi K, Sasaki S (1984). Complete amino acid sequence of amelogenin in developing bovine enamel. *Biochem Biophys Res Com* 121(2)*:*592-7.

Tamura K, Nei M (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol* 10:512-26.

Tanabe T, Aoba T, Moreno EC, Fukae M, Shimuzu M (1990). Properties of phosphorylated 32 kd nonamelogenin proteins isolated from porcine secretory enamel. *Calcif Tissue Int* 46(3)*:*205-15.

Termine JD, Torchia DA, Conn KM (1979). Enamel matrix: Structural proteins. *J Dent Res* 58(B):773-8.

Termine JD, Belcourt AB, Christner PJ, Conn KM, Nylen MU (1980). Properties of dissociatively extracted fetal tooth matrix proteins. I. Principal molecular species in developing bovine enamel. *J Biol Chem* 255(20)*:*9760-8.

Thompson JD, Higgins DG, Gibson TJ (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acid Res* 22:4673-80.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 25(24)*:*4876-82.

Throckmorton GS (1979). The effect of wear on the cheek teeth and associated dental tissues of the lizard *Uromastix aegyptius* (Agamidae). *J Morphol* 160:195-208.

Toyosawa S, O'HUigin C, Figueroa F, Tichy H, Klein J (1998). Identification and characterization of amelogenin genes in monotremes, reptiles, and amphibians. *Proc Natl Acad Sci USA* 95(22)*:*13056-61.

Toyosawa S, Fujiwara T, Ooshima T, Shintani S, Sato A, Ogawa Y, Sobue S, Ijuhin N (2000). Cloning and characterization of the human ameloblastin gene. *Gene* 256(1-2)*:*1-11.

Travis DF, Glimcher MJ (1964). The structure and organization of, and the relationships between the organic matrix and the inorganic crytals of embryonic bovine enamel. *J Cell Biol* 23:477-97.

Uchida T, Tanabe T, Fukae M, Shimizu M, Yamada M, Miake K, Kobayashi S (1991). Immunochemical and immunohistochemical studies, using antisera against porcine 25 kDa amelogenin, 89 kDa enamelin and the 13-17 kDa nonamelogenins, on immature enamel of the pig and rat. *Histochemistry* 96:129-38.

Uchida T, Fukae M, Tanabe T, *et al.* (1995). Immunochemical and immunocytochemical study of a 15 kDa non-amelogenin and related proteins in the porcine immature enamel: Proposal of a new group of enamel proteins Sheath Proteins. *Biomed Res* 16:131-140.

Valentine JW, Jablonski D, Erwin DH (1999). Fossils, molecules and embryos: new perspectives on the Cambrian explosion. *Development* 126(5)*:*851-9.

Veis A, Tompkins K, Alvares K, Wei K, Wang L, Wang XS, Brownell AG, Jengh SM, Healy KE (2000). Specific amelogenin gene splice products have signaling effects on cells in culture and in implants in vivo. *J Biol Chem* 275(52)*:*41263-72.

Vila C, Savolainen P, Maldonado JE, Amorim IR, Rice JE, Honeycutt RL, Crandall KA, Lundeberg J, Wayne RK (1997). Multiple and ancient origins of the domestic dog. *Science* 276(5319)*:*1687-9.

Wang HL, Yan ZY, Jin DY (1997). Reanalysis of published DNA sequence amplified from cretaceous dinosaur egg fossil. *Mol Biol Evol* 14(5)*:*589-91.

Watson JM, Spencer JA, Graves JA, Snead ML, Lau EC (1992). Autosomal localization of the amelogenin gene in monotremes and marsupials: implications for mammalian sex chromosome evolution. *Genomics* 14(3)*:*785-9.

Wayne RK, Nash WG, O'Brien SJ (1987a). Chromosomal evolution of the Canidae. II. Divergence from the primitive carnivore karyotype. *Cytogenet Cell Genet* 44(2-3)*:*134-41.

Wayne RK, Nash WG, O'Brien SJ (1987b). Chromosomal evolution of the Canidae. I. Species with high diploid numbers. *Cytogenet Cell Genet* 44(2-3)*:*123-33.

Wen HB, Moradian-Oldak J, Leung W, Bringas PJ, Fincham AG (1999). Microstructures of an amelogenin gel matrix. *J Struct Biol* 126(1)*:*42-51.

Westergaard B (1986). The pattern of embryonic tooth initiation in reptiles. In: Mém Mus natn Hist nat Paris (série C). DE Russell, J-P Santoro and D Sigogneau-Russell editors, pp. 55-63.

Westergaard B, Ferguson MW (1986). Development of the dentition in *Alligator mississipiensis*. Early embryonic development in the lower jaw. *J Zool Lond* 210:575-97.

Williamson MP (1994). The structure and function of proline-rich regions in proteins. *Biochem J* 297(Pt 2)*:*249-60.

Wilson DE, Reeder DM (1993). Mammal Species of the World Washington, DC USA: Smithsonian Institution Press.

Woerdeman MW (1919). Beiträge zur Entwicklungsgeschichte von Zähnen und Gebiss der Reptilien. Beitrag I. Die Anlage und Entwicklung des embryonalen Gebisses als Ganzes und seine Beziehung zur Zahnleiste. *Arch Mikrosk Anat* 92:104-192.

Wray GA, Levinton JS, Shapiro LH (1996). Molecular evidence for deep pre-Cambrian divergence among metazoan phyla. *Science* 274:568-573.

Wright JT, Deaton TG, Hall KI, Yamauchi M (1995). The mineral and protein content of enamel in amelogenesis imperfecta. *Connect Tissue Res* 32(1-4)*:*247-252.

Wright JT, Kiefer CL, Hall KI, Grubb BR (1996). Abnormal enamel development in a cystic fibrosis transgenic mouse model. *J Dent Res* 75(4)*:*966-73.

Xiao S, Zhang Y, Knoll AW (1998). Three-dimensional preservation of algae and animal embryos in a neoproterozoic phosphorite. *Nature* 391:553-8.

Yamakoshi Y, tanabe T, Fukae M, Shimizu M (1989). Amino acid sequense of porcine 25 kDa Amelogenin. In: Tooth Enamel V. RW Fearnhead editor. Yokohama.

Yamakoshi Y, Tanabe T, Fukae M, Shimizu M (1994). Porcine amelogenins. *Calcif Tissue Int* 54(1)*:*69-75.

Yamakoshi Y (1995). Carbohydrate moieties of porcine 32 kDa enamelin. *Calcif Tissue Int* 56(4)*:*323-30.

Yamakoshi Y, Pinheiro FH, Tanabe T, Fukae M, Shimizu M (1998). Sites of asparagine-linked oligosaccharides in porcine 32 kDa enamelin. *Connect Tissue Res* 39(1-3)*:*39-46.

Yamamoto K, Tsubota T, Komatsu T, Katayama A, Murase T, Kita I, Kudo T (2002). Sex Identification of Japanese Black Bear, *Ursus thibetanus japonicus*, by PCR based on Amelogenin Gene. *J Vet Med Sci* 64(6)*:*505-8.

Yamauchi K, Hamasaki S, Miyazaki K, Kikusui T, Takeuchi Y, Mori Y (2000). Sex determination based on fecal DNA analysis of the amelogenin gene in sika deer (*Cervus nippon*). *J Vet Med Sci* 62(6)*:*669-71.

Yuan ZA, Collier PM, Rosenbloom J, Gibson CW (1996). Analysis of amelogenin mRNA during bovine tooth development. *Arch Oral Biol* 41(2)*:*205-13.

Yuan ZA, Chen E, Gibson CW (2001). Model system for evaluation of alternative splicing: exon skipping. *DNA Cell Biol* 20(12)*:*807-13.

Zeichner-David M, Vides J, MacDougall M, Fincham A, Snead ML, Bessem C, Slavkin HC (1988). Biosynthesis and characterization of rabbit tooth enamel extracellular-matrix proteins. *Biochem J* 251(3)*:*631-41.

Zeichner-David M, Vo H, Tan H, Diekwisch T, Berman B, Thiemann F, Alcocer MD, Hsu P, Wang T, Eyna J, Caton J, Slavkin HC, MacDougall M (1997). Timing of the expression of enamel gene products during mouse tooth development. *Int J Dev Biol* 41(1)*:*27-38.

Zink A, Haas CJ, Reischl U, Szeimies U, Nerlich AG (2001). Molecular analysis of skeletal tuberculosis in an ancient Egyptian population. *J Med Microbiol* 50(4)*:*355-66.

Zylberberg L, Sire J-Y, Nanci A (1997). Immunodetection of amelogenin-like proteins in the ganoine of experimentally regenerating scales of *Calamoichthys calabaricus*, a primitive actinopterygian fish. *Anat Rec* 249:86-95.

CHAPITRE VI
-
ANNEXES

ANNEXE 1
Matériels & Méthodes détaillés

I. MATÉRIEL BIOLOGIQUE

(Remarque : Les noms communs et vernaculaire de ces animaux, ainsi que leur classification sont donnés en annexe 6).

Le genre *Chalcides*

Chalcides viridanus

Deux espèces du genre *Chalcides* (*Chalcides sexlineatus* et *C. viridanus*), des scinques (Sauropsida, Squamata, Scincidae), ont été choisies comme matériel biologique principal pour réaliser ma thèse. Ces espèces présentent l'avantage de pouvoir être élevées au laboratoire. L'équipe "Formation Squelettiques" UMR 8570 dans laquelle j'ai effectué ma thèse, maîtrise les conditions d'élevage de ces animaux depuis plusieurs années, et en particulier la possibilité de réduire la température l'hiver pour provoquer une sorte d'hibernation pour les animaux. Les terrariums dans lesquels se trouvent les lézards sont chauffés à l'aide de lampes qui maintiennent, le jour, une température supérieure à 20-22°C de février à juin.

A partir de juin, la température fluctue en fonction des conditions estivales, puis elle redescend de septembre à novembre, en étant toujours maintenue au-dessus de 20-22°C grâce aux lampes. A partir du mois de novembre, le chauffage est arrêté et les lézards entrent progressivement en hibernation au fur et à mesure que la température diminue, et cette période "hivernale" dure jusqu'en février. La nourriture est composée de tomates et de larves d'insectes (asticots).

Un autre avantage des *Chalcides* est de posséder des ostéodermes dont la morphologie a déjà été étudiée dans l'équipe "Formations Squelettiques".Comme je l'ai expliqué dans l'introduction, nous sommes intéressés par le développement de la couche supérieure de ces ostéodermes et, notamment, nous voulons savoir si elle contient de l'amélogénine.

Ces deux espèces sont endémiques des îles Canaries et les individus rapportés au laboratoire par Jacques Castanet (Professeur à l'Université Paris 6) ont été capturés à Ténériffe où *Chalcides sexlineatus* et *C. viridanus* sont très abondants.

La famille des scincidés à laquelle appartient le genre *Chalcides* est caractérisée par une réduction des membres plus ou moins importante en fonction des espèces. Les *Chalcides* sont omnivores, mais dans la nature ils se nourrissent principalement d'invertébrés et notamment d'insectes. Ils peuvent à l'occasion ajouter à leur régime des fruits et d'autres végétaux. De plus, ces lézards sont ovovivipares, l'éclosion des œufs a donc lieu à l'intérieur de la femelle qui peut donner naissance à plusieurs jeunes à la fois.

J'ai utilisé le foie d'un *Chalcides sexlineatus* et d'un *C. viridanus* pour extraire de l'ADN génomique et j'ai extrait les ARNm des mâchoires de ces deux espèces. De plus, plusieurs animaux ont été sacrifiés pour réaliser une étude sur le développement dentaire. Des animaux morts "naturellement" ont également été utilisés pour cette étude.

Autres Squamates

Afin de réaliser l'étude évolutive et phylogénétique du gène de l'amélogénine chez les squamates, nous avons choisi de séquencer le gène de l'amélogénine dans les principales familles de squamates.

Lacertidae

Gallotia galloti

En parallèle à mes travaux menés sur *Chalcides*, j'ai eu à ma disposition de l'ADN génomique de *Gallotia stehlini*, un lacertidé également élevé au laboratoire sur lequel Marc Girondot (actuellement Professeur à l'Université Paris XI) avait réalisé une première approche moléculaire en 1998 pour amplifier le gène de l'amélogénine (Girondot et Sire, 1998).

Les premiers résultats, positifs, avaient permis de définir une première série d'amorces pour la PCR. J'ai utilisé l'ADN de ce lézard, et obtenu ma première séquence d'amélogénine, avant de travailler sur *Chalcides sexlineatus.*

Par ailleurs, afin de compléter les études comparatives, deux *Lacerta muralis* ("lézard des murailles") ont été capturés dans la nature (espèce non-protégée) et l'ADN génomique a été extrait à partir de leur foie. De plus, j'ai extrait les ARNm d'une mâchoire de *Psammodromus algirus* un autre lacertidé capturé dans la nature, ainsi que ceux de *Gallotia gallotii* en élevage au laboratoire.

Varanidae

Marc Girondot avait également réalisé une première approche moléculaire sur *Varanus niloticus* dont il avait extrait l'ADN génomique à partir de prélèvements sanguins effectués par Vivian de Buffrénil. J'ai utilisé cet ADN pour réaliser des PCR afin d'amplifier le gène de l'amélogénine.

Serpents

Nicolas Vidal (Muséum national d'Histoire naturelle) m'a fait parvenir des échantillons de tissus conservés dans l'alcool et provenant de 3 familles différentes de serpents : *Boa constrictor* (Boidae), *Micrurus lemniscatus* (serpent corail, Elapidae), *Trimeresurus purpureomaculatus* (vipère des mangroves, Viperidae). J'ai pu extraire l'ADN génomique de ces échantillons et séquencer l'amélogénine dans le cadre de l'étude de phylogénie moléculaire et de l'analyse évolutive du gène de l'amélogénine.

Anguidae

J'ai extrait les ARNm d'une mâchoire d'orvet (*Anguis fragilis*) capturé en France, dans le cadre de l'étude phylogénétique.

Autres reptiles

Sphénodon

J'ai eu à ma disposition des tissus conservés dans l'alcool (muscles) provenant d'un *Sphenodon punctatus* (Rhynchocéphale), espèce protégée par les lois internationales et qui vit exclusivement en Nouvelle-Zélande. L'échantillon dont je me suis servi provient d'un spécimen ramené légalement en France en 1975 par Monsieur Saint-Girons pour un travail en collaboration avec certains membres de l'équipe "Formation Squelettiques". Le spécimen utilisé avait été conservé plusieurs années dans l'alcool 70° avant que je tente d'en extraire l'ADN génomique, mais sans succès. Des contacts ont été pris avec des chercheurs néo-zélandais afin d'essayer d'obtenir des tissus plus récents de cette espèce très intéressante d'un point de vue évolutif, car situé à la base des lépidosauriens.

Crocodiliens

J'ai également eu à ma disposition des jeunes crocodiles (*Crocodylus niloticus*) ainsi qu'un caïman (*Caiman crocodylus*) grâce à la collaboration scientifique de notre laboratoire avec la « Ferme aux Crocodiles » de Pierrelatte (Drôme).

J'ai pu extraire l'ADN génomique à partir de tissus provenant d'un embryon de crocodile de 28 jours, ainsi que les ARNm à partir de bourgeons dentaires provenant d'un crocodile de deux ans disséqué au laboratoire.

Par la suite, j'ai extrait les ARNm et l'ADNg d'un *Caiman crocodylus* à partir de bourgeons dentaires que nous avons disséqués au laboratoire.

Mammifères

Pour compléter les connaissances sur l'amélogénine de mammifères, j'ai effectué des PCR sur de l'ADN génomique extrait à partir de tissus appartenant à diverses espèces réparties dans différents ordres de mammifères chez lesquels l'amélogénine n'était pas connues. J'ai extrait l'ADNg de plusieurs carnivores : *Halichoerus grypus* (phoque gris), *Vulpes zerda* (fennec), *Acinonyx jubatus* (guépard). J'ai eu également à ma disposition des tissus provenant d'un hippopotame nain *(Hexaprotodon liberiensis*), d'un éléphant d'Asie (*Helephas maximus*) et d'un hérisson (*Erinaceus europaeus*, Insectivora). Enfin, j'ai pu extraire l'ADNg chez deux édentés, le paresseux (*Bradypus infuscatus*) et le tamanoir (*Tamandua tetradactyla)* dont la plupart des représentants ne possèdent que des dents à l'état embryonnaire ou ont des dents connues pour être dépourvues d'émail. J'ai pu obtenir par le Dr. W. Dabin (Muséum de la Rochelle) des tissus provenant de deux cétacés, le grand dauphin (*Tursiops truncatus*) et le rorqual commun (*Balaenoptera physalus*).

2. MÉTHODES

Méthodes de biologie moléculaire

Extraction des acides nucléiques (ADN, ARN)

Extraction d'ADNg

L'ADNg de *Chalcides sexlineatus* a été préparé à partir d'un échantillon de foie d'adulte broyé dans une solution contenant 1 ml de sarcosyl sodium 20%, 1 ml d'acétate d'ammonium 7,5 M, 12 ml l'hydrochlorure de guanidine (6 M) pour inhiber les DNases (la guanidine entre en compétition avec les bases de l'ADN génomique pour les sites actifs de l'enzyme). L'ADN est ensuite purifié en utilisant 1 volume de phénol/chloroforme, puis est précipité à l'éthanol absolu (2 volumes). L'ADN est lavé à l'éthanol 70% et il est repris enfin dans 1,5 ml de tampon TE (Tris 10 mM, pH=8 ; EDTA 1 mM). L'ADN est ensuite conservé à 4°C.

Par la suite, les extractions d'ADNg ont suivi un protocole utilisé à l'origine pour extraire de l'ADN de drosophiles et mis au point par Didier Casane et Patrick Laurenti (Maîtres de conférences respectivement à l'Université Paris XI et Paris VII). Ce protocole a l'avantage d'être rapide et simple à mettre en œuvre. Les tissus sont broyés dans 200 ml d'un tampon d'extraction chauffé à 70°C (EDTA 0,1 M ; Tris HCl 0,1 M, pH=9 ; SDS 1%). 800 ml de ce tampon d'extraction sont à nouveau ajoutés puis l'ensemble est incubé à 70°C pendant 30 minutes. On ajoute ensuite 210 μl d'acétate de potassium 5M. Le mélange est placé 30 minutes dans la glace, puis centrifugé 15 min à 13000 t/min. Le surnageant est ensuite centrifugé une seconde fois et l'ADN contenu dans le surnageant est précipité avec 0,6 volume (soit 0,6 ml) de propanol 2. On centrifuge à nouveau le mélange et le surnageant est éliminé. Le culot est lavé à l'éthanol 70° puis séché et suspendu dans 100 à 500 μl d'eau stérile ou de tampon TE (Tris 10 mM pH=8 ; EDTA 1 mM) permettant une meilleure conservation des acides nucléiques.

Enfin, pour extraire l'ADNg des mammifères, j'ai utilisé le "DNeasy® Tissue Kit" dont le principe est le suivant (Remarque : la composition de l'ensemble des tampons n'est pas indiquée par le fournisseur) :

25 mg de tissus mous (10 mg quand il s'agit de la rate) sont pesés et placés dans 180 µl d'un tampon ATL offrant des conditions favorables pour la phase suivante, c'est-à-dire l'utilisation de protéinase K (20 µl à 55°C) qui va détruire une grande partie des tissus en éliminant les protéines qu'ils contiennent. Aucun traitement par RNAse-A n'a été réalisé pour éliminer les ARNm car il n'y avait pas de possibilité d'interférence au cours des PCR puisque les messagers de l'amélogénine sont exprimés uniquement dans les dents. 200 µl de tampon AL, dont le rôle n'est pas précisé, sont ajoutés puis la solution est mélangée. 200 µl d'éthanol 100 sont alors ajoutés (précipitation de l'ADN). La mixture est ensuite placée dans une mini-colonne pour filtrer l'ADNg. La colonne est centrifugée une minute à 6000 g pour permettre à la solution de traverser le filtre (les ADNg restent fixés dessus). La colonne est ensuite lavée par 500 µl de tampon AW1 puis par 500 µl de tampon AW2 qui vont éliminer tous les résidus des tissus et des cellules sauf l'ADNg (ou les ARN). Au cours des lavages, les centrifugations sont réalisées 1 min à 6000 g (AW1) et 3 min à vitesse maximum (15300 g) (AW2) pour sécher la colonne au maximum. L'ADNg est élué avec 2x200µl de tampon d'élution (AE).

La migration de l'ADNg sur gel d'agarose se traduit par l'observation d'une bande de 20 kb confirmant la réussite de l'extraction et la non dégradation de l'ADN (si un "smir" d'ADN est observé, il peut correspondre à de l'ADNg dégradé ou la présence d'ARN si ceux-ci non pas été éliminés par l'utilisation d'une RNase).

Extraction d'ARN et fabrication d'ADN complémentaire

1ére phase : Extraction d'ARN totaux

La première phase, et probablement la plus délicate, consiste à disséquer rapidement les tissus choisis puis à les broyer dans de l'azote liquide à l'aide d'un mortier et d'un pilon.

Tout le matériel utilisé au cours de cette manipulation doit être nettoyé des RNases, des enzymes qui éliminent très rapidement l'ARN. Pour cela, les outils de dissection (pinces, ciseaux) ainsi que mortier et pilon sont chauffés à 200°C pendant 2 h dans un four. Tout le matériel consommable (cônes de Pipetman®, tubes Ependorf, tubes Falcon®) est acheté avec la garantie "RNase Free" qui signifie qu'il a subi un traitement (en général un rayonnement ionisant) qui a éliminé toute trace de cette enzyme. Certains matériels (comme la verrerie) qui ne peuvent être placés à 200°C sont irradiés pendant 10 min dans un four à UV. Enfin, les "canons" des micro-pipettes ont été nettoyés à l'eau de javel.

Pour les lézards comme *Chalcides, Gallotia, Psammodromus algirus* et *Anguis fragilis,* qui sont de taille relativement petite, une mandibule entière a été utilisée (de préférence à la mâchoire supérieure qui est plus difficile à disséquer et qui contient une plus grande proportion de tissus n'exprimant pas l'amélogénine). Les animaux ont été tués par une overdose d'anesthésique (chloroforme), puis la tête a été tranchée et la mandibule rapidement sectionnée au-dessus d'un mortier contenant de l'azote liquide de manière à protéger au maximum les ARN.

Pour des animaux de plus grande taille tels que les crocodiles et le hérisson, il a fallu procéder à une dissection des bourgeons dentaires contenant de l'émail en formation. Au fur et à mesure de leur dissection, les bourgeons étaient plongés dans un bain d'azote liquide pour éviter la lyse des ARN.

L'extraction des ARN totaux a été réalisée en utilisant le kit "RNeasy Midi" (Qiagen). (Remarque : Qiagen ne communique pas la composition des tampons cités ci-dessous). Les tissus sont broyés en fine poudre dans de l'azote liquide avant d'être repris dans un tampon de lyse (tampon RLT), pour éclater les cellules et dégrader les tissus. La quantité de tampon utilisée dépend du poids des tissus (1,9 ml de tampon RLT pour 20 à 130 mg de tissus et 3,8 mg de RLT pour 130 à 250 mg de tissus). Le mélange ainsi obtenu est passé 5 à 10 fois à travers l'aiguille d'une seringue (diamètre 1,2-0,8 mm / 18-21G) de façon à parfaitement homogénéiser le lysat. La solution est centrifugée 5 min à 3000-5000 g, puis un volume d'éthanol 70% est ajouté et le mélange est passé au vortex 5 sec, avant d'être placé sur une mini-colonne de filtration fournie par Qiagen. La colonne est centrifugée, permettant au mélange de traverser le filtre alors que les ARN qu'il contient resteront fixés à la membrane. Un tampon de lavage (RW1) est ajouté à la colonne (3,8 ml), puis celle-ci est centrifugée 5 min à 3000-5000 g. Afin de sécher la membrane et d'éliminer toute trace d'éthanol, 2,5 ml d'un second tampon de lavage (RPE) sont ajoutés à 2 reprises dans les mêmes conditions que précédemment avant centrifugation 2 min à 3000-5000 g. Enfin, les ARN totaux sont élués en appliquant 150 à 250 µl d'eau "RNase free" sur la membrane pendant 1 min avant de centrifuger la colonne 3 min à 3000-5000 g. Cette dernière étape est renouvelée une fois avec un nouveau volume d'eau "RNase free" pour permettre une élution complète des ARN.

2éme phase : Purification des ARN messagers à partir des ARN totaux.

Les ARN totaux contiennent les ARN messagers (ARNm), les ARN ribosomiques (ARNr) et des ARN de transfert (ARNt), ces deux derniers étant majoritaires. Pour fabriquer des ADNc à partir d'ARNm, il faut donc séparer les ARNm des autres ARN. Pour cela, on utilise une des propriétés fondamentales des ARNm, c'est-à-dire la présence d'une "queue poly A" du côté 3' de la séquence. En effet, en fin de synthèse des ARNm, l'ADN polymérase ajoute un nombre variable de dATP. Ceux-ci protégent un certain temps les ARN de l'attaque des RNases.

Le principe de la purification est fondé sur l'utilisation de billes de polystyrène-latex ("Oligotex") de 1,1 µm de diamètre sur lesquelles sont fixées par covalence des oligonucléotides $dC_{10}T_{30}$ qui vont fixer les "queues poly A".

Les ARN totaux sont dilués avec un volume d'eau "RNase free" dépendant de la quantité d'ARN disponible. Le tampon OBB contenant les billes Oligotex est ajouté à la solution contenant les ARN. Le mélange est incubé 3 min à 70°C (pour rompre la structure secondaire de l'ARN). L'échantillon est ensuite placé 10 min à température ambiante pour fixer les ARN aux billes (entre 20 et 30°C). Puis, il est centrifugé 2 min à une vitesse de 15300 rpm pour faire passer la solution à travers le filtre (les billes ne passent pas à travers le filtre et reste dans la colonne). Le surnageant est ensuite pipeté avec précaution. Il peut être conservé le temps de s'assurer que la liaison et l'élution des ARNm a bien eu lieu. Le culot est suspendu dans 400 µl de tampon OW2 (tampon de lavage des ARNm). Le mélange est placé sur une colonne à filtre dans un tube de micro-centrifugation et centrifugé 1 min pour faire passer la suspension à travers le filtre. L'opération est renouvelée avec la même quantité de tampon OW2. Pour permettre l'élution des ARNm (séparation des billes Oligotex), on place 20 à 100 µl (en fonction des besoins) de tampon d'élution (OEB) chauffé à 70°C sur la colonne. On mélange ce tampon avec les billes puis on centrifuge 1 min. L'opération est renouvelée pour augmenter le rendement. Les ARNm peuvent alors être conservés à –80°C jusqu'à utilisation.

3éme phase : Fabrication des ADNc à partir des ARNm.

La méthode choisie est la RT-PCR qui consiste à utiliser une polymérase particulière, la transcriptase reverse, capable de synthétiser de l'ADN à partir d'ARN, de la même manière que le font les rétrovirus. L'ADNc est synthétisé à partir des ARNm depuis la partie 3' jusqu'à la partie 5' grâce à l'utilisation d'une amorce oligo$(dT)_{18}$. En effet, les ARNm matures sont terminés dans leur partie 3' par une "queue" constituée de nombreux résidus adényl (jusqu'à 250). L'amorce se fixera donc dans cette partie polyadénylée et permettra la synthèse de tout l'ARNm jusqu'à l'extrémité 5'. Les fragments d'ADN ainsi obtenus peuvent ensuite être utilisés directement dans une réaction de PCR pour tenter d'amplifier le gène de l'amélogénine avec les amorces appropriées. Il est alors possible d'amplifier les parties de l'amélogénine situées en dehors de l'exon 6. (Remarque: dans tous les protocoles, un inhibiteur de RNase a été ajouté pour éviter la dégradation des ARN au cours de la réaction).

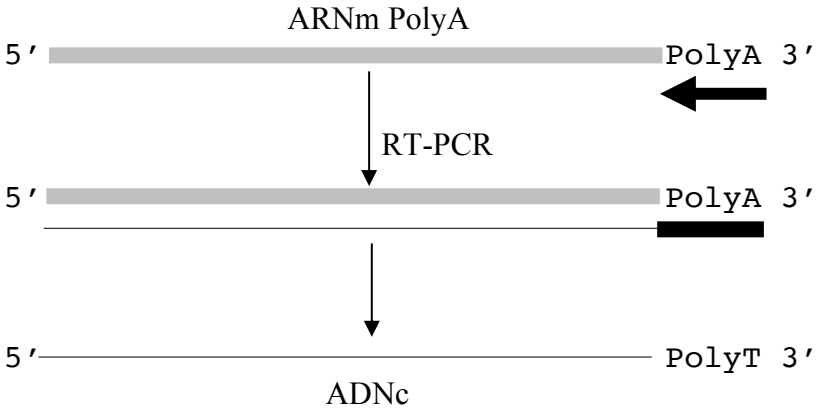
Pour réaliser ce travail, dans un premier temps j'ai utilisé le Kit "Advantage™ RT-for-PCR" (Clontech, Inc.) sur des ARNm extraits à partir de mâchoires de *Chalcides sexlineatus.* La réaction a lieu avec 12,5 µl d'ARNm (0,2 à 1 µg) auxquels sont ajoutés 1 µl d'amorce oligo$(dT)_{18}$. Le mélange est chauffé à 70°C pendant 2 min pour dénaturer les ARNm, puis les composants suivants sont ajoutés dans l'ordre: 4 µl de tampon de réaction, 1 µl de dNTP (10 mM chacun), 0,5 µl d'inhibiteur de RNase, 1 µl de reverse transcriptase MMLV dans un volume total de 6,5 µl ajusté en utilisant de l'eau distillée (traitée contre les RNases). Le mélange est incubé 1 h à 42°C pour permettre la synthèse des ADNc, puis à 94°C pendant 5 min pour arrêter la réaction de synthèse. Enfin, le mélange est dilué avec de l'eau "RNase free" dans un volume final de 100 µl avant l'utilisation en PCR.

Par la suite, j'ai utilisé le kit "ProStar First Strand RT-PCR" (Stratagene) sur des ARNm extraits à partir de mâchoires de *Chalcides viridanus.* Dans cette réaction, 50 à 100 ng d'ARNm poly(A) sont placés dans 38 µl d'eau "RNase free". 3 µl d'amorce oligo$(dT)_{18}$ sont ajoutés avant d'incuber 5 min à 65°C (dénaturer les ARNm). Le mélange est refroidi à température ambiante avant d'ajouter dans l'ordre, 5 µl de tampon de réaction 10X, 1 µl d'inhibiteur de RNase (40 U/µl), 2 µl de dNTP 100 mM et 1 µl de reverse transcriptase MMLV (50 U/µl).

J'ai également utilisé le kit "RevertAid™ First strand cDNA Synthesis" (MBI Fermentas) pour fabriquer des ADNc à partir d'ARNm de mâchoires de gallotia (*Gallotia galloti*), d'orvet (*Anguis fragilis)* et de psammodrome (*Psammodromus algirus*). Le principe

de ce système est de mélanger 10 à 500 ng d'ARNm avec 1 µl d'amorce oligo$(dT)_{18}$. Le mélange est incubé 5 min à 70°C puis refroidi rapidement sur de la glace pour dénaturer les ARNm. On ajoute ensuite dans l'ordre, 4 µl de tampon de réaction (tampon de la transcriptase reverse), 1 µl d'inhibiteur de RNase (20 U/µl) et 2 µl de dNTP (10 mM). L'ensemble est mélangé doucement puis incubé à 37°C pendant 5 min. 1 µl de transcriptase reverse M-MuLV est ajouté et le mélange est incubé à 42°C pendant 1 heure. La réaction est arrêtée en chauffant à 70°C pendant 10 min puis en refroidissant sur de la glace.

Dans les trois méthodes utilisées, la manipulation aboutit à la formation d'un ADN simple-brin (voir figure suivante). Le second brin sera fabriqué au cours du premier cycle de PCR :



Fabrication d'amorces pour la PCR

Stratégies suivies pour la fabrication des amorces de PCR

La conception d'amorces de PCR doit suivre certaines règles pour réussir à cibler une région précise de l'ADNg ou des ADNc que l'on souhaite amplifier.

La taille des amorces est variable, mais il est courant de fabriquer des amorces de 21 bases. En effet, des amorces plus petites se lieront moins solidement à l'ADN et pourront se fixer ailleurs que sur la séquence ciblée alors que des amorces plus grandes risquent de former des boucles sur elles-même. De plus, pour ces dernières les chances sont plus grandes que les nucléotides qui la constituent ne rencontrent pas de complémentaires au cours de l'hybridation.

Il faut toujours terminer l'amorce au niveau d'une seconde base de codon. En effet, le changement d'une seconde base de codon entraîne obligatoirement le changement de l'acide aminé qu'il code. Elles subissent donc plus de pression de sélection que les troisièmes bases. Il est encore plus efficace de finir l'amorce sur la troisième base d'un codon "rare", c'est-à-dire codé par un seul triplé (le Tryptophane: UGG ou la méthionine: AUG). Ces codons subissent plus de pression que les autres étant donné que le changement de l'une des trois bases transforme l'acide aminé. De plus, un acide aminé comme le tryptophane est difficile à remplacer par un autre étant donné son encombrement stérique très particulier.

Il est possible de dégénérer une amorce, c'est-à-dire que lorsque l'on hésite entre 2, 3 ou 4 nucléotides pour une position donnée, on peut choisir de fabriquer un mélange de 2, 3 ou 4 amorces qui diffèrent par un nucléotide à la position choisie. Le code international pour les diverses combinaisons de dégénérescence d'un codon est le suivant :

R= G ou A
Y= C ou T
M= A ou C
K= G ou T
S= G ou C
W= A ou T

H= A ou C ou T
B= G ou T ou C
V= G ou C ou A
D= G ou A ou T

N= A ou C ou G ou T

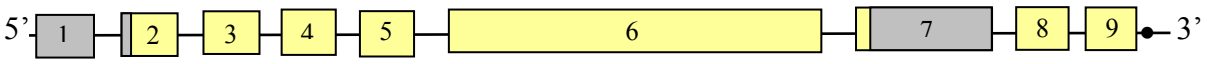
Par exemple : une séquence ANG correspond à un mélange de 4 séquences : ATG, ACG, AGG, AAG). Il faut éviter de trop dégénérer une amorce, car plus elle est dégénérée, plus elle contient un mélange d'oligonucléotides différents qui augmentent de manière exponentielle (une amorce dégénérée 2 fois avec N contient un mélange de 16 séquences différentes), donc l'oligonucléotide capable de se fixer sur la séquence ciblée sera minoritaire dans le mélange. Le risque d'amplifier de l'ADN autre que la séquence ciblée devient également plus important.

Avant de faire synthétiser un oligonucléotide, il faut vérifier que celui-ci ne risque pas de former de dimère d'amorce (c'est-à-dire se fixer avec lui-même), ou se replier sur lui-même (former une boucle), ou former un dimère avec l'autre amorce du couple. Après avoir calculé la température d'hybridation des amorces, il faut également éviter d'utiliser un couple d'amorces ayant des températures d'hybridation trop différentes.

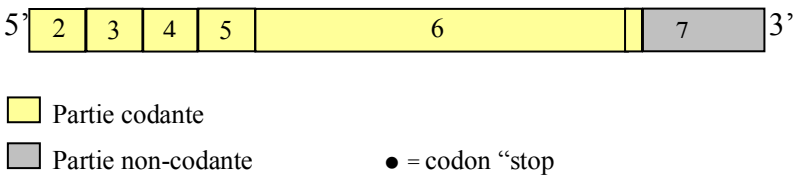
Par ailleurs, comme je l'ai expliqué dans le chapitre 1 "Introduction", le gène de l'amélogénine est composé de 9 exons. Les exons sont séparés par de longues régions, en général non codantes, appelées introns. La taille des introns est souvent supérieure à 1 kb et il est donc très difficile d'amplifier un fragment d'amélogénine situé à cheval entre 2 exons à partir d'ADNg, à moins de posséder le matériel de biologie moléculaire adéquat (polymérases et vecteurs de clonage adaptés aux longs fragments). En revanche, il est tout à fait possible de réaliser ce genre d'amplification à partir d'ADNc car ils ne possèdent pas d'introns. Pour amplifier l'ADNg et obtenir assez d'information sur la séquence d'amélogénine, je n'avais qu'un choix possible : amplifier l'exon 6 qui a une taille de 400 à 500 pb. En effet, l'amplification des autres exons de l'amélogénine, qui ont une taille moyenne de 40 à 50 pb, n'aboutirait qu'à une séquence de 1 à 10 pb seulement.

Les caractéristiques génétiques de l'amélogénine qui ont déterminé le choix des manipulations réalisées sont résumées dans les deux figures suivantes.

1/ Organisation générale du gène de l'amélogénine chez les mammifères (voir introduction) :



2/ ADNc de l'amélogénine (il s'agit d'une version complète de l'ADNc, sans épissage alternatif se terminant avec l'exon 7, majoritairement présent chez les vertébrés) :



Amorces fabriquées au cours de la thèse

Les oligonucléotides ont été fabriqués par Eurogentec et Oligogenset dans un premier temps, puis part Oligo Express (S.A.) dans un second temps (la majorité d'entre eux).

J'ai choisi, dans un premier temps, de travailler au niveau de l'exon 6 de l'amélogénine car il est plus accessible, puis de fabriquer des ADNc et d'amplifier le reste du gène en utilisant les connaissances acquises. Les premières amorces que j'ai fabriquées au début de ma thèse ont donc été construites pour amplifier l'ADNg et elles étaient situées aux extrémités 5' et 3' de l'exon 6 :

B: 5′ -GTTACGAACC**N**ATGGG**N**GG**N**TGG 3′
C: 5′ -CTTCCTCCCG**Y**TTNGT**Y**TT**R**T- 3′

avec N= A, T, G ou C. Y= T ou C. R= G ou A

(Ces amorces ont été conçues à partir de séquences d'amélogénine de mammifères, seules séquences connues en 1998).

J'ai fait ensuite synthétiser un lot d'amorces ($\mathbf{B_A}$, $\mathbf{B_G}$, $\mathbf{B_C}$, $\mathbf{B_T}$,) conçues à partir de la séquence B précédente en rajoutant un nucléotide supplémentaire du côté 3'.

Remarque: Le dernier nucléotide d'une amorce est très important dans le processus de synthèse de la polymérase. Seul l'un des quatre couples $\mathbf{B_N C}$ devrait pouvoir réamplifier les produits de PCR **BC**. L'intérêt de ce genre d'amorce est d'augmenter la spécificité des produits de PCR et d'empêcher l'amplification de contaminant **BB**. Le but est le même en ce qui concerne l'amorce $\mathbf{B_2}$, plus interne par rapport à **B** :

Ci-dessous sont présentées les séquences des amorces **B** (sens) superposées à des séquences d'amélogénine mammalienne qui ont servi de base pour les concevoir (les seules connues en 1998).

Remarque : Les séquences présentées commencent exactement au début de l'exon 6.

Sens :

	5′ 3′
HommeX	…TACCCTTCCTATGGTTACGAGCCCATGGGTGGATGGCTGCACCACCAAATCATCC…
HommeY	…TACCCTTCCTATGGTTACGAGCCCATGGGTGGATGGCTGCACCACCAAATCATCC…
Souris	…TATCCTTCCTATGGTTACGAACCCATGGGTGGATGGCTGCACCACCAAATCATCC…
Hamster	…TATCCTTCCTATGGTTACGAACCCATGGGTGGATGGCTGCACCACCAAATCATCC…
Rat	…TATCCTTCCTATGGTTACGAACCCATGGGTGGATGGCTGCACCACCAAATCATCC…
VacheX	…TATCCTTCCTATGGTTACGAACCCATGGGTGGTTGGCTGCACCACCAAATCATTC…
VacheY	…TATCCTTCCTATGGCTATGAACCCGTTGGTGGTTGGCTGCACCACCAAATAATTC…
Porc	…TACACTTCCTATGGGTACGAACCCATGGGTGGATGGCTGCACCACCAAATCATAC…
Opossum	…TATCCCTCCTATGGTTATGAACCCATGGGTGGATGGCTGCATCACCAAATCATCC…
B2	CATGGGTGG**N**TGG**Y**T**N**CA**Y**CA
B	GTTACGAACC**N**ATGGG**N**GG
B_X	GTTACGAACC**N**ATGGG**N**GGX

Par la suite, des séquences d'amélogénine provenant de deux reptiles furent connues : une séquence de caïman (*Paleosucchus palpebrosus*) (Toyosawa *et al*., 1998) et une séquence de serpent (*Elaphe quadrivirgata*) (Ishiyama *et al.*, 1998). Un nouveau jeu d'amorces (sens) a été construit en utilisant uniquement ces séquences reptiliennes car j'avais constaté une grande divergence entre les mammifères et les reptiles.

(Remarque : Les séquences présentées commencent exactement au début de l'exon 6).

Sens :

```
                5'                                                          3'
Caïman          …TATTCATCCTATGGGTATGAACCAATGGGAGGATGGCTACACCAGCCAATGTTA…
serpent         …TATCCACGTTACGGCTATGAGCCCATGGGAGGGTGGATGCACCATAATGCAGGA…
Tyr6             TATCCAYSTTATGGYTATGA
Croc                                TATGAGCCCATGGGAGGATGG
Liz                                 TATGAGCCCATGGGAGGGTGG
Caim6                         TATGGGTATGAACCAATGG
A2                            TATGGBTATGARCCHATGGGA
```

Un troisième lot d'amorces (sens et antisens) a été élaboré à partir de plusieurs séquences d'amélogénine de lézards obtenues par nos collègues et collaborateurs japonais de l'Université de Médecine Dentaire de Niigata.

Sens :

```
                5'                                                  3'
Caïman          AGACAACCG TATTCATCCTATGGGTATGAACCAATGGGAGGATGGCTACACCAG
Serpent         GGACACCAG TATCCACGTTATGGTTATGAACCAATGGGAGGATGGATGCGCCAT
Geckko          GGACACCAG TATCCACGTTACGGCTATGAGCCCATGGGAGGGTGGATGCACCAT
Iguane          GGACCCCAG TACCCACGTTATAGCTATGAACCTATGGGAGGATGGATGCACCAC
118             GGACACCAG TACCCACGTTAT
142                                   TATGAACCTATGGGAGGATGG
147                                        ACCTATGGGAGGATGGAT
150                                           TATGGGAGGATGGATGC
242                                   TATGAACCTATGGGAGGATGGATGC
```

Antisens :

```
                5'                                          3'
Caïman          CTAGAACCATGGCGGCCAATGGACAAGACCAAGCAAGAGGAAATA
Serpent         CTTGAATCATGGCCAGCTG~~~ACAAAACCAAGAAAGAAGAAGTG
Gecko           CTTGAACCATGGCCAGGAACAGATAAGACCAAGCAAGAAGAAGTG
Iguane          CTTGAATCGTGGCCAGCAGCAGACAAGACCAAGCAGGAAGAAGTG
                                         |||||||||||||||||||||
529                                      TTCTGGTTCGTTCTTCTTCAC
                                        3'                  5'
```

Amplification des ADNc

Les amplifications des ADNc ont été réalisées à partir d'amorces *sens* conçues pour s'hybrider aux exons de l'amélogénine qui se trouvent en amont de l'exon 6, c'est-à-dire les exons 2,3 et 5. Les amorces des exons 2 et 3 ont été conçue à partir des séquences publiées de caïman et de couleuvre japonaise (Toyosawa *et al*., 1998; Ishiyama *et al.*, 1998). Par contre l'amorce située dans l'exon 5, comme pour le troisième lot d'amorce destiné à l'ADNg, a été élaborée à partir de 2 séquences d'amélogénine de lézards obtenues par nos collègues japonais (Niigata) :

```
                                        EXON 2
Serpent
ATGAGAAAAATGGAGACCTGGACTTTGGTTATGTGCCTCTTAAGTACTGCATTTGCAATCCCA
CTA
Caiman
TCGAGAAAGATGGAGGGCTGGATGTTGATCACTTGCCTACTAGGTGCAACATTTGCTATACCA
TTG
exon2                        CTGGACTTTGGTTATGTGCC
exon2.2                         GACTTTGGTTATGTGCCTCTT
```

```
                                    EXON 3
Serpent   CC---------ACAACATCCTGGTTATATCAATTTCAGTTATGAG
Caiman    CCTCCCCATCCACATCATCCTGGTTATGTCAACTTCAGTTATGAG
Exon3                          CTGGTTATGTCAACTTCAGCTATGAG
Exon3.2                        CTGGTTATGTCAACTTCAGTTATGAG


                                    EXON5
Iguane    GTAATGACACCTTTAAAATGGTACCAGAGCCTAATGGGACCCCAG
Gecko     GTAATGACACCATTAAAATGGTACCAGAGCCTGATTGGACACCAG
Serpent   GTGATGACACCTTTGAAATGGTTTCAGAGCATAATTGGACACCAG
Caiman    GTGTTAACACCTTTGAAATGGTACCAGAGCCTGATGAGACAACCG
Ex5          YTWWCACCTTTGAAATGGTA
```

Afin de compléter la partie de ma thèse consacrée à la phylogénie moléculaire, j'ai fait synthétiser un lot d'amorces spécifiques des mammifères pour obtenir de nouvelles séquences d'amélogénine dans divers familles mammaliennes qui n'avait pas été étudiées jusqu'ici.

Sens :

```
                5'                                                                  3'
Homme X         …TACCCTTCCTATGGTTACGAGCCCATGGGTGGATGGCTGCACCACCAAATCATCC…
Homme Y         …TACCCTTCCTATGGTTACGAGCCCATGGGTGGATGGCTGCACCACCAAATCATCC…
Souris          …TATCCTTCCTATGGTTACGAACCCATGGGTGGATGGCTGCACCACCAAATCATCC…
Hamster         …TATCCTTCCTATGGTTACGAACCCATGGGTGGATGGCTGCACCACCAAATCATCC…
Vache X         …TATCCTTCCTATGGTTACGAACCCATGGGTGGTTGGCTGCACCACCAAATCATTC…
Vache Y         …TATCCTTCCTATGGCTATGAACCCGTTGGTGGTTGGCTGCACCACCAAATAATTC…
Porc            …TACACTTCCTATGGGTACGAACCCATGGGTGGATGGCTGCACCACCAAATCATAC…
Opossum         …TATCCCTCCTATGGTTATGAACCCATGGGTGGATGGCTGCATCACCAAATCATCC…
mam1                                TACGAACCCATGGGTGGATGGCTGC
mam127            TACCCTTCCTATGGTTACGAG
mam127B           TACCCTTCCTATGGTTACGAA
mam126          GTACCCTTCCTATGGTTACGA
mam129             CCCTTCCTATGGTTACGAGCC
```

Antisens :

```
                5'                                                                  3'
Chèvre X        …CTGACGTGCTTCTGGAAGATTGGCCAGCAACAGACAAGACCAAGCGGGAGGAAGTGGATTAA
Chèvre Y        …CTGACGTGCTTCTGGAAGATTGGCCAGCAACAGACAAGACCAAGCGGGAGGAAGTGGATTAA
Porc            …CTGACCTGCCTCTGGAAGCTTGGCCAGCAACAGACAAGACCAAGCGGGAGGAAGTGGATTAA
Vache X         …CTGACCTGCCTCTGGAAGCTTGGCCAGCAACAGACAAGACCAAGCGGGAGGAAGTGGATTAA
Vache Y         …CTGACCTCCCTCTGGAAGCCTGGCCAGCAACAGACAAGACCAAGCGGGAAGAAGTGGATTAA
Cheval Y        …CTGACCTGCCTCTGGAAGCTTGGCCAGCAACAGACAAGACCAAGCGGGAGGAAGTGGATTAA
Cheval X        …CTGACCTGCCTCTGGAAGCTTGGCCAGCAACAGACAAGACCAAGCGGGAGGAAGTGGATTAA
Homme X         …CTGATCTGACTCTGGAAGCTTGGCCATCAACAGACAAGACCAAGCGGGAGGAAGTGGATTAA
Homme Y         …CTGATCTGCATCTGGAAGCTTGGCCAGCAACAGACAAGACCAAGCAGGAGGAAGTGGATTAA
Souris          …CTGAGCTGCCTCTGGAAGCTTGGCCAGCGACAGACAAGACCAAGCGGGAAGAAGTGGATTAA
Hamster         …CTGAGCTGCCTCTGGAAGCTTGGCCAGCGACAGACAAGACCAAGCGGGAAGAAGTGGATTAA
Opossum         …CAGATATG------GAAGCTTGGCCAGCCACTGACAAGACCAAGAGGGAGGAAGTGGACTAA
                    |||||||||||||||||||||                  ||||||||||||||||||||||
mam500            3'-TAGACTGAGACCTTCGAACCG-5'
mam2                                                  3'-TTCTGGTTCGCCCTCCTTCAC-5'
mam533                                                 3'-TCTGGTTCGCCCTCCTTCACC-5'
mam501              3'-GACGGAGACCTTCGAACCGGT-5'
```

Le bilan des amorces utilisées au cours de ma thèse est présenté dans un tableau dans l'annexe 7.

L'amplification par PCR

Le principe de la PCR est le suivant : il s'agit de recopier un très grand nombre de fois ("amplifier") *in vitro* une portion de séquence comprise entre deux oligonucléotides-amorces qui sont hybridés à des régions connues. Partant de l'ADN double-brin initial et de l'extrémité 3' des amorces, la Taq polymérase, enzyme thermostable (stable à 95°C) dont l'optimum d'activité se situe vers 70-72°C, effectue les cycles de copie. Une première étape de dénaturation (entre 90 et 95°C) sépare les brins d'ADN "+" et "-". La température est ensuite abaissée brusquement jusqu'à l'optimum de d'hybridation des deux amorces (en général autour de 50-55°C). Celles-ci s'hybrident à leur cible, l'une "-" au brin "+", et l'autre "+" au brin "-", avant que l'ADN n'ait eu le temps de se renaturer. Puis la température est brusquement rehaussée à 72°C, température de synthèse de l'enzyme qui, dès lors, fabrique à partir de l'extrémité 3' libre de chacune des amorces une copie complémentaire de l'ADN initial, et ceci pour les deux brins. La synthèse terminée, deux fragments double-brin identiques au fragment de départ peuvent servir à leur tour de modèle au prochain cycle dénaturation-hybridation-synthèse. Ce cycle de température est produit ici 30 fois. (Le nombre de cycle peut être compris entre 30 et 40, parfois plus). Les brins nouvellement synthétisés lors d'un cycle servant de matrice au cycle suivant, le nombre de copies augmente exponentiellement. Les appareils de PCR ont donc comme particularité de pouvoir faire subir aux tubes dans lesquels ont lieu ces réactions des variations importantes et extrêmement rapides de température, selon des modalités programmables. (d'après Lecointre, 1993).

L'ADN qui va être amplifié peut être de l'ADNg, de l'ADNc, des produits de PCR ou une bande d'ADN prélevée sur gel d'agarose (voir "***Reprise d'ADN sur gel d'agarose*** ").

Les PCR ont été réalisées en utilisant différentes polymérases. Un "Long Template PCR System" (Boehringer Mannheim) a été utilisé dans un premier temps, puis une "Thermostable DNA Polymérase" (Advanced Biotechnologie Ltd) dans un second temps. Les conditions étaient les suivantes: Volume final de 50 µl; dNTP (mélange de dATP, dCTP, dTTP, dGTP) à 50 µM; 1 µl d'amorces à 1 nmol / 10µl (100 µM); Tampon de polymérase (type I) 1X. Les cycles de PCR suivants ont été programmés sur un thermocycleur Genius© (Techne incorporated) :

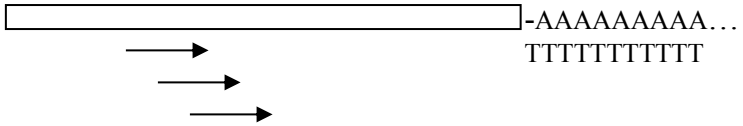
Nombre de cycles	Température	Durée
1 cycle	95°C	4 à 8 min
	95°C	1 min
30 à 40 cycles	50 à 60 °C	1 min
	72°C	1 min
1 cycle	72 °C	10 à 25 min

La température d'hybridation des amorces a été déterminée, après différents tests, entre 55°C et 60°C. Une température de 59°C fut choisie pour *Gallotia stehlini* puis 58°C pour toutes les autres PCR (voir annexe 2).

Les produits de PCR ont été séparés sur gel d'agarose 1,8% (Euromedex), dans du tampon TBE 0,5X et l'ADN a été visualisé par liaison avec du bromure d'éthidium (Bioprobe) sous lumière UV.

RACE 3' (Rapid Amplication of cDNA Ends)

Cette technique est destinée à amplifier l'extrémité 3' d'un ADNc à l'aide d'une amorce *antisens* poly(dT)$_{18}$. En effet, la transcription de l'ADNg fabrique naturellement des ARNm qui se terminent par une queue poly(A) composée de 200 à 250 A. Les ADNc fabriqués à partir de ces ARNm comportent aussi cette queue polyadénylée. Cependant, tous les ADNc possédant cette même queue, il faut employer une amorce *sens* spécifique du gène ciblé pour amplifier un ADNc particulier (ici l'amélogénine). Etant donné le manque de spécificité de l'amorce *antisens*, la première amplification ne permet jamais d'obtenir directement le fragment d'ADN recherché. Il est toujours nécessaire de prélever 1 µl de cette PCR et de la ré-amplifier à l'aide d'une amorce *sens* plus interne. A l'issue de cette PCR, un fragment d'ADN de la taille attendue apparaît, mais il est parfois nécessaire de réaliser une nouvelle PCR avec une amorce encore plus interne :

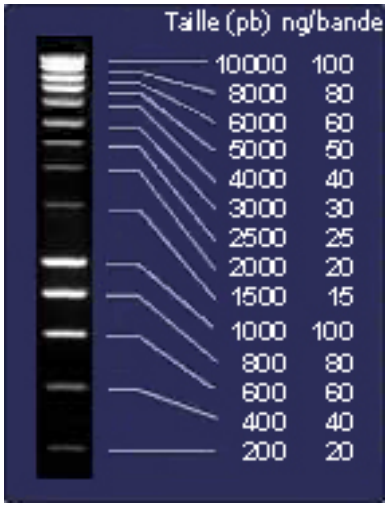


Préparation des gels d'agarose et migration de l'ADN

Pour visualiser l'ADN et séparer les produits de la PCR selon leur taille, les extraits contenants de l'ADN sont déposés sur gel d'agarose.

Les gels sont préparés en diluant une certaine quantité d'agarose (Euromedex) dans un tampon TBE (Fischer Scientific) à 0,5 X. La quantité d'agarose utilisée déterminera la concentration des gels. Pour mes expériences, j'ai utilisé des gels d'agarose à 1,8% (1,8 g pour 100 ml). Le mélange est chauffé dans un four à micro-ondes jusqu'à dissolution complète de la poudre. La solution est alors laissée refroidir à température ambiante jusqu'à environ 50-60°C, puis on ajoute 10 µl (pour 100 ml de gel) d'un intercalant de l'ADN, le bromure d'éthidium (BET), qui devient lumineux lorsqu'il est placé sous un rayonnement ultraviolet (par ses propriétés mutagènes, le BET est cancérigène).

Le gel est coulé dans un moule sur lequel un "peigne" a été disposé, permettant de creuser des puits dans lesquels l'ADN sera placé. Une fois le gel solidifié, l'ADN qui doit être déposé est mélangé avec un tampon de migration constitué (1) de glycérol, permettant d'augmenter la densité de la solution d'ADN qui se placera ainsi au fond des puits, et (2) de bleu de méthylène qui permettra de colorer en bleu la solution et ainsi de suivre la progression de la migration. J'ai utilisé 1 µl de tampon de migration pour 10 µl d'ADN.



Les gels sont ensuite photographiés grâce à des films polaroids ISO 3000/36°.

Le gel contenant l'ADN est ensuite parcouru par un courant électrique de 50 ou 100 volts (Mupid 21 de cosmo bio co., Ltd) dont le voltage détermine la vitesse de migration. Le sens du courant est choisi en fonction de l'orientation du gel, sachant que l'ADN est chargé négativement est qu'il se déplacera de l'électrode négative vers la positive. Lorsque l'ADN a suffisamment migré dans le champ électrique (le front de migration, de couleur bleue, ne doit pas sortir de l'agarose), le gel est alors placé sur une table à ultraviolet qui va permettre de visualiser l'ADN en fonction de la taille. En effet, les mailles du gel d'agarose vont freiner la migration de l'ADN proportionnellement à sa taille. Ainsi, l'ADN situé près de la zone de dépôt est le plus grand (en nombre de paires de base) alors que celui qui est le plus éloigné est le plus petit. Pour mesurer la taille des fragments d'ADN, un marqueur moléculaire a été utilisé. Le marqueur moléculaire contenant des bandes de taille connue est déposé (5 µl) en même temps que les produits de PCR. J'ai utilisé le SmartLadder Eurogentec qui permet également d'estimer la quantité d'ADN en fonction de l'intensité lumineuse de l'ADN.

Reprise d'ADN sur gel d'agarose

Lorsque la PCR génère des bandes multiples que l'on souhaite séparer ou qu'une bande jugée intéressante est trop faible et diffuse, alors il est possible de prélever sur gel d'agarose un peu d'ADN que l'on peut utiliser à nouveau dans une PCR. Pour pouvoir ré-amplifier des produits de PCR, l'ADN est directement prélevé sur le gel d'agarose. Pour cela, j'enfonce un cône de Pipetman® dans le gel à l'endroit où se trouve la bande d'intérêt, puis je trempe le cône dans un tube Ependorf contenant 10 µl d'eau distillée stérile. 1 µl du contenu de ce tube est ensuite utilisé par réaction de PCR.

Purification de produits de PCR avant clonage et séquençage

Les produits de PCR sont purifiés sur colonne à filtre QIAquick™ (QIAGEN), de manière à éliminer les amorces, les nucléotides non incorporés, l'ADN polymérase, le BET et les sels présents dans les tubes de PCR, qui pourraient perturber des réactions enzymatiques comme le clonage et le séquençage. Les produits de PCR sont mélangés à 5 volumes d'un tampon (PB) permettant une efficacité de fixation optimum de l'ADN (salinité élevée et pH ≤ 7,5) sur la membrane silica-gel placée au fond des colonnes. Cette membrane de silicone adsorbe les fragments d'ADN compris entre 70 pb et 10 kb. La solution est déposée sur une colonne et passe à travers la membrane par centrifugation à 13000 rpm pendant 1 min. La colonne est lavée dans 750 µl d'un tampon de rinçage (PE dont la composition n'est pas donnée par le fournisseur) qui va éliminer toutes les impuretés de la membrane. Puis l'ADN est séparé de la membrane par élution avec 30 à 50 µl de tampon EB (Tris Cl 10 mM, pH8,5), car cette phase est plus réalisable en conditions basiques et en faible concentration saline. Une dernière centrifugation à 13000 rpm pendant 1 min permet de récupérer l'ADN purifié dans un tube stérile.

Clonage

Le principe du clonage est d'introduire un fragment d'ADN purifié (voir 2.6) dans un vecteur (un plasmide par exemple), puis à placer le tout dans une bactérie. Ainsi, le fragment considéré est isolé et multiplié au sein d'un clone particulier. Les bactéries transformées sont en général sélectionnées par la présence d'un gène de résistance à un antibiotique sur le plasmide. Les bactéries contenant le plasmide et un insert sont ensuite sélectionnées grâce au gène *LacZ* (dans cette étude). Celui-ci est fonctionnel dans le plasmide utilisé mais pas dans la souche bactérienne. L'insertion du fragment se fait au niveau du gène *LacZ*, ce qui a pour

conséquence de le rendre non-fonctionnel. Avec un milieu contenant de l'IPTG (inducteur) (isopropylthio-ß-D-galactoside) et du Xgal (substrat) (5-bromo-4-chloro-3-indolyl-ß-D-galactoside), la β-galactosidase produite par le gène *LacZ* induit une couleur bleue (en réalité, la couleur bleue est produite par un des produits résultant de la coupure de la molécule Xgal). Ainsi, sur boîte de Pétri, les colonies bleues contiennent le plasmide sans insert (ou un insert de trop petite taille pour interrompre totalement la fonction de la protéine), alors que les colonies blanches (sur milieu contenant l'antibiotique approprié) contiennent le plasmide et un insert au niveau de *LacZ* (figure 10). Les plasmides sont ensuite extraits par des méthodes particulières de type Minipreps, Midipreps, etc.

Le clonage a été réalisé dans un premier temps grâce au système "pCR-Script Amp SK(+) electroporation competent cell cloning Kit" (Stratagene). Le plasmide utilisé dans ce cas est un pCR-script Amp SK(+) dérivant d'un Bluescript. La souche bactérienne utilisée est une *E. coli* XL1 Blue. Les produits de PCR sont tout d'abord purifiés par précipitation à l'éthanol absolu (Merck), lavés à l'éthanol 70% et re-suspendus dans un tampon TE, idéal pour l'utilisation d'enzymes par la suite. Ils sont ensuite traités avec 0,5 unités de *Pfu* polymérase et 750 µM de dNTP (dans le tampon de la *Pfu* polymérase), pendant 30 min à 37°C, de façon à créer des bouts "francs" ("blunted-end"). La réaction de ligation qui consiste à introduire les produits de PCR dans le plasmide, a été réalisée avec 1 ng de vecteur (pCR-script Amp SK+), 0,5 mM de rATP, 5 U d'enzyme de restriction *Srf* I, 4 U de T4 DNA ligase. Le rapport insert/plasmide était calculé pour être compris entre 40/1 et 100/1.

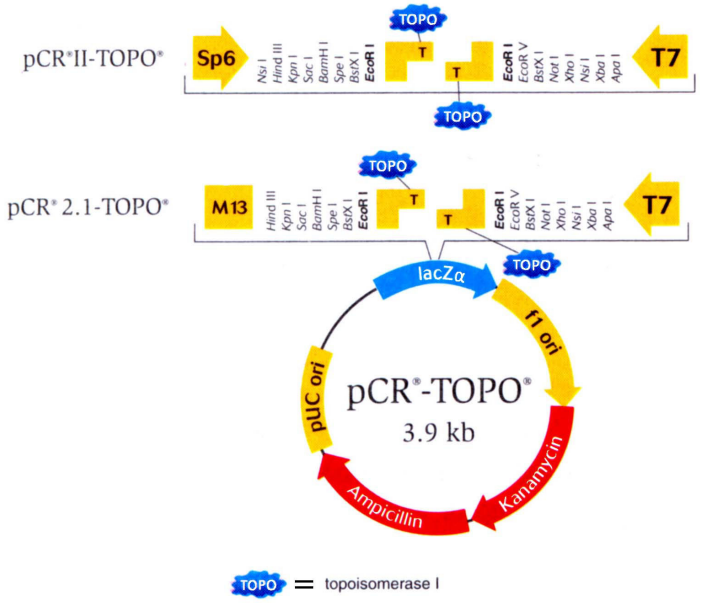


Figure 1. Schéma du vecteur pCR® 2.1 TOPO de Invitrogen™ fonctionnant sur le principe du "TA cloning". Une autre version de ce vecteur a également été utilisée: le pCR® II TOPO qui diffère du premier par la présence d'un promoteur Sp6 à 17 pb du site "M13 Reverse".

Dans une seconde phase, j'ai utilisé le système de clonage "TOPO TA Cloning® Kit " (Invitrogen™) (Figure 1). Son intérêt réside dans le fait que les produits de PCR possèdent naturellement (selon la polymérase employée) des dATP aux extrémités 3'. Ainsi, les produits de PCR et le vecteur possèdent des bouts cohésifs "A-T" qui facilitent grandement l'efficacité de ligation.

De plus, des topo isomérases (enzymes intervenant dans la topologie de l'ADN et possédant une activité ligase) vont, en 5 min et à température ambiante, intégrer le produit de PCR au plasmide. Pour que la réaction se fasse, il suffit simplement de mettre en contact 2 µl de produit de PCR, 1 µl de solution permettant d'obtenir des conditions idéales pour la réaction et 1 µl de vecteur, ce qui simplifie considérablement le protocole expérimental.

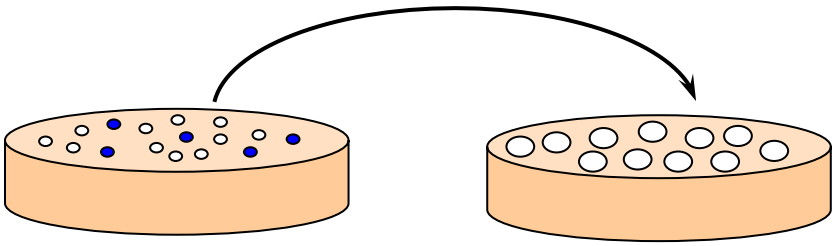
Les transformations ont été réalisées à l'aide d'une des deux méthodes suivantes:

1°) Les bactéries ont été transformées par un protocole "classique" faisant subir aux "XL1 Blue" un choc thermique: le mélange de transformation (200 µl de bactéries avec 2 à 5 µl de plasmides+insert) est incubé 30 min à 0°C, puis un choc thermique est effectué à 42°C pendant 50 sec qui va perméabiliser la paroi des bactéries; une incubation de 5 min à 0°C est ensuite réalisée et 1 ml de milieu de culture Luria Broth (Gibco BRL) est ajouté. L'ensemble est incubé pendant 45 min à 37°C puis 100 à 200 µl sont étalés sur une boîte de Pétri (50 mg/l d'Ampicilline, 20 µl d'IPTG (0,2 M) et 20 µl de X-gal 10%). Les bactéries sont rendues compétentes par un traitement de 20 min à 0°C dans un milieu FB (voir annexe 3), centrifugées à 5000 rpm pendant 10 min à 4°C, reprises dans du milieu FB, et stockées à -80°C. Par la suite, ce protocole a été modifié en fonction des indications données par Invitrogen. Ainsi, après le choc thermique, les bactéries sont reprises dans 450 µl de milieu SOC (Invitrogen), incubées 1 h à 37°C en agitation constante, puis seulement 10 à 50 µl de culture de bactéries sont étalés sur boîte de Pétri [50 mg/l d'Ampicilline, 40 µl d'IPTG (100 mM) et 40 µl de X-gal (40 mg/ml)].

2°) Les bactéries sont transformées par électroporation. L'électroporation consiste à faire subir à la solution de transformation (40 µl bactéries+plasmides) un courant électrique de 1700 V à travers les 10 mm d'épaisseur d'une cuve à électroporation (17000 V/cm). Les bactéries "XL1 Blue" sont rendues compétentes par différents lavages dans un milieu à faible force ionique (eau déionisée) (Sambrook *et al*., 1989). Les bactéries transformées sont mélangées à 960 µl de Luria Broth, incubées pendant 1 h à 37°C puis 50 à 150 µl sont étalés sur une boîte de Pétri (Ampicilline 50 mg/l, 20 µl d'IPTG 0,2M, 20 µl de X-gal 10 %).

Criblage des bactéries

Un certain nombre de clones bactériens présents sur boîte de Pétri sont repiqués de façon à obtenir une colonie de taille plus importante (au moins 5 mm de diamètre) :



La moitié de cette colonie est récupérée à l'aide d'un cône de micropipette et placée dans un microtube dans 10 µl d'eau distillée. Cette solution bactérienne est mise à bouillir à 99°C dans un thermocycleur pendant 30 min. 1 µl de cette solution est utilisé dans une réaction de PCR utilisant un couple d'amorces "universelles" (amorces qui s'hybrident à des sites présents sur l'ADN plasmidique de part et d'autres du site de clonage). Ainsi, après migration des produits de PCR, il est possible d'identifier les colonies bactériennes possédant un fragment d'ADN de taille compatible avec le fragment cloné au départ. A partir du reste des colonies présentes sur la boîte de repiquage (voir ci-dessus), les clones identifiés sont mis en culture pour une nuit dans 3 à 4 ml de milieu Luria Broth liquide avec 50 mg/l d'ampicilline pendant 12 h.

Extraction de l'ADN plasmidique des bactéries

L'extraction de l'ADN plasmidique des bactéries a été faite par Minipreps ("High Pure Plasmid Isolation Kit": Boehringer). Les bactéries cultivées pendant la nuit dans le milieu LB sont centrifugées, resuspendues dans du tampon TE (pH 8) favorisant les étapes suivantes et stabilisant l'ADN, lysées et centrifugées à 13 000 rpm pendant 10 min. Le surnageant contenant l'ADN est purifié sur filtre-tube "High Pure" (Boehringer), puis l'ADN plasmidique est élué dans 50 à 100 µl d'eau distillée ou dans du tampon TE (pH 7).

Par la suite, j'ai utilisé un protocole de Minipreps plus "classique" ne nécessitant aucun "kit". 1,5 ml de bactéries sont centrifugés à 15300 rpm pendant 5 min. Le milieu de culture est éliminé et le culot est placé en présence de 1 ml de TE à 4°C favorisant les étapes suivantes et stabilisant l'ADN (le culot n'est pas repris à ce stade). Le tube est centrifugé à nouveau (15300 rpm pendant 5 min) puis le surnageant est éliminé et le culot repris dans 100 µl de TE à 4°C. Le tube est placé sur de la glace pendant 5 min. 200 µl de NaOH 0,2 M / SDS 1% sont ensuite ajoutés pour lyser les bactéries puis les tubes sont mélangés 5 à 6 fois par simple inversion. Les tubes sont ensuite placés sur de la glace pendant 5 min. 150 µl d'acétate de potassium 3M sont ensuite ajoutés pour précipiter les protéines et les tubes sont placés sur la glace pendant 5 min (après agitation par inversion). Les tubes sont centrifugés 5 min à 15300 rpm pour culotter les protéines précipitées. Le surnageant est incubé à 37°C pendant 30 min avec 1 µl de RNase A à 20 µg/ml de façon à éliminer les ARN. Le mélange est purifié à l'aide de 450 µl de phénol/chloroforme/alcool iso-amylique. Après agitation énergique, les tubes sont centrifugés 10 min à 15300 rpm. La phase aqueuse (supérieure) est récupérée puis l'opération est renouvelée avec, cette fois, 450 µl de chloroforme. Enfin, l'ADN plasmidique est repris avec 1 ml d'éthanol froid, puis placé 1 heure à –20° C. Les tubes sont centrifugés 15 min à 15300 rpm et le culot lavé dans de l'alcool 70 (permettant une réhydratation partielle de l'ADN). Les tubes sont à nouveau centrifugés 15 min à 15300 rpm et les culots sont ensuite séchés à l'air libre pour éliminer toute trace d'alcool, avant d'être repris dans 40 µl d'H2O distillée.

J'ai également réalisé des Midipreps pour obtenir une plus grande quantité d'ADN (Kit Hybaid). Les volumes de culture bactérienne sont de 50 à 200 ml. Le protocole utilise les mêmes étapes que le "kit" Bohringer; l'ADN est filtré sur colonne "quick Flow" (Hybaid) et les plasmides sont élués dans un tampon NaCl 1,25 M, Tris 100 mM / HCl pH= 8,5.

Par la suite, si la quantité d'ADN plasmidique est suffisante, les plasmides sont utilisés directement dans des réactions de séquençage pour déterminer la séquence nucléotidique du fragment d'ADN cloné. Sinon, des réactions de PCR sont réalisées en utilisant des amorces "universelles" (M13 *Forward* et M13 *Reverse*) de façon à amplifier la région du plasmide contenant le fragment d'ADN cloné et les produits de PCR obtenus sont utilisés dans des réactions de séquençage après purification (voir 2.7).

Le séquençage de l'ADN a été réalisé dans un premier temps par le service commun de bio-systématique de l'Université Paris 6 (responsable, M. le Professeur Jean Deutch) puis par la société Génome Express (S. A.).

Histologie après inclusion en épon

Fixation des tissus et préparation à l'inclusion

Après dissection des mâchoires, les tissus sont fixés 2 h dans 10 ml d'une solution contenant (pour 100 ml) 50 ml de cacodylate de sodium 0,2 M (4°C) (voir annexe 3), 6 ml de glutaraldéhyde à 25%, 10 ml de paraformaldéhyde à 15 (voir annexe 3), complétée à 100 ml

avec de l'eau distillée. Les tissus sont ensuite déminéralisés pendant 1 semaine dans 10 ml de fixateur contenant de l'EDTA (3,72 g pour 100 ml).

Les tissus sont rincés 2 x 15 min dans le tampon cacodylate 0,1 M/sucrose (25 ml de cacodylate 0,2 M + 5 g de saccharose, complétés à 50 ml avec de l'eau distillée), puis placés dans un mélange tétroxyde d'osmium/cacodylate pendant 1 h (3 ml d'OsO_4 à 4%, 6 ml de cacodylate 0,2 M, 1 g de saccharose, complétés à 12 ml avec de l'eau distillée).

Les tissus sont alors déshydratés avec des concentrations croissantes d'éthanol (50% pendant 10 min, 70% pendant 10 min, 95% pendant 2x15 min et 100% pendant 3x1 h).

Inclusion dans l'épon

Les tissus ont ensuite été inclus dans une résine particulière, l'épon. Cette résine est préparée en mélangeant 1/3 d'une solution A et 2/3 d'une solution B. La solution A est composée de 12 ml d'Epikote 812 et de 20 ml de DDSA qui sont agités rapidement pendant 2 h à l'aide d'un agitateur magnétique. La solution B est composée de 35 ml d'Epikote 812 et de 31 ml de MNA qui sont agités dans les mêmes conditions. Le mélange des deux solutions se fait en laissant s'écouler A dans B pendant 10 sec, puis en agitant (agitateur magnétique) pendant 20 sec avant d'ajouter un accélérateur, le DMP30 (0,15 ml pour 10 ml de mélange) et laisser agiter pendant 30 sec.
Les morceaux de mâchoire sont ensuite placés dans un premier bain d'oxyde de propylène et d'alcool absolu (50/50) pendant 15 min, puis dans de l'oxyde de propylène pur pendant 2x15 min et enfin dans un mélange 50/50 d'oxyde de propylène et d'épon, pendant 1 à 3 jours. Les tissus sont alors imprégnés dans l'épon pur avec le catalyseur dans une étuve à vide pendant 7 h à une température de 37 à 40°C.
Les tissus sont alors inclus en les plaçant dans un moule rempli d'épon dans une étuve à 60°C pendant 48 h.

Coupes et coloration

Des coupes longitudinales et transversales ont été réalisées dans les mâchoires de lézard grâce à un ultramicrotome (Reichert Austria OMU3 - Leica) qui permet de réaliser des coupes semi-fines (1 à 6 μm) ou ultrafine (85 nm).Les coupes semi-fines sont placées sur des lames de verres et colorées dans 2 ml d'un mélange de bleu de toluidine 1% et de Borate de sodium 1%, pendant 40 sec. Puis les lames sont rincées à deux reprises dans de l'eau distillée, puis séchées sur une platine chauffante à 70°C. Les lames sont ensuite recouvertes d'une lamelle sur laquelle de l'épon est déposé. Le montage alors est chauffé à 60°C pendant 48h.

Etude de l'expression du gène de l'amélogénine chez Chalcides viridanus par "hybridation sur coupe".

Ce travail a été réalisé en grande partie dans le laboratoire "Développement des Tissus Dentaires - EA1892, Directeur Henry Magloire, Faculté d'Odontologie de Lyon".

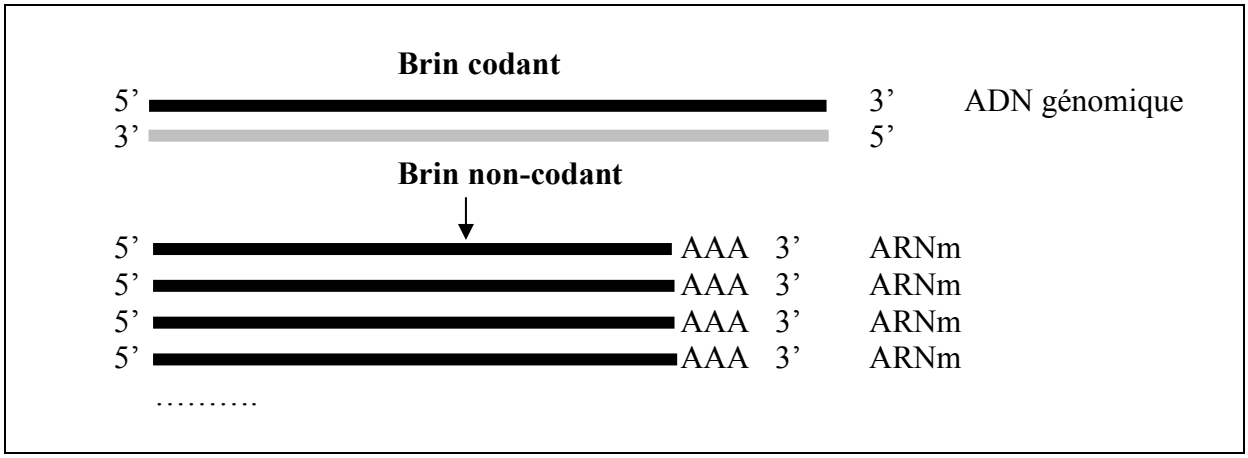
Cette étude s'est déroulée en deux étapes. Une première étape, réalisée à Paris, a consisté à fixer les tissus et à les inclure dans des blocs de paraffine. Les coupes et l'hybridation avec une sonde radioactive ont été réalisées à Lyon.

Au cours de chaque étape il a fallu travailler dans des conditions de manipulation strictes pour éviter tout contact avec des RNAses. Les instruments métalliques ont donc été chauffés à 200°C pendant 2 h, la verrerie a été lavée avec un produit décontaminant "RNase Away" (Molecular BioProducts) et rincée avec de l'eau distillée sans RNase. Les produits, poudres et tampons utilisés ont été achetés avec la mention "RNase Free".

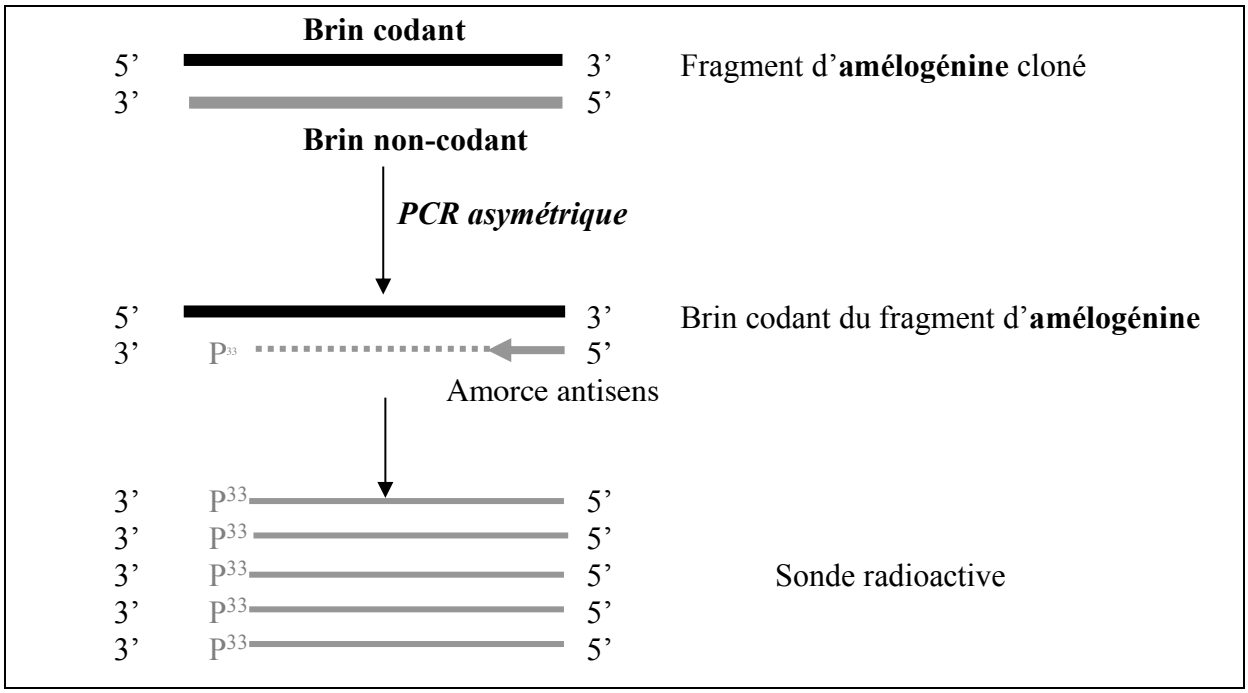
Principe de la technique

Pour réaliser une hybridation sur coupe, principe illustré dans les schémas suivants, il faut tout d'abord fabriquer une sonde d'ADN d'amélogénine en incorporant des nucléotides radioactifs (dCTP^{33}) à des brins d'ADN synthétisés au cours d'une PCR. La sonde radioactive doit-être capable de s'hybrider aux ARNm d'amélogénine qui pourraient être présents dans les cellules, à condition que les tissus aient été correctement fixés et que le gène de l'amélogénine s'y exprime. La sonde que j'ai fabriquée à partir d'un fragment d'ADN de 600 pb devait donc correspondre au brin non-codant du gène de l'amélogénine de *Chalcides viridanus* pour être complémentaire du brin codant les ARNm. Une sonde constituée du brin codant, incapable de se fixer aux ARNm, a servi de témoin négatif.

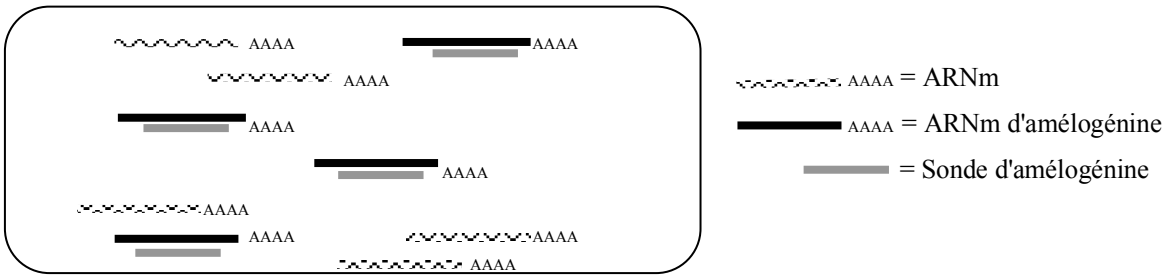
1°/ Schéma général de la transcription :



2°/ Fabrication, par PCR asymétrique, de brins non-codants radioactifs qui vont pouvoir s'hybrider avec les ARNm



3°/ Hybridation sur coupe :



Au cours de l'hybridation, la sonde radioactive (en gris) se fixe sur les brins codant des ARNm d'amélogénine présents sur la coupe (traits noirs épais). Les ARNm autres que l'amélogénine (trait ondulé) ne sont pas reconnus par la sonde.

Pour fabriquer la sonde il fallait donc disposer d'un fragment d'ADN d'amélogénine purifié (de 300 à 600 pb) et des amorces 3' et 5' bordant le fragment, amorces qui vont marquer ce fragment avec du P^{33} par PCR asymétrique (on fabrique ainsi le brin 3' = antisens et le brin 5' = sens qui nous sert de témoin), dans ce cas il faut 25 ng d'ADN par sonde.

Inclusion des tissus de Chalcides viridanus

Dans un premier temps, deux *Chalcides viridanus* âgés d'1 an ont été sacrifiés en juillet 2000 (dans une période où leur activité métabolique est maximale) et leurs mâchoires (supérieure et inférieure) ont été immédiatement disséquées. Chaque mandibule et maxillaire ont été sectionnées au niveau de la symphyse de manière à obtenir deux demi-mâchoires. Les demi-mâchoires de l'un des deux *Chalcides* ont été fixées pendant 48 h à 4°C dans un tampon phosphate 0,1 M (pH 7,4) contenant du paraformaldéhyde 4%, tandis que celles de l'autre lézard ont été fixées dans du Formoy constitué d'acide acétique à 10%, d'alcool absolu à 30% et de paraformaldéhyde à 37% (le paraformaldéhyde fixe les tissus et l'acide acétique déminéralise les dents et les os qui sont très difficiles à couper en histologie).

Les tissus ont ensuite été rincés 3 fois à 4°C pendant 1 h avec du tampon PBS. Les mâchoires ont été déshydratées successivement dans l'éthanol 50%, 70%, 95% et 100% pendant 2 x 30 min à chaque fois, puis placées pendant 12 h minimum dans du xylène. Chaque demi-mâchoire a été ensuite imprégnée dans la paraffine (Sherwood médical) pendant 3 x 1 h, dans un four à 62°C (le point de fusion de la paraffine est compris entre 55 et 57°C, il faut donc dépasser ces températures sans aller au delà de 62°C, condition qui peut détruire les polymères plastiques). Cette phase est importante pour imprégner les tissus de paraffine avant les inclusions.

Chaque demi-mâchoire a été placée dans un montage constitué de deux pièces métalliques, dans l'orientation voulue de manière à pouvoir couper les mâchoires soit transversalement soit longitudinalement. Une pince métallique est chauffée sur la flamme d'une lampe à alcool, puis elle est plongée dans le bain de paraffine et passée autour de la demi-mâchoire pour enlever les bulles d'air prisonnières de la paraffine et pour réorienter la pièce de tissu si cela est nécessaire. Lorsque la paraffine est solidifiée, le bloc ainsi formé est conservé au réfrigérateur jusqu'au moment où il sera utilisé pour réaliser des coupes histologiques.

Cette étude préliminaire m'a permis de déterminer le temps de déminéralisation au modèle *Chalcides* et de vérifier si les temps d'imprégnation dans la paraffine étaient suffisants. Il est apparu, suite à la première série de coupes histologique, que les mâchoires

étaient très insuffisamment déminéralisées. En effet, même pour les tissus qui ont été traités pendant 48 h avec du Formoy (qui contient de l'acide acétique), les blocs de paraffine étaient très difficiles à couper et les coupes de mâchoires obtenues n'étaient pas observables tant les tissus étaient abîmés. De plus, j'ai constaté, en raison de la présence de bulles d'air dans les tissus, que l'imprégnation dans la paraffine n'avait pas été suffisante.

J'ai donc modifié le protocole de manière à augmenter le temps de déminéralisation. J'ai sacrifié un nouveau *Chalcides viridanus* âgé d'un an et j'ai plongé les 4 demi-mâchoires dans un bain de Formoy pendant 48 h à 4°C, puis dans une solution de paraformaldéhyde à 4% et d'acide acétique à 1 M, pendant 24 h à 4 °C. Pour achever la déminéralisation, 2 des 4 demi-mâchoires ont été plongées dans un bain d'acide acétique 1 M pendant 24 h à 4°C tandis que les deux autres sont restées dans ce bain plus longtemps (48 et 72 h respectivement). Le reste du protocole est resté inchangé à l'exception de l'imprégnation dans la paraffine qui s'est faite pendant 2 x 24 h au lieu de 3 x 1 h.

J'ai réalisé le même type de protocole avec un échantillon de peau prélevée sur le dos de l'animal, depuis la région de la ceinture pectorale jusqu'à la ceinture pelvienne (endroit où se trouvent les ostéodermes). Le morceau de peau a été fixé et déminéralisé de la même manière que les mâchoires avec un dernier bain dans l'acide acétique 1 M de 24 h, puis il a été inclus dans un bloc de paraffine et conservé à 4°C.

Coupes histologiques

Les blocs de paraffine contenant les mâchoires et les morceaux de peau de *Chalcides viridanus* ont été placés (à Lyon) à 4°C pour augmenter la cohésion de la paraffine et ainsi faciliter les coupes. Les blocs de paraffine ont été ensuite sectionnés à l'aide d'un microtome (Reichert Austria OMU3 - Leica).

Sur chaque bloc de paraffine, une des faces a d'abord été taillée en forme de pyramide avec une lame de scalpel, il s'agit de la face choisie pour être attaquée par le microtome

J'ai choisi de sectionner 2 demi-mâchoires transversalement et les deux autres longitudinalement. Les coupes longitudinales ont débuté par la face interne (linguale) de la mâchoire et les coupes transversales par la partie antérieure. Les morceaux de peau, de forme rectangulaire, ont été coupés perpendiculairement à la surface. J'ai ensuite fixé les blocs taillés sur le microtome, parallèlement au fil du couteau. L'appareil a été réglé pour réaliser des coupes de 7 à 8 µm d'épaisseur. Les coupes obtenues sont posées délicatement sur une goutte d'eau "RNase free" que l'on a placée au 3/4 du bord interne d'une lame en verre silanisée. La silanisation consiste à placer une lame "Porte-Objets" normale ou SuperFrost® Plus (Menzel-Glaser), 2 min dans un bain contenant du 3-aminoproxyltriethoxysilane (Sigma) dilué à 2% avec de l'acétone, puis 2 x 2 min dans un bain d'eau ultra pure avant de les sécher sur un papier absorbant propre (sans RNases) pendant une nuit (les lames sont recouvertes avec du papier absorbant pour éviter les poussières).

La lame est ensuite placée quelques secondes sur une plaque chauffante pour permettre un étalement maximal des coupes de paraffine. La lame est retirée de la plaque chauffante avant que la paraffine fonde et l'eau est retirée par capillarité à l'aide de papier absorbant propre. La coupe ainsi obtenue est mise à sécher à 37°C pendant une nuit.

Les lames séchées sont ensuite préparées pour l'hybridation. On commence par les placer 3 x 7 min dans du xylène pour dissoudre la paraffine. Les coupes sont ensuite réhydratées dans des bains successifs d'éthanol de moins en moins concentrés (éthanol 100% pendant 3 min, puis éthanol 95% pendant 5 min, puis éthanol 70% pendant 5 min, puis éthanol 50% pendant 5 min). Les lames sont ensuite rincées deux fois rapidement avec de l'eau DEPC puis 2 x 5 min dans du tampon PBS (Phosphate-buffered saline). Les tissus sont fixés à nouveau en plongeant les lames 15 min dans un bain de paraformaldéhyde 4% dilué

avec du PBS puis ils sont à nouveau déshydratés avec des bains successifs dans de l'éthanol de plus en plus concentré (éthanol 50% pendant 5 min, puis éthanol 70% pendant 5 min, puis éthanol 95% pendant 5 min et enfin éthanol 100% pendant 3 min). Les lames sont séchées pendant quelques minutes à l'air libre (à l'abri de la poussière) puis traitées avec la protéinase K à 1 µg/ml (dans un tampon Tris 50 mM et EDTA 5 mM) pendant 15 min à 37°C, ce qui permet de détruire les protéines des tissus pour permettre à la sonde d'accéder plus facilement aux ARNm. Les lames sont ensuite rincées avec du PBS, d'abord rapidement puis pendant 10 min. Les tissus sont à nouveau fixés avec du paraformaldéhyde 4% (dilué dans du PBS) pendant 15 à 20 min puis les lames sont rincées 2 x 10 min dans du tampon PBS et déshydratées de la même manière que précédemment. Enfin, les lames sont séchées à l'air, à l'abri de la poussière, jusqu'à la phase de pré-hybridation.

Fabrication de la sonde

Pour obtenir un fragment d'ADN d'amélogénine qui servira de matrice pour la fabrication de la sonde radioactive, j'ai tout d'abord amplifié par PCR 0,5 µl de plasmide contenant un fragment d'ADNc d'amélogénine de *Chalcides viridanus* que j'avais cloné et fait séquencer dans une manipulation précédente. J'ai utilisé 2 amorces situées, l'une dans l'exon 2 (amorce sens = Amel-2.2) et l'autre à la fin de l'exon 6 (amorce antisens = Amel-529) (Figure 2A). J'ai obtenu un fragment d'amélogénine d'environ 600 pb qui correspond à près de 90% de la séquence (Figure 2B).

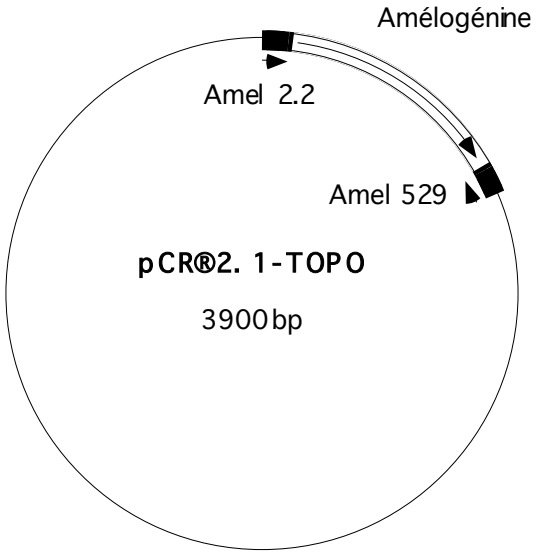


Figure 2A. Schéma de la PCR réalisée sur un plasmide, extrait précédemment d'un clone bactérien, dans le but d'obtenir un fragment d'amélogénine qui servira de matrice pour la fabrication de la sonde radioactive.

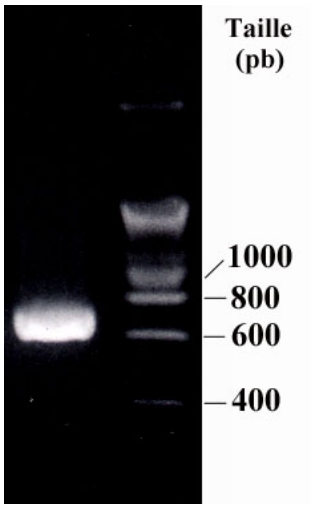


Figure 2B. Migration sur gel d'agarose à 1,8% du fragment d'amélogénine de 600 pb environ utilisé pour la fabrication de la sonde.

La PCR a été réalisée dans les conditions suivantes : 35 cycles avec une température d'hybridation de 58°C et un volume final de 50 µl. Le fragment d'ADN a été purifié sur colonne avec le système QIAquick (Qiagen).

La fabrication de la sonde nécessitant la manipulation d'un produit radioactif, le P^{33}, j'ai dû travailler dans le respect le plus strict des règles de sécurité : la radioactivité a été manipulée derrière des écrans de protection dans une pièce réservée exclusivement à cet usage

et d'où rien ne devait sortir. Le matériel consommable (cônes, gants, tubes, etc.) a été jeté dans des poubelles réservées à cet usage. Après chaque manipulation la radioactivité a été contrôlée au niveau des vêtements et des mains à l'aide d'un compteur Geiger.

J'ai préparé trois tubes pour des réactions de PCR, tubes qui ont été préalablement siliconés en y plaçant 200 µl de diméthylchlorosilane pendant quelques secondes. Deux tubes contenaient l'amorce antisens ("Amel 529") servant à fabriquer la sonde et un tube contenait l'amorce sens ("Amel 2.2") pour fabriquer une sonde qui servira de témoin négatif. En effet, cette sonde, constituée d'un brin codant, est incapable de se fixer aux ARNm et le marquage observé dans ces conditions sera un signal non-spécifique qu'il est nécessaire de connaître pour distinguer ce signal de celui provenant de l'expression de l'amélogénine. Les tubes contenaient (dans l'ordre) :

1 µl d'ADN purifié (amélogénine) à 25 ng/µl, 5 µl de tampon de polymérase, 3 µl de dATP, dGTP, dTTP à 1mM chacun, 1,5µl d'amorce à 25 µM ("Amel 529" ou "Amel 2.2"), 8,5 µl de dCTP "froid" à 10 µM, 25 µl d'eau, 5 µl de $dCTP^{33}$ radioactif, 1 µl de Taq polymérase,

Les tubes ont été placés dans un thermocycleur pour 10 cycles de PCR à une température d'hybridation de 55°C puis la réaction a été arrêtée en ajoutant 5 µl d'EDTA à 0,2 M.

Pour déterminer l'activité spécifique de la sonde, on prélève 1 µl de chaque PCR que l'on dilue 100 fois dans de l'eau distillée. 1 µl de cette solution est utilisé pour mesurer directement la radioactivité présente dans la PCR, celle qui provient à la fois de la sonde et des dNTP non-incorporés. 1 µl supplémentaire est utilisé dans une réaction appelée "TCA précipitable" qui permet de mesurer uniquement l'activité des sondes. Pour réaliser cette réaction, on commence par mélanger, à l'aide d'un vortex, 200 µl d'eau distillée et 2 µl de "DNA carrier" (10 µg/µl). L'échantillon de sonde dilué 10 fois (1µl) est ensuite ajouté et le mélange passé au vortex, puis on place dans le tube de réaction 400 µl de TCA 10 % froid avant de le passer au vortex et de laisser le tout dans la glace pendant 10 à 15 min. Grâce à une pompe à vide, le contenu de chaque tube est placé sur un filtre destiné au comptage de la radioactivité. Les filtres sont rincés avec du TCA à 10% et sont ensuite séchés. Chaque filtre est placé dans un flacon de comptage puis on ajoute 5 ml de liquide scintillant avant de mettre le tout dans un compteur de radioactivité. J'ai obtenu les valeurs de radioactivité suivantes (Tableau 1) :

Tableau 1. Activité spécifique de chacune des 3 sondes utilisées. Remarque : L'activité des 2 sondes antisens est différente car, lors de la réaction de PCR, il a manqué de $dCTP^{33}$ radioactif pour l'un des tubes (Antisens n°2). J'ai utilisé une quantité inférieure aux 5 µl de $dCTP^{33}$ requis, ce qui a diminué la quantité de radioactivité incorporée.

Sonde	Radioactivité du TCA précipitable	**Radioactivité calculée de la sonde**
Antisens n°1 (Amel-529)	2690,57 cpm	15.10^6 cpm
Antisens n°2 (Amel-529)	12228,64 cpm	7.10^6 cpm
Sens / témoin négatif (Amel-2.2)	3733,26 cpm	20.10^6 cpm

Pour utiliser les sondes en hybridation, il faut d'abord les précipiter. J'ai donc mélangé 50 µl de sonde, 25 µl d'ARNt (Sigma, 10mg/ml) et 2,5 volumes d'éthanol absolu, puis j'ai refroidi les tubes au moins 15 min à -80°C (il est aussi possible de laisser les tubes au moins 2 h à –20°C). La suite de l'expérience doit se faire au moment de la pré-hybridation des lames. A ce moment là, les sondes sont centrifugées à 14000 g pendant 20 min à 4°C, puis le culot est rincé avec de l'éthanol 70% avant d'être centrifugé à nouveau dans les mêmes conditions. L'alcool est jeté et le culot est séché à l'air libre à l'abri de la poussière puis conservé dans la glace.

Pré-hybridation des coupes

Cette phase va permettre de préparer les coupes pour l'hybridation et de saturer les sites non-spécifiques, notamment le collagène, grâce à la présence d'ADN de sperme de hareng dénaturé et natif (Sigma) (pour dénaturer l'ADN, chauffer 10 min à 100°C puis mettre dans la glace).

Auparavant, il faut préparer 1 ml de tampon $NaPO_4$ 2 M (pH 7,4) en ajoutant 685 µl de Na_2HPO_4 2 M à 315 µl de NaH_2PO_4 2 M, puis en chauffant à 60°C. Il faut également désioniser la formamide en utilisant des billes échangeuses d'ions (Sigma, "Mixed Bed Resin") dont 2 à 3 ml ont été placés dans un tube contenant 4 à 5 ml de formamide avant que le tube soit agité et placé à 4°C. Cette dernière étape doit être faite au moins 1 h avant l'utilisation de la formamide.

Il est alors possible de préparer le tampon de pré-hybridation qui va être déposé sur chacune des coupes (voir la composition en annexe 3).

50 µl du tampon d'hybridation sont déposés sur chaque lame à l'endroit où se trouvent les coupes (2 à 3 par lame), puis celle-ci est recouverte d'un morceau de parafilm. Les lames sont ensuite placées dans une boîte de lames humidifiée grâce à du papier absorbant imbibé de 5 ml d'eau distillée et fermée hermétique avec une bande autocollante. La boîte est alors placée à 37°C pendant au moins 2 h, puis les lames sont rincées rapidement avec du tampon SSC 2x (voir annexe 3) et déshydratées dans des bains successifs de 3 min d'éthanol de plus en plus concentré (50%, 70%, 95%, 100%). Les coupes sont ensuite séchées à l'air libre, à l'abri de la poussière.

Hybridation des coupes

Il faut d'abord calculer le nombre de coupes à hybrider pour déterminer la quantité de tampon d'hybridation à préparer. Pour chacune des trois sondes, j'ai utilisé 11 à 12 lames, chacune devant être traitée avec 30 à 35 µl de tampon. Le volume de tampon nécessaire dans ces conditions était de 1500 µl pour les 3 sondes. Comme l'ADN natif devait être ajouté plus tard, ce sont en réalité 1492,5 µl que j'ai préparés.

450 µl de ce mélange sont ajoutés dans les 3 tubes contenant le culot de sonde radioactive (voir *Fabrication de la sonde*), puis chaque tube est chauffé dans l'eau bouillante pendant 10 min pour dénaturer l'ADN, avant d'être placé dans la glace. L'ADN natif est alors ajouté (2,5 µl à 10 µg/µl) et les tubes sont alors mélangés dans un vortex et placés à nouveau dans la glace.

Sur chaque lame, 35 µl de sonde radioactive sont déposés puis la lame est recouverte d'une lamelle souple de 22 mm x 22 mm ("hybri-slips", Sigma). Le dextran sulfate, très visqueux, contenu dans le tampon va concentrer les molécules radioactives au niveau des coupes et le liquide ne peux donc pas s'étaler. Les lames sont ensuite placées dans une chambre humide (boîte contenant un linge humide et fermé hermétiquement), de la même

manière que pour la pré-hybridation, à 37°C pendant une nuit. Les sondes utilisées en fonction des coupes histologiques réalisées sont présentées dans le tableau 2.

Tableau 2. Nombre de lames hybridées avec les 3 sondes utilisées. La sonde Amel-2.2 est un témoin négatif qui ne peut pas s'hybrider à l'amélogénine.

	Amel-529 n°1	**Amel-529 n°2**	**Amel-2.2**
Coupe longitudinale de mâchoire (série n°1)	4	4	3
Coupe longitudinale de mâchoire (série n°2)	2	2	1
Coupe transversale de mâchoire	3	3	2
Coupe de peau	2	2	2

Une fois l'hybridation terminée, les lamelles souples qui recouvrent les coupes sont retirées. Dans ce but, quelques gouttes d'un mélange composé de SSC 2 X et de formamide à 50% sont versées sur un des cotés de chaque lamelle qui se décolle alors, ce qui permet de la saisir avec une pince. Les coupes sont alors rincées à 37°C dans un cristallisoir plongé dans un bain-marie d'eau DEPC en agitation constante, dans les solutions suivantes :

30 min dans du SSC 2 X / formamide 50%; 30 min dans du SSC 1 X / formamide 50%; 30 min dans du SSC 1 X; 30 min dans du SSC 0,5 X; 30 min dans du SSC 0,25 X (nécessaire dans le cas de sondes homologues).

Lorsque le rinçage est terminé, les coupes sont déshydratées dans des bains de 5 min d'éthanol en concentration croissante (50%, 75%, 95%, 100%), puis elles sont séchées dans une boîte à l'abri de la poussière.

Révélation des coupes

La radioactivité est révélée dans une "chambre noire" équipée d'une lampe avec un filtre rouge spécial (Kodak Wratten n°2) pour produire une lumière inactinique. J'ai utilisé une émulsion photographique IL-4 (Ilford). 10 ml d'émulsion et 5 ml d'eau sont mélangés et placés dans un bain-marie à 43°C à l'intérieur d'un récipient spécial destiné à y plonger les lames histologiques. Une première lame sans coupe est trempée deux fois pour mélanger l'eau et l'émulsion, puis chaque lame portant les coupes est trempée pendant 5 sec. Les lames sont ensuite égouttées, d'abord à l'air libre pendant 5 sec, puis en les plaçant verticalement sur un papier absorbant pendant 5 sec. Elles sont finalement placées dans une boîte à lame dans une position inclinée pendant 1h30 avant d'essuyer le dessous et de les ranger dans une boîte opaque avec des silicagels qui vont assécher le milieu. La boîte est alors placée à 4 °C jusqu'à la révélation.

Les lames sont révélées dans une "chambre noire", 1 semaine ou 2 après, en les plaçant successivement pendant 4 min dans le révélateur (D19), puis dans un bain de rinçage contenant 1 ml d'acide acétique (4 min), puis dans le fixateur (hypam-Ilford ou thiosulfate) pendant 4 min. Les coupes ont été ensuite colorées à l'Hémalin de Masson pendant 1 min, puis rincées rapidement deux fois à l'eau distillée et séchées. Les lames sont ensuite observées au photomicroscope.

Méthodes d'analyse de séquences et de phylogénie moléculaire

Les logiciels ont été utilisés sur un ordinateur MacIntosh PowerMac fonctionnant avec le système d'exploitation Mac OS 9.0.3 et muni d'un processeur PowerPC G3 de 350 MHz.

Recherche de séquences dans les banques de données

Des séquences d'amélogénine ont été recherchées sur le site Internet du NCBI (National Center for Biotechnology Information) à l'adresse suivante : http://www.ncbi.nlm.nih.gov/. J'y ai effectué des recherches de séquences nucléiques dans "Genbank" ou de séquences protéiques dans "Protein Sequence Database" avec comme mot-clef "amelogenin". Afin de vérifier si j'avais bien trouvé toutes les séquences connues d'amélogénine, j'ai effectué des recherches complémentaires au SRS ™ par l'intermédiaire de "Infobiogen" (créé par le Ministère de l'Education Nationale, de la Recherche et de la Technologie, et l'Université d'Evry Val d'Essonne, à l'adresse suivante : http://www.infobiogen.fr), qui permet de centraliser les informations de plus de 30 bases de données simultanément.

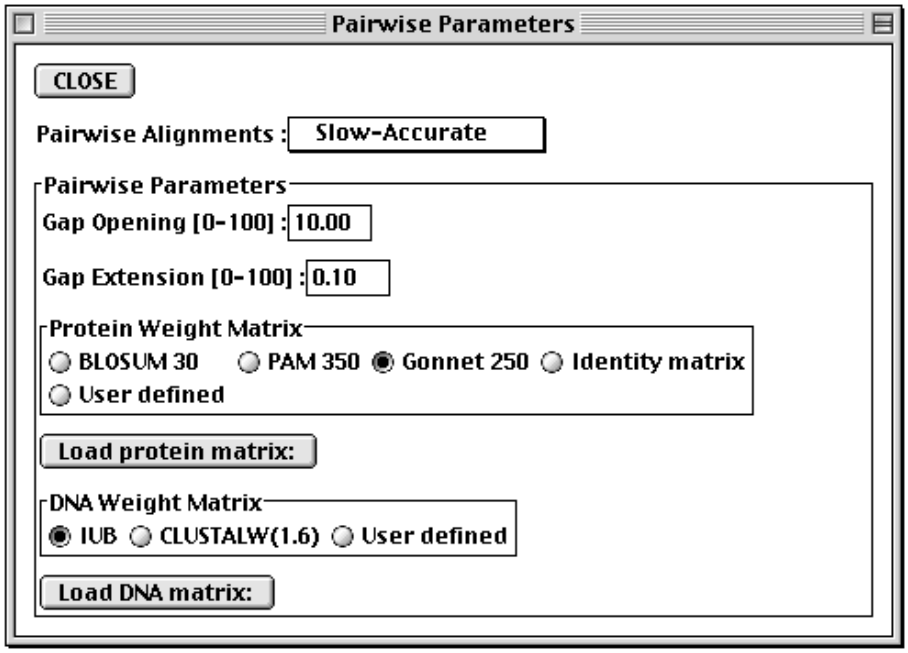
Les séquences d'amélogénine ainsi obtenues sont présentées dans un tableau en annexe 5 qui indique le numéro d'accession ainsi que la classification des taxons pour lesquels l'amélogénine a été séquencé.

DNA Strider 1.2 (© Christian Mark et C.E.A., 1991)

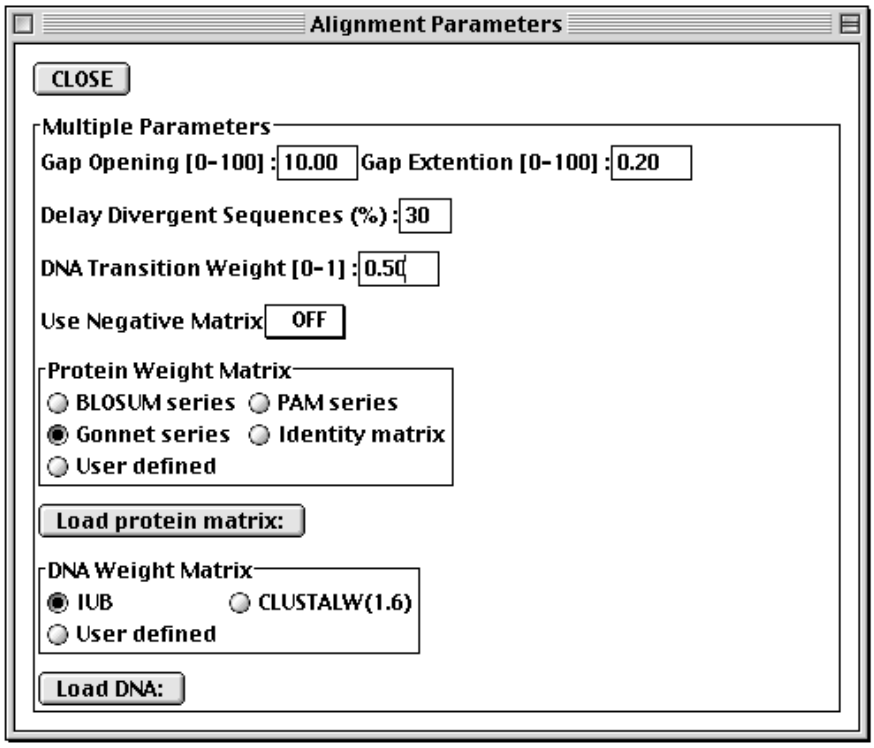
Ce logiciel de biologie moléculaire, DNA strider ©, permet de traduire les séquences d'ADN en protéines, de rechercher les phases ouvertes de lecture et de donner les cartes de restriction complète ou partielle d'une séquence d'ADN. De plus, il permet d'inverser une séquence d'ADN et/ou d'en obtenir son complémentaire. Par ailleurs, cet outil est très utile pour convertir des séquences d'ADN du format GenBank au format texte permettant ensuite de les utiliser pour réaliser des alignements de séquences par exemple.

Clustal X 1.8 (Higgins *et al.*, 1996 ; Thompson *et al.*, 1997)

Clustal est un programme d'alignement de séquences permettant de travailler avec des séquences d'ADN et des séquences protéiques. Pour aligner les séquences d'ADN, j'ai préféré choisir les réglages par défaut (Matrice IUB). En revanche, pour aligner les protéines, j'ai dû choisir entre plusieurs matrices de substitution utilisées pour pondérer le remplacement d'une lettre par une autre à une position donnée (ceci permet de prendre en compte le taux variable de mutabilité apparente d'un résidu d'acide aminé). Clustal X permet de choisir entre les matrices PAM, BLOSSUM et Gonnet :



Options disponibles pour l'alignement de séquences par paire dans Clustal X 1.8.



Options disponibles pour l'alignement de séquences multiples dans Clustal X 1.8.

La première (et plus ancienne) matrice est la série des PAM, déduites de l'observation des variabilités en AA dans des familles bien connues de protéines, avec l'observation de 1572 points de mutation dans 34 familles de protéines (Dayhoff, 1978). Une PAM 1 est la matrice qui correspond à des évènements mutationnels qui décrivent la quantité d'événements qui changeront en moyenne 1% des acides aminés. PAM 2, 3, ..., 100 et 250 sont obtenues par produit matriciel successif. Une valeur PAM petite correspond à peu d'évènements mutationnels (et sera plus adaptée pour comparer des séquences très voisines), tandis qu'une

valeur PAM élevée sera plus intéressante pour comparer des séquences plus éloignées. Une analyse statistique a montré que PAM 250 est un bon compris optimal pour la majorité des séquences, elle est aussi appelée la matrice de GONNET car elle a été recommandée par Gonnet et al. (1992). Clustal utilise les matrices PAM 350 et GONNET, pour les alignements par paire. Contrairement aux PAM, les matrices BLOSUM sont déduites des blocs de séquences alignés et regroupés par famille jusqu'à un certain seuil (BLOSUM80 correspond à des segments similaires à 80%). La matrice BLOSUM 62 est celle qui est utilisée dans BLAST par défaut. Clustal utilise quant à lui des matrices PAM 250 (Gonnet), PAM 350 et BLOSUM 30.

Etant donné la divergence existant entre les séquences que j'ai comparées, j'ai choisi la série des PAM comme matrice d'alignement multiple et la matrice PAM 350 comme option d'alignement par paire dans les paramètres principaux de Clustal. Par ailleurs, j'ai également utilisé les matrices de Gonnet pour aligner des séquences assez proches comme celles de lézards ou de mammifères.

Lorsque l'alignement multiple est effectué par Clustal, il existe alors plusieurs options que j'ai utilisées suivant la manière dont j'ai poursuivi le traitement des données.

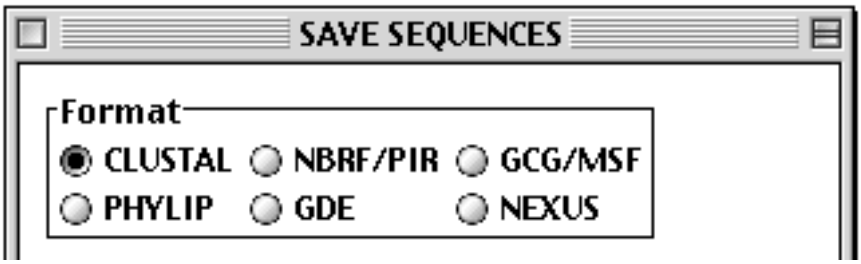
Clustal peut calculer des arbres phylogénétiques d'après la méthode phénétique mais il ne peut pas les afficher. Clustal en effet calcule des matrices de distance avec l'algorithme de *Neighbor Joining* et/ou mesure la solidité des branches par un test de *Bootstrap*. Le logiciel génère alors des fichiers ".ph" ou ".phb" qui seront décryptés ensuite par TreeView ou Njplot permettant d'afficher les arbres.

Le *Neighbor Joining* (NJ) permet, à partir d'une matrice de caractères, de calculer, à l'aide de la méthode phénétique, des arbres phylogénétiques. Le *Neighbor Joining* est la méthode de distances la plus souvent utilisée. Il existe également un algorithme plus ancien, l'UPGMA, qui impose que les distances soient ultra métriques, donc que les séquences évoluent à une vitesse constante (hypothèse d'horloge moléculaire). Étant donné que cette hypothèse est rarement confirmée, la méthode UPGMA n'est pratiquement plus utilisée dans l'étude des phylogénies. Le *Neighbor Joining* n'implique pas l'hypothèse d'horloge moléculaire. La méthode NJ consiste à calculer les longueurs des branches, telles que les distances déduites de l'arbre soient les plus proches des distances mesurées entre les séquences; et ensuite à calculer la longueur de l'arbre (la somme des longueurs de ses branches).

Deux options supplémentaires pour le calcul par *Neighbor Joining* existent dans Clustal. Il est possible, d'une part, de corriger la présence éventuelle de substitutions multiples dans le cas de séquences éloignées et, d'autre part, de ne pas inclure les insertions/délétions ("gaps") dans le calcul. Cette dernière option est importante lorsque les séquences possèdent des insertions qui peuvent fausser les calculs par accumulation d'homoplasies si elle évoluent rapidement (c'est le cas pour l'amélogénine), ou lorsque certaines séquences comparées possèdent des parties non-connues qui ne doivent pas être confondues avec des délétions.

La deuxième option, "output format option" du menu "Tree", permet de sauvegarder les données sous forme d'une matrice de distance *Phylip*. Le fichier ".dst" obtenu est alors utilisé avec BioNJ (Gascuel, 1997), un logiciel de calcul d'arbre phylogénétique par une version améliorée de l'algorithme de *Neighbor Joining*.

Enfin, comme on le voit dans la figure ci-dessous, il est possible de sauvegarder les alignements de séquences sous d'autres formats qui permettent de corriger l'alignement manuellement (avec SE-AL qui accepte le format Phylip) et de calculer des arbres avec d'autres méthodes phylogénétiques comme la cladistique (PAUP, MacClade qui acceptent le format NRBF/PIR).



BioNJ

Le logiciel BioNJ (Gascuel, 1997) calcule des arbres phylogénétiques à l'aide d'une version améliorée de l'algorithme de *Neighbor Joining*, adaptée aux distances évolutives obtenues à partir de séquences alignées. À partir d'une matrice de distance au format *Phylip* générée par Clustal, BioNJ effectue le calcul et produit un fichier permettant de visualiser l'arbre avec TreeView ou NJPlot.

PAUP 4.0 (Phylogenetic Analysis Using Parcimony and other methods)

Le logiciel PAUP 4.0 (Swofford, 1993) permet de construire des arbres phylogénétiques par les méthodes cladistiques qui utilisent le principe de parcimonie. Après avoir importé la matrice de caractères moléculaires en acides aminés ou nucléotides (voir partie précédente), il est alors possible de calculer des arbres phylogénétiques en utilisant des algorithmes heuristiques ou des algorithmes dits exacts.

Les algorithmes exacts garantissent un résultat optimal, mais augmentent les temps de calculs et nécessitent de n'avoir ni trop de taxons ni trop de caractères. La recherche des arbres dans PAUP avec les algorithmes exacts, peut se faire soit par analyse exhaustive soit par la technique du "branch and bound". « La recherche exhaustive est l'évaluation de tous les arbres possibles. Comme il y a $2x10^6$ arbres pour 9 taxons, une telle recherche n'est généralement possible qu'au-dessous de la dizaine de taxons. »… « La technique du "branch and bound" est un algorithme exact qui ne nécessite pas une recherche exhaustive. Il donne la solution optimale pour un nombre de taxons dépassant la dizaine (jusqu'à 20 à 30 taxons) selon l'implémentation de l'algorithme et la cohérence des données : plus l'homoplasie est importante, plus le temps de calcul est élevé. » (Darlu et Tassy, 1993).

Les algorithmes heuristiques permettent d'acquérir un résultat en un temps raisonnable, mais ne garantissent pas toujours d'obtenir l'arbre optimal. PAUP propose d'utiliser plusieurs algorithmes heuristiques comme le "stepwise addition" ou le "branch swapping". Alors que la première méthode est sensible à l'ordre d'introduction des taxons terminaux, la seconde méthode, dite aussi de « réarrangement des branches », tente de pallier cette difficulté en améliorant l'arbre initial par déplacement des branches. « L'expérience montre que les algorithmes de réarrangement des branches fonctionnent bien, quoique ne donnant pas toujours tous les arbres de longueur minimale. » (Darlu et Tassy, 1993).

En dehors des algorithmes précédents, la dernière version de PAUP (4.0) donne la possibilité d'utiliser la méthode dite de Quartet-Puzzling (Strimmer et von Haeseler, 1996) basée sur l'analyse de quartets de séquences. Le Quartet Puzzling utilise la méthode de maximum de vraisemblance pour reconstruire tous les quartets possibles de taxons à partir du jeu complet de données. Puis une étape de résolution ("puzzling") combine la majorité des arbres à 4 taxons calculés par maximum de vraisemblance pour construire un arbre final qui est une approximation heuristique. Au niveau des nœuds internes, il est possible de placer des

valeurs tirées de l'étape de "puzzling" et qui indiquent le nombre de fois que le nœud apparaît durant ce processus . Il a été montré que ces valeurs sont très proches des valeurs de bootstrap, au moins pour des petits jeux de données.

Dans sa dernière version, PAUP permet aussi l'analyse phénétique par les algorithmes habituels de la méthode de distance, c'est-à-dire : l'UPGMA, le Neighbor Joining et le BioNJ. Cependant, j'ai utilisé de préférence le Neighbor Joining et le BioNJ, car l'inconvénient majeur de l'UPGMA est la sensibilité de la méthode à des taux de mutations différents sur les différentes branches. L'UPGMA est d'ailleurs peu utilisé actuellement.

Par ailleurs, il a été possible de déterminer des séquences ancestrales à partir de phylogénies de référence.

TreeView1.6.6 **(**© 2000 Roderic D. M. Page)

Ce logiciel développé par Page (1996), permet d'afficher des arbres phylogénétiques à partir d'un fichier de matrice de distance résultant d'un algorithme comme le *Neighbor Joining* ou le *BioNJ*.

Il permet d'afficher les arbres calculés par Clustal X ou par BioNJ de plusieurs manières possibles : non-enraciné ou enraciné, avec les branches proportionnelles aux distances calculées (phylogramme) ou à la manière d'un cladogramme dont les branches reflètent uniquement les relations de parenté. Ce cladogramme peut être rectangulaire ou triangulaire.

Njplot (Perrière et Gouy, 1996)

Comme TreeView, ce logiciel permet d'afficher les arbres calculés par Clustal X ou BioNJ, mais dispose de moins d'options d'affichage, et ne peut afficher que des arbres dont la longueur des branches reflète les distances.

Mac Clade 4.0 (© D. R. Madisson et W. P. Madisson, 2000)

Ce logiciel, destiné prioritairement aux matrices de caractères morphologiques, permet cependant d'importer des alignements de séquences protéiques ou nucléotidiques (format Phylip) à partir desquels il est possible de déterminer des séquences consensus (majorité relative). Cette séquence consensus ne tient pas compte des "gaps" (insertions/délétions) même s'ils sont majoritaires et constituent des synapomorphies. Dans les cas où ces portions de séquences correspondent à des insertions, j'ai modifié le consensus pour inclure les gaps.

En résumé, le traitement des données moléculaires consiste en l'utilisation successive de plusieurs outils informatiques qui varient en fonction de ce que l'on veut calculer. L'ensemble est résumé dans le diagramme placé en annexe 8.

ANNEXE 2
Analyse de quelques problèmes rencontrés lors de mes travaux

Amplification du gène de l'amélogénine

Les résultats et les démarches scientifiques sont présentés ici dans un ordre chronologique pour une meilleure compréhension des problèmes rencontrés.

Au moment où j'ai commencé mes travaux, le gène de l'amélogénine n'était connu que chez quelques groupes de mammifères (tous des euthériens) assez proches les uns des autres. Il n'existait qu'une seule séquence d'amélogénine assez partielle chez les métathériens (marsupiaux). Avec des données aussi incomplètes, obtenir des résultats chez les squamates s'est avéré être une tâche difficile.

Première séquence d'amélogénine

Contrairement à ce qui était prévu (en théorie), l'amplification du gène de l'amélogénine s'est tout de suite révélée difficile car les amorces, définies à partir des séquences mammaliennes n'étaient pas assez spécifiques. Malgré ces difficultés, j'ai concentré mes efforts sur l'ADNg de Gallotia stehlini*, un lézard sur lequel j'avais obtenu, au cours de mes travaux de D.E.A, une amplification par PCR assez nette (Figure 1).*

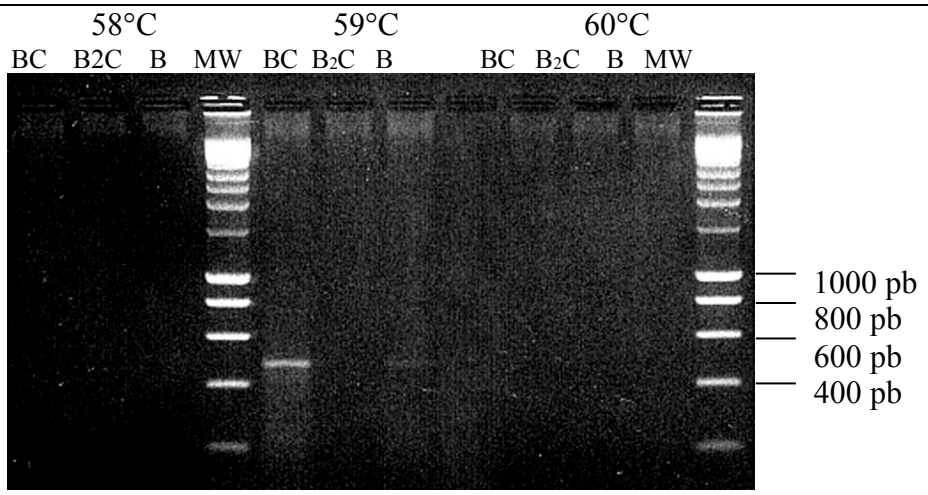


Figure 1. Amplification de l'ADN génomique de *Gallotia stehlini* à différentes température d'hybridation. L'amplification réalisée à 59°C donne une bande de 400 à 450 pb qui pouvait être de l'amélogénine (taille attendue en amplifiant l'exon 6 de l'amélogénine). Migration de 10 µl d'ADN sur gel d'agarose 1,8% avec le marqueur moléculaire (MW) Smartladder (Eurogentec).

Cependant, après une tentative de clonage, il s'est avéré que la bande d'ADN de 400-450 pb ne correspondait pas à de l'amélogénine. En contrôlant d'où venait le problème, je me suis aperçu qu'on obtenait une bande d'ADN assez faible lorsqu'une PCR était réalisée avec l'amorce *B* seule. Cette amorce était donc capable d'amplifier autre chose que l'amélogénine. J'ai supposé que de l'ADN d'amélogénine pouvait être présent, même minoritairement, dans cette amplification. C'est pourquoi j'ai utilisé un protocole sensé enrichir les produits de PCR en fragments d'amélogénine en utilisant la méthode de "nested

PCR“ qui consiste à ré-amplifier des produits de PCR avec des amorces plus internes (Figure 2).

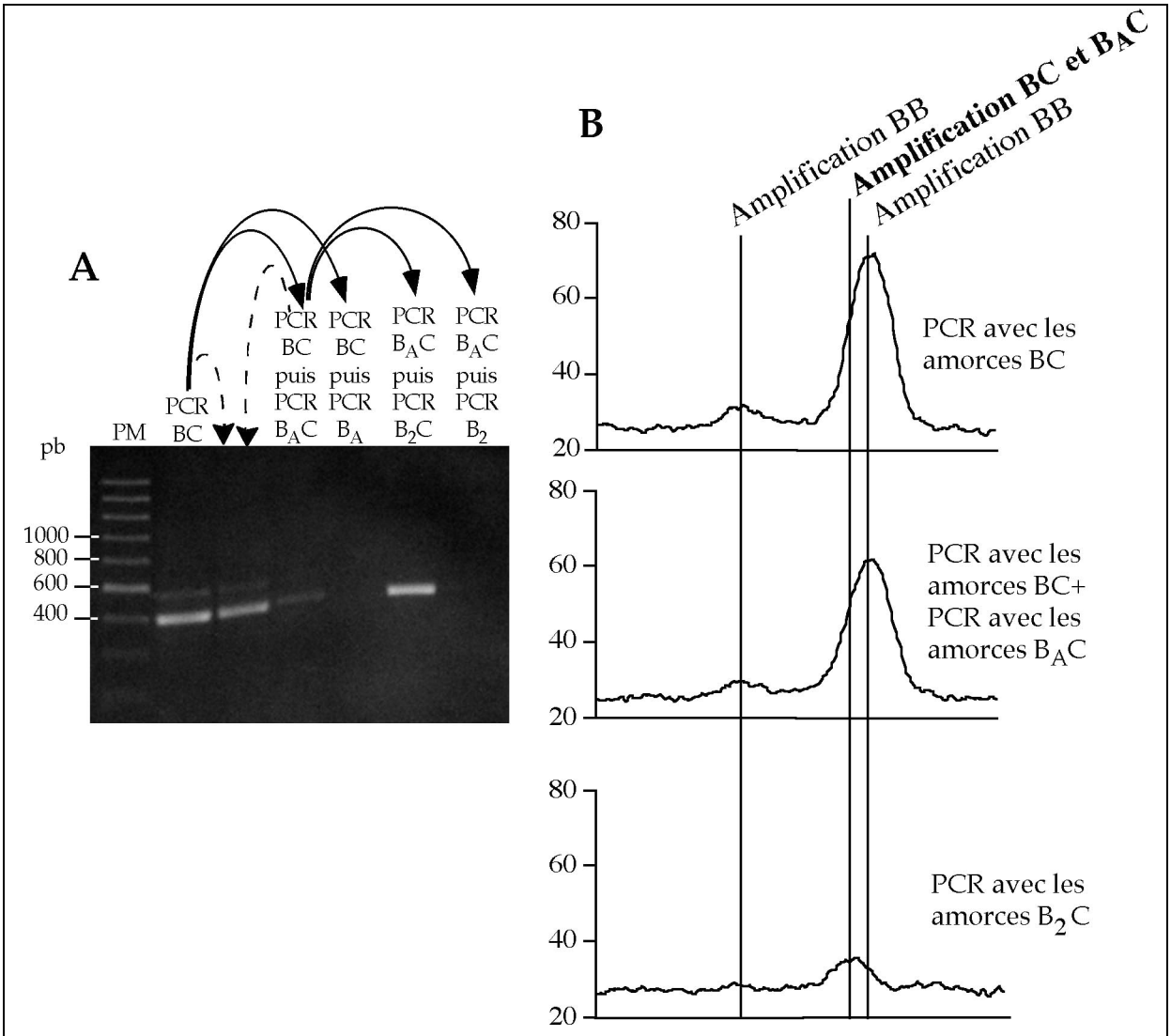


Figure 2. A. Test de trois PCR successives (température d'hybridation de 59°C) avec à chaque fois des amorces plus internes (**B**, **B$_A$**, **B$_2$**) sur l'ADN de *Gallotia stehlini*. Un contrôle **B$_A$** et **B$_2$** seul a été effectué pour les deuxièmes et troisièmes PCR afin de s'assurer que l'amplification n'était pas un artefact d'amplification de type **BB**. La migration n'a pas été très rectiligne. **B.** Profil densitométrique des puits 1, 2 et 3 montrant que les produits de PCR **BC** et **B$_A$C** se superposent à une bande artéfactuelle **BB**.

Ce protocole assez lourd m'a permis d'obtenir une séquence d'ADN qui après analyse s'est avéré être de l'amélogénine de *Gallotia stehlini*. C'est ainsi que j'ai constaté une grande divergence entre l'amélogénine de squamates et l'amélogénine de mammifères (Delgado *et al.*, 1998).

Malgré ce succès et cette première séquence du gène de l'amélogénine chez les squamates, je n'ai pas réussi à progresser avec *Chalcides sexlineatus.* En effet, la méthode de “nested PCR“ utilisée précédemment, n'a pas fonctionné avec ce scinque. De plus, la séquence de *Gallotia Stehlini* ne m'a pas permis de fabriquer des amorces plus spécifiques de l'amélogénine des squamates. En effet, les premières amorces que j'ai fabriquées étaient situées dans les régions les plus stables du gène de façon à augmenter les chances d'amplifier l'amélogénine. La séquence d'amélogénine de *Gallotia stehlini* est donc constituée de la

partie variable du gène et il n'était pas envisageable de construire de nouvelles amorces situées dans ces régions variables où une amorce aurait peu de chances de s'hybrider à l'amélogénine de *Chalcides,* même en se basant sur la séquence de *Gallotia*.

1.2. Séquence de Crocodylus niloticus

Les amorces à ma disposition étaient donc trop proches de l'amélogénine de mammifères pour réussir à amplifier facilement celle de squamates.

Cependant, deux publications m'ont permis d'avancer dans mes résultats. Fin 1998, Toyosawa *et al*. ont publié la séquence protéique de l'amélogénine de caïman (*Paleosuchus palpebrosus*) et Ishiyama *et al*. une séquence nucléotidique d'amélogénine d'un serpent (*Elaphe quadrivirgata*). Ayant à ma disposition de l'ADNg de *Crocodylus niloticus*, j'ai fabriqué une nouvelle amorce *sens* appelée *CAIM6*, située en amont de l'amorce B_2.

L'amorce *antisens* (*C*) n'ayant jamais posé de problème jusqu'ici, je n'ai donc pas cherché de la changer. J'ai ensuite réalisé des PCR sur l'ADN de crocodile à une température d'hybridation de 59°C et j'ai ainsi réussi à amplifier un fragment d'ADN d'environ 400 pb (Figure 3) que j'ai ensuite cloné.

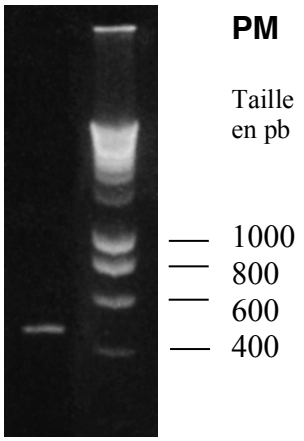


Figure 3. Résultat de la PCR sur l'ADN de crocodile obtenu par les amorces *CAIM6* et *C.*

J'ai obtenu une séquence d'ADN de *Crocodylus niloticus* que j'ai comparée plus tard à la séquence d'ADN de l'amélogénine de *Paleosuchus palpebrosus,* au moment où elle est apparue dans GenBank (numéro d'accession : AF095568). J'ai ainsi constaté que les deux séquences étaient très proches l'une de l'autre avec un pourcentage de similitude de 98,2%.

```
Caïman       …CTGATGAGACAACCGTATTCATCCTATGGGTATGAACCAATGGGAGGATGGCTACACCAG
Crocodile                              TATGGGTATGAACCAATGGGAGGATGGCTACACCAG
                                                          ****************

Caïman         CCAATGTTACCAATAGCCCAACAGCATCCACCCATCCAGACTCTCACACCTCACCACCAG
Crocodile      CCAATGTTACCATTAGCCCAACAGCATCCACCCATCCAGACTCTCACACCTCACCACCAG
               ************ **********************************************

Caïman         ATCCCCTTCCTGTCACCCCAGCACCCACTGATGCAAATGCCAGGTCCACATCAAATGATG
Crocodile      ATCCCCTTCCTGTCACCCCAGCACCCACTGATGCAAATGCCAGGTCCACAACAAATGATG

Caïman         CCTATACCACAGCAACAACCAAGCCTACAAATGCCAGTCCAAGAGCCAGTCCAACCACAG
Crocodile      CCTATACCACAGCAACAACCAAGCCTGCAAATGCCAGTCCAAGAACCAGTCCAACCACAG

Caïman         GCAGGTGAACATCCAAGCCCACCAGTGCAGCCCCAACAACCTGGGCATCCGAATCCACCG
Crocodile      CCAGGTGAACATCCAAGCCCACCAGTGCAGCCCCAACAACCTGGGCATCCAAATCCACCG

Caïman         ATGCAGCCTCAGCTACCTGGGAGTCCACACCCACCAATGCGGCCCCAGCAGCCTGGGATT
Crocodile      ATGCAGCCTCAGCTACCTGGGAGTCCACACCCACCAATGCAGCCCCAGCAACCTGGGATT

Caïman         CCAAACCCACCAATGTACCCAATGCAGCCATTGCCACCACTGCTGCCGGACATGCCACTA
Crocodile      CCAAACCCACCAATGTACCCAATGCAGCCATTGCCACCACTGCTGCCGGACATGCCACTA

Caïman         GAACCATGGCGGCCAATGGACAAGACCAAGCAAGAGGAAATAGATTAAAGAAACCAGGAA…
Crocodile      GAACCATGGCGGCCAATGGAYAARACNAARCGGGAGGAAGAAG
```

Figure 4. Alignement de la séquence de l'amélogénine de crocodile (*Crocodylus niloticus*) avec la séquence de caïman (Toyosawa *et al.,* 1998). Les différences sont indiquées en gras. Les régions encadrées représentent les amorces de PCR utilisées. On constate quelques différences entre la séquence de l'amorce C et celle du crocodile qui n'ont cependant pas gêné l'hybridation.

Séquences de serpents

L'amorce de PCR qui a permis d'amplifier l'amélogénine de *Crocodylus niloticus* (*CAIM6*) s'est avérée incapable d'amplifier l'amélogénine de *Chalcides sexlineatus*.

J'ai alors conçu un oligonucléotide (*A2*) situé à l'extrémité 5' de l'exon 6 de l'amélogénine. Bien que cette amorce ne m'ait pas permis d'obtenir de résultat positif chez *Chalcides*, j'ai entrepris de l'utiliser pour obtenir des séquences d'amélogénine chez d'autres squamates de façon à mieux connaître la partie conservée du début de l'exon 6. J'espérais ainsi être capable de fabriquer une amorce de PCR spécifique de *Chalcides* sinon de tous les squamates.

Pour ce travail, j'ai utilisé des tissus fournis par N. Vidal et appartenant à trois familles de serpents : *Micrurus lemniscatus* (serpent corail) (Elapidae), *Boa constrictor* (Boidae), *Trimeresurus purpureomaculatus* (vipère de mangrove d'Asie) (Viperidae).

En obtenant la séquence d'au moins un de ces serpents, je pouvais d'une part faire une comparaison avec la séquence d'amélogénine d'*Elaphe quadrivirgata* (Colubridae) (Ishiyama *et al.*, 1998 ; GenBank n°AF118568) et, d'autre part, cette information me permettait d'observer la variation du gène de l'amélogénine au sein d'un ensemble de familles de squamates. Après avoir extrait l'ADNg de ces serpents, j'ai donc réalisé des PCR en utilisant de nouvelles amorces (Figure 5), tout en sachant qu'il existait des différences entre la séquence des amorces et celle de l'élaphe (Figure 6).

PM	PM	PM

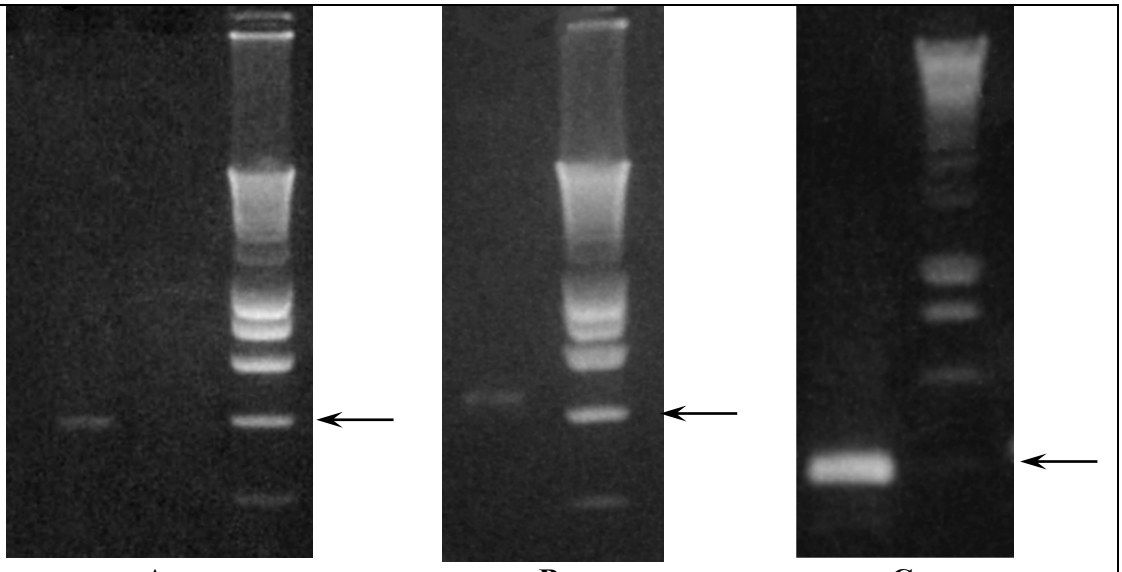


Figure 5. Résultats de l'amplification de l'ADNg de 3 serpents après migration sur gel d'agarose à 1,8%. Les PCR ont eu lieu dans les conditions suivantes : **A.** ADN de *Micrurus lemniscatus* (serpent corail) ; amorces *A2* et *C* ; température d'hybridation = 58°C ; **B.** ADN de *Boa constrictor* ; amorces *TYR6* et *C* ; température d'hybridation = 57°C ; **C.** ADN de *Trimeresurus purpureomaculatus* (vipère de mangrove), amorces *242* et *529* ; température d'hybridation = 59°C. La flèche indique la bande à 400 pb du marqueur moléculaire.

J'ai obtenu un résultat positif avec le serpent corail uniquement. Puis j'ai conçu un nouvel oligonucléotide *sens (Tyr6)* situé à l'extrémité de l'exon 6 avec lequel j'ai obtenu la séquence de l'amélogénine du *Boa constrictor*. Enfin, j'ai réussi à amplifier l'amélogénine de la vipère de mangrove d'Asie plus tard en utilisant le couple d'amorces qui me permettra d'obtenir les séquences d'amélogénine de *Chalcides* et de varan (*Varanus niloticus*) (voir 1.4 et 1.5).

Les séquences obtenues après clonage et séquençage ont été alignées avec le logiciel Clustal X (Figure 6), puis j'ai calculé les pourcentages de similitudes par rapport à l'amélogénine d'*Elaphe quadrivirgata* pour quantifier la ressemblance observée des quatre séquences.

```
Elaphe      …CACCAGTATCCACGTTACGGCTATGAGCCCATGGGAGGGTGGATGCACCATAATGCAGGA
Micrurus    ----------------TATGGGTATGAGCCCATGGGAGGGTGGATGCACCATAATTCAGGA
Boa         ------GTATCCATGTTATGGTTATGAGCCCATGGGAGGATGGATGCACCATAATGCAGGG
Vipère      ----------------------TATGAACCTATGGGAGGATGGATGCACCATAATGCAGGA
                                   *** ** ******** *************** ****

Elaphe      CCAATGATGCACCCACCTCATTTTCATGGCGTCCATGCAATACACCCAGCTCTCCATCAA
Micrurus    CCAATGACGCACCCAACTCACTTTC---------ATGCAATTCACCCAGCTCTCCATCAA
Boa         CCAGTGATGCACCCAACTCACTTTCATGGGGTCCATGCAATGCACCCAGCTCTCCATCAA
Vipère      CCAATGATGCACCCAACTCACTTTCATGGGGTCCATGCAATACACCCAGCTCTCCATCAA
            *** *** ******* **** ****         ******* ******************

Elaphe      AAGCAGCCACAACAGCTTCCTCAGAACCCACAATTGCAGCAACCAGGCCTTAACGCATTT
Micrurus    ATGCAGCCACCACAGCTCCCTCAGAACCCACATTTGCAGCAACCAGGTCCTAACCCATTT
Boa         ATGCAGCCACAACAGCCACCTCACAATCCACATTTGCAGCAACCTGGTCATAACCCATTT
Vipère      GTGCAACCACTACAGCCCCCTCAGAACCCACATTTGCAGCAACCAGGTCTTAACCCATTC
              *** **** *****  ***** ** ***** *********** ** * **** ****

Elaphe      GTTCCACTAAATGGGCATAATGCATTAATTCCACACTATCAG---------CATCCATTT
```

```
Micrurus    GTTCCACTAAATGGGCATAATACATTAATTCCACACTATCAGTCTATACAGCATCCACTT
Boa         GTCCCAATTCATGGGCATAACACATTAATTCCACACTATCAGCCTGTACAACATCCATTT
Vipère      GTTCCACTAAATGGGCATAATACATTAATTCCACACTATCAGTCTGTACAGTATCCACTT
            ** *** *  **********  ********************          ***** **

Elaphe      CAACCACAAGCAGGCGAGCATCCAATACAATCCCTATCACCAGTAAATCCA---------
Micrurus    CAACCACAAGCAGGCGAGCATCCAATGCAACCCCAACCACCAATAAATCCAAATCAATCT
Boa         CAACCACAAGCTGGGGAGCATCCAATGCAACCCCAGCCACCAGTGAATCCAAATCAATCA
Vipère      CAACCACAAATAGGCGAGCATCCAATGAAACCCCAACCACCAGTAAATCCAAATCAATCA
            *********   ** **********  ** ***   ***** * ******

Elaphe      ATGAACCCTCAGCAACCAGGGAATGCCAATCAGCCCATATACCCACTACCCCCCGTTATC
Micrurus    ATGAACCCTCAGCAACCAGGAATTGCCAATCAGCCCATGTACCCACTACCACCTCTTATC
Boa         ATGCACCCTCAGCAACCAGGGAATGCCAATCAGCCCATATATCCACTACCTCCACTTATC
Vipère      ATGAACCCTCAGCAATCAGGGAATGCCAATCAGCCCATATACCCACTACCTCCCCTTATC
            *** ********** **** * *************** ** ******** **  *****

Elaphe      CCAGATACACCTCTTGAATCATGGCCAGCTGACAAAACCAAGAAAGAAGAAGTGGATTAA
Micrurus    CCAGATACACCTCTTGAATCATGGCCAGCTGAYAARACNAARCGGGAGGAAGAAG-----
Boa         CCAGATACTCCTCTTGAATCATGGCCAGCTGAYAARACNAARCGGGAGGAAGAAG-----
Vipère      CCAGATACACCTCTTGAATCATGGCCAGCTGATAAGACCAAGCAAGAAGAAGTG------
            ******** ********************** ** *****    *******  *
```

Figure 6. Alignement des trois séquences d'amélogénine de serpents et de la séquence déjà connue d'*Elaphe quadrivirgata*. Par commodité, le *Trimeresurus* est désigné sous le nom "vipère". Le pourcentage de similitude est de 85,6% entre *Elaphe* et *Micrurus*, de 84,9% entre *Elaphe* et *Boa*, de 87,0% entre *Elaphe* et *Trimeresurus*. Les régions encadrées représentent les séquences des amorces de PCR (non-utilisées pour aligner les séquences ni pour calculer les pourcentages de similitude). Les étoiles (*) indiquent les positions où les 4 séquences sont identiques.

D'après ces résultats, il semblait donc évident que les séquences obtenues étaient de l'amélogénine de serpent. De plus, les différences observées entre ces trois séquences étaient supérieures aux erreurs de séquençage habituelles ce qui me permettait de confirmer que les 3 séquences proviennent bien de 3 serpents différents et donc que je n'ai pas amplifié deux fois le même ADNg.

Séquence de varan

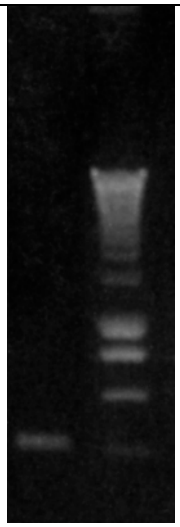
Après avoir obtenu les 2 premières séquences d'amélogénine de serpents (Micrurus et Boa) et ayant toujours des difficultés à obtenir des séquences supplémentaires de ce gène, j'ai alors entrepris de concevoir de nouvelles amorces de PCR, après avoir attentivement comparé les séquences déjà obtenues chez les reptiles et, en particulier, les squamates. J'ai déterminé les nucléotides les mieux conservés au début de l'exon 6 et les motifs les plus spécifiques des squamates. J'ai complété ces connaissances grâce à une collaboration avec une équipe de chercheurs de l'Université de Niigata qui avait obtenu la séquence d'Elaphe quadrivirgata (Ishiyama et al., 1998). Ces collègues m'ont envoyés deux séquences d'amélogénine de Gekko japonicus (Gekkonidae) et de Iguana iguana (Iguanidae) que j'ai ajouté à mes études comparatives, puis j'ai conçu un certain nombre d'amorces de 17 à 21 bases (sens et antisens), et un oligonucléotide sens de 25 bases. J'ai alors vérifié la potentialité de formation de dimères et de ponts de ces amorces.

Ces oligonucléotides ont été

Figure 7. *Amplification*

utilisés pour amplifier l'ADNg de Varanus, mais seul le couple d'amorce 242–529 m'a permis d'amplifier un fragment d'ADN de taille attendue et suffisamment concentré (Figure 5).

Le fragment d'ADN d'environ 400 pb obtenu par PCR a été cloné et la séquence obtenue comparée à l'amélogénine d'Elaphe (figure 8). Le pourcentage de similitude entre les deux séquences est de 67% (calculé sans les séquences des amorces). Cependant, ce chiffre est plus élevé dans certaines parties de la séquence, la présence d'insertions/délétions diminuant la moyenne.



de l'ADNg de varan par les amorces 242 *et* 529. *Migration sur gel d'agarose 1,8%. La flèche indique la bande de 400 pb du marqueur moléculaire.*

```
Elaphe      GGCTATGAGCCCATGGGAGGGTGGATGCACCATAATGCAGGACCAATGATGCACCCACCT
Varan       ---TATGAACCTATGGGAGGATGGATGCACCACAGTTCAGGGCCAATGATCCACCAAACT
               ***** ** ******** ********** * * **** ******** **** * **

Elaphe      CATTTTCATGGCGTCCATGCAATACACCCAGCTCTCCATCAAAAGCAGCCACAACAGCTT
Varan       CATTTTCAAGGGCTCCATCCAATGCATCCACCTCTCCACCAAATGCAGCCACAGCAGCCA
            ******** **  ***** **** ** *** ******* **** ********* ****

Elaphe      CCTCAGAACCCACAATTGCAGCAACCAGGCCTTAACGCATTTGTTCCACTAAATGGGCAT
Varan       TCTCTGAATCCACAGGTGCAGCCACCTGGGCATAGCCCATTCGTGCCAATGACTGGGCTA
             *** *** *****  ****** *** ** * ** * **** ** *** * * *****

Elaphe      AATGCATTAATTCCACACTATCAGC----------------------ATCCATTTCAA
Varan       AACAACGTGATACCGCACTATCAACCGGCTCATGCAGGACCAGCCCAACATCCCCTCCCA
            **     * ** ** ******** *                        ****  * * *

Elaphe      CCACAAGCAGGCGAGCATCCAATACAATCCCTATCACCAGTAAATCCAA---------TG
Varan       CCACATGCAGGAGAGCACCCAATGCAGCCCCCGCAGCCAGCGAACCCAAACCAGCCCATT
            ***** ***** ***** ***** **  ***     ****  ** ****         *

Elaphe      AACCCTCAGCAACCAGGGAATGCCAATCAGCCCATATA---------CCCACTACCCCCC
Varan       TATCCCCAGCGACCAGGGAATCCCAGCCAGCCAATGTATCCAATTCACCCACTGCCTCCG
             * ** **** ********** ***  ***** ** **         ****** ** **

Elaphe      GTTATCCCAGATACACCTCTTGAATCATGGCCAGCTGACAAAACCAAGAAAGAAGAAGTG
Varan       TTAATCCCAGATACACCTCTTGAATCGTGGCCAGTTGACAAGACCAAGCAAGAAGAAGTG
             * ********************** ******* ****** ****** ***********
```

Figure 8*. Alignement par Clustal X de la séquence de* Varanus niloticus *avec la séquence d'amélogénine d'*Elaphe quadrivirgata. *On constate la présence d'un certain nombre de "gaps" entre les deux séquences.*

Séquences de Chalcides

Ayant constaté que le couple *242 – 529* permet d'amplifier l'amélogénine de varan, j'ai donc entrepris de l'utiliser pour tenter d'obtenir une séquence d'amélogénine de *Chalcides*. Pour cette amplification, j'avais à ma disposition de l'ADNg d'un *Chalcides sexlineatus* et d'un *C. viridanus.* N'ayant pas obtenu précédemment de résultat positif avec

ces ADNg, j'avais également extrait les ARNm d'une mâchoire de *Chalcides sexlineatus.* Malheureusement, les PCR qui ont suivi ont été sans succès également, faute de connaissance suffisante de l'amélogénine chez les reptiles. Plus tard, j'ai entrepris de concevoir de nouvelles amorces situées dans les exons 2, 3 et 5 de l'amélogénine et d'extraire à nouveau des ADNc d'un *C. viridanus* craignant que les anciens ADNc soient périmés (1 an s'était écoulé).

Le couple *242 – 529* a été testé avec succès sur l'ADNg de *Chalcides viridanus*, puis sur les ADNc de *C. viridanus et C. sexlineatus* (pour vérifier leur pérennité).

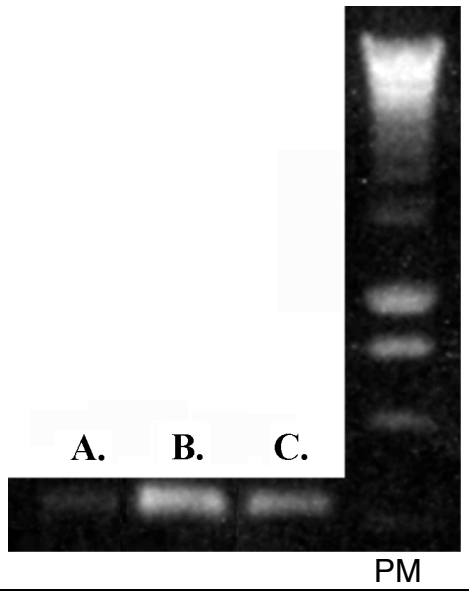


Figure 9. Amplification de l'ADN de *Chalcides* par le couple d'amorces *242 – 529* ; température d'hybridation = 59°C ; nombre de cycles = 38. **A.** PCR réalisée avec l'ADNc de *Chalcides viridanus* ; **B.** PCR réalisée avec l'ADNc de *Chalcides sexlineatus ;* **C.** PCR réalisée sur l'ADNg de *Chalcides viridanus*.
La flèche indique la bande de 400 pb du marqueur moléculaire.

L'ADN correspondant aux lignes B et C (figure 9) a été cloné puis la séquence nucléotidique correspondante déterminée. L'ADNc de Chalcides viridanus (ligne A) n'a pas été cloné car la séquence obtenue serait identique à celle de l'ADN de la ligne C. J'ai décidé d'amplifier plus tard cet ADNc en utilisant des amorces situées dans les autres exons de l'amélogénine pour obtenir le reste du gène.

En alignant ces séquences obtenues avec celle de Gallotia stehlini (squamate le plus proche parent d'après la phylogénie connue (Estes et al., 1988, Harris et al., 1999)) on constate de nombreuses ressemblances entre les deux familles de squamates : 65,5% entre C. sexlineatus et G. stehlini, 70,3% entre C. viridanus et G. stehlini et 97% entre les deux Chalcides (différence due principalement à une insertion/délétion de 6 nucléotides) (Figure 10).

Puisque les extraits obtenus à partir d'ARNm de mâchoire contiennent bien des ADNc d'amélogénine, de nouvelles amorces *sens* dans les exons 5, 3 et 2 appelées respectivement *Ex5*, *Ex3.2* et *Ex2* ont été fabriquées pour pouvoir amplifier la majorité du gène. Pour cela j'ai utilisé les séquences connues de ces exons chez les squamates (*Elaphe*, *Iguana*, *Gekko*).

Des PCR avec ces amorces *sens* et l'amorce *antisens 529* ont été réalisées avec comme résultat des fragments d'ADN de 400 à 500 pb (Figure 11) qui ont été clonés et dont la séquence a été déterminée (Figure 12).

```
Chalci. vir.   TATGAACCTATGGGAGGATGGATGCGCCACCATTCAGGGCCAATGCTGCCGCAGCTGCCTC
Chalci._sex.   TATGAACCTATGGGAGGATGGATGCGCCACCATTCAGGGCCAATGCTGCCGCAGCTGCCTC
Gallotia               CATGGGTGGNTGGYTNCAYCACCATGCAGGGCCAACGATGCACCAAGCTCATT

Chalci. vir.   ATCAAAGTATCTATCAAAGCCTTCATCCAATGCAAACTCCTCTCCATCAACTGCATCCCCT
Chalci._sex.   ATCAGAGTATCTATCAAAGCCTTCATCCAATGCAAACTCCTCTCCATCAACTGCATCCCCT
Gallotia       ATCAAA---TGCACCCACCCCTCCATCAAATGCTG------CCACAGCAGC--CACCATTT

Chalci. vir.   GCCACCACAAAACCCATCAATGCAACAACCTGCACATCACCCATTTGTGCCAATGACTGGG
Chalci. sex.   GCAACCACAAAACCCATCAGTGCAACAAACTGCACATCACCCATTTGTGCCAATGACTGGG
Gallotia       GC-ACATTCAAACCCACAAATGCCGCAACCTGGTCATGGCCCATATATGCCTATGACTGGG

Chalci. vir.   CAAAATTCACTAAATCCACACTTTCAACCAGTTCATGCAGCACCAGCCCAACACCCACTTG
Chalci. sex.   CAAAATTCACTAAATCCACACTTTCAACCAGTTCATGC------AGCCCAACACCCACTTG
Gallotia       CAAAACACAATTATGCCACAGTACCAACCAGGTCATGCAGGATCAGCTCATCACTCAATTC

Chalci. vir.   CACCACACACAGGACAGCTCCCGGTGCAGCCCCAGCAACCAGATCATTCAAACCATCCAGT
Chalci. sex.   CACCACACACAGGACAGCTCCCAGTGCAGCCCCAGCAACAAGATCATTCAAACCATCCAGT
Gallotia       CACAACATGCAGGAGAGCAACCAATGCACCCCCTGCCACCAAGTAGCCCACACGAGCCAAT

Chalci. vir.  ACAACACCAGCAACCAGCTGATCCAATCCAGCCAATGTACCCAGTTCCACAACTGCCACCAA
Chalci. sex.  ACAACACCAGCAACCAGCTGATCCAATCCAGCCAATGTACCCAGTTCCACAACTGCCACCAA
Gallotia      GCACCCTCAGCAACCAGGGGGTCCAAACTACGCAATGTACCCAGAGCCCGTCTTGCGTCCAG

Chalci. vir.  TGATCCCAGATTTGCCTCTTGAACCATGGCCAGCAGCAGACAAGACCAAGCAAGAAGAAGTG
Chalci. sex.  TGATCCCAGATTTGCCTCTTGAACCATGGCCAGCAGCAGACAAGACCAAGCAAGAAGAAGTG
Gallotia      TGCTACCAGATACACCCCTTGAATCATGGCCACCAGTCGAYAARACNAARCGGGAGGAAGAA
```

Figure 10. Alignement des 2 séquences de *Chalcides* et de celle de *Gallotia* par le logiciel Clustal X. Les séquences des amorces sont encadrées.

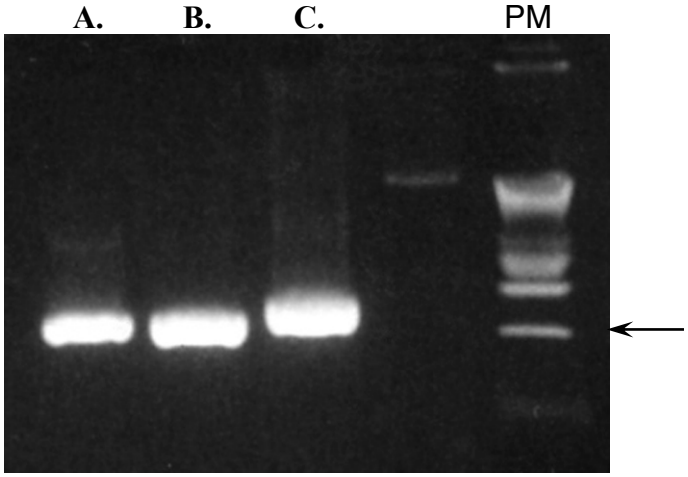


Figure 11. Amplification des ADNc de *Chalcides* par des amorces dans les exons 2, 3 et 5 de l'amélogénine. **A.** ADNc de *Chalcides sexlineatus* ; amorces *Ex5-529.* **B.** ADNc de *Chalcides viridanus*; amorces *Ex3-529.* **C.** ADNc de *Chalcides viridanus*; amorces *Ex2-529.* La flèche indique la bande de 400 pb du marqueur moléculaire.

```
C. viridanus   ---------------------------------------------- --------------
Exon5 (sex.)   ---------------------------------------------- --------------
Exon3 (sex.)   ---------------------------------------------- --------------
Exon2 (vir.)   CTGGACTTTGGTTATGTGCCTTTTAAGTGCCACATTTGCTATTCCA CTGCCGCCTCATGC
                                    EXON 2            ↑            EXON 3



C. viridanus   ----------------------------------- -------------------------
```

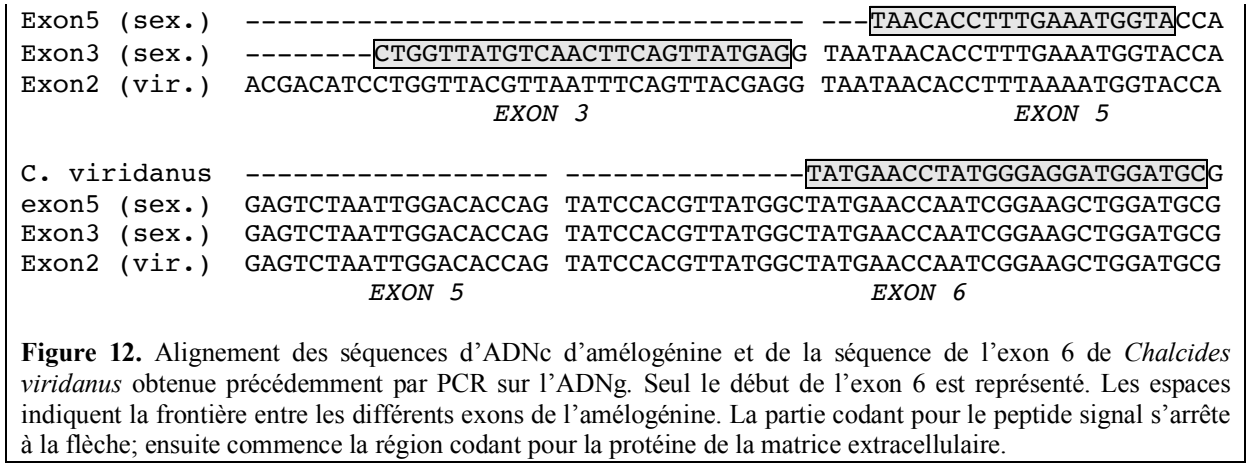


Figure 12. Alignement des séquences d'ADNc d'amélogénine et de la séquence de l'exon 6 de *Chalcides viridanus* obtenue précédemment par PCR sur l'ADNg. Seul le début de l'exon 6 est représenté. Les espaces indiquent la frontière entre les différents exons de l'amélogénine. La partie codant pour le peptide signal s'arrête à la flèche; ensuite commence la région codant pour la protéine de la matrice extracellulaire.

A ce stade, j'avais donc obtenu chez *Chalcides* la moitié de la séquence de l'exon 2, l'exon 3 et l'exon 5 entiers, et la majorité de l'exon 6 moins les 21 pb de l'amorce de PCR *antisens* (*529*). La région codante du gène de l'amélogénine de *Chalcides* était donc presque entièrement connue, tout au moins la partie constituant la protéine extracellulaire qui intervient dans la minéralisation de l'émail. En effet, bien que la région codante de l'amélogénine commence 9 pb après le début de l'exon 2, celui-ci code majoritairement pour le peptide signal chargé de diriger la protéine vers l'extérieur de la cellule (voir le chapitre *Introduction*). Une fois sa tache remplie, le peptide signal est alors éliminé de la protéine bien avant que celle-ci ne se retrouve dans la matrice extracellulaire. Il me restait donc à connaître la fin de l'exon 6 et l'exon 7 entier de l'amélogénine de *Chalcides,* le dernier exon codant seulement pour un acide aminé.

Pour obtenir cette partie du gène, j'ai utilisé la technique de la RACE 3' utilisant une amorce *antisens* poly-T censée se fixer à la queue poly-A qui termine naturellement tous les ARNm et les ADNc. La spécificité de la PCR résulte donc des amorces *sens*

J'ai réalisé, sur l'ADNc de *Chalcides sexlineatus*, une PCR à l'aide d'une amorce *antisens* poly(dT)$_{18}$ et de l'amorce *sens* (*118*), puis j'ai ré-amplifié 1 µl de cette PCR avec une amorce *sens* plus interne (*142*) et toujours l'amorce *antisens* poly(dT). En effet, le manque de spécificité de l'amorce poly(dT) ne permet pas d'amplifier directement l'ADNc visé, il est donc nécessaire d'amplifier les produit de PCR 1 ou 2 fois pour obtenir un fragment visible sur gel d'agarose. Dans le cas présent, une seconde PCR seulement à été nécessaire.

J'ai obtenu la séquence suivante (le codon stop "TAA" est souligné et la partie codante de l'exon 7 est grisé ; tout ce qui suit constitue la partie non-codante (exon 7 et éléments de régulation de l'ARNm).

…CAA CTG CCA CCA ATG ATC CCA GAT TTG CCT CTT GAA CCA TGG CCA
GCA GCA GAC AAG ACC AAG CAA GAG GAA GTG GAT TAA
AGAGACAAAGAAATCAGAATAATGAAAATAAATATTTTA
GGCATTTTCATTTGATTTCAAGAACCCATCATCCTACTGCAGTCAGCAATGCATGTTTTCAT
TAATCCAGGTTCTTGGAGAATGCTGTAGCTAGAAAGCAAACAAGTCCAAAAGTAGCAAATAA
TAAAACCTGTTCTAATCAATAAAAAAAAAAAAAAAAAA

Séquences d'ADNc d'autres lézards

Après la mise au point de l'extraction des ARN des mâchoires de *Chalcides,* j'ai réussi à extraire ceux de deux lézards capturés en été (le psammodrome (*Psammodromus algirus)* et l'orvet (*Anguis fragilis*)) ainsi que ceux d'un Lacertidae en élevage au laboratoire (*Gallotia*

gallotii). J'ai transformé ces ARNm en ADNc et amplifié les ADNc par PCR avec les amorces *EX2* et *529* grâce auxquelles j'ai réussi à amplifier précédemment l'ADNc de *Chalcides.* Après migration, j'ai obtenu un fragment d'ADN de taille identique à celui de *Chalcides* (voir ci-dessus). Après clonage et séquençage, j'ai donc obtenu les séquences suivantes que j'ai alignées avec l'ADNc de *Chalcides* :

```
Orvet          CTGGACTTTGGTTATGTGCCTCTTAAGTACCACTTTTGCTCTTCCGCTGCC---------AC
Gallotia       CTGGACTTTGGTTATGTGCCTCTTAAGCACTACATTTGCTATTCCACTGCC---------A
Chalcides      CTGGACTTTGGTTATGTGCCTTTTAAGTGCCACATTTGCTATTCCACTGCCGCCTCATGCA
Psammodrome    CTGGACTTTGGTTATGTGCCTCTTAAGCACTACATTTGCTATTCCATTGCC---------A

Orvet          AACATCCTGGTTATATCAATTTCAGTTATGAGGTAATGACACCTTTAAAATGGTACCAGAGC
Gallotia       AACATCCTGGCTATATCAATTTCAGTTACGAGGTAATGACACCTTTGAAGTGGTACCAGAAA
Chalcides      GACATCCTGGTTACGTTAATTTCAGTTACGAGGTAATAACACCTTTAAAATGGTACCAGAGT
Psammodrome    AACATCCTGGCTATATCAATTTCAGTTACGAGGTAATGACACCTTTAAAGTGGTACCAGAAA

Orvet            CTAATTGGCCATCAGTATCCACGGTATGGCTATGAACCCATGGGAGGATGGATGCACCAC
Gallotia         CTAATGGGACCCCAGTATCCACATTATGCCTATGAACCCATGGGAGGATGGATGCACCAC
Chalcides        CTAATTGGACACCAGTATCCACGTTATGGCTATGAACCAATCGGAAGCTGGATGCGCCAC
Psammodrome      CTAATGGGACCCCAGTATCCACATTACTCCTATGAGCCCATGGGAGGATGGATGCACCAC

Orvet            AGTACAGGGCCAGTGACGCACCAAACTTACTTTCAAG------------GGCATTCA-
Gallotia         CATGCAGGGCCAACGATGCACCAAGCTCATTATCAAA------------TGCACCCAC
Chalcides        CATTCAGGGCCAATGCTGCCGCAGCTGCCTCATCAAAG---------TATCTATCAAA
Psammodrome      CACGCAGGACCATCAGTGCACCAAGCTCATTATC------------CAATGCACCCAC

Orvet            -----------ATGCAG------CCCCAGCAGCCAC------------CCCTGAACCCAC
Gallotia         CCCTCCATCAAATGCTG------CCACAGCAGCCAC---CGTTTGCACATCCAAACCCAC
Chalcides        GCCTTCATCCAATGCAAACTCCTCTCCATCAACTGCATCCCCTGCCACCACAAAACCCAT
Psammodrome      CCCTCCATCAAATGCTG------CCACAGCAGCCAC---ATTTTGCACATTCAAACCCAC

Orvet            AC------ATGCAGCATCCTGGGCATAACTCATACAGGTCAATGACTGGACAAGATATAT
Gallotia         AA------ATGCCGCAACCTGGTCATGGCTCATATATGCCGATGACTGGGCAAAACACAA
Chalcides        CA------ATGCAACAACCTGCACATCAC-CATTTGTGCCAATGACTGGGCAAAATTCAC
Psammodrome      AA------ATGCCGCAACCTGGTCATGGCCCATATATGCCAATGACTGGTCAAAATCCAA

Orvet            TAATGCCA------------GCTCACGCAGGACCAGTCCAACATCCGCCTCCACCACACG
Gallotia         TTATGCCACAGTACCAACCAGGTCATGCAAGACCAGCTCATCACTCAATTCCACAACATG
Chalcides        TAAATCCACACTTTCAACCAGTTCATGCAGCACCAGCCCAACACCCACTTGCACCACACA
Psammodrome      TAATGCCACAGTACCAACCAGGTCATGCAGGACCAGTTCATCACTCAATTCCACAACATG

Orvet            TGGGAGAGTTCCCAATGCACCCCCAGCAACCAGGGAACCCAAATCAGCCAATGCACCCTC
Gallotia         CAGGAGAGCAACCAATGCACCCCCTGCCACCAAGTAGCCCACACCAGCCAATGAACCCTC
Chalcides        CAGGACAGCTCCCGGTGCAGCCCCAGCAACCAGATCATTCAAACCATCCAGTACAACACC
Psammodrome      TAGGAGAGCAACCAATGCACCCCCTACCACCAAGTAGTCCAAACCAGCCAATGCACCCTC

Orvet            AGCAACCAGGGAATCCTAACCAGCCAATGTATCCAATTCAGCCACTGCCTCCACTGATCC
Gallotia         TGCAACCAGGGAATCCGAACCACGCAATGTACCCAGTGCCTGTCTTGCCTCCAGTGATAC
Chalcides        AGCAACCAGCTGATCCAATCCAGCCAATGTACCCAGTTCCACAACTGCCACCAATGATCC
Psammodrome      AGCAACCAGGGAATCCAAACCACGCAATGTACCCAGGGCCTGTCTTGCCTCCAGTGCTAC

Orvet            CAGATACACCTCTTGAATCATGGCCACCAGCTGACAAGACCAAGCAAGAAGAAGTG
Gallotia         CAGATACACCCCTTGAATCATGGCCACCAGTCGACAAGACCAAGCAAGAAGAAGTG
Chalcides        CAGATTTGCCTCTTGAACCATGGCCAGCAGCAGACAAGACCAAGCAAGAAGAAGTG
Psammodrome      CAGATACACCCCTTGAATCATGGCCACCAGTCGACAAGACCAAGCAAGAAGAAGTG
```

Aléas et problèmes expérimentaux

Pendant une longue période de ma thèse, les résultats ont été difficiles à obtenir et il est logique, à la vue des séquences obtenues, de s'intéresser aux causes de ces échecs.

La cause la plus importante des problèmes rencontrés est la non-spécificité des amorces. Dès le départ, mais je ne l'ai su que plus tard, les amorces dégénérées fabriquées à partir des seules séquences de mammifères connues présentaient un certain nombre de divergences avec les séquences des squamates. Seules les connaissances de l'amélogénine d'un crocodile et d'un serpent m'ont permis d'avancer dans la conception d'amorces plus spécifiques. La séquence de *Chalcides* a été la plus difficile à obtenir et, malheureusement, elle constituait l'objectif principal de mes recherches. *A posteriori*, après avoir obtenu la séquence de l'ADNc de l'amélogénine de *Chalcides*, j'ai compris d'où provenaient mes difficultés. En effet, l'amélogénine de ce lézard possède des substitutions, que l'on ne retrouve pas chez les autres squamates, et qui sont situées dans la région conservée de l'exon 6 dans laquelle les amorces de PCR ont été construites pour amplifier l'ADNg (Figure 13).

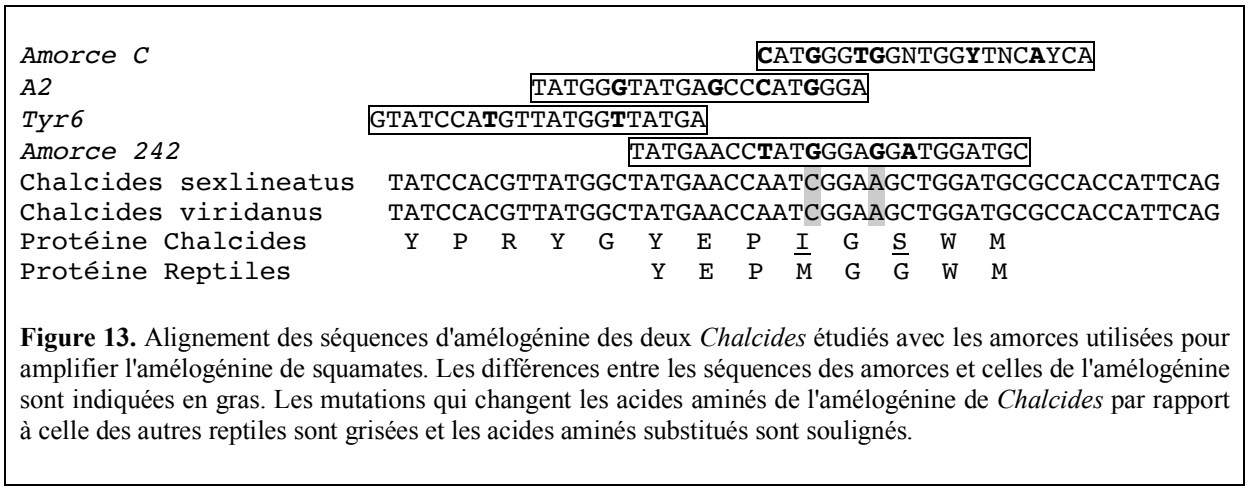


Figure 13. Alignement des séquences d'amélogénine des deux *Chalcides* étudiés avec les amorces utilisées pour amplifier l'amélogénine de squamates. Les différences entre les séquences des amorces et celles de l'amélogénine sont indiquées en gras. Les mutations qui changent les acides aminés de l'amélogénine de *Chalcides* par rapport à celle des autres reptiles sont grisées et les acides aminés substitués sont soulignés.

Traduite en protéine, on constate que la séquence de l'amélogénine de *Chalcides* est unique dans cette région comparée à celle des autres lézards et à celle de tous les vertébrés connus (voir chapitre I). Ce fait, totalement imprévisible au départ, explique les problèmes pour amplifier l'amélogénine de *Chalcides* car ces mutations s'ajoutaient aux différences résultant du manque de connaissance du gène chez les squamates.

Un autre problème plus général est lié au fait que les amorces de PCR trouvent parfois un environnement favorable ailleurs dans le génome pour s'hybrider. Certains sites sur l'ADNg peuvent alors concurrencer le site de fixation sur l'amélogénine. Après clonage et séquençage, on obtient une séquence d'ADN parasite. Cela est d'autant plus difficile à éviter que souvent les fragments d'ADN parasite amplifiés avaient une taille proche de celle attendue pour le fragment d'amélogénine (400 pb). Parfois, une seule amorce pouvait suffire pour amplifier de l'ADN parasite en se fixant ailleurs dans le génome et en trouvant un site antiparallèle non loin de là. Ce dernier cas est plus difficile à contourner que le premier, car s'il est possible d'utiliser un couple d'amorce plus interne pour ré-amplifier uniquement l'amélogénine présent dans les produits de PCR, parfois ce duo n'amplifie rien. Il faut alors se résoudre à utiliser uniquement une seule amorce plus interne en la combinant avec une des deux amorces de départ, ce qui conduit souvent à réutiliser l'amorce qui produit de l'ADN parasite !

Une amorce peut aussi former un dimère en trouvant des sites de fixation sur un autre exemplaire d'elle-même. On obtient alors, après amplification et clonage, une séquence d'ADN parasite constituée par la répétition de la séquence de l'amorce (assez jolie à voir mais très frustrant !). Pour éviter ce problème, je l'ai appris après quelques erreurs de ce genre, il

suffit de faire bouillir l'amorce pendant quelques minutes et de la plonger dans la glace pour la dénaturer avant de l'utiliser.

Une autre cause courante d'erreur est la contamination par de l'ADN étranger. Si de l'ADN contaminant est présent dans un des composants de la PCR, celui-ci peut d'autant mieux être amplifié qu'il concurrence l'ADN cible. Le gène de l'amélogénine d'un ADN génomique contaminant a, en principe, autant de chance d'être amplifié que l'amélogénine de l'ADN introduit dans la réaction. Dans ce cas, ce sont les amorces qui détermineront quel ADNg sera amplifié en fonction de leur affinité pour l'un ou l'autre. En revanche, lorsque c'est un fragment d'amélogénine qui contamine le tube de PCR, il n'y a pratiquement aucune chance de ne pas l'amplifier. C'est ainsi qu'au cours de ma thèse, j'ai obtenu la séquence de mon propre gène de l'amélogénine, ainsi que le gène de l'amélogénine du crocodile que j'avais déjà cloné un mois auparavant ! Lorsque cet événement se produit, il ne reste qu'à réaliser des PCR test sans ADN dans lesquels il faut introduire un seul des produits utilisés dans la réaction de PCR de départ, le reste étant neuf ou non contaminé. Cela permet de détecter et d'écarter un tube contaminé.

ANNEXE 3
Tampons et solutions

Cacodylate de sodium 0,2 M
Pour 100 ml de solution, peser 4,28 g de cacodylate de sodium. Dissoudre la poudre dans 80 ml d'eau distillée puis compléter à 100 ml. Le pH est de 8,3 et doit être ajusté à 7,4 avec de l'HCl 0,2 M (1 ml de HCl fumant à 37% dans 60 ml d'eau).

Eau DEPC (sans RNase)
Préparée sous une hotte aspirante en additionnant 2 l d'eau distillée, 1 ml d'éthanol 100% et 2 ml de DEPC, puis en agitant 1 h avant d'autoclaver à 120°C pendant 20 min.

FB (milieu de culture pour bactéries)
KCl 7,4 g/l ; $CaCl_2$ 7,5 g/l ; Kacétate 10mM, ; Glycerol 11,5 % ; pH=6,2.

Formoy
Acide acétique à 10%, d'alcool absolu à 30% et paraformaldéhyde à 37%.

Oligotex (solution contenant les billes se fixant aux ARNm poly A^+ dans le Kit Qiagen). Tris Cl 10 mM (pH 7,5), NaCl 500 mM, EDTA 1 mM, SDS 0,1%, NaN_3 (sodium azide) 0,1%.

OW2
(Tampon de lavage utilisé pour purifier les ARNm poly A^+).
Tris Cl 10 mM (pH 7,5), NaCl 150 mM, EDTA 1 mM, SDS 0,1%.

OEB
Tampon d'élution permettant de séparer les billes Oligotex des ARNm poly A^+.
Tris Cl 10 mM (pH 7,5).

Paraformaldéhyde 15% (PFA)
Préparer de la soude normale (2 pastilles de 0,2 g dans 10 ml d'eau). Dans un erlenmeyer de 25 ml, mettre 1,5 g de PFA et 10 ml d'eau distillée. Faire chauffer l'erlenmeyer dans un bain-marie d'eau distillée à 70°C (utiliser un thermomètre pour vérifier), puis ajouter 2 gouttes de soude dans la PFA. Faire refroidir rapidement l'erlen contenant la PFA sous l'eau du robinet. Le préparer extemporanément car il ne se conserve pas longtemps.

PBS
8 g de NaCl ; 0,2 g de KCl ; 1,44 g de Na_2HPO_4 et 0,24 g de KH_2PO_4 dans 800 ml d'eau distillée ; pH=7,4 ajusté avec de l'HCl. La solution est complétée à 1 l avec de l'eau distillée puis autoclavée 20 min à 120°C.

SSC 1 X
150 mM de NaCl, 15 mM de trisodium citrate à pH = 7.

TBE 1X
Tris base 0,89 M ; acide borique 0,89 M ; EDTA 0,02 M ; pH= 8,3 à 25°C.

TE
Tris base 10 mM ; EDTA 1 mM ; pH= 7 avec HCl 6 M.

Tampon de pré-hybridation (pour 2 litres de tampon utilisés en "Hybridation sur coupe") :

- 1000 µl de formamide (100%) désionisée
- 20 µl de $NaPO_4$ 2 M, pH 7,4
- 200 µl de SSC 20 X
- 50 µl d'EDTA 0,2 M
- 50 µl de Denhart 100 X
- 40 µl de DNA dénaturé (10 µg/µl) } 250 µg/µl d'ADN au total
- 10 µl de DNA natif (10 µg/µl) } 250 µg/µl d'ADN au total
- 100 µl d'ARNt (10 µg/µl)
- 530 µl d'eau ultra pure

Tampon d'hybridation (utilisé en "Hybridation sur coupe")

- 750 µl de formamide 100% désionisée
- 15 µl de $NaPO_4$ 2 M, pH 7,4
- 150 µl de SSC 20 X
- 37,5 µl d'EDTA 0,2 M
- 37,5 µl de Denhart 100 X
- 30 µl d'ADN dénaturé (10 µg/µl)
- (7,5 ml d'ADN natif ajouté plus tard)
- 375 µl de dextran sulfate (0,4 µg/µl)
- 97,5 µl d'eau ultra pure

ANNEXE 4 - Séquence de Chalcides sexlineatus obtenue à partir de PCR sur des ADNc. La technique de RACE 3' a permis d'obtenir la partie non codante de l'exon 7 jusqu'à la queue poly A. Les exons sont séparés par une ligne verticale.

CTGGACTTTGGTTATGTGC

```
CTT TTA AGT GCC ACA TTT GCT ATT CCA|CTG CCG CCT CAT GCA CGA
 L   L   S   A   T   F   A   I   P | L   P   P   H   A   R
                                 2 3
CAT CCT GGT TAC GTT AAT TTC AGT TAC GAG|GTA ATA ACA CCT TTA
 H   P   G   Y   V   N   F   S   Y   E | V   I   T   P   L
                                   3 5
AAA TGG TAC CAG AGT CTA ATT GGA CAC CAG|TAT CCA CGT TAT GGC
 K   W   Y   Q   S   L   I   G   H   Q | Y   P   R   Y   G
                                   5 6
TAT GAA CCA ATC GGA AGC TGG ATG CGC CAC CAT TCA GGG CCA ATG
 Y   E   P   I   G   S   W   M   R   H   H   S   G   P   M

CTG CCG CAG CTG CCT CAT CAA AGT ATC TAT CAA AGC CTT CAT CCA
 L   P   Q   L   P   H   Q   S   I   Y   Q   S   L   H   P

ATG CAA ACT CCT CTC CAT CAA CTG CAT CCC CTG CCA CCA CAA AAC
 M   Q   T   P   L   H   Q   L   H   P   L   P   P   Q   N

CCA TCA ATG CAA CAA CCT GCA CAT CAC CCA TTT GTG CCA ATG ACT
 P   S   M   Q   Q   P   A   H   H   P   F   V   P   M   T

GGG CAA AAT TCA CTA AAT CCA CAC TTT CAA CCA GTT CAT GCA GCA
 G   Q   N   S   L   N   P   H   F   Q   P   V   H   A   A

CCA GCC CAA CAC CCA CTT GCA CCA CAC ACA GGA CAG CTC CCG GTG
 P   A   Q   H   P   L   A   P   H   T   G   Q   L   P   V

CAG CCC CAG CAA CCA GAT CAT TCA AAC CAT CCA GTA CAA CAC CAG
 Q   P   Q   Q   P   D   H   S   N   H   P   V   Q   H   Q

CAA CCA GCT GAT CCA ATC CAG CCA ATG TAC CCA GTT CCA CAA CTG
 Q   P   A   D   P   I   Q   P   M   Y   P   V   P   Q   L

CCA CCA ATG ATC CCA GAT TTG CCT CTT GAA CCA TGG CCA GCA GCA
 P   P   M   I   P   D   L   P   L   E   P   W   P   A   A

GAC AAG ACC AAG CAA GAG GAA GTG|GAT TAA
 D   K   T   K   Q   E   E   V | D   *
                               6 7
```

```
AGAGACAAAGAAATCAGAATAATGAAAATAAATATTTTAGGCATTTTCATTTGATTTCA
AGAACCCATCATCCTACTGCAGTCAGCAATGCATGTTTTCATTAATCCAGGTTCTTGGAGAAT
GCTGTAGCTAGAAAGCAAACAAGTCCAAAAGTAGCAAATAATAAAACCTGTTCTAATCAATAA
AAAAAAAAAAAAAAAAAA
```

ANNEXE 5 – Séquences d'amélogénine obtenues dans GenBank

Nom	Genre espèce	Classe	Ordre	Famille	N° Accession
Xénope -clone 1	*Xenopus laevis*	Amphibia	Anura	Pipidae	AF095569
Xénope -clone 2	*Xenopus laevis*	Amphibia	Anura	Pipidae	AF095570
Caïman nain	*Paleosuchus palpebrosus*	Sauropsida	Archosauria	Crocodylidae	AF095568
Couleuvre japonaise	*Elaphe quadrivirgata*	Sauropsida	Squamata	Colubridae	AF118568
Vache – chrom. Y	*Bos taurus*	Mammalia	Cetartiodactyla	Bovidae	M63500.
Vache – chrom. X	*Bos taurus*	Mammalia	Cetartiodactyla	Bovidae	M63499
Chèvre – chrom. X	*Capra hircus*	Mammalia	Cetartiodactyla	Bovidae	AF215889
Chèvre – chrom. Y	*Capra hircus*	Mammalia	Cetartiodactyla	Bovidae	AF215890
Bouquetin – chrom. X	*Capricornis crispus*	Mammalia	Cetartiodactyla	Bovidae	AB073741
Bouquetin – chrom. Y	*Capricornis crispus*	Mammalia	Cetartiodactyla	Bovidae	AB073742
Bouquetin – chrom. Y2	*Capricornis crispus*	Mammalia	Cetartiodactyla	Bovidae	AB073743
Cerf Sika– chrom. X	*Cervus nippon*	Mammalia	Cetartiodactyla	Cervidae	AB028027
Cerf Sika – chrom. Y	*Cervus nippon*	Mammalia	Cetartiodactyla	Cervidae	AB028028
Cochon	*Sus scrofa*	Mammalia	Cetartiodactyla	Suidae	U43406
Cheval – chrom. X	*Equus caballus*	Mammalia	Perissodactyla	Equidae	AB032193
Cheval – chrom. Y	*Equus caballus*	Mammalia	Perissodactyla	Equidae	AB032194
Chien domestique	*Canis familiaris*	Mammalia	Carnivora	Canidae	AB080686
Loup	*Canis lupus*	Mammalia	Carnivora	Canidae	AB080687
Rat	*Rattus norvegicus*	Mammalia	Rodentia	Muridae	U67130
Souris	*Mus musculus*	Mammalia	Rodentia	Muridae	D31769
Hamster	*Mesocricetus auratus*	Mammalia	Rodentia	Muridae	AF005245
Cobaye	*Cavia porcellus*	Mammalia	Rodentia	Muridae	AJ012200
Homme – chrom. Y	*Homo sapiens*	Mammalia	Primates	Hominidae	M86933
Homme – chrom. X	*Homo sapiens*	Mammalia	Primates	Hominidae	M86932
Opossum	*Monodelphis domestica*	Mammalia	Metatheria	Didelphidae	U43407
Echidné australien	*Tachyglossus aculeatus*	Mammalia	Monotremata	Tachyglossidae	AF095567
Ornithorynque	*Ornithorhynchus anatinus*	Mammalia	Monotremata	Ornithorhynchidae	AF095566

ANNEXE 6 – Bilan des séquences d'amélogénine de reptiles clonées.

Nom	Genre espèces	Classe	Ordre	Famille	ADN utilisé	Exons Séquencés	Tissus source de l'ADN
Scinque des Canaries	*Chalcides sexlineatus*	Sauropsida	Squamata	Scincidae	ADNc	2, 3, 5, 6,7	Mandibule
"	*Chalcides viridanus*				ADNc	2, 3, 5, 6	Mandibule
Lézard des Canaries	*Gallotia stehlini*	Sauropsida	Squamata	Lacertidae	ADNg	6	Foie
"	*Gallotia galloti*	Sauropsida	Squamata	Lacertidae	ADNc	2, 3, 5, 6	Mandibule
Psammodrome algire	*Psammodromus algirus*	Sauropsida	Squamata	Lacertidae	ADNc	2, 3, 5, 6	Mandibule
Lézard des murailles	*Lacerta muralis*	Sauropsida	Squamata	Lacertidae	ADNg	6	Foie
Orvet	*Anguis fragilis*	Sauropsida	Squamata	Anguidae	ADNc	2, 3, 5, 6	Mandibule
Serpent corail	*Micrurus lemniscatus*	Sauropsida	Squamata	Elapidae	ADNg	6	Muscles
Boa	*Boa constrictor*	Sauropsida	Squamata	Boidae	ADNg	6	Muscles
Vipère des mangroves	*Trimeresurus purpureomaculatus*	Sauropsida	Squamata	Viperidae	ADNg	6	Muscles
Varan	*Varanus niloticus*	Sauropsida	Squamata	Varanidae	ADNg	6	/
Crocodile	*Crocodylus niloticus*	Sauropsida	Archosauria	Crocodylidae	ADNg	6	Embryon
Caiman	*Caiman crocodylus*	Sauropsida	Archosauria	Crocodylidae	ADNg	6	Foie
Gecko	*Gekko Japonicus*	Sauropsida	Squamata	Gekkonidae	ADNc	1, 2, 3, 5, 6,7	
Iguane	*Iguana iguana*	Sauropsida	Squamata	Iguanidae	ADNc	1, 2, 3, 5, 6,7	
Agame	*Leiolepis belliana*	Sauropsida	Squamata	Agamidae	ADNg	6	

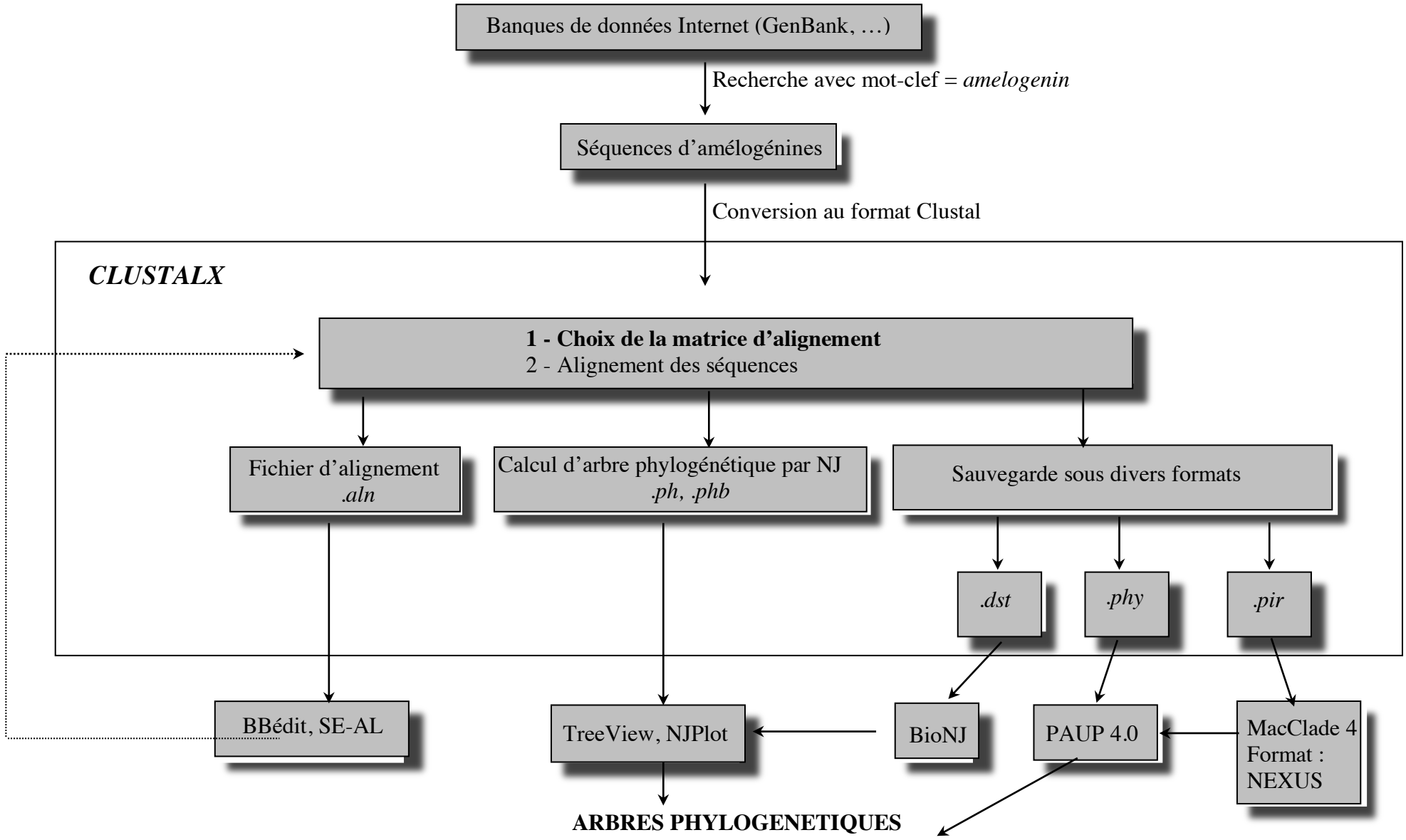
Les noms en grisé représentent les séquences obtenues en collaboration avec le docteur Ishiyama (Department of Histology and Oral Microbiology, The Nippon Dental University School of Dentistry at Niigata, Niigata, JAPAN).

ANNEXE 6 (suite)– Bilan des séquences d'amélogénine de mammifères clonées.

Nom	Genre espèces	Classe	Ordre	Famille	ADN utilisé	Exons Séquencés	Tissus source de l'ADN
Phoque gris[1]	*Halichoerusus grypus*	Mammalia	Carnivora	Phocidae	ADNg	6	/
Fennec[2]	*Vulpes zerda*	Mammalia	Carnivora	Canidae	ADNg	6	Foie
Guépard[2]	**Acinonyx jubatus**	Mammalia	Carnivora	Felidae	ADNg	6	Pancréas
Mégachiroptère[3]	*Cynopterus brachyotis*	Mammalia	Chiroptera	Dermoptera	ADNg	6	?
Hérisson[4]	**Erinaceus europaeus**	*Mammalia*	Insectivora	Erinaceidae	ADNg	6	Foie
Zèbre de Grévy[2]	*Equus grever*	Mammalia	Perissodactyla	Equidae	ADNg	6	Foie
Hippopotame nain[2]	*Hexaprotodon liberiensis*	Mammalia	Artiodactyla	Hippopotamidae	ADNg	6	Rate
Eléphant d'Asie[2]	*Elephas maximus*	Mammalia	Proboscidea	Elephantidae	ADNg	6	Rate
Dauphin[5]	*Tursiops truncatus*	Mammalia	Cetacea	Delphinidae	ADNg	6	Rein
Galéopithèque tacheté[3] (ou "lémur volant")	*Cynocephalus variegatus*	Mammalia	Dermoptera	Cynocephalidae	ADNg	6	?
Toupaye[3]	*Tupaia glis*	Mammalia	Scandentia	Tupaiidae	ADNg	6	?
Paresseux[3]	*Bradypus infuscatus*	Mammalia	Edentata	Bradypodidae	ADNg	6	?
Tamanoir[3]	*Tamandua tetradactyla*	Mammalia	Edentata	Myrmecophagidae	ADNg	6	?
Pangolin[3]	*Manis javanica*	Mammalia	pholidota	Manidae	ADNg	6	?

[1] Tissus obtenus auprès de Marc Girondot (Professeur à Paris XI).
[2] Tissus obtenus auprès du Muséum national d'Histoire naturelle.
[3] Tissus obtenus auprès du zoo de Vincennes (Muséum national d'Histoire naturelle).
[4] Tissus obtenus auprès d'un individu capturé et sacrifié au laboratoire.
[5] Tissus obtenus auprès du "Centre de Recherche sur les Mammifères Marin" de La Rochelle.

ANNEXE 7 – Amorces de PCR utilisées pour amplifier l'amélogénine de reptiles (1ère moitié) et de mammifères (2ème moitié)

Nom de l'amorce	Séquence **5' –** **3'**	Orientation	**Exon**	**Emplacement sur l'exon**
AML2	CTGGACTTTGGTTATGTGCC	S	E2	5'
AML2.2	GACTTTGGTTATGTGCCTCTT	S	E2	5'
AML3	CTGGTTATGTCAACTTCAGCTATGAG	S	E3	3'
AML3.2	CTGGTTATGTCAACTTCAGTTATGAG	S	E3	3'
AML5	YTWWCACCTTTGAAATGGTA	S	E5	5'
AML_LIZ	TATGAGCCCATGGGAGGGTGG	S	E6	5'
AML_CROC	TATGAGCCCATGGGAGGATGG	S	E6	5'
AML_CAIM	TATGGGTATGAACCAATGG	S	E6	5'
AMLA2	TATGGBTATGARCCHATGGGA	S	E6	5'
AMLTyr6	GTATCCAYSTTATGGYTATGA	S	E6	5'
AML242	TATGAACCTATGGGAGGATGGATGC	S	E6	5'
AMLC	CTTCCTCCCGYTTNGTYTTRT	A	E6	3'
AML62	CAYTTCYTCYYKCTTGGTYTT	A	E6	3'
AML529	CACTTCTTCTTGCTTGGTCTT	A	E6	3'
AMLmam1	TACGAACCCATGGGTGGATGGCTGC	S	E6	5'
AMLmam2	CACTTCCTCCCGCTTGGTCTT	A	E6	3'
AML126	GTACCCTTCCTATGGTTACGA	S	E6	5'
AML127	TACCCTTCCTATGGTTACGAG	S	E6	5'
AML127B	TACCCTTCCTATGGTTACGAA	S	E6	5'
AML129	CCCTTCCTATGGTTACGAGCC	S	E6	5'
AML500	GCCAAGCTTCCAGAGTCAGAT	A	E6	3'
AML501	TGGCCAAGCTTCCAGAGGCAG	A	E6	3'
AML533	CCACTTCCTCCCGCTTGGTCT	A	E6	3'
AMLmam242	ACGAACCCATGGGTGGATGGCTGCA	S	E6	5'
AMLmam62	CTCTTCYTCCCGCTTGGTCTT	A	E6	3'
AMLmam5	GTGCTTACCCCTCTGAAGTGG	S	E5	3'
AMLmam5b	ACCCCTCTGAAGTGGTACCAG	S	E5	5'
AMLmam6	GCACGGGAATGATTTGGTGGT	A	E6	5'
AMLmam6c	GTTGGGACAGCACGGGGATGA	A	E6	5'

ANNEXE 8 - Diagramme des outils bio-informatiques utilisés