Cells Tissues Organs

Cells Tissues Organs 2011;194:108–112
DOI: 10.1159/000324225

Published online: May 13, 2011

The SCPP Gene Family and the Complexity of Hard Tissues in Vertebrates

Kazuhiko Kawasaki

Department of Anthropology, Pennsylvania State University, University Park, Pennsylvania, Pa., USA

Key Words
Gene duplication · Vertebrate evolution · Biomineralization · Enamel · Extracellular matrix protein

Abstract
Diverse hard tissues constituted a tooth-like skeletal element in extinct jawless vertebrates. Today, similar tissues are found in our teeth. These tissues mineralize in the extracellular matrix and involve various macromolecules. Among these molecules are secretory calcium-binding phosphoproteins (SCPPs) coded by genes that arose by duplication. Although the repertoire of SCPPs may vary in different lineages, some SCPPs are unusually acidic and are thought to participate in the mineralization of a collagenous matrix, principally either bone or dentin. Other SCPPs are rich in Pro and Gln (P/Q) and are employed to form the tooth surface. In tetrapods, the tooth surface is usually covered with enamel which develops in a matrix comprised of P/Q-rich SCPPs. By contrast, the tooth surface tissue in teleosts is called enameloid and it forms in a dentin-like collagenous matrix. Despite the difference in their matrix, both enamel and enameloid mature into hypermineralized inorganic tissues. Notably, some P/Q-rich SCPP genes are primarily expressed at this stage and their proteins localize between the tooth surface and overlying dental epithelium. Moreover, an orthologous gene is used for maturation of these 2 different tissues. These findings suggest distinct roles of acidic and P/Q-rich SCPPs during the evolution of hard tissues. Acidic SCPPs initially regulated the mineralization of bone, dentin, or a similar ancient collagenous tissue through interaction with calcium ions. P/Q-rich SCPPs arose next and originally assembled a structure or a space that facilitated the hypermineralization of dentin or a dentin-like tissue. Subsequently, some P/Q-rich SCPPs were coopted for the mineralizing enamel matrix. More recently, however, many SCPP genes were lost in toothless birds and mammals. Thus, it appears that, in vertebrates, the phenotypic complexity of hard tissues correlates with gain and loss of SCPP genes.

Introduction

Paleozoic jawless vertebrates developed a tooth-like dermal skeletal element comprised of diverse hard tissues [Donoghue et al., 2006]. Similar tissues are found in our teeth today. In humans, like most toothed tetrapods, the tooth consists of 3 principal hard tissues: enamel, dentin, and bone [Hall, 2005]. Formation of these tissues involves 2 classes of macromolecules [Kawasaki et al., 2009]. One class includes fibrillar collagens, secreted protein acidic and rich in cysteine (SPARC), and many proteoglycans.

Abbreviations used in this paper

ECM	extracellular matrix
P/Q	Pro and Gln
SCPP	secretory calcium-binding phosphoprotein

KARGER

Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

1422–6405/11/1944–0108$38.00/0

Accessible online at:
www.karger.com/cto

Dr. Kazuhiko Kawasaki
Department of Anthropology, Pennsylvania State University
409 Carpenter Building
University Park, PA 16802 (USA)
Tel. +1 814 863 0032, E-Mail kuk2@psu.edu

These molecules generate collagenous extracellular matrix (ECM) for bone and dentin as well as soft tissues. The other class is more specific to mineralized tissues. This class comprises γ-carboxyglutamic acid containing proteins and secretory calcium-binding phosphoproteins (SCPPs). In particular, SCPPs are used for the mineralization of collagenous bone and dentin as well as noncollagenous enamel [Kawasaki and Weiss, 2003].

We have previously reported that SCPP genes originated from the SPARC-like 1 *(SPARCL1)* gene that was derived from *SPARC* near the origin of mineralized skeleton in vertebrates [Kawasaki et al., 2004; Kawasaki et al., 2007]. Since then, many SCPP genes have arisen by duplication. Some SCPPs are unusually acidic (>25% Asp, Glu, or phospho-Ser). These SCPPs have been known as the small integrin-binding ligand N-linked glycoproteins (SIBLINGs) [Fisher et al., 2001]. However, I refer to these proteins as acidic SCPPs in order to acknowledge the evolutionary relationship with the other class, Pro and Gln (P/Q)-rich SCPPs (typically >20% P/Q). In mammals, expression of SCPP genes predominates in mineralized tissues, mammary glands, or salivary glands [Kawasaki and Weiss, 2006]. For tissue mineralization, acidic SCPPs are principally secreted from mesenchymally derived osteoblasts, osteocytes, and/or odontoblasts and are used for bone and dentin [George and Veis, 2008]. By contrast, P/Q-rich SCPPs are primarily deposited by dental epithelial cells (ameloblasts) and are employed to form the tooth surface [Kawasaki and Weiss, 2008]. In this report, I update phylogenetic distributions of SCPP genes. The result suggests that the degree of complexity of vertebrate hard tissues correlates with gain and loss of SCPP genes.

Phylogenetic Distribution of SCPP Genes

SCPP Genes in the Human Genome

In the human genome, 23 distinct SCPP genes have been identified (fig. 1a). Five acidic SCPP genes are clustered with both *SPARCL1* and the SCPP-Pro-Gln-rich 1 *(SCPPPQ1)* gene, and 16 other P/Q-rich SCPP genes form a distinct cluster on the same chromosome 4. The only exception is the P/Q-rich amelogenin *(AMEL)* gene located on chromosomes X and Y. Among these SCPP genes, 10 P/Q-rich genes have the termination codon in the penultimate exon. With the exception of *SCPPPQ1*, these genes are all localized between the α_{S1}-casein *(CSN1S1)* gene and the κ-casein *(CSN3)* gene at one end of the P/Q-rich SCPP gene cluster (fig. 1a). The configuration of these genes suggests an original genomic linkage between the 2 separate SCPP gene clusters. This view is supported by the arrangement of frog and fugu SCPP genes, as described below.

AMEL, ameloblastin (AMBN), and enamelin (ENAM) are secreted by ameloblasts and constitute the matrix of mineralizing enamel [Simmer et al., 2010]. It has been shown that their genes are decaying in toothless or enameless mammals [Meredith et al., 2010]. As the deposition of the enamel matrix decreases, ameloblasts initiate secretion of SCPPPQ1, amelotin (AMTN), and odontogenic ameloblasts associated (ODAM) [Iwasaki et al., 2005; Moffatt et al., 2006a; Moffatt et al., 2006b; Moffatt et al., 2008]. Differences in these 2 types of P/Q-rich SCPPs will be discussed below. Among acidic SCPPs, secreted phosphoprotein 1 (SPP1) is rich in the bone matrix, whereas dentin sialophosphoprotein (DSPP) is rich in the dentin matrix. These and other SCPPs have been recently reviewed in detail [Kawasaki and Weiss, 2008].

SCPP Genes in the Lizard and Bird Genomes

In the lizard genome, 5 acidic SCPP genes are arranged in a manner similar to their mammalian orthologs (fig. 1b). However, an additional *DSPP* was recently identified between the dentin matrix acidic protein 1 *(DMP1)* gene and the integrin-binding sialoprotein *(IBSP)* gene [McKnight and Fisher, 2009]. Both *DSPP* genes code for a large acidic region including many potential phosphor-Ser residues at the 3′ half, a characteristic of this gene. Six P/Q-rich SCPP genes, including the follicular dendritic cell secreted peptide *(FDCSP)* and *SCPPPQ1*, have been identified in the lizard genome. Their mammalian orthologs are all known to be expressed in dental tissues [Nakamura et al., 2005; Kawasaki and Weiss, 2008].

Since birds became toothless in the Late Cretaceous, several SCPP genes, including *DSPP*, *AMEL*, and *ENAM*, became pseudogenes [Sire et al., 2008; McKnight and Fisher, 2009; Al-Hashimi et al., 2010]. Moreover, no functional P/Q-rich SCPP genes have been identified in birds. The poor SCPP repertoire in birds illustrates critical roles of the SCPP gene family in tooth formation.

SCPP Genes in the Frog Genome

In the frog genome, P/Q-rich *AMBN*, *ENAM*, and *ODAM* are clustered with 5 acidic SCPP genes (fig. 1d). The organization of this gene cluster indicates that many P/Q-rich and acidic SCPP genes initially formed a single cluster. Subsequently, the order of *AMBN-ENAM* was inverted in an ancestor of the frog. Independently, in the amniote lineage, the original cluster split into 2 clusters and only *SCPPPQ1* remains clustered with acidic SCPP genes. This view is corroborated by the arrange-

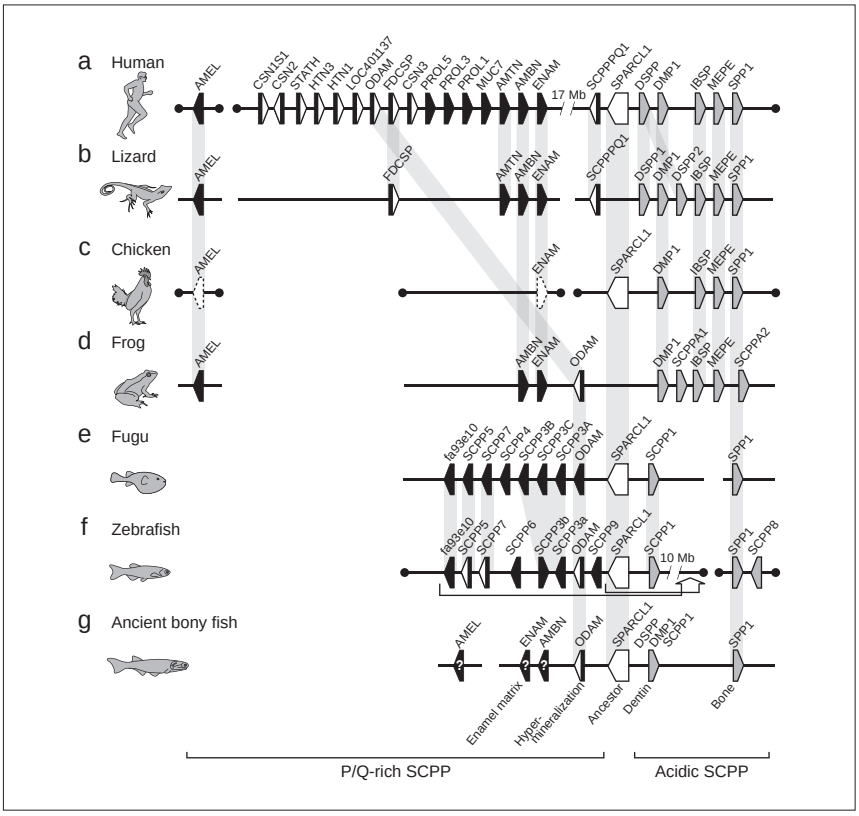

Fig. 1. The SCPP gene cluster in the genome of humans (**a**), lizard *(Anolis carolinensis)* (**b**), chicken (**c**), frog *(Xenopus tropicalis)* (**d**), fugu *(Takifugu rubripes)* (**e**), zebrafish (**f**), and ancient bony fish (common ancestor of actinopterygians and sarcopterygians) (**g**). Each pentagon illustrates a gene and the transcriptional direction. P/Q-rich SCPP genes, acidic SCPP genes, and *SPARCL1* are shown in different grey scales. P/Q-rich SCPP genes possessing the termination codon in the penultimate exon (hence, the last exon is entirely untranslated) are shown with a white tail (this exon is missing in small fugu genes). Pseudogenes for *AMEL* and *ENAM* in the chicken genome are shown in a dashed outline. Two lizard *DSPP* genes *(DSPP1* and *DSPP2)* are located in scaffold 8 (3,535,612–3,538,885 and 3,580,701–3,584,421) of the genomic sequence version anoCar1. *SPARCL1* has not been identified in the lizard genome or the frog genome. Orthologs are linked with grey bands. Gene symbols had been summarized previously [Kawasaki and Weiss, 2008]. To avoid confusions, gene and protein symbols are all capitalized, with a few exceptions in both figure 1 and the text.

ment of SCPP genes in the fugu genome, as described below.

The chromosomal location (unlike the initial report) and the expression pattern of the frog SCPP acidic 2 (*SCPPA2*) gene are both similar to those of amniote *SPP1* [Kawasaki, 2009]. However, no significant sequence similarity has been detected between these 2 genes even though zebrafish SPP1 shows the highest sequence similarity to mammalian SPP1 among nonteleost proteins deposited in GenBank. Moreover, whereas *SPP1* consists of 7 or 8 exons coding for 250–300 residues, *SCPPA2* comprises only 4 exons encoding 136 residues. It is possible that *SCPPA2* is orthologous to *SPP1* if *SCPPA2* had an extremely high evolutionary rate. However, *SCPPA2* could be a paralog of *SPP1* if *SCPPA2* was duplicated from *SPP1* and took over functions in bone and if the original *SPP1* was subsequently lost from the genome. To date, the orthology between *SCPPA2* and *SPP1* has been unclear. In the frog, strong expression of *DMP1* was detected in odontoblasts [Kawasaki, 2009]. I have now identified the matrix extracellular phosphoglycoprotein *(MEPE)* gene in expressed sequence tag records of the skin (accession number DT434807).

SCPP Genes in Teleost Genomes

In the fugu genome, 8 P/Q-rich SCPP genes are clustered with acidic *SCPP1* (fig. 1e), corroborating the initial genomic linkage of the 2 classes of SCPP genes. However, this cluster is divided in the zebrafish genome (fig. 1f), and *SPP1* (and *SCPP8*) is separated from the SCPP gene cluster in teleosts. Among these genes, only *SPP1* and *ODAM* have been identified as the tetrapod ortholog. *SCPP1* shows strong expression in odontoblasts [Kawasaki, 2009], similar to mammalian *DSPP* and frog *DMP1*. However, no significant sequence similarity has been detected between *SCPP1* and the other 2 genes. Further, whereas *SCPP1* consists of 9 exons coding for 198–242 residues, *DMP1* comprises 6 (mammals, caiman, and frog) or 4 (chicken) exons encoding 413 or more residues [Toyosawa et al., 1999], and mammalian *DSPP* comprises 5 exons encoding 585 or more residues. To date, we have been unable to show that *SCPP1* is the teleost ortholog of *DMP1* or *DSPP*. *fa93e10 (dkey-22i16.3)* also belongs to the SCPP gene family. This gene is expressed in the basal layer of the distal fin epithelium but not in adjacent osteoblasts [Goldsmith et al., 2003]. No similar gene has been identified in tetrapods.

P/Q-Rich SCPP Genes and Hypermineralized Tissues

In the maturation stage of enamel formation, P/Q-rich SCPPs constituting the initial matrix are removed from enamel and the protein space is further mineralized. Enamel thus matures into a hypermineralized tissue [Simmer et al., 2010]. In teleosts, the tooth surface is covered with a different hypermineralized tissue called enameloid. Teleost enameloid forms in a matrix, deposited by both inner dental epithelial cells (equivalent to ameloblasts) and odontoblasts [Sasagawa, 1995]. Expression analysis of various zebrafish genes in these cells revealed that enameloid and dentin form in a similar collagenous ECM [Kawasaki, 2009]. After the enameloid matrix begins mineralization, the IDE cells initiate expression of both *ODAM* and *SCPP9*. Similar late expression of *ODAM*, *AMTN,* and *SCPPPQ1* was also detected in ameloblasts in rodents [Iwasaki et al., 2005; Moffatt et al., 2006b; Moffatt et al., 2008]. In particular, ODAM and AMTN were shown to localize between maturing enamel and ameloblasts. These findings suggest that some P/Q-rich SCPPs assemble a structure or a space between the hypermineralizing tooth surface and overlying epithelium. This space is probably important for providing cell-tooth adhesion and for the inhibition of spontaneous calcification to facilitate the transportation of ions and organic molecules between tooth and dental epithelium [Takano, 1979; Al Kawas and Warshawsky, 2008]. Thus, I refer to these proteins as *hypermineralization* SCPPs and distinguish them from *enamel matrix* SCPPs.

ODAM likely plays a role in the hypermineralization of both tetrapod enamel and teleost enameloid. This finding suggests that ancient hypermineralization machinery that arose in their common ancestor descended to these 2 tissues. Enamel forms in a noncollagenous ECM consisting of P/Q-rich SCPPs, whereas enameloid and dentin mineralize in a similar collagenous matrix in teleosts. The similarities and differences in these tissues suggest that teleost enameloid originated as a hypermineralized dentin and that enamel evolved from a hypermineralized tissue that is similar to enameloid in modern teleosts. Thus, enamel matrix SCPPs are apparently derived from a hypermineralization SCPP.

Acidic and P/Q-Rich SCPP Genes and Diverse Mineralized Tissues

Many SCPP genes arose in specific lineages and evolved new functions [Kawasaki et al., 2005]. To date, only *SPP1* and *ODAM* have been identified clearly in both tetrapods and teleosts, and each ortholog shows a similar expression pattern in both clades [Kawasaki, 2009]. However, mammalian *DSPP*, frog *DMP1*, and teleost *SCPP1* all show strong expression in odontoblasts. Similarly, frog *SCPPA2* and *SPP1* in other lineages are highly expressed in osteoblasts. Notably, teleost *SCPP1*, frog *DMP1*, and mammalian *DSPP* are located upstream in the acidic SCPP gene cluster, whereas frog *SCPPA2* and amniote *SPP1* are downstream (fig. 1). The similar positions of these genes suggest their close relationship; these genes could have arisen by local tandem duplication, if they are not orthologous. It is thus possible to reconstruct an ancient SCPP gene cluster in the common ancestor of tetrapods and teleosts (fig. 1g). In this cluster, at least 4 genes, acidic *SPP1* mainly used in bone among mineralized tissues, another acidic SCPP gene principally employed in dentin, *SPARCL1*, and hypermineralization P/Q-rich *ODAM*, are arranged in this order. No enamel matrix P/Q-rich SCPP genes have been confirmed in actinopterygians but, if they exist, these genes sit at the distal end of the P/Q-rich SCPP gene cluster.

We have previously reported that *SPARCL1* has an unusually acidic calcium-binding domain that descended to SCPPs [Kawasaki et al., 2004]. This domain is known to adopt an unfolded structure due to electrostatic repulsion [Hambrock et al., 2003]. These characteristics are common to SPP1. Further, *SPP1* is conserved in both tetrapods and teleosts. It is thus likely that the initial SCPP was SPP1 or a similar acidic SCPP. Unusually acidic proteins are employed in biomineralization widely in eukaryotes, illustrating a critical role of interactions between acidic proteins and calcium ions in mineralization [Marin and Luquet, 2007]. Hence, ancient SPP1 or an SPP1-like acidic SCPP could have been used for the mineralization of bone, dentin, or a similar collagenous tissue found in early jawless vertebrates [Donoghue et al., 2006]. Subsequent duplication of such an ancient acidic SCPP gene may have been important for the differentiation of matrix-matrix and matrix-cell interactions, and the differentiation of these interactions in turn led to diversification of an ancient mineralized tissue into bone and dentin in modern vertebrates.

P/Q-rich proteins also adopt unfolded structures. Unlike acidic proteins, however, P/Q-rich proteins facilitate protein-protein interactions or self-assembly through hydrogen bonds or hydrophobic interactions [Williamson, 1994]. Thus, in general, P/Q-rich proteins appear to be more suitable than acidic proteins for creating a space between the hypermineralizing tissue surface and overlying

epithelium and for assembling a mineralizing matrix. Indeed, in the sea urchin, framework proteins for spicule mineralization also contain P/Q-rich or Pro-rich repeats [Wilt and Ettensohn, 2007]. Apparently, the innovation of P/Q-rich SCPPs was important for the origin of enameloid and enamel.

In summary, I propose that different hard tissues arose by modifications of old tissues and that some duplicate SCPP genes were used principally in modified new tissues. In this process, unusually acidic SCPPs and P/Q-rich SCPPs played distinct roles. More recently, however, many SCPP genes were lost in toothless birds and mammals. It thus appears that the phenotypic complexity of vertebrate hard tissues correlates with gain and loss in members of the SCPP gene family.

Acknowledgements

I am grateful to Prof. Kenneth M. Weiss and Dr. Anne V. Buchanan at Penn State University for their generous comments and encouragement. Financial support from the National Science Foundation (grant BCS0725227, BCS0343442) and the Penn State Evan Pugh Professors Research Fund to Prof. Kenneth Weiss is gratefully acknowledged.

References

Al-Hashimi, N., A.G. Lafont, S. Delgado, K. Kawasaki, J.Y. Sire (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.

Al Kawas, S., H. Warshawsky (2008) Ultrastructure and composition of basement membrane separating mature ameloblasts from enamel. Arch Oral Biol *53:* 310–317.

Donoghue, P.C., I.J. Sansom, J.P. Downs (2006) Early evolution of vertebrate skeletal tissues and cellular interactions, and the canalization of skeletal development. J Exp Zool B Mol Dev Evol *306B:* 278–294.

Fisher, L.W., D.A. Torchia, B. Fohr, M.F. Young, N.S. Fedarko (2001) Flexible structures of SIBLING proteins, bone sialoprotein, and osteopontin. Biochem Biophys Res Commun *280:* 460–465.

George, A., A. Veis (2008) Phosphorylated proteins and control over apatite nucleation, crystal growth, and inhibition. Chem Rev *108:* 4670–4693.

Goldsmith, M.I., S. Fisher, R. Waterman, S.L. Johnson (2003) Saltatory control of isometric growth in the zebrafish caudal fin is disrupted in long fin and rapunzel mutants. Dev Biol *259:* 303–317.

Hall, B.K. (2005) Bones and Cartilage: Developmental and Evolutionary Skeletal Biology. San Diego, Elsevier.

Hambrock, H.O., D.P. Nitsche, U. Hansen, P. Bruckner, M. Paulsson, P. Maurer, U. Hartmann (2003) SC1/hevin: an extracellular calcium-modulated protein that binds collagen I. J Biol Chem *278:* 11351–11358.

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

Kawasaki, K. (2009) The SCPP gene repertoire in bony vertebrates and graded differences in mineralized tissues. Dev Genes Evol *219:* 147–157.

Kawasaki, K., A.V. Buchanan, K.M. Weiss (2007) Gene duplication and the evolution of vertebrate skeletal mineralization. Cells Tissues Organs *186:* 7–24.

Kawasaki, K., A.V. Buchanan, K.M. Weiss (2009) Biomineralization in humans: making the hard choices in life. Annu Rev Genet *43:* 119–142.

Kawasaki, K., T. Suzuki, K.M. Weiss (2004) Genetic basis for the evolution of vertebrate mineralized tissue. Proc Natl Acad Sci USA *101:* 11356–11361.

Kawasaki, K., T. Suzuki, K.M. Weiss (2005) Phenogenetic drift in evolution: the changing genetic basis of vertebrate teeth. Proc Natl Acad Sci USA *102:* 18063–18068.

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

Kawasaki, K., K.M. Weiss (2006) Evolutionary genetics of vertebrate tissue mineralization: the origin and evolution of the secretory calcium-binding phosphoprotein family. J Exp Zool B Mol Dev Evol *306:* 295–316.

Kawasaki, K., K.M. Weiss (2008) SCPP gene evolution and the dental mineralization continuum. J Dent Res *87:* 520–531.

Marin, F., G. Luquet (2007) Unusually acidic proteins in biomineralization; in Bäuerlein E. (ed): Handbook of Biomineralization. Weinheim, Wiley-VCH, vol 1, pp 273–290.

McKnight, D.A., L.W. Fisher (2009) Molecular evolution of dentin phosphoprotein among toothed and toothless animals. BMC Evol Biol *9:* 299.

Meredith, R.W., J. Gatesy, J. Cheng, M.S. Springer (2011) Pseudogenization of the tooth gene enamelysin (MMP20) in the common ancestor of extant baleen whales. Proc Biol Sci, *278:* 993–1002.

Moffatt, P., C.E. Smith, R. Sooknanan, R. St-Arnaud, A. Nanci (2006a) Identification of secreted and membrane proteins in the rat incisor enamel organ using a signal-trap screening approach. Eur J Oral Sci *114(suppl 1):* 139–146.

Moffatt, P., C.E. Smith, R. St-Arnaud, A. Nanci (2008) Characterization of Apin, a secreted protein highly expressed in tooth-associated epithelia. J Cell Biochem *103:* 941–956.

Moffatt, P., C.E. Smith, R. St-Arnaud, D. Simmons, J.T. Wright, A. Nanci (2006b) 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.

Nakamura, S., T. Terashima, T. Yoshida, S. Iseki, Y. Takano, I. Ishikawa, T. Shinomura (2005) Identification of genes preferentially expressed in periodontal ligament: specific expression of a novel secreted protein, FDC-SP. Biochem Biophys Res Commun *338:* 1197–1203.

Sasagawa, I. (1995) Fine structure of tooth germs during the formation of enameloid matrix in *Tilapia nilotica*, a teleost fish. Arch Oral Biol *40:* 801–814.

Simmer, J.P., P. Papagerakis, C.E. Smith, D.C. Fisher, A.N. Rountrey, L. Zheng, J.C. Hu (2010) Regulation of dental enamel shape and hardness. J Dent Res *89:* 1024–1038.

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

Takano, Y. (1979) Cytochemical studies of ameloblasts and the surface layer of enamel of the rat incisor at the maturation stage. Arch Histol Jpn *42:* 11–32.

Toyosawa, S., C. O'hUigin, J. Klein (1999) The dentin matrix protein 1 gene of prototherian and metatherian mammals. J Mol Evol *48:* 160–167.

Williamson, M.P. (1994) The structure and function of proline-rich regions in proteins. Biochem J *297:* 249–260.

Wilt, F.H., C.A. Ettensohn (2007) The morphogenesis and biomineralization of the sea urchin larval skeleton; in: Baeuerlein E. (ed): Handbook of biomineralization, Weinheim, Wiley-VCH, vol 1, pp 183–210.