Available online at www.sciencedirect.com
ScienceDirect
Molecular Phylogenetics and Evolution 47 (2008) 865–869
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

Short Communication

Amelogenin, the major protein of tooth enamel: A new phylogenetic marker for ordinal mammal relationships

Sidney Delgado [a], Nicolas Vidal [b], Géraldine Veron [c], Jean-Yves Sire [a,*]

[a] *UMR 7138, Equipe "Evolution et développement du squelette", Université Pierre & Marie Curie—Paris 6, 7 quai St-Bernard, Case 05, 75252 Paris, France*
[b] *UMR 7138, Equipe "Phylogénie", Muséum national d'Histoire naturelle, Paris, France*
[c] *UMR 5202, Unité "Origine, Structure et Evolution de la Biodiversité", Muséum national d'Histoire naturelle, Paris, France*

Received 3 May 2007; revised 14 January 2008; accepted 23 January 2008
Available online 2 February 2008

1. Introduction

Teeth and their tissues, dentin and enamel, have a long, well-defined history. Their origin was traced back to the extra-oral dermal skeleton of early jawless vertebrates, approximately 500 million years ago, mya (see reviews in Huysseune and Sire, 1998; Smith and Coates, 2000; Sire and Huysseune, 2003). Once recruited into the mouth in early gnathostomes, circa 450 mya, teeth were subjected to strong selective pressure due to their crucial function. This explains why teeth, and particularly their developmental processes, organization and structural components, were conserved nearly unchanged through geological times.

In mammals, as in most tetrapod taxa, teeth are covered by a thick and highly mineralized, protective tissue, enamel. The amelogenin gene (*AMEL*) encodes the major protein of enamel (90% of the organic matrix). Recent molecular analyses have brought insights into the evolutionary pattern of *AMEL* in mammals (Delgado et al. 2005) and have shown that the history of this protein could have started by the end of the Precambrian period (Sire et al., 2007). Comparative studies of *AMEL* in mammals, reptiles and amphibians have revealed highly conserved residues located at the C- and N-terminal regions and have indicated that a large part of the hydrophobic, central region of the molecule, encoded by the largest exon 6, was more variable (Ishiyama et al., 1998; Toyosawa et al., 1998; Delgado et al., 2005). Because *AMEL* is X-linked in many mammal lineages, the gene as a whole is predicted to be particularly strongly conserved (under Ohno's rule in general, Ohno, 1967, and because of the X's bias toward transmission through the slowly mutating female mammal germline, Li et al., 2002). In eutherians, *AMEL* was shown to span an ancient pseudoautosomal boundary on the X-chromosome, exon 6 being a formerly pseudoautosomal segment of the gene (Iwase et al., 2003). This additional stringency at this particular location may have reinforced the conservation of exon 6 sequence because recombination has been shown to have little effect on the rate of sequence divergence in this pseudoautosomal boundary among humans and great apes (Yi et al., 2004). This possibility has been also discussed in a recent article (Richard et al., 2007).

Both functional constraints and sequence variation indicate that *AMEL*, and particularly the variable region, could contain a useful phylogenetic signal for deep cladogenetic events, even if exon 6, the only exon easily retrieved using PCR, is rather short (approximately 400 bp). We have therefore tested the utility of this region of *AMEL* for inferring a mammalian phylogeny above the family level.

Comparative genomic data from mammals have accumulated rapidly in the recent past and have contributed significantly to resolving long-standing phylogenetic controversies. Mitochondrial then nuclear DNA sequence analyses revealed new interordinal mammalian relationships (e.g., Springer et al., 1997; Stanhope et al., 1998; Madsen et al., 2001; Murphy et al., 2001; Delsuc et al., 2002; Waddell and Shelley, 2003). Four superordinal eutherian clades are recognized: Laurasiatheria (six orders: Cetartiodactyla, Perissodactyla, Carnivora, Pholidota, Chiroptera and Eulipotyphla), Euarchontoglires (five orders: Primata, Dermoptera, Scandentia, Rodentia and

* Corresponding author. Fax: +33 1 44 27 35 72.
E-mail address: sire@ccr.jussieu.fr (J.-Y. Sire).

1055-7903/$ - see front matter
doi:10.1016/j.ympev.2008.01.025

Lagomorpha), Xenarthra, and Afrotheria (six orders: Macroscelidea, Afrosoricida, Tubulindentata, Sirenia, Hyracoidea, Proboscidea). Eighteen eutherian (placentals) orders were defined, to which are added the metatherian order (marsupials) and the prototherian order (monotremes) to give a total of 20 orders encompassing all extant mammalian species (Waddell and Shelley, 2003).

Most of the recent molecular phylogenies confirm these relationships (e.g., Amrine-Madsen et al., 2003; Springer et al., 2003; Hallström et al., 2007; Murphy et al., 2007). However, controversies still persist both among molecular phylogenies and when comparing these data to evolutionary relationships based on morphology. This is particular pertinent to several superordinal eutherian relationships such as between Afrotheria, Xenarthra and Boreoeutheria (Euarchontoglires + Laurasiatheria), although monophyly of Afrotheria was recently supported by morphological features (Sanchez-Villagra et al., 2005; Tabuce et al., 2007). It is clear, however, that more nuclear data are required as early placental divergences may have been compressed in time (Kriegs et al., 2006; but see Bininda-Emonds et al., 2007 for further discussion on the diversification of today's mammals).

In the present study, we have used 55 sequences of the amelogenin exon 6 from species representative of all main mammalian lineages. We show that *AMEL* exon 6 is an additional efficient marker for ordinal mammal relationships.

2. Material and methods

The species and accession numbers of *AMEL* sequences used in this study are listed in Table 1. Eighteen sequences were found in databases. The other sequences were obtained from genomic DNA extracted from either frozen or ethanol-preserved soft tissues (kidney, liver, spleen, skin) using the DNeasy Tissue System kit (Qiagen). The source of material is indicated in "Acknowledgments" section.

AMEL exon 6 was amplified using the following primers: *Mam1* (*sense*: 5′-TACGAACCATGGGTGGATGGC TGC-3′) or *Mam3* (*sense*: 5′-TACCCTTCCTATGGTTAC GAG-3′) to hybridize the 5′ region, and *Mam2* (*antisense*: 5′-CACTTCCTCCCGCTTGGTCTT-3′) or *Mam4* (*antisense*: 5′-GCCAAGCTTCCAGAGTCAGAT-3′) to hybridize the 3′ region.

Amplification was performed in 38 cycles, each cycle comprising: 1 min denaturation at 94 °C, 1 min annealing at 59 °C and 1 min extension at 72 °C. The final extension was for 30 min at 72 °C. Sequencing of PCR products was done by Genome Express S.A.

Sequences were aligned manually using the editor Se-Al software (Rambaut, 1996) and amino acid properties were used. Resulting gaps were treated as missing data in all analyses. The 5′ (36 first bp) and 3′ (21 last bp) regions of exon 6 are highly conserved and were deleted. This resulted in 567 sites for 55 taxa (322 variable sites, 203 of which are informative). The alignment is available upon request.

Table 1
Species studied (55 taxa)

Human	Hominidae	*Homo sapiens*
Orangutan	Hominidae	*Pongo pygmaeus*
Rhesus monkey	Cercopithecidae	*Macaca mulatta*
Squirrel monkey	Cebidae	*Saimiri boliviensis*
Marmoset	Cebidae	*Callithrix jacchus*
Ring-tailed lemur	Lemuridae	*Lemur catta*
Bushbaby	Galagidae	*Otolemur garnettii*
Tree shrew	Tupaiidae	*Tupaia belangeri*
Flying lemur	Cynocephalidae	*Cynocephalus variegatus*
Mouse	Muridae	*Mus musculus*
Rat	Muridae	*Rattus norvegicus*
Hamster	Muridae	*Mesocricetus auratus*
Guinea pig	Caviidae	*Cavia porcellus*
Squirrel	Sciuridae	*Spermophilus tridecemlineatus*
Cow	Bovidae	*Bos taurus*
Goat	Bovidae	*Capra hircus*
Japanese serow	Bovidae	*Capricornis crispus*
Pig	Suidae	*Sus scrofa*
Hippopotamus	Hippopotamidae	*Hexaprotodon liberiensis*
Dolphin	Delphinidae	*Tursiops truncatus*
Porpoise	Phocoenidae	*Phocoena phocoena*
Horse	Equidae	*Equus caballus*
Tapir	Tapiridae	*Tapirus terrestris*
Rhinoceros	Rhinocerotidae	*Ceratotherium simum*
Dog	Canidae	*Canis familiaris*
Black bear	Ursidae	*Ursus americanus*
Panda	Ursidae	*Ailuropoda melanoleuca*
Gray seal	Phocidae	*Halichoerus grypus*
Sea lion	Otariidae	*Otaria byronia*
Cat	Felidae	*Felis catus*
Tiger	Felidae	*Panthera tigris*
Cheetah	Felidae	*Acinonyx jubatus*
Pangolin	Manidae	*Manis javanica*
Fruit bat	Pteropodidae	*Cynopterus brachyotis*
Flying fox	Pteropodidae	*Pteropus vampyrus*
Roundleaf bat	Rhinolophidae	*Hipposideros ater*
Microbat	Vespertilionidae	*Myotis lucifugus*
Hedgehog	Erinaceidae	*Erinaceus europaeus*
Shrew	Soricidae	*Sorex araneus*
Armadillo	Dasypodidae	*Dasypus novemcinctus*
Tamandua	Myrmecophagidae	*Tamandua tetradactyla*
Three-toed sloth	Bradypodidae	*Bradypus infuscatus*
Two-toed sloth	Megalonychidae	*Choloepus hoffmanni*
African elephant	Elephantidae	*Loxodonta africana*
Tenrec	Tenrecidae	*Echinops telfairi*
Golden mole	Chrysochloridae	*Chrysochloris asiatica*
Aardvark	Orycteropidae	*Orycteropus afer*
Hyrax	Procaviidae	*Procavia capensis*
Elephant shrew	Macroscelididae	*Elephantulus edwardii*
Manatee	Trichechidae	*Trichechus manatus*
Opossum	Didelphidae	*Monodelphis domestica*
Aquatic opossum	Didelphidae	*Chironectes minimus*
Wallaby	Macropodidae	*Macropus eugenii*
Echidna	Tachyglossidae	*Tachyglossus aculeatus*
Platypus	Ornithorhynchidae	*Ornithorhynchus anatinus*

The entries in the table are ordered alphanumerically (by Accession No: EU168848–EU168899).

We built phylogenies using probabilistic approaches with Maximum Likelihood (ML) and Bayesian methods of inference. ML analyses were performed with PAUP*4 (Swofford, 1998). Bayesian analyses were performed with MrBayes 3.1 (Ronquist and Huelsenbeck, 2003). For both approaches an appropriate model of sequence evolution

was inferred from the data themselves using ModelTest (Posada and Crandall, 1998). The model selected (Akaike Information Criterion) was the TrN+G model with substitution parameters as A–C/A–T/C–G = 1, A–G = 4.1029, and C–T = 2.5407, base frequencies as A = 0.2398, C = 0.4136, G = 0.1682 and T = 0.1785, and a Γ parameter of 0.707. Bayesian analyses were run with model parameters estimated as part of the Bayesian analyses, and the best-fit model as inferred by Modeltest. ML results are presented under the form of a bootstrap consensus tree (1000 replicates, NJ starting tree with NNI branch swapping) which is considered to be a reliable estimate of phylogeny. Bayesian analyses were performed by running 2,000,000 generations in four chains, saving the current tree every 100 generations. The last 18,000 trees were used to construct a 50% majority-rule consensus tree.

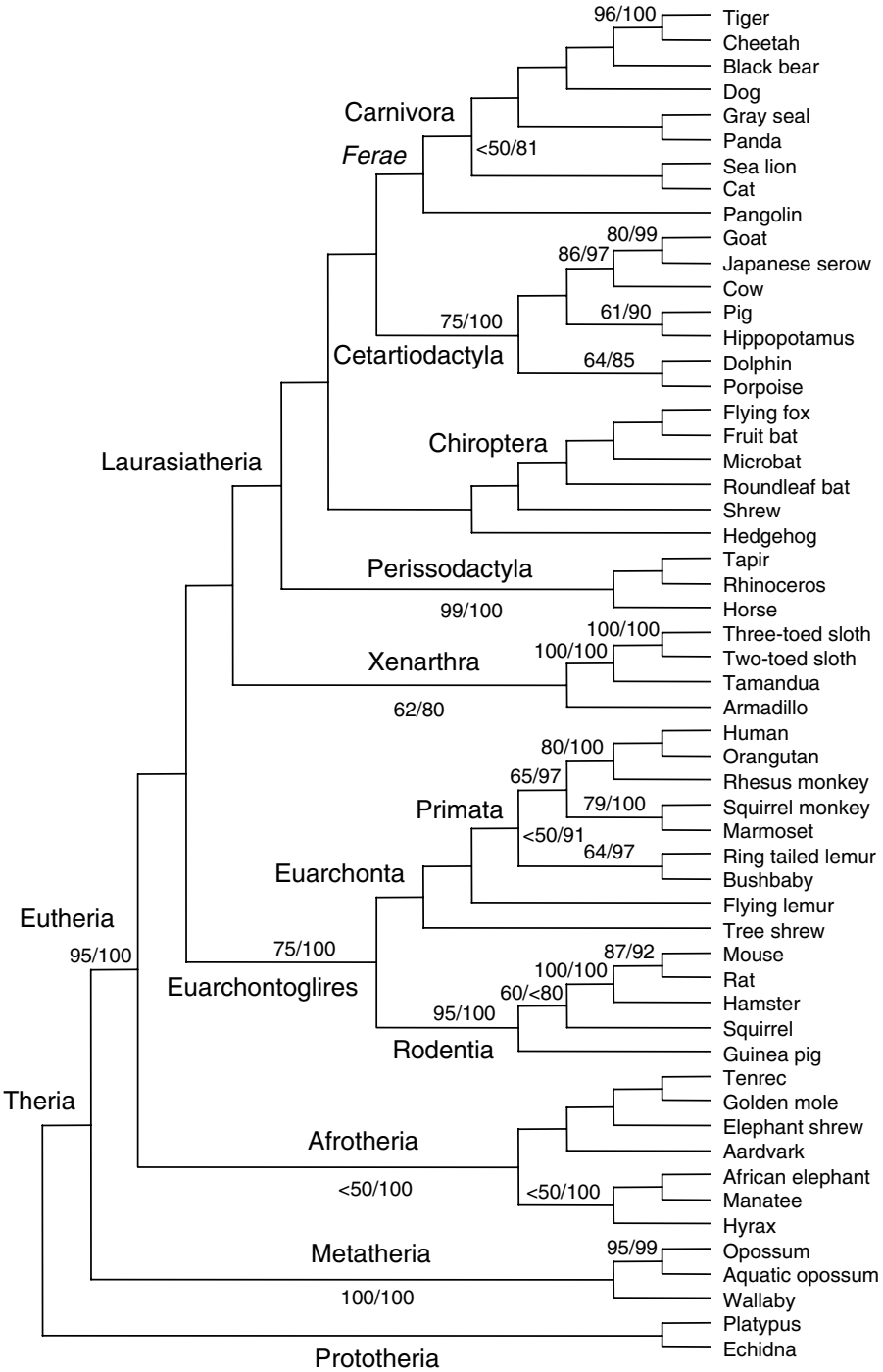

Fig. 1. Phylogenetic relationships of mammals based on *AMEL* sequences (bootstrap ML consensus tree). ML bootstrap values above 50% are shown, followed by Bayesian posterior probabilities above 80%.

3. Results and discussion

The 50% majority-rule ML bootstrap consensus tree is shown in Fig. 1. As expected based on analyses from one portion of gene only (less than 600 sites), most basal nodes show low robustness values. However, remarkable congruence with previously published molecular mammalian phylogenies is obtained.

Our analyses based on *AMEL* exon 6 from 55 species recovered the marsupial and eutherian clades with high support values.

The four superordinal clades of eutherian mammals, Laurasiatheria, Euarchontoglires, Xenarthra and Afrotheria, recognized by, e.g., Madsen et al. (2001), Murphy et al. (2001), Amrine-Madsen et al. (2003), and Nishihara et al. (2006) are all identified. Moreover, seven eutherian orders are identified as monophyletic: Carnivora, Cetartiodactyla, Chiroptera, Perissodactyla, Primata, Rodentia and Afrosoricida.

Several other higher-level mammalian relationships are found congruent with the most recent eutherian mammal phylogenies cited above: Ferae: Carnivora and Pholidota; Euarchonta: Primata, Dermoptera and Scandentia; Tethytheria: Sirenia and Proboscidea; Paenungulata: Hyracoidea and Tethytheria; Afroinsectivora: Afrosoricida and Macroscelidea; Afroinsectiphilia: Afroinsectivora and Tubulidentata.

This tree shows only a few discrepancies with other molecular phylogenies based on large concatenated DNA sequences: for instance, Cetartiodactyla is identified as Cetacea + Artiodactyla, but Hippopotamidae + Cetacea is not retrieved; also, a recent article by Hallström et al. (2007) strongly supports Xenarthra as the sister lineage to Afrotheria, reinforcing the clade 'Atlantogenata' already proposed by Delsuc et al. (2002).

These results indicate that *AMEL* exon 6, although composed of approximately 400 bp, is a very efficient phylogenetic marker for higher-level mammalian relationships that could be added to the current large data sets of DNA sequences.

Acknowledgments

We are greatly indebted to the following colleagues and laboratory collections for providing material for this work. Gray seal, tamandua, tapir, rhinoceros, manatee come from the Université de Montpellier 2, France (UMR 5554, Dr F. Catzeflis); tiger, cheetah, panda, pygmy hippopotamus, from the Zoo de Vincennes, Muséum national d'Histoire naturelle, France (Dr F. Ollivet and A. Lécu); dolphin and porpoise from the Muséum de la Rochelle, France (Dr W. Dabin); sea lion, black bear, flying lemur, tree shrew, fruit bat, roundleaf bat, shrew, three-toed sloth, tamandua, pangolin, hedgehog, tenrec, aquatic opossum, wallaby from the Laboratoire Mammifères et Oiseaux, Muséum national d'Histoire naturelle, France; African elephant, golden mole, elephant shrew, aardvark, hyrax, manatee from the University of Stellenbosch, South Africa (Dr. T.J. Robinson).

References

Amrine-Madsen, H., Koepfli, K.-P., Wayne, R.K., Springer, M.S., 2003. A new phylogenetic marker, apolipoprotein B, provides compelling evidence for eutherian relationships. Mol. Phylogent. Evol. 28, 225–240.

Bininda-Emonds, O.R.P., Cardillo, M., Jones, K.E., MacPhee, R.D.E., Beck, R.M.D., Grenyer, R., Price, S.A., Vos, R.A., Gittleman, J.L., Purvis, A., 2007. The delayed rise of present-day mammals. Nature 446, 507–512.

Delgado, S., Girondot, M., Sire, J.-Y., 2005. Molecular evolution of amelogenin in mammals. J. Mol. Evol. 60, 12–30.

Delsuc, F., Scally, M., Madsen, O., Stanhope, M.J., de Jong, W.W., 2002. Molecular phylogeny of living xenarthrans and the impact of character and taxon sampling on the placental tree rooting. Mol. Biol. Evol. 19, 1656–1671.

Hallström, B., Kullberg, M., Nilsson, M., Janke, A., 2007. Phylogenomic data analyses provide evidence that Xenarthra and Afrotheria are sistergroups. Mol. Biol. 24, 2059–2068.

Huysseune, A., Sire, J.-Y., 1998. Evolution of patterns and processes in teeth and tooth-related tissues in non-mammalian vertebrates. Eur. J. Oral Sci. 106 (suppl. 1), 437–481.

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, 467–474.

Iwase, M., Satta, Y., Hirai, Y., Hirai, H., Imai, H., Takahata, N., 2003. The amelogenin loci span an ancient pseudoautosomal boundary in diverse mammalian species. Proc. Natl. Acad. Sci. USA 100, 5258–5263.

Kriegs, J.O., Churakov, G., Kiefmann, M., Jordan, U., Brosius, J., Schmitz, J., 2006. Retroposed elements as archives for the evolutionary history of placental mammals. PLoS Biol. 4, e91.

Li, W.-H., Yi, S., Makova, K.D., 2002. Male-driven evolution. Curr. Opin. Genet. Dev. 12, 650–656.

Madsen, O., Scally, M., Douady, C.J., Kao, D.J., Debry, R.W., Adkins, R., Amrine, H.M., Stanhope, M.J., de Jong, W.W., Springer, M.S., 2001. Parallel adaptive radiations in two major clades of placental mammals. Nature 409, 610–614.

Murphy, W.J., Elzirik, E., Johnson, W.E., Zhang, Y.P., Ryder, O.A., O'Brien, S.J., 2001. Molecular phylogenetics and the origin of placental mammals. Nature 409, 614–618.

Murphy, W.J., Pringle, T.H., Crider, T.A., Springer, M.S., Miller, W., 2007. Using genomic data to unravel the root of the placental mammal phylogeny. Genome Res. 17, 413–421.

Nishihara, H., Hasegawa, M., Okada, N., 2006. Pegasoferae, an unexpected mammalian clade revealed by tracking ancient retrosposon insertions. Proc. Natl. Acad. Sci. USA 103, 9929–9934.

Ohno, S., 1967. Sex chromosomes and sex-linked genesMonographs on Endocrinology, vol. 1. Springer-Verlag, Berlin.

Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14, 818–819.

Rambaut, A., 1996. Se-Al. <http://evolve.zoo.ox.ac.uk/Se-Al/Se-Al.html>.

Richard, B., Delgado, S., Gorry, P., Sire, J.-Y., 2007. A study of polymorphism in human *AMELX*. Arch. Oral Biol. 52, 1026–1031.

Ronquist, F., Huelsenbeck, J.P., 2003. Mr Bayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.

Sanchez-Villagra, M.R., Narita, Y., Kuratani, S., 2005. Thoracolumbar vertebral number: the first skeletal synapomorphy for afrotherian mammals. Syst. Biodiversity, 1–7.

Sire, J.-Y., Davit-Béal, T., Delgado, S., Gu, X., 2007. The origin and evolution of enamel mineralization genes. Cells Tissues Organs 186, 25–48.

Sire, J.-Y., Huysseune, A., 2003. Formation of skeletal and dental tissues in fish: a comparative and evolutionary approach. Biol. Rev. 78, 219–249.

Smith, M.M., Coates, M.I., 2000. Evolutionary origins of teeth and jaws: developmental models and phylogenetic patterns. In: Teaford, M.F.,

Smith, Ferguson, M.W.J. (Eds.), Development, Function and Evolution of Teeth. Cambridge University Press, Cambridge, MA, pp. 133–151.

Springer, M.S., Cleven, G.C., Madsen, O., de Jong, W.W., Waddell, V.G., Amrine, H.M., Stanhope, M.J., 1997. Endemic African mammals shake the phylogenetic tree. Nature 388, 61–64.

Springer, M.S., Murphy, W.J., Eizirik, E., O'Brien, S.J., 2003. Placental mammal diversification and the Cretaceous–Tertiary boundary. Proc. Natl. Acad. Sci. USA 100, 1056–1061.

Stanhope, H.M., Waddell, V.G., Madsen, O., de Jong, W.W., Hedges, S.B., Cleven, G.C., Kao, D., Springer, M.S., 1998. Molecular support for multiple origins of insectivora and for a new order of endemic African insectivore mammals. Proc. Natl. Acad. Sci. USA 95, 9967–9972.

Swofford, D.L., 1998. PAUP*: phylogenetic analysis using parsimony (*and other methods), version 4. Sinauer Associates, Sunderland, Mass.

Tabuce, R., Marivaux, L., Adaci, M., Bensalah, M., Hartenberger, J.-L., Mahboubi, M., Mebrouk, F., Tafforeau, P., Jaeger, J.-J., 2007. Early Tertiary mammals from North Africa reinforce the molecular Afrotheria clade. Proc. R. Soc. Lond. B 274, 1159–1166.

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, 13056–13061.

Waddell, P.J., Shelley, S., 2003. Evaluating placental inter-ordinal phylogenies with novel sequences including RAG1, gamma-fibrinogen, ND6, and mt-tRNA, plus MCMC-driven nucleotide, amino acid, and codon models. Mol. Phylogenet. Evol. 28, 197–224.

Yi, S., Summers, T.J., Pearson, N.M., Li, W.-H., 2004. Recombination has little effect on the rate of sequence divergence in pseudoautosomal boundary 1 among humans and great apes. Genome Res. 14, 37–43.