Evolutionary Analysis Suggests That AMTN is Enamel-specific and a Candidate for AI

B. Gasse, J. Silvent and J.-Y. Sire

J DENT RES 2012 91: 1085 originally published online 11 September 2012
DOI: 10.1177/0022034512460551

RESEARCH REPORTS
Biological

B. Gasse, J. Silvent, and J.-Y. Sire*

Université Pierre et Marie Curie, UMR 7138, EDS research group, 7 quai St-Bernard, Case 5, 75005 Paris, France; *corresponding author, jean-yves.sire@upmc.fr

J Dent Res 91(11):1085-1089, 2012

Evolutionary Analysis Suggests That AMTN is Enamel-specific and a Candidate for AI

ABSTRACT

Molecular evolutionary analysis is an efficient method to predict and/or validate amino acid substitutions that could lead to a genetic disease and to highlight residues and motifs that could play an important role in the protein structure and/or function. We have applied such analysis to amelotin (AMTN), a recently identified enamel protein in the rat, mouse, and humans. An *in silico* search for *AMTN* provided 42 new mammalian sequences that were added to the 3 published sequences with which we performed the analysis using a dataset representative of all lineages (*circa* 220 million years of evolution), including 2 enamel-less species, sloth and armadillo. During evolution, of the 209 residues of human AMTN, 17 were unchanged and 34 had conserved their chemical properties. Substituting these important residues could lead to amelogenesis imperfecta (AI). Also, AMTN possesses a well-conserved signal peptide, 2 conserved motifs whose function is certainly important but unknown, and a putative phosphorylation site (SXE). In addition, the sequences of the 2 enamel-less species display mutations revealing that *AMTN* underwent pseudogenization, which suggests that AMTN is an enamel-specific protein.

KEY WORDS: enamel matrix proteins, amelogenesis imperfecta, mammals, pseudogene, *in silico* analysis, genetic disease.

DOI: 10.1177/0022034512460551

Received April 12, 2012; Last revision July 10, 2012; Accepted July 19, 2012

A supplemental appendix to this article is published electronically only at http://jdr.sagepub.com/supplemental.

INTRODUCTION

Amelogenesis imperfecta (AI) is a genetic disease characterized by abnormal formation and/or mineralization of enamel, which is prone to wear and breakage in both primary and permanent dentitions. Without other non-oral syndromic phenotypes, these disorders are mostly due to mutations of genes encoding functionally important proteins involved in enamel formation (Witkop, 1988; Hart *et al*., 2002; Aldred *et al.*, 2003; Wright *et al.*, 2011). However, in some cases the genetic cause of AI is not found, and other candidate genes have to be unmasked (Wright *et al.*, 2011). Among the currently identified genes responsible for AI, 2 encode enamel matrix proteins (EMP) — amelogenin (AMELX) and enamelin (ENAM) — and 2 encode enamel proteases — matrix metalloproteinase 20 (MMP20) and kallikrein 4 (KLK4). This statement suggests that the other EMP genes, ameloblastin (*AMBN*), odontogenic ameloblast-associated (*ODAM*), and amelotin (*AMTN*), could also be candidates for AI.

Amelotin (*AMTN*) was cloned first in the mouse (Iwasaki *et al*., 2005) then in the rat (Moffatt *et al*., 2006). In humans, *AMTN* is localized on chromosome 4, close to *AMBN* and *ENAM*, in a locus that was associated with AI. *In situ* hybridization and immunohistochemistry experiments revealed the temporal and spatial localization of transcripts and protein during rodent odontogenesis, but the role of AMTN remains unclear (Iwasaki *et al.*, 2005; Moffatt *et al*., 2006; Trueb *et al*., 2007; Gao *et al.*, 2010; Nishio *et al*., 2010; Somogyi-Ganss *et al*., 2011). It has been suggested, however, that AMTN could play a role at the ameloblast-enamel interface during the late stages of enamel mineralization.

Evolutionary analysis is an efficient method for (i) predicting and/or validating residue substitutions that could lead to a genetic disease and (ii) highlighting amino acids and motifs that play an important role in the protein structure and/or function. We successfully applied this method to AMELX and ENAM (Delgado *et al.*, 2005, 2007; Al-Hashimi *et al*., 2009). Here, we have targeted AMTN using 43 sequences representing approximately 220 million years (Ma) of mammalian evolution (Madsen, 2009). We also studied the AMTN sequence in 2 enamel-less species, to learn whether enamel loss in the ancestors resulted in the inactivation of the encoding gene.

MATERIALS & METHODS

Sequence Dataset and Alignment

Three published (mouse, rat, and human) and 42 unpublished mammalian *AMTN*s (*i.e.*, computer-predicted from sequenced genomes) were extracted from databases. We completed some sequences by exploring the *NCBI Trace Archives* repository. In total, 45 nucleotide sequences, representative of all

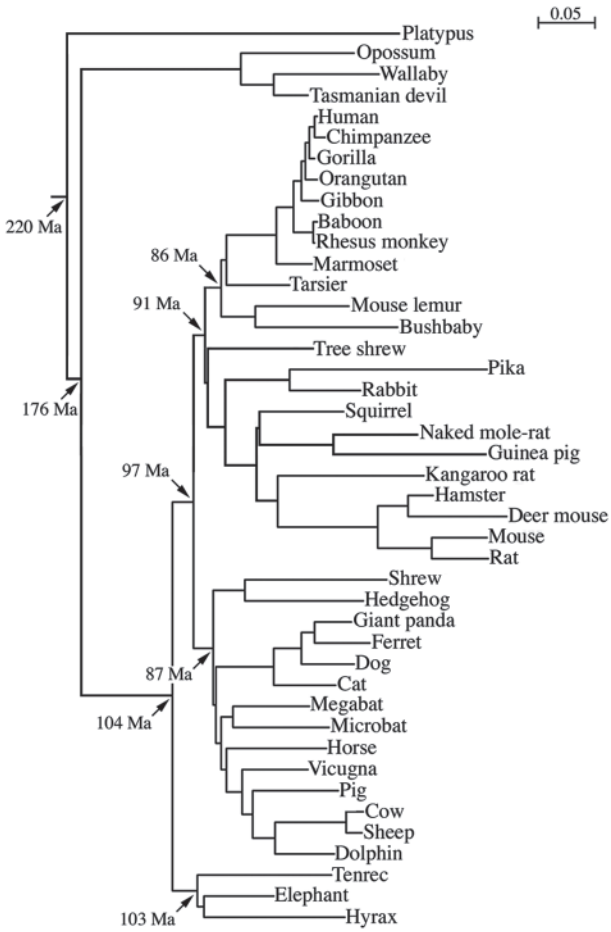

Figure 1. Distance tree of the 43 mammalian AMTNs obtained with *HyPhy* software. The branch length illustrates the evolutionary rate for each taxon. When the evolutionary rate of AMTN in a lineage leading to a taxon was high, this is represented with a long branch (*e.g.*, in platypus, pika, guinea pig) while a slow evolutionary rate is indicated with a short branch (*e.g.*, primates). The scale is indicated at the top. Divergence dates (Ma = million years) from Madsen (2009) and Murphy and Eizirik (2009).

mammalian lineages, were obtained (Appendix 1). Nucleotide sequences were translated into protein sequences and aligned with *Clustal X 2.0* and *Se-Al v2.0a11*. Human AMTN was chosen as the reference sequence in our alignment, which resulted in a 225-amino-acid sequence, when indels were included (Appendix 2). Only 0.6% of the data were missing. The 2 xenarthran sequences (armadillo and sloth) were sequenced *de novo* (see primers in Appendix 3) and deposited in Genbank (Accession numbers JX523656 and JX523657, respectively). These sequences were analyzed separately because they showed variations probably related to a lack of enamel in both species.

Evolutionary Analyses

The appropriate substitution model for our dataset — TN93 (Tamura and Nei, 1993) — was defined with an online automatic model selection tool (http://www.datamonkey.org: Delport *et al.*, 2010).

To calculate the evolutionary distances among the 43 *AMTN* sequences and to test their relevance for the analysis, we established a distance matrix using the *HYpothesis testing using PHYlogenies* (*HyPhy*) software (http://www.datam0nk3y.org/hyphy/doku.php: Pond and Frost, 2005), and taking into account (i) a recent phylogeny (Meredith *et al.*, 2011a) using *MacClade 4.08* (http://macclade.org) (Maddison and Maddison, 2005), (ii) the maximum likelihood method, and (iii) the TN93 substitution model. A distance tree was obtained, as was the putative ancestral AMTN sequence at each evolutionary node.

Sloth and armadillo AMTN sequences were aligned with the ancestral placental sequences.

A sliding window analysis of nucleotide sequence variability was conducted with *HyPhy* to identify strong functional constraints acting on *AMTN*. The Ln of substitution rate *per* site along each branch of the mammalian phylogeny was estimated, for a window of 15 bp with an overlap of 5 bp between windows.

The selective pressures acting at each site of *AMTN* were identified with the codon-based *SLAC* method (*Single Likelihood Ancestor Counting*) in *HyPhy*. With TN93, this method determines non-synonymous substitutions (dN) in the alignment and compares them with the number of non-synonymous changes expected at random in the absence of selective pressure.

We also used this method to determine the positions subjected to positive selection. One codon is considered positively selected when $dN > dS$. The test assumes that, under a neutral hypothesis, the probability for a random substitution to be synonymous is $p = ES/(ES+EN)$, where EN and ES represent the expected number of normalized non-synonymous and synonymous substitutions. The result is considered significant when $p < 0.1$.

Post-translationally modified sites on human and placental ancestral sequences were searched in the *Prosite* database (http://prosite.expasy.org: Sigrist *et al.*, 2010). Putative signal peptides were analyzed with SignalP 4.0 (http://www.cbs.dtu.dk/services/SignalP: Petersen *et al.*, 2011).

RESULTS

Sequence Comparison

The sequences of the 42 new *AMTN* genes were validated through alignment with the mouse, rat, and human sequences, including the intron splicing sites. Only 3 sequences — mouse lemur, hyrax, and wallaby — were slightly incomplete, but with no consequence for our analyses (Appendix 2). The gene structure of *AMTN* with 9 exons, the first being non-coding (not shown), is conserved in the 45 species, with the exception of the platypus, in which exons 4 and 5 were not found in the sequenced genome. The only variation concerns the C-terminal region, in which the stop codon is present at the end of the region encoded by exon 8 in baboon, macaque, and kangaroo rat AMTN. Each exon has a similar length in each species, with a few exceptions resulting from insertions or deletions (indels) that have occurred independently. Most indels are located in the large and variable exon 8. Insertions are short (1 or 2 residues), except for 5 amino acids inserted in dolphin AMTN. Deletions are larger (*e.g.*, rabbit, pika, bushbaby). These indels contribute to sequence length variations from 195 residues in rabbit to 217 in dolphin (Appendix 2). In the primate lineage, the first 3

residues encoded by exon 8 were lost in the common ancestor of the old world monkeys (Catarrhini).

Evolutionary Analyses

The distance tree indicates that the evolutionary rate of *AMTN* is relatively constant in all lineages, which means that mutations accumulate at the same rate, with a few exceptions: the high rate in platypus and rodent *AMTNs*, and the low rate in primates, except in lemurs (Fig. 1).

The analysis of purifying selection revealed 17 conserved positions, *i.e.*, in which the amino acids were not changed during more than 220 Ma of mammalian evolution (Fig. 2; Appendix 2). These residues are of prime importance for AMTN structure and/or function. In addition, 34 positions are conservative, *i.e.*, they can be substituted, but only with an amino acid that does not change its functionality. If the chemical property of the replaced residue differs from that of the former one in that position, the substitution may result in a disorder. These 51 important positions are mainly located in regions encoded by exon 2 (10 positions out of 18), exon 4 (11 out of 22), exon 6 (7 out of 12), and exon 7 (6 out of 9). The other positions (158 in humans) are considered variable, *i.e.*, the amino acid can be substituted with another residue without consequence for the protein.

The intronic sequences that border the exons are well-conserved, which indicates their crucial importance for intron splicing, and which, again, would lead to a disorder when changed (Appendix 4).

The analysis of the putative untranslated 5′ *AMTN* sequences (not shown) did not reveal the presence of additional translation initiation sites (ATG) that would be in a correct context. The length of the putative signal peptide (SP) was the same (16 residues) for each AMTN sequence, and *SignalP 4.0* analysis indicated that the SP was correct. In each sequence, the cleavage site is predicted to occur before the 2 last residues encoded by exon 2 (Fig. 2; Appendix 2).

The sliding window analysis conducted with *HyPhy* identified several regions with strong functional constraints located in exons 2, 4, 6, and 7 (Fig. 3A). The selective pressures acting at each position were identified in the same regions with the *SLAC* method (Fig. 3B).

Therefore, the 3 methods of evolutionary analysis (purifying selection, sliding window, and *SLAC*) pointed to the same, functionally important residues, most of them being regrouped into 4 short, conserved regions of the protein: the hydrophobic region [LL(*FC*)LLG] of the SP; the region coded by the first half of exon 4 [V(*F*)PSL(*SL*)IPLTQM]; the region encoded by the 3′ end of exon 5 and most of exon 6 [HVLPI(*F*)VT(*QL*)G]; and the putative phosphorylation site in exon 7 [GT(*IL*)SSEE]. These conserved regions were not found referenced in the *Prosite* database as protein domains, and screening the AMTN sequences for post-translationally modified sites did not reveal conserved sites of importance, except the putative, well-conserved phosphorylation site cited above. Several conservative residues are distributed along the AMTN sequence as, for instance, 6 prolines (P) in the

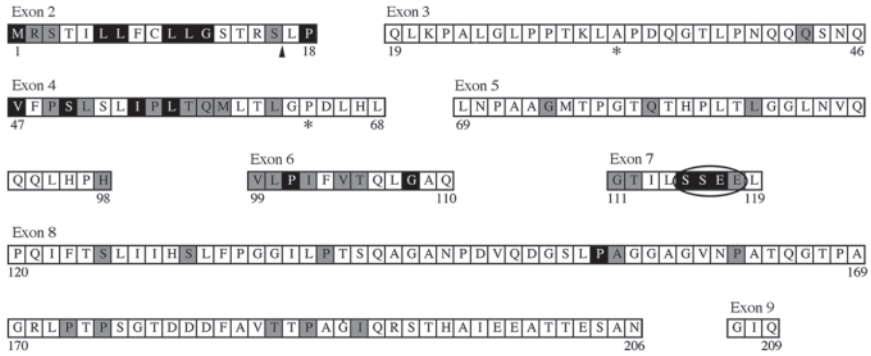

Figure 2. Amino acid sequence of human AMTN displaying the selected residues during mammalian evolution. The sequence is composed of 209 residues, including the putative signal peptide encoded by exon 2. The arrowhead points to the cleavage site of the signal peptide. Conserved positions, *i.e.*, subjected to purifying selection during 220 Ma of mammalian evolution, are represented on a black background. Residues on a gray background are conservative positions where the amino acid can be replaced by a residue possessing the same properties. In humans, we predict that substitutions of unchanged and conservative residues could lead to amelogenesis imperfecta. The 2 positively selected residues identified in our analysis are indicated with an asterisk. Several positions are well-conserved in the signal peptide, and 2 conserved motifs are identified in the AMTN regions encoded by exon 4 and exon 6. The putative phosphorylation site (SXE) is well-conserved (surrounded by an oval).

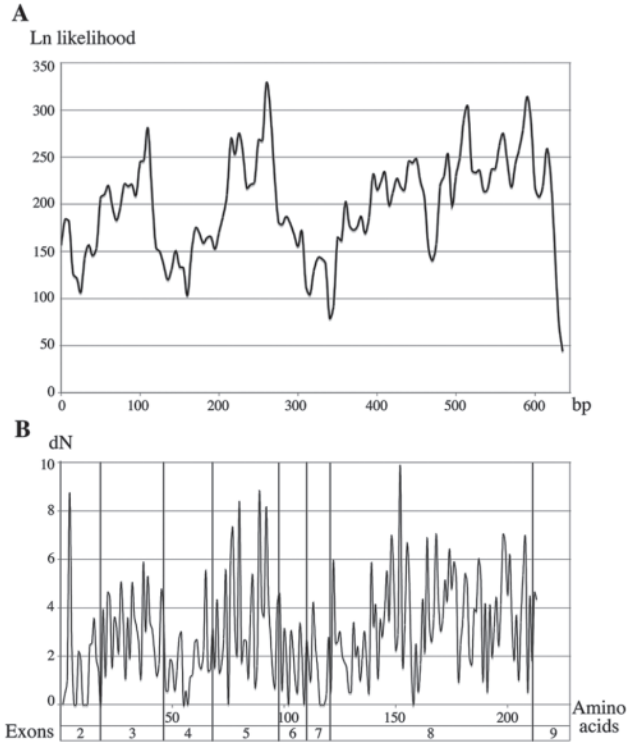

Figure 3. Substitution rate analysis of 43 mammalian AMTNs. **(A)** Ln of substitution rate *per* site along each branch of the mammalian phylogeny. Low substitution rates reflect unchanged base pairs. The peaks oriented downward indicate regions under strong functional constraints. **(B)** Analysis of non-synonymous substitution rate (dN) in the full-length amino acid AMTN sequences.

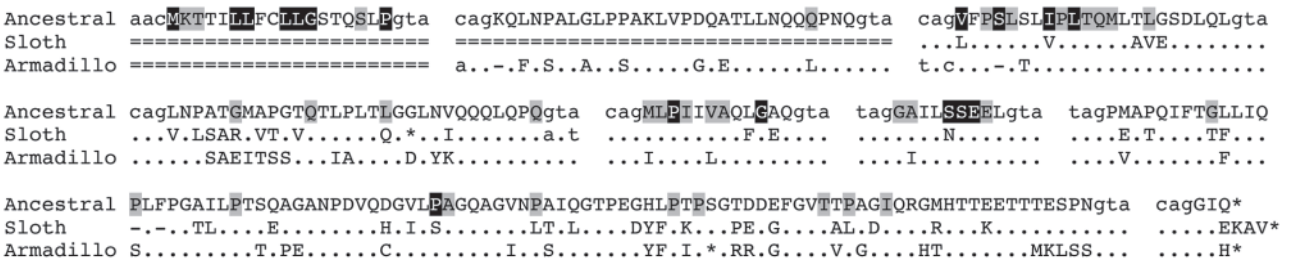

Figure 4. Alignment of the partial AMTN amino acid sequences of sloth (*Choloepus hoffmanii*) and armadillo (*Dasypus novemcinctus*), 2 xenarthran enamel-less species, with the ancestral AMTN sequence of placentals calculated with *HyPhy*. The crucial nucleotides of splice donor and acceptor sites are indicated in lower case letters. Important amino acids for AMTN function and/or structure are indicated in black (unchanged positions) and grey (conservative positions) background on the ancestral sequence (see Fig. 2). Both xenarthran AMTN sequences display several substitutions on important positions. In sloth AMTN, a single-nucleotide deletion occurred in exon 5, leading to a reading frameshift and to a premature stop codon generating a truncated protein. In addition, the splice acceptor site of exon 5 is mutated, and 1 amino acid of the conserved putative phosphorylation site in exon 7 is substituted. In armadillo AMTN, the splice donor site of exon 4 was mutated. Moreover, a reading frameshift in the C-terminal region of exon 8 generated a premature stop codon. Taken together, these observations indicate a relaxed functional pressure on these sequences, and thus pseudogenization. (.) residue identical to the ancestral AMTN residue; (-) indel; (=) unknown amino acid; * stop codon.

sequence encoded by exon 8; one of them, P^{54}, was unchanged during mammalian evolution.

In addition, our analysis revealed 2 sites that were putatively positively selected (Fig. 2): position 32, in which the ancestral valine is replaced with alanine in carnivores and in primates; and position 64, in which the ancestral serine is replaced with a proline in simiiform primates.

Enamel-less Mammals

The alignment of sloth and armadillo sequences with the ancestral AMTN of placental mammals clearly indicates that mutations have occurred in several positions that were found important for AMTN function and in some splice sites (Fig. 4). In addition, sloth *AMTN* displays a reading frameshift in exon 5 and armadillo *AMTN* in exon 8, both resulting in a premature stop codon. Our analysis indicated that these mutations have occurred independently in the 2 species lineages.

DISCUSSION

Amelotin is Possibly a Candidate for Amelogenesis Imperfecta

This study extends our knowledge from 3 to 45 mammalian AMTN sequences. The well-conserved gene structure indicates that this organization was present in the last common mammalian ancestor, and that the protein function was already defined 220 Ma (Madsen, 2009). Given the important function of the region encoded by exon 4 (see below), the lack of exons 4 and 5 in platypus *AMTN* is surprising. Either AMTN function is ensured by another enamel matrix protein or, in this particular region, platypus genome assembly is not correct. Re-sequencing this region or sequencing *AMTN* in the echidnas, the sister lineage of platypus in monotremes, could answer this question.

The distance tree does not reveal particular features. The constant substitution rate in most lineages indicates that AMTN evolved similarly and suggests the presence of variable positions, which explains the high evolutionary rates in species displaying rapid generation times (*e.g.*, rodents).

Our study highlighted 4 specific regions. The signal peptide sequence displayed an identical length, and 5 residues of the hydrophobic region were unchanged, indicating a strong functional constraint in AMTN secretion. The SP of the other EMPs studied, AMELX and ENAM, were also conserved (Delgado *et al*., 2007; Al-Hashimi *et al*., 2009), and this characteristic was shared by all members of the secretory calcium-binding phosphoprotein (SCPP) family (Kawasaki and Weiss, 2003). The putative casein kinase II phosphorylation site (SXE), already identified in previous studies, was also unchanged in the region encoded by exon 7 (SSE, positions 115-117 in human AMTN), which strongly suggests the crucial importance of this site. The presence of a SXE motif is also a characteristic of most SCPPs, in which, however, it is generally located in the N-terminal region (Kawasaki and Weiss, 2003). This indicates that the 2 phosphorylation sites are not homologous. Understanding the evolutionary origin of AMTN among SCPPs would reveal whether this site was inherited from an ancestral protein, from which it derived after duplication, or was created in this region, then recruited for its function. Two other conserved motifs were highlighted in our analyses; they certainly play an important functional role, but they are not referenced as known protein domains in databases.

In total, 51 positions (out of 209 in the human sequence) were found conserved or conservative during 220 Ma of mammalian evolution (purifying selection). These positions are supposed to be critical for the biological activity of AMTN. They were indicated on the human sequence and could be used as predictors of AMTN function and used as a proxy in case AMTN is involved in AI. Intron splicing sites were also well-conserved. The 158 non-conservative, variable positions identify regions that are less critical for AMTN function. Except for a few positions, the region encoded by exon 8 largely accounts for this variability. In particular, the C-terminal extremity is highly variable and, therefore, does not appear to be required for the protein function. Indeed, in some species, a stop codon occurs prematurely, leading to a slightly different 3′UTR region and, in some cases, to the non-translation of exon 9.

In contrast to the other EMPs studied, in the *AMTN* sequence there were neither amino acid repeats (as in *AMELX*: Delgado

et al., 2005; Sire *et al*., 2007) nor exon repeats (*ENAM, AMBN*: Sire *et al*., 2007; Al-Hashimi *et al*., 2009), which could indicate possible functional adaptation. However, in some lineages, the AMTN sequence may undergo subtle changes that could be revealed through positive selection, as for the 2 positions identified in this study.

Therefore, *AMTN* could be a candidate for AI and should be included as a target gene in the diagnosis of this genetic disease. AI inheritance is probably autosomal-recessive as for *ENAM, MMP20*, and *KLK4* (Wright *et al*., 2011). Two copies of the altered gene need to be inherited from the parents, a probability that is weak, except for in-bred families and/or frequent substitutions of important AMTN residues in human populations.

Amelotin is Possibly Enamel-specific

The 2 xenarthrans studied lacked enamel, which was lost more than 40 Ma and farther back in time in sloth than in armadillo lineages (Davit-Béal *et al*., 2009; Meredith *et al*., 2009). Such species are true experiments in nature, allowing us to understand the fate of enamel-related protein genes. There is strong evidence that the 2 xenarthran *AMTN*s are subjected to pseudogenization, and that the invalidation started after the divergence of the 2 lineages from their common ancestor, which possessed enamel and undoubtedly functional AMTN. After *AMELX* (Sire *et al*., 2008), *ENAM* and *AMBN* (Deméré *et al*., 2008; Meredith *et al*., 2009; Al-Hashimi *et al*., 2010), and *MMP20* (Meredith *et al*., 2011b), *AMTN* is the fifth enamel-related protein gene that is being invalidated in enamel-less or edentulous species. AMTN function is therefore predicted to be enamel-specific.

ACKNOWLEDGMENTS

We thank the following colleagues for the generous gift of xenarthran tissues (*via* Drs. F. Delsuc and F. Catseflis, ISEM, Université Montpellier II, France): Dr. John Trupkiewicz (Philadelphia Zoo, PA 33694, USA) for *Choloepus hoffmanni* and Dr. W. Jim Loughry (Valdosta State University, Georgia, USA) for *Dasypus novemcinctus*. This study was financially supported by grants from CNRS and Université Pierre & Marie Curie to UMR7138, and by the 2008-2011 National "Programme Hospitalier de Recherche Clinique" (PHRC) from the French Ministry of Health, *via* the "Hôpitaux Universitaires de Strasbourg". The authors declare no potential conflicts of interest with respect to the authorship and/or publication of this article.

REFERENCES

Al Hashimi N, Sire JY, Delgado S (2009). Evolutionary analysis of mammalian enamelin, the largest enamel protein, supports a crucial role for the 32 kDa peptide and reveals selective adaptation in rodents and primates. *J Mol Evol* 69:635-656.

Al Hashimi N, Lafont AG, Delgado S, Kawasaki K, Sire JY (2010). The enamelin genes in lizard, crocodile and frog, and the pseudogene in the chicken provide new insights on enamelin evolution in tetrapods. *Mol Biol Evol* 27:2078-2094.

Aldred MJ, Savarirayan R, Crawford PJ (2003). Amelogenesis imperfecta: a classification and catalogue for the 21st century. *Oral Dis* 9:19-23.

Davit-Béal T, Tucker A, Sire JY (2009). The loss of ability to form teeth and enamel in living tetrapod taxa: genetic data and morphological adaptations. *J Anat* 214:477-501.

Delgado S, Girondot M, Sire JY (2005). Molecular evolution of amelogenin in mammals. *J Mol Evol* 60:12-30.

Delgado S, Ishiyama M, Sire JY (2007). Validation of amelogenesis imperfecta inferred from amelogenin evolution. *J Dent Res* 86:326-330.

Delport W, Scheffler K, Botha G, Gravenor MB, Muse SV, Kosakovsky Pond SL (2010). CodonTest: modeling amino acid substitution preferences in coding sequences. *PLoS Comput Biol* 6:e1000885.

Deméré TA, McGowen MR, Berta A, Gatesy J (2008). Morphological and molecular evidence for a stepwise evolutionary transition from teeth to baleen in mysticete whales. *Syst Biol* 57:15-37.

Gao Y, Wang W, Sun Y, Zhang J, Li D, Wei Y, *et al*. (2010). Distribution of amelotin in mouse tooth development. *Anat Rec* 293:135-140.

Hart PS, Aldred M, Crawford P, Wright N, Hart T, Wright JT (2002). Amelogenesis imperfecta phenotype-genotype correlations with two amelogenin gene mutations. *Arch Oral Biol* 47:261-265.

Iwasaki K, Bajenova E, Somogyi-Ganss E, Miller M, Nguyen V, Nourkeyhani H, *et al*. (2005). Amelotin — a novel secreted, ameloblast-specific protein. *J Dent Res* 84:1127-1132.

Kawasaki K, Weiss KM (2003). Mineralized tissue and vertebrate evolution: the secretory calcium-binding phosphoprotein gene cluster. *Proc Natl Acad Sci USA* 100:4060-4065.

Maddison DR, Maddison WP (2005). MacClade4 Version 4.08. Sunderland, MA, USA: Sinauer Associates, Inc.

Madsen O (2009). Mammals. In: The time tree of life. Hedges SB, Kumar S, editors. New York, NY: Oxford University Press, pp. 459-461.

Meredith RW, Gatesy J, Murphy WJ, Ryder OA, Springer MS (2009). Molecular decay of the tooth gene enamelin (ENAM) mirrors the loss of enamel in the fossil record of placental mammals. *PLoS Genet* 5:e1000634.

Meredith RW, Janecka JE, Gatesy J, Ryder OA, Fisher CA, Teeling EC, *et al*. (2011a). Impacts of the Cretaceous terrestrial revolution and KPg extinction on mammal diversification. *Science* 334:521-524.

Meredith RW, Gatesy J, Cheng J, Springer MS (2011b). Pseudogenization of the tooth gene enamelysin (*MMP20*) in the common ancestor of extant baleen whales. *Proc Biol Sci* 278:993-1002.

Moffatt P, Smith CE, St-Arnaud R, Simmons D, Wright T, Nanci A (2006). Cloning of rat amelotin and localization of the protein to the basal lamina of maturation stage ameloblasts and junctional epithelium. *Biochem J* 399:37-46.

Murphy WJ, Eizirik E (2009). Placental mammals (Eutheria). In : The time tree of life. Hedges SB, Kumar S, editors. New York, NY: Oxford University Press, pp. 471-474.

Nishio C, Wazen R, Kuroda S, Moffatt P, Nanci A (2010). Expression pattern of odontogenic ameloblast-associated and amelotin during formation and regeneration of the junctional epithelium. *Eur Cells Mater* 20:393-402.

Petersen TN, Brunak S, von Heijne G, Nielsen H (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* 8:785-786.

Pond SL, Frost SD (2005). Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics* 21:2531-2533.

Sigrist CJ, Cerutti L, de Castro E, Langendijk-Genevaux PS, Bulliard V, Bairoch A, *et al*. (2010). PROSITE, a protein domain database for functional characterization and annotation. *Nucleic Acids Res* 38:D161-D166.

Sire JY, Davit-Béal T, Delgado S, Gu X (2007). The origin and evolution of enamel mineralization genes. *Cells Tissues Organs* 186:25-48.

Sire JY, Delgado S, Girondot M (2008). Hen's teeth with enamel cap: from dream to impossibility. *BMC Evol Biol* 8:246.

Somogyi-Ganss E, Nakayama Y, Iwasaki K, Nakano Y, Stolf D, McKee MD, *et al*. (2011). Comparative temporospatial expression profiling of murine amelotin protein during amelogenesis. *Cells Tissues Organs* 195:535-549.

Tamura K, Nei M (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol* 10:512-526.

Trueb B, Taeschler S, Schild C, Lang NP (2007). Expression of phosphoproteins and damelotin in teeth. *Int J Mol Med* 19:49-54.

Witkop CJ Jr (1988). Amelogenesis imperfecta, dentinogenesis imperfecta and dentin dysplasia revisited: problems in classification. *J Oral Pathol* 17:547-553.

Wright JT, Torain M, Long K, Seow K, Crawford P, Aldred MJ, *et al*. (2011). Amelogenesis imperfecta: genotype-phenotype studies in 71 families. *Cells Tissues Organs* 194:279-283.