Biochem Genet (2013) 51:413–425
DOI 10.1007/s10528-013-9574-0

Conserved or Lost: Molecular Evolution of the Key Gene *GULO* in Vertebrate Vitamin C Biosynthesis

Hongwen Yang

Received: 10 June 2012 / Accepted: 21 September 2012 / Published online: 12 February 2013

Abstract L-gulono-gamma-lactone oxidase (GULO) catalyzes the final step in vertebrate vitamin C biosynthesis. Vitamin C-incapable vertebrates lack the *GULO* gene. Gene structure and phylogenetic analyses showed that vertebrate *GULO* genes are 64–95% identical at the amino acid level and consist of 11 conserved exons. *GULO* pseudogenes have multiple indel mutations and premature stop codons in higher primates, guinea pigs, and some bats. No *GULO*-like sequences were identified in teleost fishes. During animal *GULO* evolution, exon F was subdivided into F1 and F2. Additional *GULO* retropseudogenes were identified in dogs, cats, and giant pandas. *GULO*-flanking genome regions acquired frequent segment translocations and inversions during vertebrate evolution. Purifying selection was detected across vertebrate *GULO* genes ($d_N/d_S = 0.069$), except for some positively selected sites identified in sharks and frogs. These positive sites demonstrated little functional significance when mapped onto the three-dimensional GULO protein structure. Vertebrate *GULO* genes are conserved except for those that are lost.

Keywords Vertebrates · *GULO* loss · Synteny analysis · Selective pressure

Introduction

Vitamin C (L-ascorbic acid, AsA) is an indispensable, recycled reducing nutrient that acts as an antioxidant and is involved in animal collagen biosynthesis (Drouin

Electronic supplementary material The online version of this article (doi: 10.1007/s10528-013-9574-0) contains supplementary material, which is available to authorized users.

H. Yang (✉)
Biology Department, Zhangzhou Normal University, 36 Xian Qian Zhi Street, Zhangzhou 363000, Fujian, China
e-mail: hwyang2002@sina.com

et al. 2011). AsA is biosynthesized from glucose in animal livers or kidneys, with the terminal step being catalyzed by L-gulono-gamma-lactone oxidase (GULO, EC 1.1.3.8). As an important antioxidant, AsA requirement and the correlated GULO activity often vary depending on different oxidative stress levels and are further correlated with varying energy requirements in animals. In addition, animal GULO activity also fluctuates according to dietary AsA intake. For example, GULO activity in rabbits during the winter is higher than in the summer, indicating that less AsA is obtained from grass or that more energy is needed for rabbits in winter. Decreasing GULO activity was observed in rabbits fed indoors, implying that less energy, lower oxidation level, and thus smaller antioxidative capacity are required for indoor feeding than for outdoor feeding (Jenness et al. 1978). Sturgeons have higher *GULO* expression levels during the embryonic stage than during any other developmental period, implying that more AsA is required for early development (Akbarzadeh et al. 2011). A similar situation was observed in other vertebrate species (Jenness et al. 1984).

During vertebrate evolution, some species, including anthropoid primates (Nishikimi et al. 1994), guinea pigs (Nishikimi et al. 1992), teleost fishes (Dabrowski 1990), some bats (Cui et al. 2011), and certain passerine birds (Chaudhuri and Chatterjee 1969), lost their ability to synthesize vitamin C, mainly through the loss of their *GULO* genes. Often, the antioxidative role of AsA in these *GULO*-less vertebrates is replaced by other advanced reducing nutrients or by reducing systems such as SOD, vitamin E, and uric acid (Nandi et al. 1997; Nakata and Maeda 2002; Spagnuolo et al. 2001). The role of AsA in collagen biosynthesis, however, is irreplaceable. As a result, these species are prone to scurvy (Maeda et al. 2000), osteoporosis, and fractures (Togari et al. 1995), as well as arteriosclerosis (Nakata and Maeda 2002), when they cannot take in enough dietary vitamin C. To date, the genetic and genomic aspects of this gene loss remain unclear, as do the phylogeny and molecular evolution of authentic *GULO*s. Here, the gene structure and synteny of vertebrate *GULO* genes were analyzed, along with selective pressure identification, to provide some clues about these issues.

Materials and Methods

Data Collection and Gene Structure Analysis

In the NCBI genome database and ENSEMBL (release December 2011), 54 animal genomes, from sponges to humans, were searched for *GULO*-like sequences. First, some identified *GULO* coding sequences were collected via searches for keywords such as "GULO" and "L-gulono gamma lactone oxidase." Second, for animals in which *GULO* was unidentified or lost, Blastn and tBlastp searches were performed against their genome sequences or protein databases, using porcine *GULO* as a query to gather putative *GULO* orthologs. The exon–intron structures of the examined *GULO* genes were identified via comparison of their cDNA and gDNA sequences. Similarities among animal *GULO* genes, both in their exon–intron

structure and nucleotide (or amino acid) sequences, were combined with the subsequent synteny information to confirm these newly identified *GULO* orthologs. *GULO* pseudoexons were assumed to be lost when no valid hit was obtained from Blastn against sequenced syntenic regions with matched porcine *GULO* exon sequences as queries.

Phylogenetic and Synteny Analyses

Protein sequences of newly identified *GULO* genes were deduced from their transcripts, and those of *GULO* pseudogenes were deduced from their in-frame alignment with the porcine *GULO* coding sequence, after recovery of indels and deletion of premature stop codons. These GULO protein sequences were then aligned using Clustal X 2.1 and used to construct a phylogenetic tree by the neighbor-joining method in software Mega 5.0 (http://www.megasoftware.net/index.php). The reliability of each node in the phylogenetic tree was assessed by bootstrapping with 1,000 replications. Synteny information for vertebrate *GULO*s, represented by zebrafish (bony fishes), western clawed frog (amphibians), anole lizard (reptiles), turkey (birds), human and rabbit (mammals), was collected from their annotated genome regions in ENSEMBL. Synteny analysis of cartilaginous fishes and lancelets was impractical because of their incomplete genome sequences or invalid genome annotations.

Selective Pressure Identification

Identification of selective pressure on vertebrate *GULO* genes was conducted using the CodeML program in PAML. In CodeML, the site model (model = 0, NSsites = 1 or 2; neutral M1a versus positive M2a) was used to identify the common positively selected sites and the d_N/d_S ratio (ω) across the whole set of vertebrate *GULO*s. The branch-site model (model = 2, NSsites = 2; neutral MA1 versus positive MA) was used to identify selective pressure on each vertebrate lineage (Yang 2007). To ascertain the potential influence of positively selected sites on GULO functions, the UCSF Chimera package (Pettersen et al. 2004) was used to map them onto the three-dimensional (3D) structure (PDB:2VFS) of alditol oxidase (EC 1.1.3.41) (Forneris et al. 2008), which is highly similar to the GULO protein. The alditol oxidase protein shares common domains with GULO proteins, including the cofactor FAD-binding domain and the ALO (D-arabinono-1,4-lactone oxidase) motif, highly conserved FAD-binding and ligand-binding residues, as well as similar substrates and the same catalytic mechanism (Supplementary Fig. 1) (Yamashita et al. 2000). It is interesting that the shared ALO domain (EC 1.1.3.37) itself can also catalyze the final step in AsA biosynthesis from L-gulono-1,4-lactone or D-arabinono-γ-lactone (Sauer et al. 2004; Biyani and Madhubala 2011). The amino acid residues interacting with cofactors and ligands were also identified by Forneris et al. (2008).

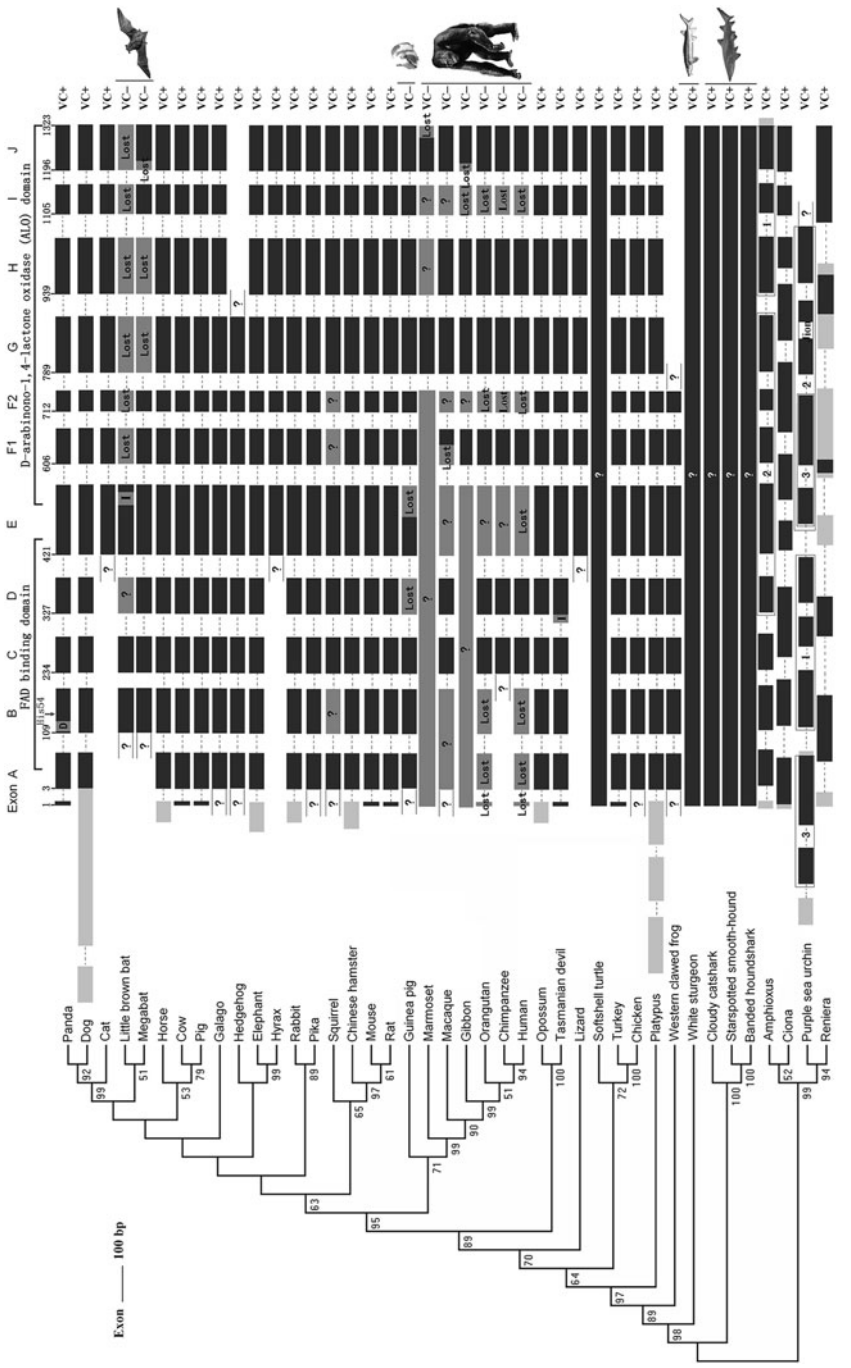
Fig. 1 Phylogenetic relationships and exon–intron structures of animal *GULO* genes (pseudogenes). The ▶ simplified phylogenetic tree (*left*) was constructed from GULO protein sequences using the neighbor-joining method in Mega 5.0 with 1,000 replicates. Amino acid sequences of *GULO* pseudogenes in primates, guinea pigs, and some bats were deduced from their in-frame DNA sequence alignment with the porcine *GULO* coding sequence after recovery of indel mutations. Bootstrap values beyond 50 are shown below the nodes. The branches align with the *GULO* gene structures (*right*) for each species. Common names of animals refer to species listed with their loci in Supplementary Table 1. The positions of the 11 exons conserved in the vertebrate genomes are identified by *capital letters* (A, B, C, D, E, F1, F2, G, H, I, J) above the domain designations. Gene structures: homologous exons (*black*), nonhomologous exons (*light gray*), unknown sequences (*dark gray* marked "?"), sequence deletions (*dark gray* marked "D"), sequence insertions (*dark gray* marked "I"), exons that have accumulated too many mutations to be recognized (*dark gray* marked "Lost"), unknown transcription (*unshaded* marked "?"), unknown exon–intron structures (*black with white* "?"), homologous joined exons in invertebrates (*boxes* numbered 1, 2, or 3). *His54 arrow,* above the FAD-binding domain, indicates the conserved histidine covalently binding cofactor FAD. *VC* indicates the animal's ability to synthesize AsA (yes VC+, no VC−). Exons are drawn to scale; introns (*dotted lines* between exons) are not

Results and Discussion

Conserved Exon–Intron Structures and High Sequence Identity of Vertebrate *GULO* Genes

Of the 54 animal species examined, 39 were identified as possessing a *GULO* gene or its pseudogenes (Supplementary Table 1). Among the 15 *GULO*-less animals, some (such as teleost fishes) have really lost their *GULO* genes, but others (sheep and dolphins) may possess authentic genes that cannot be identified at present because of their incomplete genome sequences. The well-supported phylogenetic tree (bootstrap values for most nodes are over 70%) generally conforms to common knowledge of animal phylogeny, including placental mammals, nonplacental mammals, birds, reptiles, amphibians, cartilaginous fishes, lancelets, and sea squirts, as well as other invertebrates such as the purple sea urchin and sponges (Fig. 1). The range of sequence identity among GULO proteins varies from 64–95% for vertebrates to 47–95% for chordates and 29% between pigs and sponges. Besides high sequence similarities, vertebrate *GULO*s are also highly conserved in their exon–intron structures. All identified vertebrate *GULO* orthologs share 11 conserved exons (A, B, C, D, E, F1, F2, G, H, I, J). In addition, the *GULO* gene structure of amphioxus, a close relative of vertebrates, is similar to that of vertebrate species. The urochordate sea squirt and achordate sponge both possess *GULO* gene structures that are distinct from those in vertebrates and amphioxus (Fig. 1).

Of the conserved *GULO* exons, A, B, C, and D form the FAD-binding domain, in which a completely conserved histidine residue covalently binds the cofactor FAD (Fig. 1). The remaining seven conserved exons constitute another essential ALO domain, in which several conserved residues are responsible for binding ligands. Apart from the 11 highly conserved *GULO* exons, other upstream exons in various species show no homology, and the 11 conserved exons share little sequence similarity with each other or with their upstream additional exons (data not shown). This suggests that no exon duplication has occurred during vertebrate *GULO* evolution, at least for those with identified gene structures. Some indels were also identified in the 11 conserved exons, such as a deletion of four codons in exon B of


Exon A
B FAD binding domain
C
D
E
F1 F2 D-arabinono-1,4-lactone oxidase (ALO) domain
G
H
I
J
VC+
VC−
Lost
Exon —— 100 bp
Panda
Dog
Cat
Little brown bat
Megabat
Horse
Cow
Pig
Galago
Hedgehog
Elephant
Hyrax
Rabbit
Pika
Squirrel
Chinese hamster
Mouse
Rat
Guinea pig
Marmoset
Macaque
Gibbon
Orangutan
Chimpanzee
Human
Opossum
Tasmanian devil
Lizard
Softshell turtle
Turkey
Chicken
Platypus
Western clawed frog
White sturgeon
Cloudy catshark
Starspotted smooth-hound
Banded houndshark
Amphioxus
Ciona
Purple sea urchin
Reniera

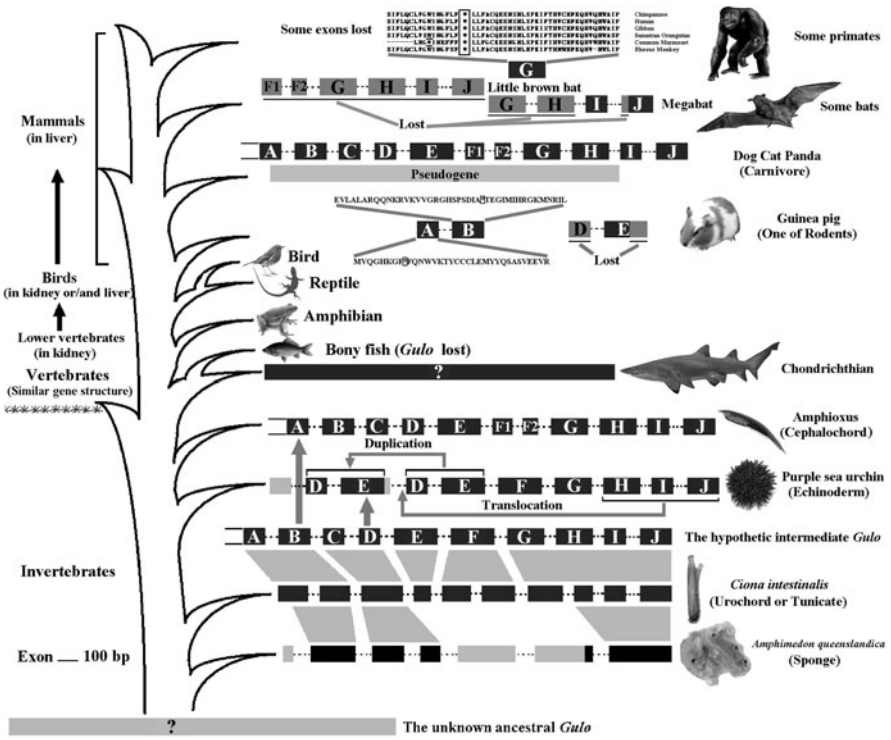


Fig. 2 Putative evolutionary mode of animal *GULO* genes (pseudogenes). *Top* Deduced amino acid sequences from *GULO* pseudoexons in primates (*boxed asterisks,* in-frame premature stop codons). Conserved vertebrate exons and other gene structures: homologous exons are *shaded black*, with conserved exon indicated by *white capital letter*; nonhomologous exons are *shaded light gray*. In bats and guinea pig, lost exons are *shaded dark gray* with *black capital letter* (*boxed asterisks* within guinea pig amino acid sequence, in-frame premature stop codons). Exons are drawn to scale; introns (*dotted lines* between exons) are not

the giant panda *GULO* and an insertion of three codons in exon D of the Tasmanian devil *GULO* (Fig. 1).

Evolution and Loss of Vertebrate *GULO* Genes

Several AsA-deficient vertebrate species depend on their diet to supply this vitamin. In humans, guinea pigs, and some bats, the loss of the *GULO* gene is the genetic reason for their inability to manufacture AsA (Nishikimi et al. 1992, 1994; Cui et al. 2011). Here, the presence of each of the 11 conserved *GULO* exons was examined in these AsA-dependent animals (Figs. 1 and 2).

Higher primates, including the human, orangutan, chimpanzee, gibbon, and macaque, have lost nearly half of the 11 exons. In the completely sequenced human genome, only six recognizable exons have been retained and are interspersed with numerous substitutions, multiple indels, and premature stop codons; the other five exons have disappeared without a trace, including exon B encoding a part of the FAD-binding domain. Among primate species, anthropoids are not the only lineage that depends on

dietary vitamin C. Tarsiers (lower members of the Haplorrhini suborder) also possess no *GULO*-like sequences in their genome, corresponding to the lack of GULO enzyme activity in the liver (Pollock and Mullin 1987). Another member of the Haplorrhini, the marmoset, cannot biosynthesize AsA because of the loss of the 3′ part of exon J from its *GULO* pseudogene. In contrast, the bushbaby, a lower member of the Strepsirrhini suborder, has an intact functional *GULO* gene, consistent with its ability to biosynthesize AsA (Drouin et al. 2011). In the two primate suborders, all the Strepsirrhini, such as bushbabies, have an intact *GULO* gene and can biosynthesize AsA themselves, while all species belonging to the higher suborder Haplorrhini lack *GULO* (Pollock and Mullin 1987). In this study, most *GULO* pseudogenes identified among the Haplorrhini demonstrated similar mutation patterns, with the loss of exons A, B, E, F2, and I (Fig. 1). This suggests that the *GULO* gene was inactivated in the common ancestor of all Haplorrhini primates when it diverged from the Strepsirrhini about 63 million years ago (MYA), which is almost consistent with the inactivation dates of primate *GULO* genes (about 61 MYA) calculated from the comparison between functional and nonfunctional *GULO* sequences (Lachapelle and Drouin 2011).

Chiropteran species, such as the little brown bat and the megabat, have also lost some *GULO* exons belonging to the ALO domain; however, those encoding the FAD-binding domain are retained (Fig. 1). All other bats, except *Rousettus leschenaultii* and *Hipposideros armiger,* possess no active *GULO,* regardless of whether their diets are rich or poor in AsA (Birney et al. 1976). Thus, the loss of bat *GULO* genes seems to be a result of neutral selection. Concerning this neutral selection, some bats, such as *R. leschenaultii* and *H. armiger,* regained active *GULO* but at a low expression level because of random mutations in the promoter of this gene. It has been speculated that the two newly reactivated *GULO* genes were likely to be silenced again and begin their pseudogenization (Cui et al. 2011). These repeated losses and gains of the *GULO* gene were also observed in some passerine birds (Martinez del Rio 1997), implying a relaxed, neutral selection, free of dietary AsA supplement, acting on *GULO* evolution in bats and passerine birds.

Compared with the highly mutated *GULO* pseudogenes in higher primates and bats, the gene of the guinea pig is relatively intact, having lost only exon D and the 3′ part of exon E, accompanied by premature stop codons in exons A and B, during *GULO* pseudogenization (Fig. 2). The small number of mutations in guinea pigs suggests a more recent inactivation date (about 14 MYA; Lachapelle and Drouin 2011). In addition, Blastn searches using the pig and white sturgeon *GULO* coding sequences as queries detected no *GULO*-like sequences in bony fishes such as cod, fugu, medaka, tilapia, and stickleback, partly because of their unfinished genome sequencing. It is certain, however, that *GULO* has been lost entirely from the zebrafish genome, one of the completely sequenced bony fishes.

The authentic functional *GULO* orthologs in chordate species share a highly conserved gene structure and similar protein sequences, except for the urochordate sea squirt. It is interesting that the purple sea urchin, an achordate echinoderm, possesses *GULO* exons (D, E, F, G, H, I, and J) similar to those in the vertebrate-amphioxus clade, although in a rearranged pattern (the last three, H, I, and J, are translocated upstream of exon D, and a duplicon of D and E is inserted upstream of the newly generated translocation locus of exons H, I, and J). Another difference is that the intact exon F in the purple sea urchin is split into two novel *GULO* exons,

F1 and F2, in lancelet-vertebrate clade (e.g., amphioxus; Fig. 2). The similarities in gene structure and conserved protein sequences between purple sea urchin *GULO* and the amphioxus orthologs suggest that they may have originated from a common ancestral *GULO* gene in an intermediate achordate species. In addition, the carnivorous mammals dog, cat, and giant panda cluster closely in a branch in the GULO phylogenetic tree, and each of them has an intronless *GULO* retropseudogene in a location different from that of the authentic gene. For example, compared with the sole *GULO* gene (pseudogene) on chromosomes 8 in humans, 14 in pigs, and 15 in rats, the dog *GULO* gene and its retropseudogene have been mapped on chromosomes 25 and 33, respectively (data not shown).

High Frequency of Translocations and Inversions in Vertebrate *GULO*-Flanking Genome Regions

Synteny analysis is a valid way to confirm orthologs. Vertebrate genome regions flanking *GULO* or its pseudogenes (*ΨGULO*) are conserved in an anchor gene layout

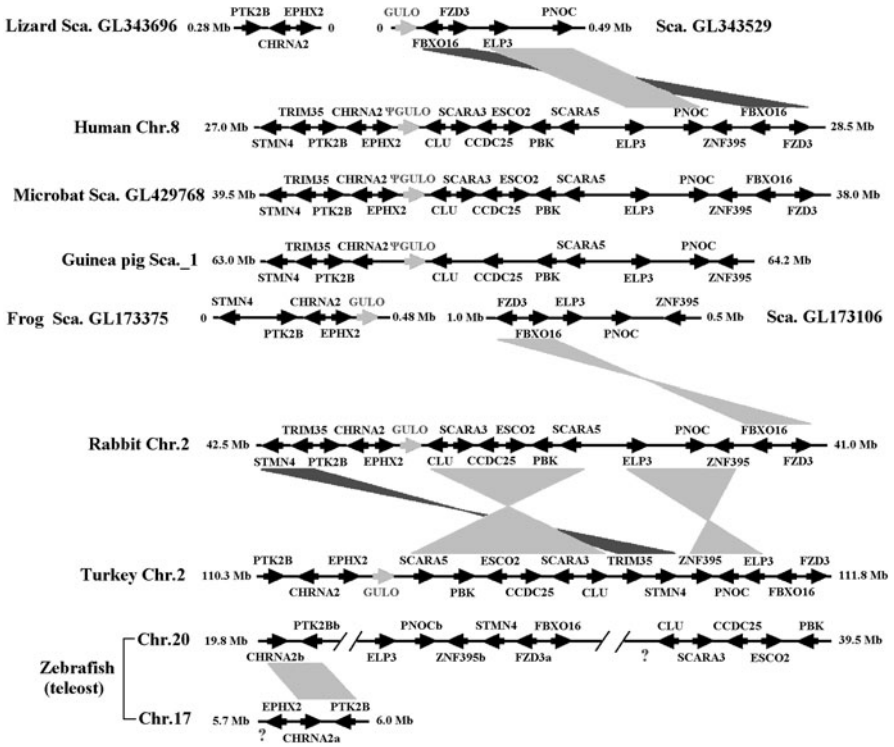


Fig. 3 Genomic region flanking *GULO* genes (pseudogenes) in representative vertebrate species. The *shaded shapes* connecting chromosomes indicate the positions of micro-rearrangements, such as translocated segments or local inversions, between species. *GULO* genes and their pseudogenes (*ΨGULO*) are *shaded light gray*, anchor genes are *black*. *Arrows* indicate the orientation of gene transcription. *?* Probable locus of a former *GULO* gene now lost in zebrafish. Gene size and intergenic spacing are not drawn to scale

(Fig. 3), partly confirming the *GULO* sequences from the Blast analysis. Genome regions neighboring mammalian *GULO* (represented here by human, microbat, guinea pig, and rabbit) are highly conserved in both gene layout and transcription orientations; however, multiple genome micro-rearrangement events, such as segment translocation and inversion, have occurred in *GULO*-flanking regions during vertebrate evolution. For example, compared with mammal genome regions flanking the *GULO* gene, the *TRIM35-STMN4* segment is inversely translocated upstream of *ELP3* on chromosome 2 in the turkey (birds), accompanied by local inversions of both the downstream segment *ELP3-PNOC-ZNF395* and the upstream segment *CLU-SCARA3-CCDC25-ESCO2-PBK-SCARA5*. The anchor *SCARA5* was immediately adjacent to the turkey *GULO* gene, whereas *CLU* is the neighbor of *GULO* in mammals. These frequent genome micro-rearrangements flanking the *GULO* locus are observed in other tetrapod vertebrates, such as lizards (reptiles) and clawed frogs (amphibians) (Fig. 3). Such frequent segment translocations and inversions suggest the genetic instability of the *GULO*-flanking genome regions during vertebrate evolution. It remains unknown as to which vertebrate lineage most resembles the arrangement of *GULO*-neighboring genes in their common ancestor, because of the lack of complete genome sequences in cartilaginous fishes and valid genome annotations in lancelets.

Teleost fishes cannot biosynthesize AsA themselves, and this study detected no *GULO*-like sequences in bony fish genomes. Further synteny analysis confirmed the complete loss of teleost *GULO*s, at least in the completely sequenced zebrafish. Most orthologs of mammalian *GULO*-neighboring genes are scattered in a 20 Mb *GULO*-less segment on chromosome 20 in the zebrafish, including *CLU*, one of the two adjacent anchors of the mammalian *GULO* gene. *EPHX2,* the other close neighbor of the mammalian *GULO* gene, is located on chromosome 17 in the zebrafish and is closely linked with another copy of *CHRNA2-PTK2B,* a cluster in the *GULO*-less segment on chromosome 20 (Fig. 3). Such segment duplication in bony fishes is ubiquitous, and it has been speculated that it originated from the teleost-specific 3R whole genome duplication (WGD) (Jaillon et al. 2004). The subsequent long period of rapid genome evolution in the remote teleost ancestor, after *GULO* ceased to function, accounts for the complete loss of this key gene in all present-day bony fishes (Ravi and Venkatesh 2008). The possibility cannot be ruled out, however, that homologous recombination events between similar sequences following the teleost-specific 3R WGD may also have led to the entire loss of teleost *GULO* pseudogenes, because recombination between repetitive elements is often involved in DNA segment loss (Deininger and Batzer 1999).

Positive Selection with Little Influence on Vertebrate GULO Function

The site model analysis demonstrated that all the vertebrate *GULO* genes examined here had endured strict purifying selection ($\omega = 0.069$), with no positively selected sites identified, suggesting constant functional constraints on vertebrate *GULO* genes. In the more precise branch-site model analysis, compared with the purifying selection on the background (all vertebrate *GULO* genes examined), some lineages (foreground) were detected to have undergone significant positive selection,

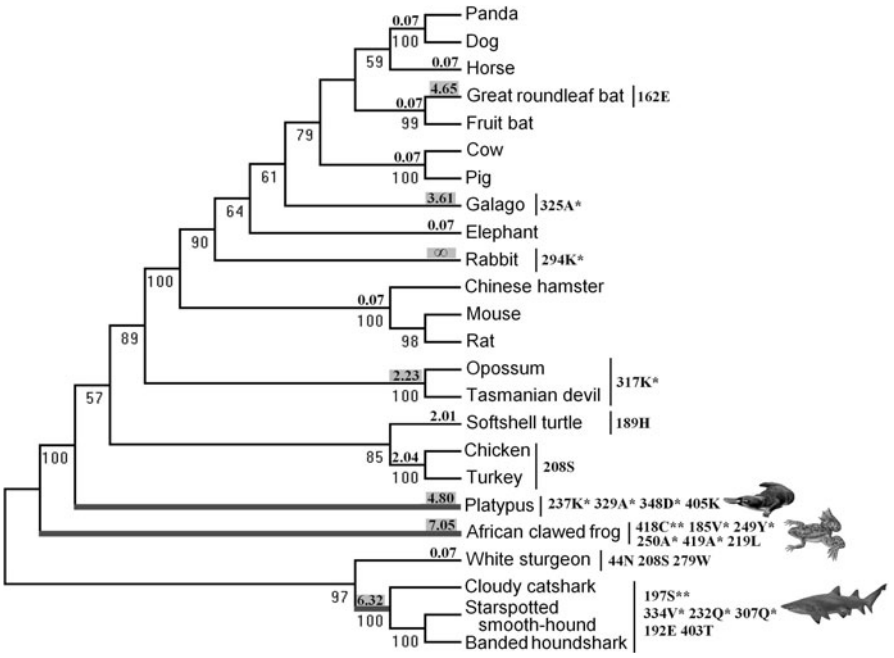


Fig. 4 Selective pressure on vertebrate *GULO* genes. The phylogenetic tree was constructed from manually edited *GULO* coding sequences using the neighbor-joining method in Mega 5.0. Supporting bootstrap values beyond 50 are shown below the nodes; ω (d_N/d_S) for each branch is above the node. *Gray shading* of the ω value indicates a branch with significant positively selected sites (posterior probability > 95%). *Thickened branch lines* mark lineages with $\omega > 4$. *Vertical lines* following animal names indicate positively selected sites with posterior probability > 85% within those genes; the sites are identified to the right of the vertical lines, with * for posterior probability beyond 95 and ** for 99%. The positively selected amino acids refer to those in the porcine GULO protein. The tree topology is similar to that from Clustal X 2.1, which was also used in branch-site model analysis (Supplementary Fig. 2). Common names of animals refer to species listed with their loci in Supplementary Table 1

especially in platypuses, clawed frogs, and sharks ($\omega > 4.0$). Some positively selected sites with high posterior probability (> 95 or 99%) in these clades were also identified using Bayes empirical Bayes analysis (Fig. 4; Supplementary Table 2). These positive sites were unique to each positive clade, consistent with the purifying selection results from the site model analysis, in which no common positive site across all vertebrate *GULO* genes was identified.

These positively selected sites were mapped onto the 3D structure of the alditol oxidase protein to ascertain whether they have functional significance or simply contribute to phylogenetic classification (Fig. 5). In alditol oxidase, which contains ligands and cofactor FAD, the conserved amino acids His54 and Ser106 are responsible for binding FAD covalently and noncovalently, respectively, and the conserved β-sheets 286 ELQSEY 291, 320 EIR 322, 343 HFT 345, and 373 WGK 375 interact with ligands and oxygen. Unfortunately, none of the positive GULO sites from the positive lineages was positioned in the cofactor-binding or ligand-binding regions, implying that they are involved in GULO functions other than

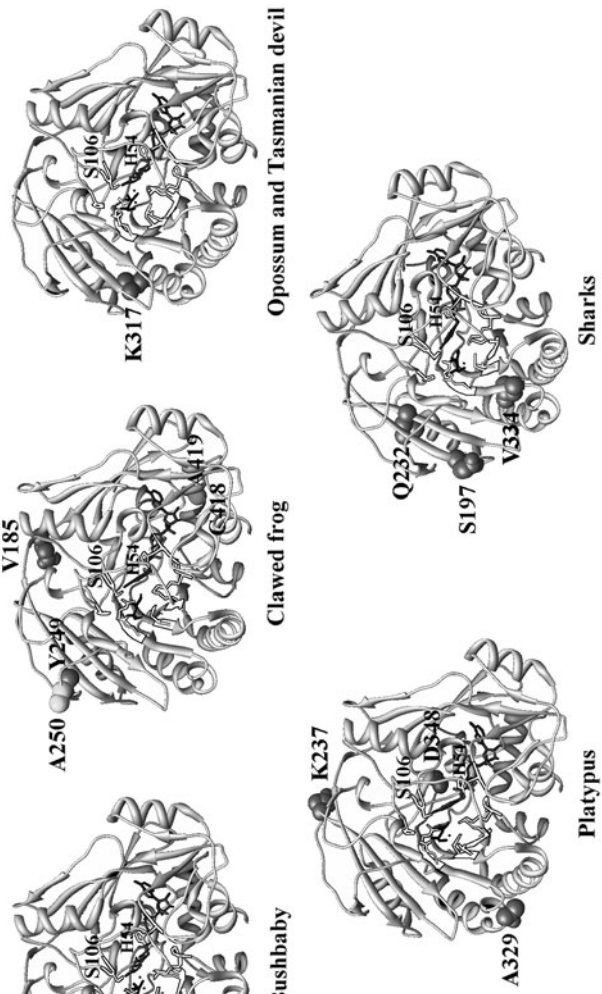


Fig. 5 Positively selected sites on the 3D structure of GULO proteins. The 3D protein structure is from alditol oxidase which shares high identity of amino acid sequences with GULO proteins, as well as the same motifs and a similar catalytic mechanism (Supplementary Fig. 1). The ligands and cofactor FAD are in *black*. The hollow amino acid residues indicate those involved in ligand binding or cofactor FAD binding, including conserved amino acids His54 and Ser106 and the conserved β-sheets. The "H54" in *black* is the conserved histidine covalently binding FAD. Amino acids in spheres indicate positively selected sites from branch-site model analysis in PAML

ligand-binding or cofactor-binding, or that they act as phylogenetic markers of different vertebrate *GULO* genes.

Vertebrate *GULO* genes are conserved at both the gene level and the genome level. They share high sequence identity and similar exon–intron structures. The 11 conserved exons have endured strict negative selection during evolution. The constitution of their neighboring genome regions is also conserved in spite of frequent local translocation and inversion events.

Among the vertebrate species that cannot produce AsA, teleost fishes and higher primates both lost their *GULO* genes in remote ancestors; most bats and guinea pigs also lost their *GULO* genes during evolution. The evolutionary cause for this loss remains controversial, either by dietary AsA supplement or by another, unknown, neutral selection. For the same reason, it is hard to predict whether some species are on the way to losing their *GULO* genes and the capacity to produce AsA. Addressing these issues needs further study and more evidence.

Acknowledgments This project was supported by the Natural Science Foundation of Fujian Province, China (No. 2008F3111) and the Initial Research Fund (No. 2006L20775) from Zhangzhou Normal University to Dr. Hongwen Yang.

References

Akbarzadeh A, Farahmand H, Mahjoubi F, Nematollahi MA, Leskinen P, Rytkonen K, Nikinmaa M (2011) The transcription of L-gulono-gamma-lactone oxidase, a key enzyme for biosynthesis of ascorbate, during development of Persian sturgeon *Acipenser persicus*. Comp Biochem Physiol B 158:282–288

Birney EC, Jenness R, Ayaz KM (1976) Inability of bats to synthesise L-ascorbic acid. Nature 260:626–628

Biyani N, Madhubala R (2011) *Leishmania donovani* encodes a functional enzyme involved in vitamin C biosynthesis: arabino-1,4-lactone oxidase. Mol Biochem Parasitol 180:76–85

Chaudhuri CR, Chatterjee IB (1969) L-ascorbic acid synthesis in birds: phylogenetic trend. Science 164:435–436

Cui J, Yuan X, Wang L, Jones G, Zhang S (2011) Recent loss of vitamin C biosynthesis ability in bats. PLoS ONE 6:e27114

Dabrowski K (1990) Gulonolactone oxidase is missing in teleost fish: the direct spectrophotometric assay. Biol Chem Hoppe Seyler 371:207–214

Deininger PL, Batzer MA (1999) Alu repeats and human disease. Mol Genet Metab 67:183–193

Drouin G, Godin JR, Pagé B (2011) The genetics of vitamin C loss in vertebrates. Curr Genomics 12:371–378

Forneris F, Heuts DP, Delvecchio M, Rovida S, Fraaije MW, Mattevi A (2008) Structural analysis of the catalytic mechanism and stereoselectivity in *Streptomyces coelicolor* alditol oxidase. Biochemistry 47:978–985

Jaillon O, Aury JM, Brunet F, Petit JL, Stange-Thomann N, Mauceli E, Bouneau L, Fischer C, Ozouf-Costaz C, Bernot A, Nicaud S, Jaffe D, Fisher S, Lutfalla G, Dossat C, Segurens B, Dasilva C, Salanoubat M, Levy M, Boudet N, Castellano S, Anthouard V, Jubin C, Castelli V, Katinka M, Vacherie B, Biémont C, Skalli Z, Cattolico L, Poulain J, De Berardinis V, Cruaud C, Duprat S, Brottier P, Coutanceau JP, Gouzy J, Parra G, Lardier G, Chapple C, McKernan KJ, McEwan P, Bosak S, Kellis M, Volff JN, Guigó R, Zody MC, Mesirov J, Lindblad-Toh K, Birren B, Nusbaum C, Kahn D, Robinson-Rechavi M, Laudet V, Schachter V, Quétier F, Saurin W, Scarpelli C, Wincker P, Lander ES, Weissenbach J, Roest Crollius H (2004) Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. Nature 431:946–957

Jenness R, Birney EC, Ayaz KL (1978) Ascorbic acid and L-gulonolactone oxidase in lagomorphs. Comp Biochem Physiol B 61:395–399

Jenness R, Birney EC, Ayaz KL, Buzzell DM (1984) Ontogenetic development of L-gulonolactone oxidase activity in several vertebrates. Comp Biochem Physiol B 78:167–173
Lachapelle MY, Drouin G (2011) Inactivation dates of the human and guinea pig vitamin C genes. Genetica 139:199–207
Maeda N, Hagihara H, Nakata Y, Hiller S, Wilder J, Reddick R (2000) Aortic wall damage in mice unable to synthesize ascorbic acid. Proc Natl Acad Sci USA 97:841–846
Martinez del Rio C (1997) Can passerines synthesize vitamin C? Auk 114:513–516
Nakata Y, Maeda N (2002) Vulnerable atherosclerotic plaque morphology in apolipoprotein E-deficient mice unable to make ascorbic acid. Circulation 105:1485–1490
Nandi A, Mukhopadhyay CK, Ghosh MK, Chattopadhyay DJ, Chatterjee IB (1997) Evolutionary significance of vitamin C biosynthesis in terrestrial vertebrates. Free Radic Biol Med 22:1047–1054
Nishikimi M, Kawai T, Yagi K (1992) Guinea pigs possess a highly mutated gene for L-gulono-gamma-lactone oxidase, the key enzyme for L-ascorbic acid biosynthesis missing in this species. J Biol Chem 267:21967–21972
Nishikimi M, Fukuyama R, Minoshima S, Shimizu N, Yagi K (1994) Cloning and chromosomal mapping of the human nonfunctional gene for L-gulono-gamma-lactone oxidase, the enzyme for L-ascorbic acid biosynthesis missing in man. J Biol Chem 269:13685–13688
Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF Chimera–a visualization system for exploratory research and analysis. J Comput Chem 25:1605–1612
Pollock JI, Mullin RJ (1987) Vitamin C biosynthesis in prosimians: evidence for the anthropoid affinity of *Tarsius*. Am J Phys Anthropol 73:65–70
Ravi V, Venkatesh B (2008) Rapidly evolving fish genomes and teleost diversity. Curr Opin Genet Dev 18:544–550
Sauer M, Branduardi P, Valli M, Porro D (2004) Production of L-ascorbic acid by metabolically engineered *Saccharomyces cerevisiae* and *Zygosaccharomyces bailii*. Appl Environ Microbiol 70:6086–6091
Spagnuolo MS, Cigliano L, Balestrieri M, Porta A, Abrescia P (2001) Synthesis of ascorbate and urate in the ovary of water buffalo. Free Radic Res 35:233–243
Togari A, Arai M, Nakagawa S, Banno A, Aoki M, Matsumoto S (1995) Alteration of bone status with ascorbic acid deficiency in ODS (osteogenic disorder Shionogi) rats. Jpn J Pharmacol 68:255–261
Yamashita M, Omura H, Okamoto E, Furuya Y, Yabuuchi M, Fukahi K, Murooka Y (2000) Isolation, characterization, and molecular cloning of a thermostable xylitol oxidase from *Streptomyces* sp. IKD472. J Biosci Bioeng 89:350–360
Yang Z (2007) PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol 24:1586–1591