

The Evolution of Transglutaminases Underlies the Origin and Loss of Cornified Skin Appendages in Vertebrates

Attila Placido Sachslehner,^{1,†} Marta Surbek,^{1,†} Karin Brigit Holthaus,¹ Julia Steinbinder,¹ Bahar Golabi,¹ Claudia Hess,² and Leopold Eckhart ^{1,*}

¹Department of Dermatology, Medical University of Vienna, 1090 Vienna, Austria

²Clinic for Poultry and Fish Medicine, Department for Farm Animals and Veterinary Public Health, University of Veterinary Medicine Vienna, 1210 Vienna, Austria

[†]These authors contributed equally to this work.

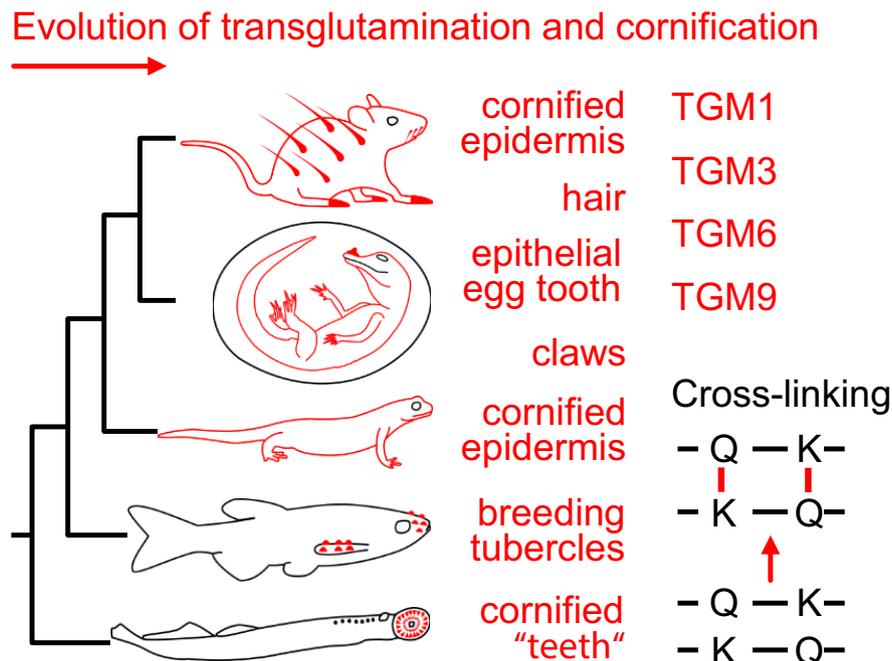
*Corresponding author: E-mail: leopold.eckhart@meduniwien.ac.at.

Associate editor: Guang Yang

Abstract

Transglutaminases (TGMs) cross-link proteins by introducing covalent bonds between glutamine and lysine residues. These cross-links are essential for epithelial cornification which enables tetrapods to live on land. Here, we investigated which evolutionary adaptations of vertebrates were associated with specific changes in the family of TGM genes. We determined the catalog of TGMs in the main clades of vertebrates, performed a comprehensive phylogenetic analysis of TGMs, and localized the distribution of selected TGMs in tissues. Our data suggest that TGM1 is the phylogenetically oldest epithelial TGM, with orthologs being expressed in the cornified teeth of the lamprey, a basal vertebrate. Gene duplications led to the origin of TGM10 in stem vertebrates, the origin of TGM2 in jawed vertebrates, and an increasing number of epithelium-associated TGM genes in the lineage leading to terrestrial vertebrates. TGM9 is expressed in the epithelial egg tooth, and its evolutionary origin in stem amniotes coincided with the evolution of embryonic development in eggs that are surrounded by a protective shell. Conversely, viviparous mammals have lost both the epithelial egg tooth and TGM9. TGM3 and TGM6 evolved as regulators of cornification in hair follicles and underwent pseudogenization upon the evolutionary loss of hair in cetaceans. Taken together, this study reveals the gain and loss of vertebrate TGM genes in association with the evolution of cornified skin appendages and suggests an important role of TGM9 in the evolution of amniotes.

Graphical abstract



Received: February 13, 2024. Revised: April 11, 2024. Accepted: May 20, 2024

© The Author(s) 2024. Published by Oxford University Press on behalf of Society for Molecular Biology and Evolution.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Open Access

Key words: transglutaminase, gene family, gene loss, hair, water–land transition.

Introduction

Transglutaminases (TGMs) are calcium-dependent enzymes which introduce covalent bonds (N ϵ -(γ -glutamyl) lysine bridges) between lysine and glutamine residues of proteins (Greenberg et al. 1991). This process, known as transglutamination, allows the formation of large macromolecular structures which have been implicated in blood coagulation, apoptotic cell death, and the cornification of epithelia (Eckert et al. 2014). In addition, TGMs can catalyze the deamidation of glutamine residues and the linkage of amines to glutamine residues of proteins, known as transamidation (Lorand and Iismaa 2019). TGM genes have been identified and partially characterized in a broad range of metazoans ranging from the fruit fly (Shibata and Kawabata 2018) to humans (Greenberg et al. 1991). However, the evolution of TGMs has remained incompletely understood. In the present study, we utilized comparative genomics to infer the evolutionary history of TGMs in vertebrates, focusing on the evolutionary lineage leading to humans.

Humans have nine TGM genes, namely *TGM1* through *TGM7*, *F13A1*, and *EPB42* with different expression patterns (Eckert et al. 2014). *TGM2* is expressed in a wide variety of cell types, whereas expression of *F13A1* (coagulation factor XIII A chain) is confined to macrophages (Beckers et al. 2017). *EPB42* (erythrocyte membrane protein band 4.2) is expressed as a catalytically inactive TGM family protein in erythrocytes (Korsgren et al. 1990), and *TGM4* is specific for the prostate (Dubink et al. 1996). Similarly, the expression of *TGM1*, *TGM3*, and *TGM5* is confined to stratified epithelia, such as the epidermis of the skin, skin appendages, and the esophageal epithelium. Epithelial cells, known as keratinocytes, of the epidermis, hair, and nails undergo cornification, that is, a conversion into dead, mechanically stable and interconnected cell remnants, known as corneocytes (Candi et al. 2005; Yamane et al. 2010; Surbek et al. 2023). Protein cross-linking is an essential step of cornification (Eckhart et al. 2013), and the importance of TGMs for cornification is highlighted by the effects of TGM gene mutations in knockout mouse models and human patients. Deletion of *Tgm1* in mice leads to a skin barrier defect and perinatal death due to excessive transepidermal water loss (Matsuki et al. 1998; Kuramoto et al. 2002). Humans can partially compensate inactivating mutations of *TGM1*, but develop severe ichthyosis (Rice et al. 2003; Zhang et al. 2019). Mutations of *TGM3* and *TGM5* cause deformations of hair fibers and acral peeling skin syndrome, respectively (Cassidy et al. 2005; Ü Basmanav et al. 2016). *TGM6* and *TGM7* are the least characterized human TGMs, with contradictory data on expression being available in the literature (Grenard et al. 2001; Thomas et al. 2013) and in the Genotype-Tissue Expression database (<https://www.gtexportal.org/>, last accessed on 2024 February 9). In addition to orthologs of human TGMs, a gene termed *TGM8* was identified in zebrafish (Deasey et al. 2012).

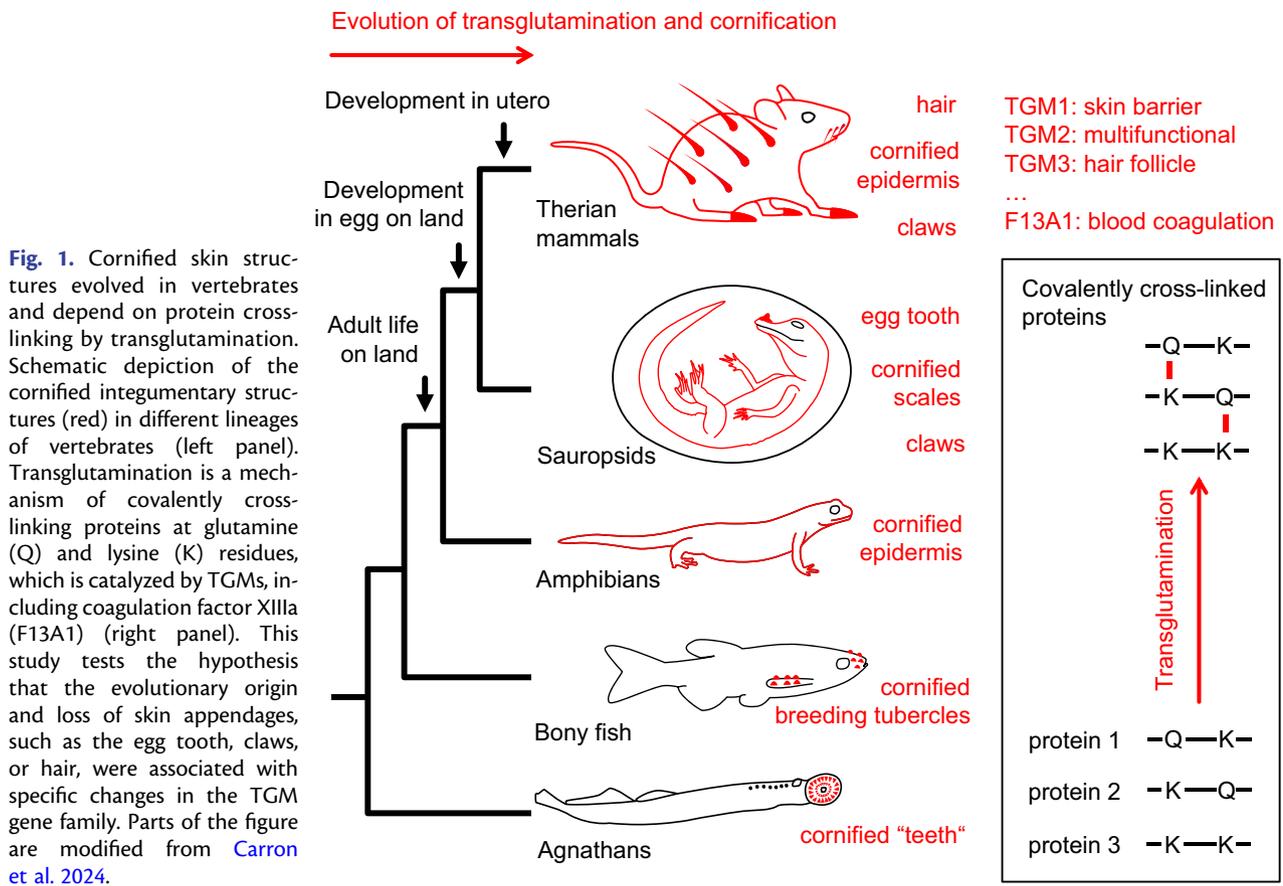
Cornified integumentary structures are not specific for mammals, but also form in other vertebrates (Wu et al. 2004; Akat et al. 2022; Alibardi 2022). The cornified layer of the epidermis, the stratum corneum, is conserved in all land-dwelling vertebrates (Alibardi 2003; Candi et al. 2005; Akat et al. 2022). The stratum corneum consists of a single layer of dead cells in amphibians and many layers thereof in amniotes (Bereiter-Hahn et al. 1986). Epidermal cornification is diminished or modified at the subcellular level in mammals that have secondarily acquired a fully aquatic lifestyle (cetaceans and sirenians) (Ehrlich et al. 2019; Menon et al. 2022). Hard skin appendages, such as hair, claws, and feathers, are cornified by transglutamination (Sachslehner et al. 2023) and by the formation of disulfide bonds (Strasser et al. 2015; Harland and Plowman 2018; Ehrlich et al. 2020). The cornified epithelial egg tooth, also known as caruncle, is conserved in many but not all amniotes (Mlitz et al. 2021; Fenelon et al. 2023). Various species of teleost fish have cornified breeding tubercles (Fischer et al. 2014), and the most basal vertebrates, the agnathans (lamprey and hagfish), have cornified epithelial teeth (Fig. 1) (Rice et al. 1994; Zaccone et al. 1995; Alibardi and Segalla 2011). The distribution of cornification among vertebrate taxa suggests that TGMs are active in phylogenetically diverse clades. However, it is unknown whether and how the evolution of cornified structures was linked to the evolution of the TGM gene family.

Here, we determined the repertoire of TGM genes in phylogenetically and phenotypically diverse vertebrates and reconstructed the evolutionary history of TGMs in different lineages. Our results point to key roles of TGM gene innovation and degeneration in the evolution of skin appendages that were critical for evolutionary transitions of vertebrates.

Results

The Phylogenetically Basal *TGM1* and *F13A1* Genes are Conserved in Vertebrates, Whereas *TGM4* has been Lost in At Least Four Lineages

We identified the complete set of TGM genes in representatives from major groups of vertebrates (Chordata, Olfactores, Vertebrata) and subjected them to a phylogenetic analysis (supplementary table S1 and supplementary fig. S1, Supplementary Material online). TGM genes of a tunicate (Chordata, Olfactores, Tunicata) and lancelet (Chordata, Cephalochordata, Leptocardii) were included for comparison. Besides tunicate TGM genes, which do not cluster together with any TGM of the other species, three distinct clades of TGMs (Fig. 2) were identified: (i) the clade comprising *TGM1* and *F13A1*, (ii) *TGM4*, and (iii) the clade comprising *TGM2* through *TGM8*, *EPB42*, and two genes that we tentatively name *TGM9* and *TGM10*. A comparison of the exon–intron structures showed that, in agreement with a



previous report (Phillips et al. 1992), *TGM1* and *F13A1* have an intron that is missing in all other TGM genes (supplementary fig. S2, Supplementary Material online), supporting the hypothesis that *TGM1* and *F13A1* are closely related. *TGM4* does not have a homologous intron but contains another unique intron (supplementary fig. S2, Supplementary Material online).

Homologs of *TGM1* (supplementary figs. S3 and S4 and supplementary table S2, Supplementary Material online) and *F13A1* (supplementary fig. S5 and supplementary table S3, Supplementary Material online) exist in all species investigated, except the tunicates. *TGM1* and *F13A1*-like homologs of the lancelet form the outgroup to the vertebrate *TGM1/F13A1* genes (Fig. 2). Sea lamprey, zebrafish, and coelacanth have more than one copy of *TGM1*, indicating lineage-specific gene duplications (supplementary fig. S4, Supplementary Material online). By mass spectrometry-based proteomics, we detected *TGM1*-like proteins in the cornified epithelial teeth of the sea lamprey (supplementary table S4, Supplementary Material online), suggesting that orthologs of mammalian *TGM1* contribute to epithelial cornification in agnathans.

TGM4 is expressed predominantly in the prostate of humans and mice (Dubbink et al. 1996; Tian et al. 2009). Our comparative genomics and phylogenetic analyses revealed orthologs of *TGM4* in sauropsids, amphibians, the coelacanth, and actinopterygians (ray-finned fish), but not in sharks (supplementary fig. S6 and supplementary table S5, Supplementary Material online), indicating that *TGM4* has

originated in stem Osteichthyes (bony fish including tetrapods) and not at the emergence of land vertebrates, as suggested in a previous study (Tian et al. 2009). *TGM4* was independently duplicated in the gray birchir (*Polypterus senegalus*), the diploid frog *Xenopus tropicalis*, and the green anole lizard (*Anolis carolinensis*) and has been lost or pseudogenized in the lungfish (supplementary fig. S6, Supplementary Material online) and at least three groups of mammals, namely marsupials, Paenungulata (a subclade of Afrotheria, comprising elephants and sirenians), and Laurasiatheria (supplementary fig. S6, Supplementary Material online). The origin of *TGM4* in fish and its loss in several taxa of mammals argue against a strict association of *TGM4* with a function in the prostate, because this organ is confined to mammals and conserved among them.

The Previously Uncharacterized Gene *TGM10* is Present in Fish and has been Lost Upon the Water-to-Land Transition of Tetrapods

According to molecular phylogenetics (Fig. 2), the largest clade of TGMs is subdivided into two subclades, of which the first one comprises *TGM2* and a *TGM2*-like gene (Lisetto et al. 2023) which we tentatively name *TGM10*, and the second subclade comprises the orthologs of human *TGM3*, *TGM5*, *TGM6*, *TGM7*, *EPB42*, fish *TGM8*, and another uncharacterized gene, tentatively named *TGM9*. For further analysis of homology, we performed a synteny analysis of TGM gene loci.

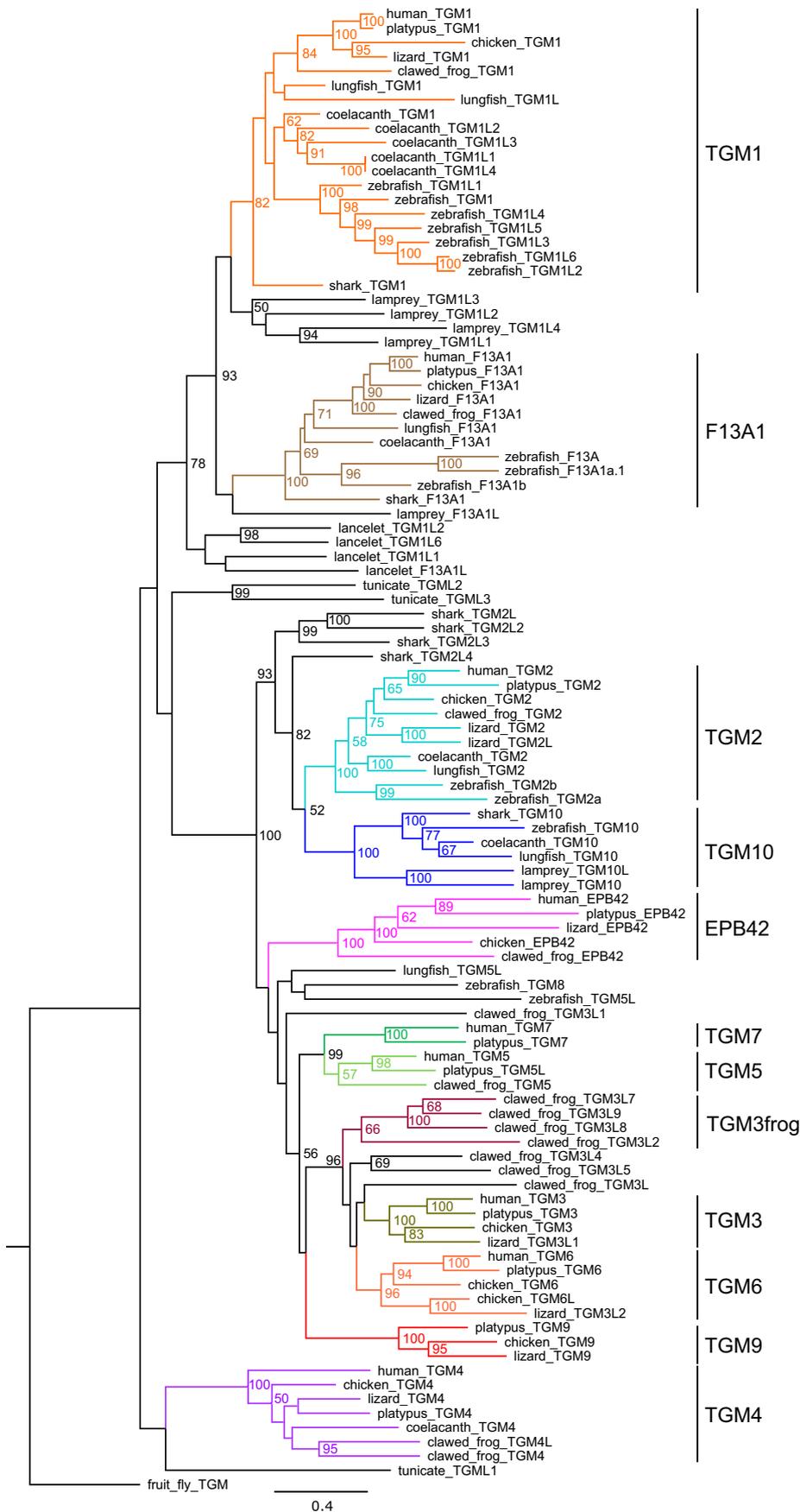


Fig. 2. Phylogenetic analysis of TGMs of vertebrates. A phylogenetic tree of TGMs was inferred using the maximum likelihood method. TGMs of species from the main clades of vertebrates were included, with TGMs from a tunicate, the lancelet, and the fruit fly as outgroups. Colors highlight groups of TGMs that are discussed in the text. Bootstrap values above 50 are shown. The scale bar indicates amino acid substitutions per site. Species: Human (*Homo sapiens*), platypus (*Ornithorhynchus anatinus*), chicken (*Gallus gallus*), lizard (*A. carolinensis*), clawed frog (*X. tropicalis*), lungfish (*Protopterus annectens*), coelacanth (*L. chalumnae*), zebrafish (*Danio rerio*), shark (*Carcharodon carcharias*), lamprey (*P. marinus*), tunicate (*Ciona intestinalis*), lancelet (*Branchiostoma floridae*), and fruit fly (*Drosophila melanogaster*).

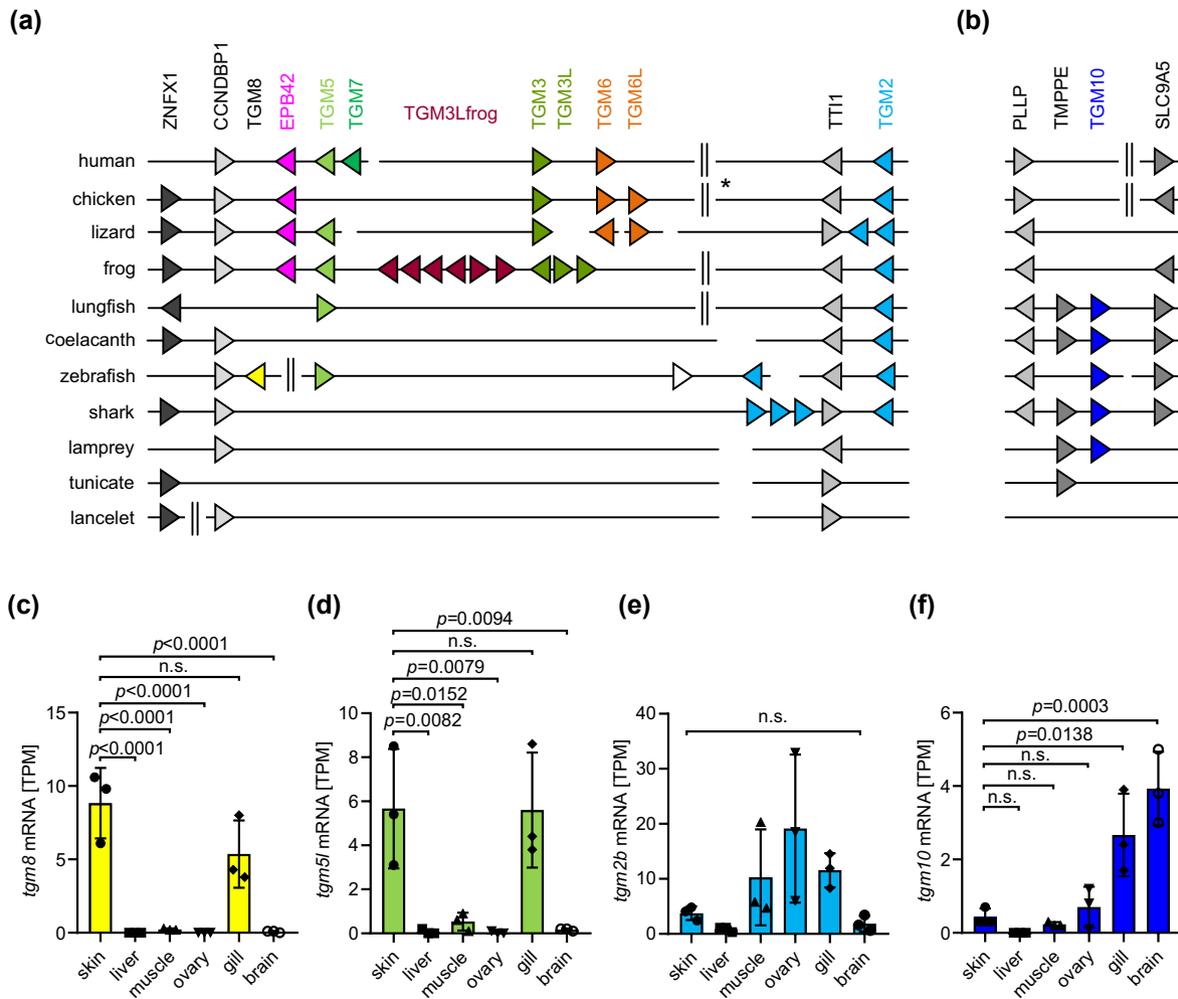


Fig. 3. Synteny analysis of *TGM2*-related genes in vertebrates. a) Schematic depiction of gene clusters comprised of *TGM2* and *TGM* genes closely related to *TGM2* (Fig. 2). Genes are shown as triangles pointing in the direction of transcription. Horizontal lines indicate chromosomes or sequence scaffolds. Because of space limitations, not all genes are shown. Vertical double lines indicate gaps in the display where several genes are not shown. An asterisk besides the vertical double lines indicates a complex rearrangement of the left and right segments of the chicken gene locus. b) Locus of the *TGM10* gene in different species. Note that *TGM10* is not present in tunicates and terrestrial vertebrates. c to f) Expression of zebrafish *TGMs* in tissues. The colors of bars correspond to the colors of the genes in panel (a). RNA-seq data were analyzed for the mRNA levels of *tgm8* (c), *tgm5l*, which is orthologous to human *TGM5* (d), *tgm2b*, which is one of two orthologs of human *TGM2* (e) and *tgm10* (f). *P*-values were calculated by ANOVA. n.s., not significant. Species: Human (*H. sapiens*), chicken (*Gallus gallus*), lizard (*A. carolinensis*), frog (*X. tropicalis*), lungfish (*P. annectens*), coelacanth (*L. chalumnae*), zebrafish (*D. rerio*), shark (*C. carcharias*), lamprey (*P. marinus*), tunicate (*C. intestinalis*), and lancelet (*B. floridae*).

In most species investigated, the genes *TGM2*, *TGM3*, *TGM5*, *TGM6*, *TGM7*, *EPB42*, and, if present, *TGM8* are clustered at one chromosomal locus (Fig. 3a, supplementary tables S6 and S7, Supplementary Material online). Deviations from this organization are likely due to rearrangements of genes or chromosomal segments. *TGM9* (see the last section of Results) and *TGM10* (Fig. 3b) were found at loci lacking synteny with the aforementioned *TGM* gene cluster. Importantly, a basal vertebrate, the lamprey, has a *TGM10* gene but lacks *TGM* genes at the loci corresponding to those of *TGM2* and *TGM9*, suggesting that *TGM10* originated prior to all other *TGMs* of this clade. Of note, an additional *TGM10*-like gene at another locus of the lamprey genome is likely the product of a lamprey-specific *TGM10* gene duplication (supplementary fig. S7, Supplementary

Material online). *TGM10* is conserved in fish including the lungfish but not in tetrapods, indicating that this gene was lost upon the evolutionary transition to life on land.

A cluster of *TGM2* and other *TGM* genes is located next to *CCNDBP1* (*cyclin D1-binding protein 1*) in jawed vertebrates. In zebrafish, *TGM8* localizes to this cluster, confirming the association in molecular phylogenetics (Fig. 2). Analysis of tissue transcriptomes (supplementary table S8, Supplementary Material online) showed that *TGM8* is predominantly expressed in the skin of the zebrafish, with *TGM1*, *TGM1L1*, *TGM5-like* (*TGM5L*), and *F13A1B* being other skin-enhanced *TGMs* (Fig. 3, c to f, supplementary fig. S8, Supplementary Material online). A monophyletic subcluster of *TGM3*-like genes (Fig. 2) is present in the frog (Fig. 3a), indicating lineage-specific gene duplications.

EPB42, which is specifically expressed in erythrocytes (Wang et al. 2020), is present in tetrapods but not in fish, suggesting an evolutionary origin at the time of the water-to-land transition. Residues critical for catalytic TGM activity are conserved in EPB42 orthologs of some but not all amphibians and sauropsids and in none of the mammalian species investigated (supplementary fig. S9, Supplementary Material online), suggesting that EPB42 was originally an active enzyme which lost its activity in distinct lineages, including the one leading to humans.

Mammalian TGM6 is Associated with Hair Follicles and has been Lost in Cetaceans

The TGM2-related gene cluster contains at least one gene, namely TGM3, that is strongly expressed in epithelial cells of hair follicles (Thibaut et al. 2009; Eckert et al. 2014; Ü Basmanav et al. 2016; Harland and Plowman 2018). The functions of the TGM3-related gene TGM6 (Fig. 2) are not known. We hypothesized that TGMs with a predominant role in hair follicles may have become dispensable in mammals that have entirely lost their hair coat. Thus, we investigated the TGM genes of cetaceans (whales and dolphins), which are the only clade of mammals in which hair has completely disappeared. We found that cetaceans lack protein-coding TGM3 through TGM7 genes (Fig. 4a). In contrast to TGM4, which is also absent in the closest land-dwelling relatives of cetaceans (Fig. 4a), TGM3, TGM5, TGM6, and TGM7 have been lost specifically in cetaceans. The loss of TGM5 was reported previously (Sharma et al. 2018). Detailed sequence analysis of TGM6 showed that, besides other mutations, a frameshift mutation in exon 5 is shared among all cetaceans investigated (Fig. 4b), indicating that TGM6 underwent pseudogenization in stem cetaceans. By analyzing publicly available gene expression (supplementary tables S9 and S10, Supplementary Material online) and hair proteome data (supplementary tables S11, Supplementary Material online) from mouse studies (Joost et al. 2020; Sukseree et al. 2024), we found evidence for expression of TGM6 in hair keratinocytes (Fig. 4c, supplementary tables S9 and S11, Supplementary Material online). Of note, TGM3 and TGM5 are also expressed in human and mouse tissues other than hair follicles (Fig. 4c, list of “other expression sites”), suggesting that the loss of these genes in cetaceans occurred despite pleiotropic functions of the TGMs. We conclude that the loss of hair follicles, possibly together with modifications of other TGM-associated structures, allowed the pseudogenization of TGM6 and the contraction of the entire TGM gene repertoire in cetaceans.

Origin and Loss of TGM9 Paralleled the Evolution of the Epithelial Egg Tooth in Amniotes and the Evolution of Claws in Squamates

Our phylogenetic analysis (Fig. 2) revealed that TGM9 is most closely related to TGM3, TGM5, TGM6, and TGM7. TGM9 is present in the genomes of turtles, crocodylians, birds, and the platypus, a monotreme mammal, where it

is flanked by the conserved genes *TMX3* (thioredoxin-related transmembrane protein 3) and *DSEL* (dermatan sulfate epimerase-like) (Fig. 5a). The species distribution of TGM9 orthologs suggests that TGM9 has originated, presumably by duplication of a TGM3/TGM5-like gene (Fig. 2), in stem amniotes and that it has been lost or pseudogenized independently in at least three clades of amniotes, namely therian mammals, snakes, and worm lizards (Fig. 5, a and b, supplementary table S12, Supplementary Material online).

We noticed that all of the TGM9-negative amniotes lack an epithelial egg tooth, which is a cornified structure that evolved in stem amniotes to facilitate hatching from the egg after completion of embryonic development. To test the hypothesis that the epithelial egg tooth is associated with TGM9, we localized TGM9 mRNA in chicken embryos isolated from eggs (Fig. 5c). TGM9 mRNA was prominently detected by in situ hybridization in the epithelial cells that give rise to the egg tooth, but only at minute amounts in the beak and the periderm, a transient epithelium that covers the egg tooth and epidermis during development (Fig. 5c). TGM9 co-localized with TGM activity (Fig. 5d). Reverse transcription polymerase chain reaction (RT-PCR) confirmed the expression of TGM9 in the egg tooth and revealed additional expression of TGM9 in the wings and back skin where feathers were developing (Fig. 5e). To substantiate these data, we determined the proteome of the epithelial egg tooth and detected the TGM9 protein in three biological replicates there of (PRIDE dataset: PXD048875, accession number TGM9: XP_040519905.1). Reanalysis of a publicly available proteomic dataset showed that TGM9 is also present in developing feathers of the pied flycatcher (*Ficedula hypoleuca*) (protein accession number Ensembl: ENSTGUP00000009578, corresponding to GenBank: XP_030121789.1) (Leskinen et al. 2012). In the anole lizard (*A. carolinensis*), TGM9 mRNA was detected by RT-PCR in the claws (Fig. 5f), whereas internal organs were TGM9-negative.

These data suggest that TGM9 plays roles of in diverse cornified skin appendages of amniotes and that the loss of the epithelial egg tooth has been a prerequisite for the degeneration of the TGM9 gene. Additional constraints on the evolution of TGM9 appear to arise from its roles in claws, so that claw-bearing squamates have retained the TGM9 gene after the loss the epithelial egg tooth, whereas limbless and therefore also clawless squamates (snakes and worm lizards) have lost TGM9.

Discussion

The results of this study suggest a new model for the evolution of TGMs in vertebrates (Fig. 6a). TGM1 and F13A1 are the phylogenetically basal TGM genes, which evolved from a single ancestral gene prior to the emergence of vertebrates. A subsequent gene duplication in stem vertebrates gave rise to TGM10 from which the ancestor of TGM2 evolved at a new locus. Subsequent gene duplications starting from TGM2 led to the evolution of TGM3 through TGM8 and EPB42 during the diversification of vertebrates. TGM9 arose from the TGM2-like gene cluster by

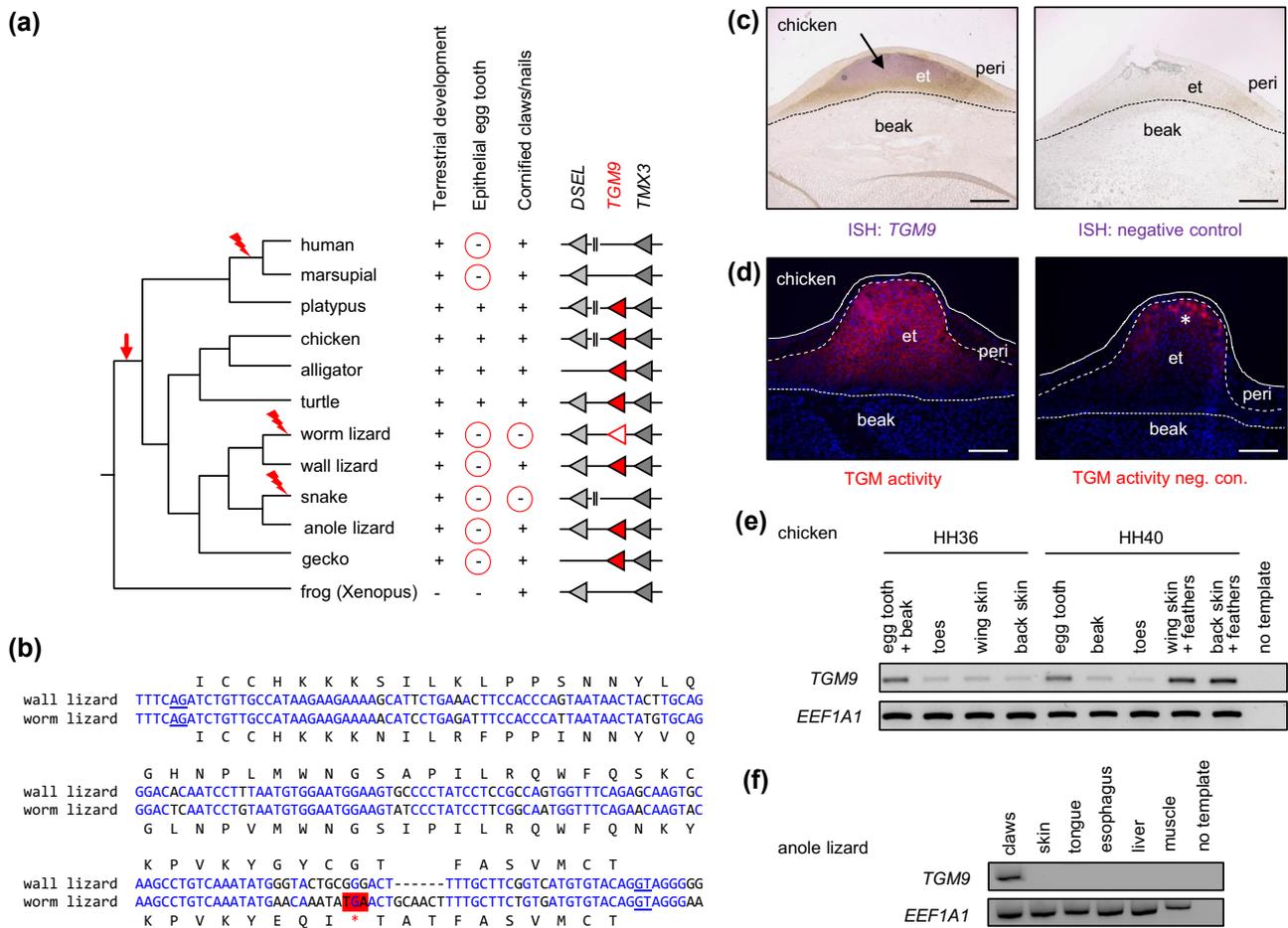


Fig. 5. *TGM9* is expressed in the epithelial egg tooth of chicken embryos and in claws of lizards. a) Evolution of *TGM9* in relation to the evolution of cornified skin appendages. The locus of *TGM9* and its neighboring genes was compared between diverse species of tetrapods (right panel). The relationship of species is shown by a phylogenetic tree (left) on which the inferred origin (arrow) and loss of *TGM9* (flash symbol) are indicated. Traits that are correlated with the conservation of *TGM9* are shown for each lineage, whereby “+” and “-” indicate the presence or absence of the respective trait. Red circles highlight the evolutionary loss of traits. Species: human (*H. sapiens*), marsupial (Tasmanian devil, *Sarcophilus harrisii*), platypus (*O. anatinus*), chicken (*Gallus gallus*), alligator (*Alligator sinensis*), turtle (*Mauremys reevesii*), worm lizard (*Rhineura floridana*), wall lizard (*Podarcis muralis*), snake (*Pantherophis guttatus*), anole lizard (*A. carolinensis*), gecko (*Gekko japonicus*), and frog (*X. tropicalis*). b) Nucleotide sequence alignment of exon 6 of the *TGM9* genes of the wall lizard and the legless worm lizard. Splicing signals (underlined) and further intronic sequences are included in the alignment. Sequence identity is highlighted by blue fonts. A premature stop codon is highlighted by red shading. c) mRNA in situ hybridization of *TGM9* in the chicken egg tooth at developmental stage HH36. An arrow points to the strongest signal in the epithelial egg tooth (et). In the negative control reaction (right panel), the antisense probe was replaced by the corresponding sense probe. Scale bars, 200 μ m. peri, periderm. d) TGM activity (red) was localized by using a fluorescently labeled (red) substrate on a cryosection of the chicken beak and egg tooth at developmental stage HH36. In the negative control (neg. con.) reaction (right panel), calcium ions were replaced by EDTA. Unspecific signals are indicated by an asterisk. Scale bars, 200 μ m. e) RT-PCR analysis of *TGM9* expression in chicken embryos at developmental stages HH36 and HH40. f) RT-PCR analysis of *TGM9* expression in the adult anole lizard (*A. carolinensis*). The house-keeping gene *EEF1A1* was amplified to confirm the integrity of tissue cDNAs in panels e) and f).

Other TGMs are expressed either predominantly in nonepithelial tissues or at similar levels in epithelial and nonepithelial tissue. Thus, the role of TGMs in epithelial cornification was either an ancestral trait of vertebrate TGMs with roles in nonepithelial tissues evolving for F13A1, TGM10, TGM2, and possibly TGM4, or TGM-mediated epithelial cornification evolved at least twice during the diversification of TGMs. Importantly, the evolution of epithelial TGMs was constrained by pleiotropic functions of TGMs, correlating with the expression of individual TGMs at more than one epithelial site, and influenced by redundancy of TGMs due to expression of more than one TGM at

particular epithelial sites. We propose that, in general, pleiotropy favors conservation of a gene, whereas redundancy favors gene loss. Our findings of TGM gene losses in cetaceans (Fig. 4) and loss of *TGM9* in therian mammals and legless squamates (Fig. 5) indicate that phenotypic adaptations can be associated with the degeneration of genes even when the genes have pleiotropic roles in other phylogenetic lineages. These data support the concept that the conservative effects of pleiotropy are limited by gene redundancy and parallel adaptations of multiple body sites (Hecker et al. 2017).

TGM1 is the phylogenetically oldest epithelial TGM (Fig. 6). Previous studies have shown expression of *TGM1*

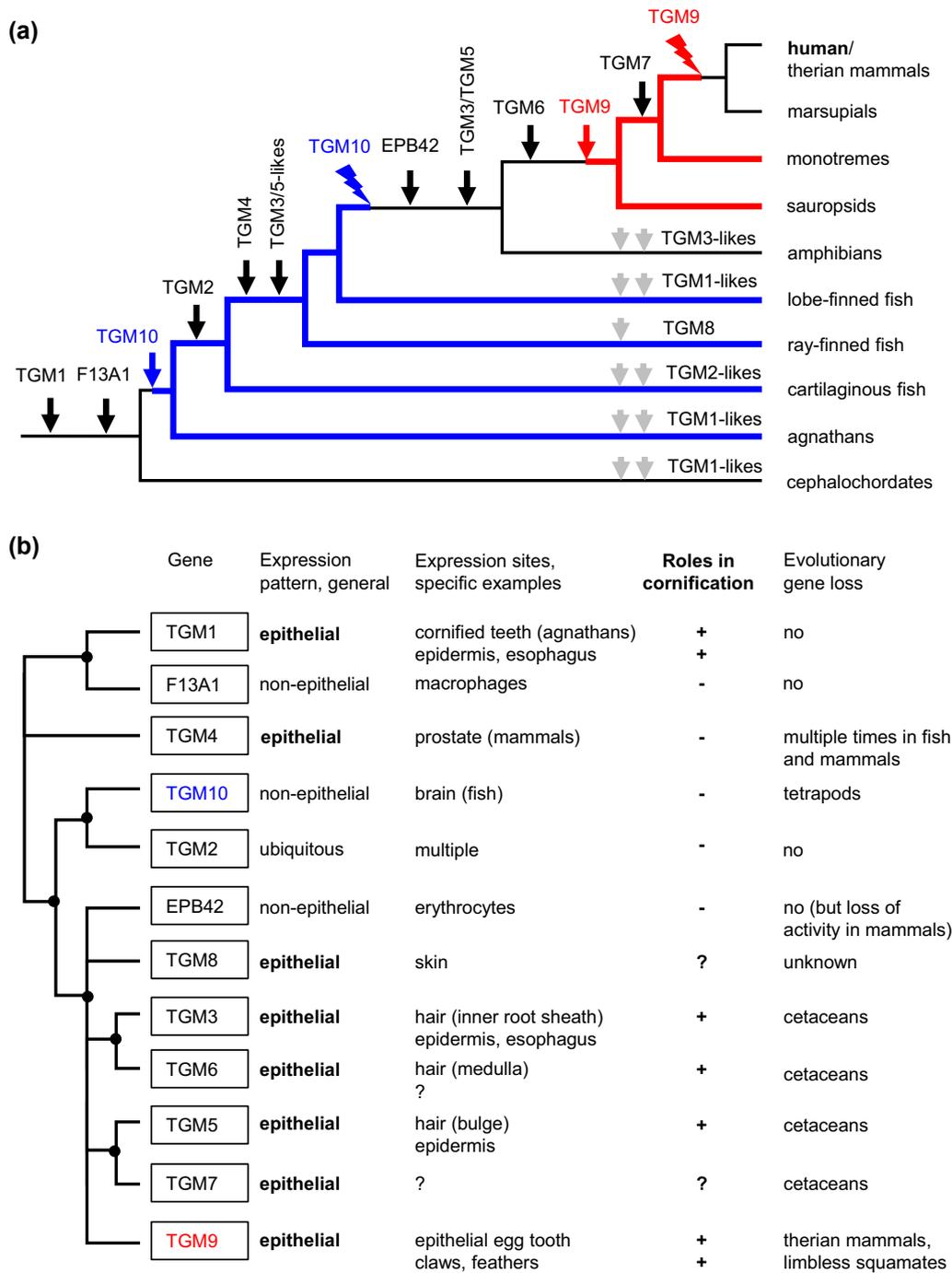


Fig. 6. Gene duplications and gene loss have shaped the repertoire of TGMs in humans and other vertebrates. a) The evolution of the TGM gene family is schematically depicted on a phylogenetic tree focusing on the lineage leading to humans. The origin (vertical arrows) and loss (flash symbols) of TGM genes are indicated at positions inferred from the species distribution of TGMs in extant species. The presence of *TGM10* and *TGM9* is indicated by blue and red colors of branches, respectively. b) Summary of functional diversification of TGMs in vertebrates. The phylogenetic relationships of TGMs are indicated by a simplified tree (left side) in which well-supported nodes (Fig. 2) are marked.

in the epidermis of mammals, birds, and ray-finned fish (Rodríguez Cruz et al. 2017). Here, we added proteomic evidence for the presence of *TGM1* orthologs in the cornified epithelial teeth of the sea lamprey, which diverged from the aforementioned *TGM1*-positive taxa more than 500 million years ago (Kumar et al. 2022). The group of *TGM3/TGM5*-like genes evolved more recently than *TGM1* and was not as strictly conserved as *TGM1*. Upon

the evolutionary transition to life in water and loss of hair in cetaceans, *TGM3*, *TGM5*, *TGM6*, and *TGM7* were pseudogenized or lost (Fig. 6b). The loss of functional *TGM6* in cetaceans and the evidence for expression of *TGM6* in hair follicles, particularly in the medulla of growing hair fibers (Supplementary table S9, Supplementary Material online), suggest that *TGM6* resembles its presumable parental gene, *TGM3*, with regard to a primary function in hair

(John et al. 2012). It is worth mentioning that a role of *TGM6* in the central nervous system, as discussed in the literature (Thomas et al. 2013), is not supported by the distribution of *TGM6* mRNA in human tissues, which is limited to the skin (<https://www.gtexportal.org/home/gene/TGM6>), last accessed on 2024 February 9), and a hypothetical link between mutations of *TGM6* and spinocerebellar ataxia (Wang et al. 2010) was refuted by a recent genetic study (Chen et al. 2020).

Our phylogenetic and synteny analyses define *TGM9* as a gene that evolved in amniotes prior to the divergence of the lineages leading to sauropsids and mammals. Previous studies of chicken keratinocytes cultured in vitro had detected *TGM9* mRNA (Couteaudier et al. 2015; Lachner et al. 2021), which was then named protein—glutamine gamma-glutamyltransferase 5-like according to the GenBank annotation or TG5 (TGM5, ENSGALG00000013762). The data of the current study show that *TGM9* is not orthologous to *TGM5* and, more importantly, reveal expression of *TGM9* in the epithelial egg tooth (Fig. 5, c and e), which is a cornified skin appendage that has played an essential role in the evolutionary transition of tetrapods to fully terrestrial life (Fenelon et al. 2023). The egg tooth must have co-evolved with the egg shell in stem amniotes to facilitate hatching. However, the epithelial egg tooth was lost in specific subgroups of amniotes, namely squamate reptiles, in which the epithelial egg tooth was functionally substituted by a calcified “real” tooth, and in therian mammals, which have replaced embryonic development *in ovo* by development *in utero*. Of note, monotremes have both an epithelial egg tooth (caruncle) and a calcified egg tooth, which, however, differs histologically from the egg tooth of squamates (Fenelon et al. 2023). We could confirm the presence of a complete *TGM9* gene in the platypus (Fig. 5; supplementary table S1, Supplementary Material online), but we found only one exon of *TGM9* in the current genome sequence assembly of the echidna (*Tachyglossus aculeatus*) (GenBank accession number NC_052070.1, nucleotides 28415425 to 28415602). Thus, the evolution of *TGM9* in monotremes requires further investigation. Based on gene expression evidence, we propose that *TGM9* contributes to the cornification of the epithelial egg tooth and thereby helps to establish the hardness required for its function in opening the egg shell. The loss of *TGM9* in viviparous mammals supports the primary association of *TGM9* with the epithelial egg tooth, whereas a role of *TGM9* in the claws of squamates is suggested by the conservation of *TGM9* in claw-bearing lizards and geckos (Fig. 5a).

Conclusion

In conclusion, the evolution of the TGM gene family was linked to important adaptations of the skin epithelium in vertebrates and, indirectly, to major changes in the embryonic development of tetrapods. Future studies will explore the possible coevolution of TGMs with their substrate proteins and the evolution of the mechanisms that regulate the expression of established and newly identified TGMs in vertebrates.

Materials and Methods

Comparative Genomics and Sequence Alignments

For the comparative genomic analysis, the genome sequences of human (GCF_000001405.40, International Human Genome Sequencing Consortium), Tasmanian devil (GCF_902635505.1, Stammnitz et al. 2023), platypus (GCF_004115215.2, Zhou et al. 2021), chicken (GCF_016699485.2, Vertebrate Genomes Project), alligator (GCF_000455745.1, Wan et al. 2013), turtle (GCF_016161935.1, Liu et al. 2022), worm lizard (GCF_030035675.1, Vertebrate Genomes Project), wall lizard (GCF_004329235.1, Andrade et al. 2019), corn snake (GCF_029531705.1), green anole lizard (GCF_000090745.1, Alföldi et al. 2011), Japanese gecko (GCF_001447785.1, Kim et al. 2016), frog (GCF_000004195.4, Mitros et al. 2019), lungfish (GCF_019279795.1, Wang et al. 2021), coelacanth (GCF_000225785.1, Amemiya et al. 2013), zebrafish (GCF_000002035.6, Howe et al. 2013), great white shark (GCF_017639515.1, Vertebrate Genomes Project), sea lamprey (GCF_010993605.1, Vertebrate Genomes Project), tunicate (GCF_000224145.3, Satou et al. 2008), and lancelet (GCF_000003815.2, Simakov et al. 2020) were investigated. TGM sequences were collected from NCBI GenBank. The chicken TGM1 homolog was identified in a previous study (Sachslehner, Surbek, et al. 2021). For several genes, the predictions of exon sequences were corrected according to the results of tBLASTn searches with other TGMs as queries (supplementary table S1, Supplementary Material online). *TGM1* and *TGM1L4* of the coelacanth (*Latimeria chalumnae*) lacked large parts of their coding sequences in the annotations available in GenBank. For identification of the complete coding sequences, RNA-seq data of pooled tissues of the coelacanth were downloaded from NCBI GenBank (accession number: SRX112771, run accession: SRR391920) with the prefetch (version 3.0.6) package of SRA Toolkit (<https://github.com/ncbi/sra-tools>, last accessed on 2024 January 18) and converted to the fastq format with fastq-dump (version 3.0.6) package of SRA Toolkit. A quality check was performed with FastQC (version 0.12.1, <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>, last accessed on 2024 January 18). The transcriptome was assembled with Trinity (Grabherr et al. 2011, version 2.15.1), and coding regions were predicted with TransDecoder (version 5.7.1, <https://github.com/TransDecoder/TransDecoder>, last accessed on 2024 January 18). The complete *TGM1* and *TGM1L4* coding sequences were identified via BLASTp (Altschul et al. 1990, version 2.14.0+). The assembled transcriptome is available at <https://doi.org/10.5281/zenodo.10619730>.

Molecular Phylogenetics

Multiple sequence alignments for phylogeny were created and manually trimmed with aliview (Larsson 2014). After trimming, the amino acid sequences contained only the residues that span from the “Transglut_N superfamily” domain to the second “Transglut_C superfamily” domain (supplementary fig. S10, Supplementary Material online). Prottest (version 3.0) (Abascal et al. 2005; Darriba et al.

2011) was used to calculate the amino acid substitution model. All available matrices and models with rate variation among sites were included. The Akaike information criterion was used to assess the likelihood of the predicted models (Akaike 1998). The best suited amino acid substitution model for the TGM phylogeny was LG (Le and Gascuel 2008). Maximum likelihood analysis was performed using PhyML (version 20120412, <https://github.com/stephaneguindon/phyml>, last accessed on 2024 February 9). Tree topology, branch length, and rate parameters were optimizing according to a published approach (Guindon and Gascuel 2003). Phylogenetic trees were visualized with FigTree (<http://tree.bio.ed.ac.uk/software/figtree/>, last accessed on 2024 February 2). Phylogenetic trees were edited with Inkscape (version 1.0.0.0; <https://inkscape.org/de/>, accessed on 2024 January 18).

Animals and Tissue Samples

Tissue samples were prepared from 21-d-old specific pathogen-free (SPF) chickens and chicken embryos from fertilized SPF eggs (VALO BioMedia, Osterholz-Scharmbeck, Germany) on days 9, 10, and 18 of incubation, corresponding to Hamburger and Hamilton stages 35, 36, and 44, respectively (Hamburger and Hamilton 1951). Samples were obtained from untreated chickens which were maintained in a trial approved by the ethics and animal welfare committee of University of Veterinary Medicine, Vienna, Austria, and the Austrian Federal Ministry of Education, Science and Research (license number BMBWF GZ: GZ-2021-0.842.250). Sea lamprey (*Petromyzon marinus*) tissue was generously provided by the Museum of Natural History Vienna (inventory number: NMW-63577). cDNAs from green anole lizards were available from a previous study (Strasser et al. 2014).

RNA Extraction and RT-PCR

Tissue samples were placed in RNA-later (Invitrogen) immediately after dissection and stored at 4 °C overnight and subsequently at −80 °C. RNA isolation was performed with Trizol (VWR) from homogenized tissues according to published protocols (Chomczynski 1993). cDNA synthesis was performed with the iScript® cDNA synthesis kit (Quantabio). cDNA was PCR-amplified with Dream Taq DNA polymerase (Thermo Scientific) using the following primers (Microsynth, Switzerland): chicken TGM9 (forward: 5′-AGCGTCCCTATTCTTCAGCA-3′, reverse: 5′-TAGCTTGTCCCTGCCACAGA-3′, product 222 bp, Fig 5e; forward: 5′-AGTTTTCTGTGACGCTTGGC-3′, reverse: 5′-ACAGAAGAATTCACCAGGG-3′, product 157 bp, Fig 5d), chicken eef1a1 (forward: 5′-GCCCCGAAGTTCCTGAAATC-3′, reverse: 5′-GGCCTTGATGACACCAACAG-3′, product 153 bp, Fig 5, d and e), anole lizard TGM9 (forward: 5′-GAGTGGCAGCAGTCTCAATG-3′, reverse: 5′-GCACCTCCTTCTCCAGATGT-3′, product 400 bp, Fig 5e), anole lizard eef1a1 (forward: 5′-TTGCCACACTGCCATATTG-3′, reverse: 5′-CGCTTTCTGTCAACTGCCT-3′, product 250 bp, Fig 5e).

Analysis of Zebrafish RNA-Seq Data

Publicly available zebrafish sequence read archives (SRA) were downloaded from NCBI GenBank (supplementary table S8, Supplementary Material online) with prefetch (version 3.0.10) package of SRA Toolkit (<https://github.com/ncbi/sra-tools>, last accessed on 2024 January 18) and converted to fastq files with fastq-dump (version 3.0.10) package of sratoolkit. Quality check was done with FastQC (version 0.12.1, <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>, last accessed on 2024 January 18). Expression analysis was performed with Salmon (version 1.10.1, Patro et al. 2017). The reference transcriptome was concatenated with the corresponding genome to create a decoy aware reference index, which avoids mapping of genomic reads, before the calculation of transcripts per million. Statistical significance was calculated with Analysis of Variance (ANOVA).

Histology, TGM In Situ Activity Labeling

An in situ TGM activity assay combined with immunofluorescence was performed based on a published protocol (Sachslehner et al. 2023). The samples were encircled with a liquid blocker (Daido Sangyo Co. Ltd.) and subsequently incubated with 2% bovine serum albumin (BSA, Sigma-Aldrich) in phosphate-buffered saline (PBS) with 0.05% Tween at room temperature for 30 min. Subsequently, the samples were incubated with 5 μM Alexa-fluor555-cadaverine (Thermo Fisher Scientific) in 0.1 M Tris-HCl pH 7.4 with either 5 mM CaCl₂ to facilitate TGM activity or 5 mM EDTA to suppress TGM activity (negative control) for 2 h at room temperature under protection from light. The reaction was stopped by incubating the samples in 25 mM EDTA in PBS for 5 min after which the samples were rinsed with PBS. Afterward, the sections were rinsed with PBS and mounted with Permafluor (Thermo Fisher Scientific). Sections were studied with an Olympus BX63 light microscope, and images were taken with an Olympus UC-90 camera using the cellSens Dimensions software (version 1.16).

Preparation of Riboprobes

A riboprobe template that anneals at the beginning of the coding region of chicken TGM9 was amplified by RT-PCR from chicken skin cDNA using the following primer pair: TGM9 (forward: 5′-CCTGAAGTGCCTTCCAAGTGC-3′ and reverse: 5′-GCACAGTGATGAGGTGTTTGG-3′), further ligated into a pGEM-T Easy vector (Promega). Plasmids containing an insert were amplified in *E. coli* (L2001, Promega), and plasmid DNA was isolated with the QIAprep spin miniprep kit 250 (Qiagen). The plasmids were sequenced (Microsynth) using SP6 primers to determine the orientation of the inserts. The plasmids that contained the insert in antisense and sense orientation were linearized via PCR with M13 primers (forward: 5′-CGCCAGGGTTTTCCAGTACAGAC-3′ and reverse: 5′-CAGGAAACAGCTATGAC-3′). The insert contained a T7 RNA polymerase-binding site and was used as

template for in vitro transcription with the DIG RNA labeling mix (Roche). T7 RNA polymerase was used for transcription of antisense and sense probes. RNA probes were precipitated with 4 M LiCl (Sigma-Aldrich) and ethanol overnight at -20°C , followed by subsequent washing steps with 70% and 100% ethanol and resuspension in RNase-free water. The riboprobes were stored at -80°C .

mRNA In Situ Hybridization

Sections were prepared with a Microm HM 335E microtome at a thickness of 5 μm and collected in a water bath set to 42°C on super-adhesive slides (Menzel). mRNA in situ hybridization was performed on tissue sections of formalin-fixed paraffin-embedded chicken beaks of stage HH36 (Hamburger and Hamilton 1951) based on published protocols (Mlitz et al. 2021; Sachslehner, Zieger, et al. 2021). Briefly, paraffin sections were melted at 58°C for 1 h, deparaffinized in xylene (Fisher Chemical) for 20 min, and further hydrated in a descending ethanol series. The samples were treated with 20 $\mu\text{g}/\text{ml}$ proteinase K for 10 min at 37°C . Specimens were subsequently incubated for 10 min in 1% triethanolamine (Sigma-Aldrich) in PBS, for 5 min in 1% triethanolamine with 0.15% acetic anhydride (Prolabo), and for 5 min in 1% triethanolamine with 0.3% acetic anhydride to neutralize charged probe binding. Samples were postfixed with 4% paraformaldehyde in PBS for 40 min followed by dehydration in an ascending ethanol series. Samples were dried at room temperature for 1 h and prehybridized in hybridization buffer at 60°C for 2 h. Corresponding sense and antisense riboprobes were prepared at a concentration of 2 $\text{ng}/\mu\text{l}$ and preheated in hybridization buffer (4 M urea, VWR; 5x saline sodium citrate (SSC), Roth; 100 $\mu\text{g}/\text{ml}$ heparin, Sigma-Aldrich; 5 mM EDTA, Roth; 1x Denhardt's block reagent, Sigma-Aldrich; 100 $\mu\text{g}/\text{ml}$ salmon sperm DNA, Agilent; 5% dextran sulfate, Sigma-Aldrich) at 85°C for 10 min. Hybridization took place overnight at 60°C . To remove unbound probes, samples were washed thrice with 4 M urea and 4x SSC, once with a solution containing 4 M urea and 2x SSC, and once with 4 M urea and 1x SSC at 58°C for 15 min each, followed by one wash with 1x SSC at 37°C for 15 min. Specimens were incubated for 2 h in 0.1 M maleic acid buffer (MAB, Sigma-Aldrich), pH 7.5 containing 2% BSA to prevent nonspecific antidigoxigenin antibody binding. Afterward, specimens were incubated with an antidigoxigenin antibody conjugated to the alkaline phosphatase enzyme (1:5,000, Roche) in 2% MAB block solution overnight at 4°C . A buffer suitable for alkaline phosphatase enzyme activity (0.5 M Tris pH 9.5; 0.5 M NaCl) was used, and samples were washed twice with alkaline phosphatase buffer for 15 min at room temperature. The staining was developed with alkaline phosphatase buffer supplemented with 0.05 M MgCl_2 , 3.75 $\mu\text{l}/\text{ml}$ 5-bromo-4-chloro-3-indolyl phosphate (Roche), and 5 $\mu\text{l}/\text{ml}$ nitroblue tetrazolium (Roche). Sections were mounted with Aquatex (Sigma-Aldrich), and photographs were taken with an Olympus UC-90 camera on an Olympus BX63 light microscope.

Mass Spectrometry-Based Proteomics

The beaks of stage HH44 chicken embryos were dissected and dried for some minutes at room temperature. The egg tooth became clear after the drying step and was removed from the beak with a pointy scalpel. The epidermis of the upper beak was removed with forceps. Samples were frozen in liquid nitrogen and stored at -80°C until lysate preparation. Lamprey teeth were obtained from a museum sample that was stored in ethanol at room temperature. The specimens were placed in 200 μl of lysis buffer, made of 30 mM Tris, 7 M urea (VWR), 2 M thiourea (Sigma-Aldrich), and 4% CHAPSO (Pierce). 0.2 M dithiothreitol (DTT) was added to the samples. After incubation at 70°C for 3 h, the samples were homogenized with a homogenizer (Precellys, VWR) and centrifuged at $18,000 \times g$ for 15 min at 4°C , and the supernatant was collected. The pellet was sonicated with a sonicator (Hielscher Ultrasound Technology) twice for 30 s at an amplitude of 100, and centrifugation was repeated as before. The supernatant was pooled with the one from the initial homogenization and stored at -80°C until analysis was started. The proteomic analysis and database search were performed like in a previous study (Cruz-Bustos et al. 2023) with the following modification: After protein reduction with 200 mM DTT (37°C , 30 min), the proteins were alkylated with 500 mM iodoacetamide (Sigma) at 37°C for 30 min. The database for the lamprey (NCBI_Petromyzon_marinus_tx7757_230919.fasta) and the chicken (NCBI_Gallus_gallus_tx9031_230919.fasta) was downloaded from *crap.fasta* (<https://www.thegpm.org/crap/>, last accessed on 2024 February 7).

Supplementary Material

Supplementary material is available at *Molecular Biology and Evolution* online.

Acknowledgments

The authors thank Dr. Anja Palandačić from the Museum of Natural History Vienna for providing sea lamprey tissue and Prof. Thomas Rattei for providing virtual workspace on the Vienna Life Science Cluster. This research was supported by resources of the VetCore Facility Proteomics of the University of Veterinary Medicine Vienna. M.S. is a recipient of a DOC fellowship of the Austrian Academy of Sciences at the Department of Dermatology, Medical University of Vienna.

Funding

This research was funded in part by the Austrian Science Fund (FWF): grant-DOI: 10.55776/P32777 and grant-DOI: 10.55776/P36596. For Open Access purposes, the authors have applied a CC BY public copyright license to any author accepted manuscript version arising from this submission.

Conflict of interest: The authors declare no conflict of interest.

Data Availability

Proteome data were deposited in the PRIDE database under the accession numbers PXD048875 (chicken egg tooth) and PXD048873 (lamprey teeth and skin). The assembled reads of the coelacanth transcriptome (NCBI read archive accession: SRX112771) were deposited at Zenodo (doi: <https://doi.org/10.5281/zenodo.10619730>).

References

- Abascal F, Zardoya R, Posada D. ProtTest: selection of best-fit models of protein evolution. *Bioinformatics*. 2005;21(9):2104–2105. <https://doi.org/10.1093/bioinformatics/bti263>.
- Akaike H. Information theory and an extension of the maximum likelihood principle. In: Parzen E, et al. editors. *Selected papers of Hirotugu Akaike*. New York: Springer; 1998. p. 199–213.
- Akat E, Yenmiş M, Pombal MA, Molist P, Megias M, Arman S, Vesely M, Anderson R, Ayaz D. Comparison of vertebrate skin structure at class level: a review. *Anat Rec (Hoboken)*. 2022;305(12):3543–3608. <https://doi.org/10.1002/ar.24908>.
- Alföldi J, Di Palma F, Grabherr M, Williams C, Kong L, Mauceli E, Russell P, Lowe CB, Glor RE, Jaffe JD, et al. The genome of the green anole lizard and a comparative analysis with birds and mammals. *Nature*. 2011;477(7366):587–591. <https://doi.org/10.1038/nature10390>.
- Alibardi L. Adaptation to the land: the skin of reptiles in comparison to that of amphibians and endotherm amniotes. *J Exp Zool B Mol Dev Evol*. 2003;298(1):12–41. <https://doi.org/10.1002/jez.b.24>.
- Alibardi L. Vertebrate keratinization evolved into cornification mainly due to transglutaminase and sulfhydryl oxidase activities on epidermal proteins: An immunohistochemical survey. *The Anatomical Record*. 2022;305(2):333–358. <https://doi.org/10.1002/ar.24705>.
- Alibardi L, Segalla A. The process of cornification in the horny teeth of the lamprey involves proteins in the keratin range and other keratin-associated proteins. *Zool Stud*. 2011;50(4):416–425.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 1990;215(3):403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2).
- Amemiya CT, Alföldi J, Lee AP, Fan S, Philippe H, Maccallum I, Braasch I, Manousaki T, Schneider I, Rohner N, et al. The African coelacanth genome provides insights into tetrapod evolution. *Nature*. 2013;496(7445):311–316. <https://doi.org/10.1038/nature12027>.
- Andrade P, Pinho C, Pérez I de Lanuza G, Afonso S, Brejcha J, Rubin CJ, Wallerman O, Pereira P, Sabatino SJ, Bellati A, et al. Regulatory changes in pterin and carotenoid genes underlie balanced color polymorphisms in the wall lizard. *Proc Natl Acad Sci U S A*. 2019;116(12):5633–5642. <https://doi.org/10.1073/pnas.1820320116>.
- Beckers CML, Simpson KR, Griffin KJ, Brown JM, Cheah LT, Smith KA, Vacher J, Cordell PA, Kearney MT, Grant PJ, et al. Cre/lox studies identify resident macrophages as the major source of circulating coagulation factor XIII-A. *Arterioscler Thromb Vasc Biol*. 2017;37(8):1494–1502. <https://doi.org/10.1161/ATVBAHA.117.309271>.
- Bereiter-Hahn J, Matoltsy AG, Richards KS. *Biology of the integument*. Berlin: Springer-Verlag; 1986.
- Candi E, Schmidt R, Melino G. The cornified envelope: a model of cell death in the skin. *Nat Rev Mol Cell Biol*. 2005;6(4):328–340. <https://doi.org/10.1038/nrm1619>.
- Carron M, Sachslehner AP, Cicekdağ MB, Bruggeman I, Demuyneck S, Golabi B, De Baere E, Declercq W, Tschachler E, Vlemminckx K, et al. Evolutionary origin of Hoxc13-dependent skin appendages in amphibians. *Nat Commun*. 2024;15(1):2328. <https://doi.org/10.1038/s41467-024-46373-x>.
- Cassidy AJ, van Steensel MA, Steijlen PM, van Geel M, van der Velden J, Morley SM, Terrinoni A, Melino G, McLean WHI. A homozygous missense mutation in TGM5 abolishes epidermal transglutaminase 5 activity and causes acral peeling skin syndrome. *Am J Hum Genet*. 2005;77(6):909–917. <https://doi.org/10.1086/497707>.
- Chen Y, Wu D, Luo B, Zhao G, Wang K. TGM6 I517w is not a pathogenic variant for spinocerebellar ataxia type 35. *Neurol Genet*. 2020;6(3):e424. <https://doi.org/10.1212/NXG.0000000000000424>.
- Chomczynski P. A reagent for the single-step simultaneous isolation of RNA, DNA and proteins from cell and tissue samples. *Biotechniques*. 1993;15(3):532–534, 536–537.
- Couteaudier M, Trapp-Fragner L, Auger N, Courvoisier K, Pain B, Denesvre C, Vautherot J-F. Derivation of keratinocytes from chicken embryonic stem cells: establishment and characterization of differentiated proliferative cell populations. *Stem Cell Res*. 2015;14(2):224–237. <https://doi.org/10.1016/j.scr.2015.01.002>.
- Cruz-Bustos T, Dolezal M, Feix AS, Ruttkowski B, Hummel K, Razzazi-Fazeli E, Joachim A. Unravelling the sexual developmental biology of *Cystoisospora suis*, a model for comparative coccidian parasite studies. *Front Cell Infect Microbiol*. 2023;13:1271731. <https://doi.org/10.3389/fcimb.2023.1271731>.
- Darriba D, Taboada GL, Doallo R, Posada D. ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics*. 2011;27(8):1164–1165. <https://doi.org/10.1093/bioinformatics/btr088>.
- Deasey S, Grichenko O, Du S, Nurminskaya M. Characterization of the transglutaminase gene family in zebrafish and in vivo analysis of transglutaminase-dependent bone mineralization. *Amino Acids*. 2012;42(2-3):1065–1075. <https://doi.org/10.1007/s00726-011-1021-0>.
- Dubbink HJ, Verkaik NS, Faber PW, Trapman J, Schröder FH, Romijn JC. Tissue specific and androgen-regulated expression of human prostate-specific transglutaminase. *Biochem J*. 1996;315(3):901–908. <https://doi.org/10.1042/bj3150901>.
- Eckert RL, Kaartinen MT, Nurminskaya M, Belkin AM, Colak G, Johnson GV, Mehta K. Transglutaminase regulation of cell function. *Physiol Rev*. 2014;94(2):383–417. <https://doi.org/10.1152/physrev.00019.2013>.
- Eckhart L, Lippens S, Tschachler E, Declercq W. Cell death by cornification. *Biochim Biophys Acta*. 2013;1833(12):3471–3480. <https://doi.org/10.1016/j.bbamcr.2013.06.010>.
- Ehrlich F, Fischer H, Langbein L, Praetzel-Wunder S, Ebner B, Figlak K, Weissenbacher A, Sipos W, Tschachler E, Eckhart L. Differential evolution of the epidermal keratin cytoskeleton in terrestrial and aquatic mammals. *Mol Biol Evol*. 2019;36(2):328–340. <https://doi.org/10.1093/molbev/msy214>.
- Ehrlich F, Lachner J, Hermann M, Tschachler E, Eckhart L. Convergent evolution of cysteine-rich keratins in hard skin appendages of terrestrial vertebrates. *Mol Biol Evol*. 2020;37(4):982–993. <https://doi.org/10.1093/molbev/msz279>.
- Fenelon JC, Bennetts A, Anthwal N, Pyne M, Johnston SD, Evans AR, Tucker AS, Renfree MB. Getting out of a mammalian egg: the egg tooth and caruncle of the echidna. *Dev Biol*. 2023;495:8–18. <https://doi.org/10.1016/j.ydbio.2022.12.005>.
- Fischer B, Metzger M, Richardson R, Knyphausen P, Ramezani T, Franzen R, Schmelzer E, Bloch W, Carney TJ, Hammerschmidt M. P53 and TAp63 promote keratinocyte proliferation and differentiation in breeding tubercles of the zebrafish. *PLoS Genet*. 2014;10(1):e1004048. <https://doi.org/10.1371/journal.pgen.1004048>.
- Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol*. 2011;29(7):644–652. <https://doi.org/10.1038/nbt.1883>.
- Greenberg CS, Birckbichler PJ, Rice RH. Transglutaminases: multifunctional cross-linking enzymes that stabilize tissues. *FASEB J*. 1991;5(15):3071–3077. <https://doi.org/10.1096/fasebj.5.15.1683845>.
- Grenard P, Bates MK, Aeschlimann D. Evolution of transglutaminase genes: identification of a transglutaminase gene cluster on human chromosome 15q15. Structure of the gene encoding

- transglutaminase X and a novel gene family member, transglutaminase Z. *J Biol Chem*. 2001;**276**(35):33066–33078. <https://doi.org/10.1074/jbc.M102553200>.
- Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol*. 2003;**52**(5):696–704. <https://doi.org/10.1080/10635150390235520>.
- Hamburger V, Hamilton HL. A series of normal stages in the development of the chick embryo. *J Morphol*. 1951;**88**(1):49–92. <https://doi.org/10.1002/jmor.1050880104>.
- Harland DP, Plowman JE. Development of hair fibres. *Adv Exp Med Biol*. 2018;**1054**:109–154. https://doi.org/10.1007/978-981-10-8195-8_10.
- Hecker N, Sharma V, Hiller M. Transition to an aquatic habitat permitted the repeated loss of the pleiotropic KLK8 gene in mammals. *Genome Biol Evol*. 2017;**9**(11):3179–3188. <https://doi.org/10.1093/gbe/evx239>.
- Howe K, Clark MD, Torroja CF, Torrance J, Berthelot C, Muffato M, Collins JE, Humphray S, McLaren K, Matthews L, et al. The zebrafish reference genome sequence and its relationship to the human genome. *Nature*. 2013;**496**(7446):498–503. <https://doi.org/10.1038/nature12111>.
- John S, Thiebach L, Frie C, Mokkapat S, Bechtel M, Nischt R, Rosser-Davies S, Paulsson M, Smyth N. Epidermal transglutaminase (TGase 3) is required for proper hair development, but not the formation of the epidermal barrier. *PLoS One*. 2012;**7**(4):e34252. <https://doi.org/10.1371/journal.pone.0034252>.
- Joost S, Annusver K, Jacob T, Sun X, Dalessandri T, Sivan U, Sequeira I, Sandberg R, Kasper M. The molecular anatomy of mouse skin during hair growth and rest. *Cell Stem Cell*. 2020;**26**(3):441–457.e7. <https://doi.org/10.1016/j.stem.2020.01.012>.
- Kim I-H, Park J, Cheon K-S, Lee H-J, Kim J-K, Park D. Complete mitochondrial genome of Schlegel's Japanese gecko *Gekko japonicus* (Squamata: Gekkonidae). *Mitochondrial DNA A DNA Mapp Seq Anal*. 2016;**27**(5):3684–3686. <https://doi.org/10.3109/19401736.2015.1079855>.
- Korsgren C, Lawler J, Lambert S, Speicher D, Cohen CM. Complete amino acid sequence and homologies of human erythrocyte membrane protein band 4.2. *Proc Natl Acad Sci U S A*. 1990;**87**(2):613–617. <https://doi.org/10.1073/pnas.87.2.613>.
- Kumar S, Suleski M, Craig JM, Kasprovic AE, Sanderford M, Li M, Stecher G, Hedges SB. TimeTree 5: an expanded resource for species divergence times. *Mol Biol Evol*. 2022;**39**(8):msac174. <https://doi.org/10.1093/molbev/msac174>.
- Kuramoto N, Takizawa T, Takizawa T, Matsuki M, Morioka H, Robinson JM, Yamanishi K. Development of ichthyosiform skin compensates for defective permeability barrier function in mice lacking transglutaminase 1. *J Clin Invest*. 2002;**109**(2):243–250. <https://doi.org/10.1172/JCI13563>.
- Lachner J, Derdak S, Mlitz V, Wagner T, Holthaus KB, Ehrlich F, Mildner M, Tschachler E, Eckhart L. An in vitro model of avian skin reveals evolutionarily conserved transcriptional regulation of epidermal barrier formation. *J Invest Dermatol*. 2021;**141**(12):2829–2837. <https://doi.org/10.1016/j.jid.2021.04.029>.
- Larsson A. AliView: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics*. 2014;**30**(22):3276–3278. <https://doi.org/10.1093/bioinformatics/btu531>.
- Le SQ, Gascuel O. An improved general amino acid replacement matrix. *Mol Biol Evol*. 2008;**25**(7):1307–1320. <https://doi.org/10.1093/molbev/msn067>.
- Leskinen PK, Laaksonen T, Ruuskanen S, Primmer CR, Leder EH. The proteomics of feather development in pied flycatchers (*Ficedula hypoleuca*) with different plumage coloration. *Mol Ecol*. 2012;**21**(23):5762–5777. <https://doi.org/10.1111/mec.12073>.
- Lisetto M, Fattorini M, Lanza A, Gerdol M, Griffin M, Wang Z, Ferrara F, Sblattero D. Biochemical and functional characterization of the three zebrafish transglutaminases 2. *Int J Mol Sci*. 2023;**24**(15):12041. <https://doi.org/10.3390/ijms241512041>.
- Liu J, Liu S, Zheng K, Tang M, Gu L, Young J, Wang Z, Qiu Y, Dong J, Gu S, et al. Chromosome-level genome assembly of the Chinese three-keeled pond turtle (*Mauremys reevesii*) provides insights into freshwater adaptation. *Mol Ecol Resour*. 2022;**22**(4):1596–1605. <https://doi.org/10.1111/1755-0998.13563>.
- Lorand L, Iismaa SE. Transglutaminase diseases: from biochemistry to the bedside. *FASEB J*. 2019;**33**(1):3–12. <https://doi.org/10.1096/fj.201801544R>.
- Matsuki M, Yamashita F, Ishida-Yamamoto A, Yamada K, Kinoshita C, Fushiki S, Ueda E, Morishima Y, Tabata K, Yasuno H, et al. Defective stratum corneum and early neonatal death in mice lacking the gene for transglutaminase 1 (keratinocyte transglutaminase). *Proc Natl Acad Sci U S A*. 1998;**95**(3):1044–1049. <https://doi.org/10.1073/pnas.95.3.1044>.
- Menon GK, Elias PM, Wakefield JS, Crumrine D. Cetacean epidermal specialization: a review. *Anat Histol Embryol*. 2022;**51**(5):563–575. <https://doi.org/10.1111/ahc.12829>.
- Mitros T, Lyons JB, Session AM, Jenkins J, Shu S, Kwon T, Lane M, Ng C, Grammer TC, Khokha MK, et al. A chromosome-scale genome assembly and dense genetic map for *Xenopus tropicalis*. *Dev Biol*. 2019;**452**(1):8–20. <https://doi.org/10.1016/j.ydbio.2019.03.015>.
- Mlitz V, Hermann M, Buchberger M, Tschachler E, Eckhart L. The trichohyalin-like protein scaffoldin is expressed in the multi-layered periderm during development of avian beak and egg tooth. *Genes (Basel)*. 2021;**12**(2):248. <https://doi.org/10.3390/genes12020248>.
- Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods*. 2017;**14**(4):417–419. <https://doi.org/10.1038/nmeth.4197>.
- Phillips MA, Stewart BE, Rice RH. Genomic structure of keratinocyte transglutaminase. Recruitment of new exon for modified function. *J Biol Chem*. 1992;**267**(4):2282–2286. [https://doi.org/10.1016/S0021-9258\(18\)45875-9](https://doi.org/10.1016/S0021-9258(18)45875-9).
- Rice RH, Crumrine D, Hohl D, Munro CS, Elias PM. Cross-linked envelopes in nail plate in lamellar ichthyosis. *Br J Dermatol*. 2003;**149**(5):1050–1054. <https://doi.org/10.1111/j.1365-2133.2003.05510.x>.
- Rice RH, Wong VJ, Pinkerton KE. Ultrastructural visualization of cross-linked protein features in epidermal appendages. *J Cell Sci*. 1994;**107**(7):1985–1992. <https://doi.org/10.1242/jcs.107.7.1985>.
- Rodriguez Cruz SI, Phillips MA, Kultz D, Rice RH. Tgm1-like transglutaminases in tilapia (*Oreochromis mossambicus*). *PLoS One*. 2017;**12**(5):e0177016. <https://doi.org/10.1371/journal.pone.0177016>.
- Sachslehner AP, Surbek M, Golabi B, Geiselhofer M, Jäger K, Hess C, Kuchler U, Gruber R, Eckhart L. Transglutaminase activity is conserved in stratified epithelia and skin appendages of mammals and birds. *Int J Mol Sci*. 2023;**24**(3):2193. <https://doi.org/10.3390/ijms24032193>.
- Sachslehner AP, Surbek M, Lachner J, Paudel S, Eckhart L. Identification of chicken transglutaminase 1 and in situ localization of transglutaminase activity in avian skin and esophagus. *Genes (Basel)*. 2021;**12**(10):1565. <https://doi.org/10.3390/genes12101565>.
- Sachslehner A, Zieger E, Calcino A, Wanninger A. HES and Mox genes are expressed during early mesoderm formation in a mollusk with putative ancestral features. *Sci Rep*. 2021;**11**(1):18030. <https://doi.org/10.1038/s41598-021-96711-y>.
- Satou Y, Mineta K, Ogasawara M, Sasakura Y, Shoguchi E, Ueno K, Yamada L, Matsumoto J, Wasserscheid J, Dewar K, et al. Improved genome assembly and evidence-based global gene model set for the chordate *Ciona intestinalis*: new insight into intron and operon populations. *Genome Biol*. 2008;**9**(10):R152. <https://doi.org/10.1186/gb-2008-9-10-r152>.
- Sharma V, Hecker N, Roscito JG, Foerster L, Langer BE, Hiller M. A genomics approach reveals insights into the importance of gene losses for mammalian adaptations. *Nat Commun*. 2018;**9**(1):1215. <https://doi.org/10.1038/s41467-018-03667-1>.
- Shibata T, Kawabata SI. Pluripotency and a secretion mechanism of *Drosophila* transglutaminase. *J Biochem*. 2018;**163**(3):165–176. <https://doi.org/10.1093/jb/mvx059>.
- Simakov O, Marlétaz F, Yue J-X, O'Connell B, Jenkins J, Brandt A, Calef R, Tung C-H, Huang T-K, Schmutz J, et al. Deeply

- conserved syntenies resolves early events in vertebrate evolution. *Nat Ecol Evol.* 2020;**4**(6):820–830. <https://doi.org/10.1038/s41559-020-1156-z>.
- Stammnitz MR, Gori K, Kwon YM, Harry E, Martin FJ, Billis K, Cheng Y, Baez-Ortega A, Chow W, Comte S, et al. The evolution of two transmissible cancers in Tasmanian devils. *Science.* 2023;**380**(6642):283–293. <https://doi.org/10.1126/science.abq6453>.
- Strasser B, Mlitz V, Hermann M, Rice RH, Eigenheer RA, Alibardi L, Tschachler E, Eckhart L. Evolutionary origin and diversification of epidermal barrier proteins in amniotes. *Mol Biol Evol.* 2014;**31**(12):3194–3205. <https://doi.org/10.1093/molbev/msu251>.
- Strasser B, Mlitz V, Hermann M, Tschachler E, Eckhart L. Convergent evolution of cysteine-rich proteins in feathers and hair. *BMC Evol Biol.* 2015;**15**:82. <https://doi.org/10.1186/s12862-015-0360-y>.
- Sukseree S, Karim N, Jaeger K, Zhong S, Rossiter H, Nagelreiter IM, Gruber F, Tschachler E, Rice RH, Eckhart L. Autophagy controls the protein composition of hair shafts. *J Invest Dermatol.* 2024;**144**(1):170–173.e4. <https://doi.org/10.1016/j.jid.2023.06.199>.
- Surbek M, Van de Steene T, Sachslehner AP, Golabi B, Griss J, Eyckerman S, Gevaert K, Eckhart L. Cornification of keratinocytes is associated with differential changes in the catalytic activity and the immunoreactivity of transglutaminase-1. *Sci Rep.* 2023;**13**(1):21550. <https://doi.org/10.1038/s41598-023-48856-1>.
- Thibaut S, Cavusoglu N, de Becker E, Zerbib F, Bednarczyk A, Schaeffer C, van Dorsselaer A, Bernard BA. Transglutaminase-3 enzyme: a putative actor in human hair shaft scaffolding? *J Invest Dermatol.* 2009;**129**(2):449–459. <https://doi.org/10.1038/jid.2008.231>.
- Thomas H, Beck K, Adamczyk M, Aeschlimann P, Langley M, Oita RC, Thiebach L, Hils M, Aeschlimann D. Transglutaminase 6: a protein associated with central nervous system development and motor function. *Amino Acids.* 2013;**44**(1):161–177. <https://doi.org/10.1007/s00726-011-1091-z>.
- Tian X, Pascal G, Fouchécourt S, Pontarotti P, Monget P. Gene birth, death, and divergence: the different scenarios of reproduction-related gene evolution. *Biol Reprod.* 2009;**80**(4):616–621. <https://doi.org/10.1095/biolreprod.108.073684>.
- Ü Basmanav FB, Cau L, Tafazzoli A, Méchin MC, Wolf S, Romano MT, Valentin F, Wiegmann H, Huchenq A, Kandil R, et al. Mutations in three genes encoding proteins involved in hair shaft formation cause uncombable hair syndrome. *Am J Hum Genet.* 2016;**99**(6):1292–1304. <https://doi.org/10.1016/j.ajhg.2016.10.004>.
- Wan Q-H, Pan S-K, Hu L, Zhu Y, Xu P-W, Xia J-Q, Chen H, He G-Y, He J, Ni X-W, et al. Genome analysis and signature discovery for diving and sensory properties of the endangered Chinese alligator. *Cell Res.* 2013;**23**(9):1091–1105. <https://doi.org/10.1038/cr.2013.104>.
- Wang K, Wang J, Zhu C, Yang L, Ren Y, Ruan J, Fan G, Hu J, Xu W, Bi X, et al. African lungfish genome sheds light on the vertebrate water-to-land transition. *Cell.* 2021;**184**(5):1362–1376.e18. <https://doi.org/10.1016/j.cell.2021.01.047>.
- Wang JL, Yang X, Xia K, Hu ZM, Weng L, Jin X, Jiang H, Zhang P, Shen L, Guo JF, et al. TGM6 identified as a novel causative gene of spinocerebellar ataxias using exome sequencing. *Brain.* 2010;**133**(12):3510–3518. <https://doi.org/10.1093/brain/awq323>.
- Wang X, Zhang A, Huang M, Chen L, Hu Q, Lu Y, Cheng L. Genetic and clinical characteristics of patients with hereditary spherocytosis in Hubei province of China. *Front Genet.* 2020;**11**:953. <https://doi.org/10.3389/fgene.2020.00953>.
- Wu P, Hou L, Plikus M, Hughes M, Scehnet J, Suksaweang S, Widelitz R, Jiang TX, Chuong CM. Evo-Devo of amniote integuments and appendages. *Int J Dev Biol.* 2004;**48**(2-3):249–270. <https://doi.org/10.1387/ijdb.041825pw>.
- Yamane A, Fukui M, Sugimura Y, Itoh M, Alea MP, Thomas V, El Alaoui S, Akiyama M, Hitomi K. Identification of a preferred substrate peptide for transglutaminase 3 and detection of in situ activity in skin and hair follicles. *FEBS J.* 2010;**277**(17):3564–3574. <https://doi.org/10.1111/j.1742-4658.2010.07765.x>.
- Zaccone G, Howie AJ, Mauceri A, Fasulo S, Lo Cascio P, Youson JH. Distribution patterns of cytokeratins in epidermis and horny teeth of the adult sea lamprey, *Petromyzon marinus*. *Folia Histochem Cytobiol.* 1995;**33**(2):69–75.
- Zhang H, Ericsson M, Weström S, Vahlquist A, Virtanen M, Törmä H. Patients with congenital ichthyosis and TGM1 mutations overexpress other ARCI genes in the skin: part of a barrier repair response? *Exp Dermatol.* 2019;**28**(10):1164–1171. <https://doi.org/10.1111/exd.13813>.
- Zhou Y, Shearwin-Whyatt L, Li J, Song Z, Hayakawa T, Stevens D, Fenelon JC, Peel E, Cheng Y, Pajpach F, et al. Platypus and echidna genomes reveal mammalian biology and evolution. *Nature.* 2021;**592**(7856):756–762. <https://doi.org/10.1038/s41586-020-03039-0>.