Evolution of *trappin* genes in mammals

Akira Kato1*, Alejandro P Rooney2, Yutaka Furutani1,3, Shigehisa Hirose1

Abstract

**Background:** Trappin is a multifunctional host-defense peptide that has antiproteolytic, antiinflammatory, and antimicrobial activities. The numbers and compositions of *trappin* paralogs vary among mammalian species: human and sheep have a single *trappin-2* gene; mouse and rat have no *trappin* gene; pig and cow have multiple *trappin* genes; and guinea pig has a *trappin* gene and two other derivative genes. Independent duplications of *trappin* genes in pig and cow were observed recently after the species were separated. To determine whether these *trappin* gene duplications are restricted only to certain mammalian lineages, we analyzed recently-developed genome databases for the presence of duplicate *trappin* genes.

**Results:** The database analyses revealed that: 1) duplicated *trappin* multigenes were found recently in the nine-banded armadillo; 2) duplicated two *trappin* genes had been found in the Afrotherian species (elephant, tenrec, and hyrax) since ancient days; 3) a single *trappin-2* gene was found in various eutherian species; and 4) no typical *trappin* gene has been found in chicken, zebra finch, and opossum. Bayesian analysis estimated the date of the duplication of *trappin* genes in the Afrotheria, guinea pig, armadillo, cow, and pig to be 244, 35, 11, 13, and 3 million-years ago, respectively. The coding regions of *trappin* multigenes of almadillo, bovine, and pig evolved much faster than the noncoding exons, introns, and the flanking regions, showing that these genes have undergone accelerated evolution, and positive Darwinian selection was observed in pig-specific *trappin* paralogs.

**Conclusion:** These results suggest that trappin is an eutherian-specific molecule and eutherian genomes have the potential to form *trappin* multigenes.

Background

Trappins are a family of small secretory proteins that possess an N-terminal transglutaminase-substrate (TGS) domain and a C-terminal whey acidic protein (WAP) domain [1]. The TGS domain consists of repeats of six semi-conserved amino acids, KGQDPV, that act as anchoring regions. In this case, the lysine or glutamine residues of these regions are cross-linked with extracellular-matrix proteins by the action of transglutaminases, which helps trappin molecules to become concentrated at the site of action [2-4]. In contrast, the WAP domain is a four-disulfide core region and is defined by eight conserved cysteine residues. The WAP domain of trappin shows anti-proteolytic [4-6] and antimicrobial [7-9] activities that allow it to act as an innate immune defense molecule. In fact, trappin-2 displays antibacterial activities against Gram-positive and Gram-negative bacteria [7-9]; it also has antifungal activity [9], and the antimicrobial activity is independent of its antiprotease function [9]. The most well characterized trappin is human trappin-2, which is also known as elafin, skin-derived antileukoproteinase (SKALP), elastase-specific inhibitor (ESI), or protease inhibitor 3 (PI3) [1,10]. It has strong inhibitory activity against leukocyte and pancreatic elastases and protease 3 [4-6], and shows anti-inflammatory activity [11] as well. The antiproteolytic and antimicrobial activities of trappin-2 are quite similar to those of secretory leukocyte protease inhibitor (SLPI) [12,13], which consists of two WAP domains with the second WAP domain being highly homologous to the WAP domain of trappin-2. Trappin-2 is expressed in the trachea, lung, gut, epidermis, esophagus, vagina, and oral epithelia [2,4]. In these tissues, the expression is induced by proinflammatory cytokines, such as interleukin-1 (IL-1) and tumor necrosis factor (TNF)-α [14,15].

The number of *trappin* genes varies among mammalian species. For example, humans and sheep have a single *trappin-2* gene [16,17], while pigs have at least six: *trappin-1*, *trappin-2*, *trappin-3*, *trappin-7*, *trappin-8*, *trappin-9*...
and trappin-9 [18,19]. At the other extreme are the mouse and rat, which lack trappin genes entirely [20], though the guinea pig has genes for trappin-12 and its derivatives caltrin II and seminal vesicle secretory protein (SVP), which lack TGS- and WAP-coding regions, respectively [21,22]. Despite the variance in copy number between the different mammalian lineages, all trappin genes are encoded by three exons. Exon 1 encodes a signal peptide, exon 2 codes for a TGS- and WAP-domains, and exon 3 encodes a 3′ untranslated region [18,19,23]. While the exonic organization is highly conserved among various mammalian lineages, there is variation in the number of six-amino-acid repeats in the TGS domain [18,19]. Due to a point mutation of splitting site, guinea pig trappin-12 exceptionally lacks intron 2, which is present at the 3′ noncoding region of the trappin gene [22]. A short interspersed element (SINE) is found in intron 2 of the trappin genes of the pig, wart hog, and collared peccary [18,19].

While we have mentioned several species that possess multiple trappin genes, it is not known if (1) these are exceptional cases or (2) trappin genes normally exist as a multigene family. In an attempt to find the answers, we analyzed genome databases developed by the Mammalian Genome Project http://www.broad.mit.edu/mammals/ and identified six trappin genes from the nine-banded armadillo (Dasypus novemcinctus) genome. The nine-banded armadillo belongs to the taxonomic order Xenarthra. Because this lineage is believed to be one of the most ancient lineages of placental mammals [24], the analyses of armadillo trappin genes are quite interesting because the duplication and evolution of armadillo trappin genes are expected to have occurred independently from other species. In contrast, we identified a single trappin-2 gene from the genome databases of many species including the chimpanzee, rhesus macaque, bushbaby, dog, cat, horse, cow, European shrew, European hedgehog, megabat, and microbat. This fact suggests that trappin-2 is the ancestral form of trappin genes, and trappin-null species such as mouse and rat are exceptional. Finally, we identified anciently duplicated trappin-18 gene in Afrotheria such as the elephant (Loxodonta africana), tenrec (Echinops telfairi), and hyrax (Procavia capensis), and trappin-related genes in chicken and opossum, suggesting that the gene family originated as far back as more than 100 million years ago.

Results
Identification of trappin, SLPI, and trappin-related genes in eutherian mammals, opossum, platypus, chicken, and zebra finch
To estimate the origin of trappin genes, we analyzed the genome databases of eutherian mammals, opossum, platypus, birds (chicken and zebra finch), Xenopus, fish (Danio rerio, Takifugu rubripes, Tetraodon nigroviridis, Gasterosteus aculeatus, and Oryzias latipes), sea squirts (Ciona intestinalis and Ciona savignyi), insects (Drosophila melanogaster, Anopheles gambiae, and Aedes aegypti), and Caenorhabditis elegans. Typical trappin genes were identified only in the mammalian species (Figure 1C-D). A single homologous gene was identified in chicken, zebra finch, and opossum, and multiple homologous genes were identified in platypus, (Figure 1A-B). The other species did not show any homologous genes to trappins except for other WAP-coding genes with a low homology (data not shown).

The presence of the SLPI gene was also analyzed using the genome databases, and a single orthologous gene was identified in all mammalian species except for the guinea pig and the rabbit (Figure 1A). In contrast, there are no clear direct orthologs for the SLPI gene in the genome databases of chicken, zebra finch, and opossum, and the trappin-homologous genes are also the most homologous to SLPI. In platypus, the above-mentioned trappin-homologous genes encode two-WAP-domain proteins and may be the paralogs of the mammalian SLPI gene.

The trappin-related genes in chicken, zebra finch, and opossum have a single WAP-coding region but lack a TGS-coding region. Only the WAP-coding region is shown in alignment with those of mammalian trappin and the SLPI genes (Figure 1B). The catalytically important Met residue (an asterisk in Figure 1, A and 1C) is conserved in the opossum and platypus genes but not in the chicken and zebra finch genes.

Platypus SLPI genes show a stronger identity with trappin-2 (67%) than SLPI (55%) in the deduced amino acid sequences. The phylogenetic analysis for trappin, mammalian SLPI, platypus SLPI, and trappin-related genes of chicken and opossum is shown in Figure 2A. Trappin-related genes of chicken and opossum and platypus SLPI genes are not clearly categorized as trappin or SLPI.

Identification of trappin-2 genes from various eutherian mammals
We analyzed genome databases for various eutherian mammals including human, chimpanzee, rhesus macaque, bushbaby, mouse, rat, rabbit, dog, cat, cow, European shrew, European hedgehog, megabat, nine-banded armadillo, sloth, elephant, hyrax, and tenrec. Only mouse, rat, and rabbit lack trappin genes in
Figure 1 WAP domain of the trappin gene family. WAP domains of SLPI (A), trappin-related genes (B), trappin-2 (C), and the other species-specific trappin paralogs (D) are shown. Conserved and semiconserved residues among trappin-2 genes are indicated by light gray. Eight conserved Cys residues constituting the WAP motif signature sequence are shaded in black. The asterisk indicates the catalytically important Met residue. The variable region, which is thought to determine the specificity of WAP motifs, is boxed.
their genome databases, but the other species have at least one trappin gene in their genome databases (Figure 3). All the newly identified trappin genes consist of three exons like previously analyzed trappin genes: exon 1 encodes a signal peptide, exon 2 encodes the TGS and WAP domains, and exon 3 encodes the 3′ untranslated region (data not shown). We aligned the amino acid sequences of the WAP domains of those genes, and categorized them into two groups: the first group contained the catalytically important Met residue (asterisk in Figure 1C) and was named trappin-2 (Met rule); and the second group lacked the Met residue and was named according to the order of discovery. Most animals have a single trappin-2 gene (Figures 1C and 3).

This finding suggests that the trappin-2 gene is the ancestral form of the trappin genes. According to this definition, the previously reported porcine trappin-8 should also be renamed trappin-2b as it also has the Met residue. Tenrec trappin-18 also has the Met residue at the catalytic site. However, the phylogenetic analyses using nucleotide sequences of the noncoding regions (introns, exon 3 and 5′ and 3′ flanking regions) clarified that this gene is closely related with trappin-18 genes of elephant and hyrax. Therefore we call this gene tenrec trappin-18 as an exception to the Met rule.

The average nonsynonymous and synonymous distances were calculated on 150 bp WAP-coding regions among the trappin-2 genes of the various species,
trappin paralogs except for trappin-2, and all trappin genes (Table 1). In the WAP-coding region, the rate of nonsynonymous substitutions is lower than that of synonymous substitutions among trappin-2. The rates of nonsynonymous and synonymous substitutions are similar in the WAP-coding region among trappin paralogs except for trappin-2. These results suggest that the purifying selection is operating on the trappin-2 gene in various eutherian mammals.

Identification of novel trappin multigene families in armadillo and Afrotheria (elephant, tenrec, and hyrax)

Database analyses demonstrated the presence of six trappin genes in nine-banded armadillo (Dasypus novemcinctus), which were named trappin-2 and trappins-13-17 (Figures 1 and 3). Afrotherian species such as the elephant (Loxodonta africana), tenrec (Echinops telfairi), and hyrax (Procavia capensis) had two trappin genes, which were named trappin-2 and trappin-18 (Figures 1 and 3). We also found two novel trappin paralogs from the bovine genome database, and named them trappin-19 and trappin-20 (Figures 1D and 3).

Phylogenetic analyses of the noncoding regions of trappin genes from several mammalian species are shown in Figure 2B. All armadillo trappins-13-17 genes form a single branch with armadillo trappin-2 gene. Bovine trappin-19 and trappin-20 also share the same branch with bovine trappin-2. These results suggest that those genes are recently duplicated species specific.
paralogs. On the other hand, Afrotherian trappin-18 is divided near the root, suggesting that trappin-18 duplicated much earlier.

**Estimations of the dates for the duplication of trappin multigenes**

A linearized tree was constructed by using the nucleotide sequences of trappin multigenes and the dates of the duplications were calculated with MEGA software. When the divergence time between Primate and Artiodactyla (96.2 Mya) was used as a reference point, the date of the duplications of trappin genes of pig, cow, armadillo, guinea pig, and Afrotheria (elephant, hyrax, and tenrec) were calculated as 7.0, 8.8, 15.9, 79.0, and 161 Mya, respectively (asterisks in Figure 3).

We next calculated the date of duplications individually for each species using the taxon pair that was most closely related to the node of interest as a reference point. We found that some of the trappin gene subfamilies were relatively young. For instance, when the divergence time between sheep and cow (183.3 Mya) was used as a calibration point, the date of the duplication events giving rise to the pig and bovine trappin gene families were calculated as 7.8 and 9.7 Mya, respectively. Similarly, when the divergence time between human and armadillo (96.2 Mya) was used as a reference point, the date of the duplication event giving rise to the armadillo trappin gene family was calculated as 15.5 Mya. On the other hand, certain trappin gene subfamilies appear to be more ancient. For example, when the divergence time between primate and rodent (61.7 Mya) was used as a reference point, the date the guinea pig trappin gene subfamily was estimated to have originated 55.2 Mya, and when the divergence time between elephant and tenrec (48.6 Mya) was used as a reference point, the Afrotherian trappin trappin gene subfamily was calculated to have originated 91.9 Mya (double asterisks in Figure 3).

We also estimated divergence times using the Bayesian method implemented in the BEAST software package [25]. The date of the duplications of trappin genes of pig, cow, armadillo, guinea pig, and Afrotheria were calculated as 3.3, 12.6, 11.4, 34.7, and 244 Mya (arrows in Figure 3), respectively, when the nucleotide sequences of trappin genes were used for the calculation. When amino acid sequences were used for the calculation, the date of the duplications of trappin genes of pig, cow, armadillo, and Afrotheria were calculated as 49.7, 18.3, 50.9, and 154 Mya (arrowheads in Figure 3), respectively. Thus, the dates calculated for nucleotide and amino acid sequence data are very different. The former are similar to the dates generated using the linearized tree method in pig, cow, and armadillo, and the latter are similar to the dates generated using the linearized tree method in Afrotheria. In pig, cow, and armadillo, the protein-coding regions of trappin multigenes have evolved rapidly (Table 2), and the dates calculated for amino acid sequence data showed larger values. In Afrotheria, since amino acid sequences evolve more slowly than nucleotide sequences and are, therefore, less prone to homoplasy over the long evolutionary time frames being considered, it is reasonable to suspect that the dates generated using nucleotide sequence data are not as reliable as those generated from acid sequence data for this particular study.

**Synteny analyses around trappin genes**

It has been demonstrated that the trappin-2 is mapped on the WAP four-disulphide core (WFDC) domain locus which contains a number of WFDC genes [17]. Moreover, the conserved synteny of WFDC loci has been studied for primates, rodents, and the dog [20,26]. To extend these studies, we analyzed the genes neighboring trappins in species that were included in our study. Most scaffolds containing trappin genes were too short to analyze. However, we could analyze genes neighboring to horse trappin-2, megabat trappin-2, bovine trappins, hyrax trappin-18 and elephant trappins (Figure 4A). As reported previously for the dog WFDC locus, the horse, megabat, and bovine trappin genes flanked WFDC5 in opposed directions, whereas WFDC12 and WFDC15 were not found between WFDC5 and trappin(s). The bovine trappin genes were tandemly arrayed, suggesting that these genes arose by tandem gene duplication. In contrast, hyrax and elephant trappin-18 flanked WFDC5 in the same direction and was mapped to the same locus as WFDC12. Harr Plot analyses demonstrated that human WFDC12 is highly homologous to the 5'-flanking region, exon 1, and intron 1 of hyrax trappin-18 (Figure 4C). Similar homology was observed between human WFDC12 and elephant trappin-18 (Figure 4D).

Recently, Hurle et al. found that primate trappin-2 contains a pseudogene for WFDC12 in intron 1, and suggested that trappin-2 and WFDC12 have a common ancestral gene [26]. All trappin genes contained a pseudogene for WFDC12 in intron 1 (data not shown) except for Afrotherian trappin-18, which codes for a WFDC12-like peptide in intron 1 (Figure 4B).

### Table 1 Purifying selection of trappin-2 genes.

|            | Trappin-2 | Other trappin paralogs | All trappin genes |
|------------|-----------|------------------------|-------------------|
| dn         | 0.190 ± 0.041 | 0.455 ± 0.092 | 0.369 ± 0.071 |
| ds         | 0.333 ± 0.057 | 0.388 ± 0.066 | 0.377 ± 0.054 |
| Dn/ds      | 0.57 | 1.17 | 0.98 |

Average non-synonymous and synonymous distances were calculated on 150 bp WAP-coding regions among trappin-2 genes of various species, the other trappin paralogs, and all trappin genes.
Accelerated evolution of TGS and WAP coding region

trappin multigenes in armadillo, cow, and pig

and positive selection of the WAP-coding region of pig

trappin paralogs

The average distances of the 5′-flanking region, exon 1, intron 1, exon 2, intron 2, exon 3, and 3′-flanking region among trappin multigenes for each species were calculated (Table 2, line A1-A7). In armadillo trappins, the average Jukes-Cantor (JC) distance between the exon 2 regions was 0.235 (Table 2, line A4), which is 4.4 times higher than that between the non-coding regions (0.054; Table 2, line B1). When we calculated the average Tamura-Nei (TN) distances with gamma correction, the value between exon 2 (0.442; Table 2, line A4) was also 5.6 times higher than that between the non-coding regions (0.079; Table 2, line B1). Fisher’s exact test using the numbers of varied sites and common sites between the exon 2 regions (39 varied sites in 201 common sites) and those between the non-coding regions (87 varied sites in 1691 common sites) demonstrated that the difference is significant \( P < 0.01 \). A similar difference was not observed in the other regions.

In cow, the average distances between the exon 2 regions were 0.167 (JC method) and 0.299 (NJ method) \( < 0.01 \). In contrast, there was no significant difference between the exon 2 regions (87 varied sites in 1691 common sites) and those between the non-coding regions calculated by the JC and TN methods, respectively (Table 2, line B1; \( P < 0.01 \), Fisher’s exact test). In pig, the average distances between the exon 2 regions (Table 2, line A4) were 12 and 16 times higher than those between the non-coding regions when calculated by the JC and TN methods, respectively (Table 2, line B1; \( P < 0.01 \)). In contrast, there was no significant difference between the average distances between the exon 2 regions (Table 2, line A4) and those between the non-coding regions (Table 2, line B1) of elephant, hyrax, and guinea pig trappin genes. In pig, the average distances...
Figure 4 Evolutional relationship between trappin-18 and WFDC12. (A) Schematic illustration of conserved synteny around trappin(s). Illustrations of human, mouse, rat, and dog loci were generated based on the works by Clauss et al. [20] and Hurle et al. [26]. Black and gray arrows indicate genes and pseudo genes, respectively. KCNS, potassium voltage-gated channel, member 1 (B) The WAP domains of WFDC12 and the WFDC12-like peptide encoded by intron 1 region of trappin-18. Conserved and semiconserved residues are indicated by light gray. Eight conserved Cys residues constituting the WAP motif signature sequence are shaded in black. (C) Harr plot analyses of the hyrax trappin-18 in comparison with the hyrax trappin-2 and human WFDC12. (D) Harr plot analyses of the elephant trappin-18 in comparison with the elephant trappin-2 and human WFDC12. Exons are indicated by boxes. WAP-coding regions are represented by black boxes. Black circles indicate regions encoding WFDC12 like peptide.
between the exon 1 regions and between the intron 2 regions of different genes (Table 2, lines A2 and A5) were also higher than those of the non-coding regions (Table 2, line B1) (P < 0.01).

Next, we calculated distance values for synonymous substitutions per site (ds) and non-synonymous substitutions per site (dn) for the signal peptide (pre peptide), TGS, and WAP coding regions (Table 2, line C1-E3), and compared against the average distance of the non-coding regions. In armadillo, dn of the TGS coding domain (P < 0.05) and both ds and dn of the WAP coding domain (P < 0.01) were significantly higher than the average distance of the non-coding regions. In cow, only dn of the TGS and WAP coding regions were significantly higher than the average distance of the non-coding regions (P < 0.01). In pig, both dn and ds of the signal peptide, TGS, and WAP coding regions were higher than the average distance of the non-coding regions.

To examine the presence of positive Darwinian selection, we compared dn and ds of the rapidly-evolved coding regions using Fisher’s test. When the average values in each species were used for the analyses, we could not detect any statistically significant difference between dn and ds of all the trappin multigens. However, the pairwise comparison matrix of dn and ds on the paralogs of each species demonstrated a dn/ds rate of 3.3 for porcine trappin-2 vs. trappin-3, 4.9 for porcine trappin-2 vs. trappin-9, and 5.2 for porcine trappin-8 vs. trappin-9 (P < 0.05) (Table 3). Because trappin-2 is the most conservative gene within the trappin family and porcine trappin-8 is the closest homolog of porcine trappin-2, these data indicate the positive Darwinian selection of porcine trappin-3 and trappin-9. Although the dn/ds rate between bovine trappin-5 and other bovine paralogs are as high as 2.8-5.8, the differences were not statistically significant (P = 0.07) (Table 3).

### Evaluation of the quality of genomic sequence with low coverage

The nucleotide substitutions between seven known cDNAs and corresponding exons in the genome databases were calculated and shown in Table 4. Some of these substitutions may have occurred as a result of sequencing errors or site-specific polymorphism within each species. Yet, we can still infer that the average rates of sequencing errors are lower than the substitution rates. In low coverage genomic sequences of armadillo, rabbit, cat, and elephant, the average substitution rates of the seven genes were 0.10-0.49%. In high coverage genomic sequences of cow and human, the average substitution rates of the seven genes were 0.23 and 0.06%, respectively. These estimates are not substantially different, suggesting that artifacts due to errors or polymorphism are negligible. In the case for armadillo trappin genes, for example, the 129-bp WAP-coding regions and the ~2.1-kb entire genes may contain less than 0.4 and 6-base sequence errors, respectively. Among armadillo trappin multigens, the WAP-coding regions and the ~2.1-kb entire genes have 8-50 and 96-165-base substitutions, respectively. Thus, the sequencing errors of genomic sequences with low coverage appear to be negligible, although the caveat remains that we can not negate a small number of possible errors.

### Discussion

#### Origin of trappin gene

Computer analyses of genome databases revealed that typical trappin is a eutherian mammalian specific gene. The typical trappin genes were found only in eutherian mammals and not other species including Xenopus, fish, sea squirt, insects, and C. elegans. The trappin-related genes were found in chicken and opossum. The computer analyses also showed that most eutherian mammalian species have a single SLPI gene, and platypus has multiple SLPI genes. The trappin-related genes of those animals and platypus SLPI genes show strong similarity with trappin in the WAP domain only, but all the other regions have no significant homology. Therefore, these genes may relate with the ancestral WAP domain of trappin. Interestingly, platypus SLPI showed higher

### Table 3 Positive selection of species-specific trappin paralogs.

|        | pTr-2 | pTr-8 | pTr-8 |
|--------|-------|-------|-------|
| pTr-1  | 1.5   | 1.3   | 1.3   |
| pTr-3  | 3.3*  | 3.8   | 3.0   |
| pTr-9  | 4.9*  | 5.2*  | 2.4   |
| pTr-7  | 1.5   | 1.5   | 2.3   |
|        | 1.5   | 1.5   | 2.1   |

Pairwise comparison matrices of dn/ds rates are shown. p, pig; b, bovine; a, armadillo; Tr, trappin; * P < 0.05.
homology to the WAP domain of trappin-2 than that of mammalian SLPI. This strongly suggests that the WAP domain of trappin and SLPI share a common ancestor.

Trappin is a protein that consists of TGS and WAP domains. SLPI is a two WAP-domain protein. The second WAP domain of SLPI and the WAP domain of trappin-2 are quite similar in their amino acid sequences and functions such as antiproteolytic and antimicrobial activities, suggesting an ancestral relationship. However, except for the WAP-coding regions, there is no significant homology in the nucleotide sequences between trappin and SLPI. Interestingly, trappin genes are known to have weak but significant similarity with other TGS genes in introns, TGS-coding, and noncoding regions (Figure 5) [27]. This mosaic pattern of homology in trappin genes indicates that trappin genes originated from the TGS gene and obtained a WAP domain possibly by exon-shuffling. Afrotherian trappin-18 codes for a WFDC12-like peptide in intron 1 region, and the other trappins contain a pseudogene for WFDC12 in intron 1. These results support the hypothesis of Hurle et al. [26] that trappin and WFDC12 are derived from a common ancestral gene which codes for both trappin and WFDC12.

Evolution of trappin genes in eutherian mammals

Nineteen species of eutherian mammals were analyzed by a search for the presence of trappin genes within their genome databases, and the results were combined with those of previous experimental analyses of human [17,23], pig [18,19], wart hog [19], collared peccary [19], cow [28], sheep [16], and guinea pig [21,22]. In total, we could compare the trappin genes from 24 eutherian mammals (Figure 3). Within the 24 species analyzed, we could isolate trappin genes from 21 species. A single trappin-2 gene was found in 11 species, and multiple trappin genes were found in 8 species. These results indicate that trappin-2 is the most common and is an ancestral form while the other trappins are specie-specific paralogs. We could not find trappin genes in three mammalian species: mouse, rat, and rabbit. Our experimental analyses (data not shown) and the integrity of the genome databases of mouse and rat suggest that mouse and rat lack trappin genes in their genome [20]. In mouse and rat, other WAP-motif containing proteins such as SLPI and SWAMs may compensate the function.

Table 4 Nucleotide substitutions between known cDNA and corresponding genomic sequences

|              | Armadillo (2 × coverage) | Rabbit (2 × coverage) | Cat (1.87 × coverage) | Elephant (2 × coverage) | Cow (7 × coverage) | Human (GRCh37) |
|--------------|--------------------------|-----------------------|-----------------------|-------------------------|-------------------|----------------|
| SDHA         | 1/899                    | 14/543                | 1/976                 | 5/1222                  | ND                | 3/1650         |
| MDH2         | 2/713                    | 2/306                 | 7/878                 | 1/626                   | 0/835             | 0/912          |
| ATP5B        | 0/876                    | 2/1237                | 4/1072                | 0/926                   | 0/950             | 0/1335         |
| GAPDH        | 6/413                    | ND                    | 2/683                 | 0/542                   | 1/674             | 0/691          |
| SDH8         | 0/476                    | 1/468                 | 0/586                 | ND                      | 1/551             | 0/591          |
| CS           | 0/646                    | 0/672                 | 2/564                 | 2/677                   | 2/653             | 0/673          |
| IDH1         | 4/477                    | 0/666                 | 0/740                 | 1/1077                  | 7/1064            | 1/1114         |
| total        | 13/4500                  | 19/3892               | 16/5499               | 9/5070                  | 11/4727           | 4/6966         |
| substitution rate (%) | 0.29                  | 0.49                   | 0.29                  | 0.10                   | 0.23              | 0.06           |
| identity (%) | 99.7                    | 99.5                   | 99.7                  | 99.9                   | 99.8              | 99.9           |

Numbers show the nucleotide substitutions per the length of the sequences used for the analysis. The scaffold, contig, or accession numbers of the sequences are shown in Supplementary Table S1 (additional file 1). ND, no data.
of trappin. In the case of rabbit, it is not certain whether rabbit really lacks trappin genes or rabbit has a trappin gene that has not yet been analyzed by the genome project.

By computer analyses of genome databases, we found that the nine-banded armadillo as well as pig and cow have recently-duplicated trappin multigene. The computer analyses of bovine genome databases also revealed two novel trappin paralogs and the sequences of the introns and flanking regions, which enabled the detailed evolutionary analyses of bovine trappin multigens. As previously reported porcine trappin multigens, the WAP-coding regions of the trappin multigens of armadillo and cow were shown to have evolved under accelerated evolution. Only dn was accelerated in the WAP coding regions of bovine trappins, and both dn and ds of the WAP coding regions were accelerated in armadillo and porcine trappins. The accelerated substitutions of non-synonymous sites of WAP-coding regions may be explained by positive Darwinian selection or relaxation of functional constraints, because we observed statistically significant positive selection of the WAP coding regions of porcine trappin-3 and trappin-9 but no significant difference between dn and ds of other trappins (Table 3). However, the question why synonymous substitutions are also accelerated can not be interpreted simply by the existence of positive Darwinian selection or relaxation of functional constraint. The mechanism whereby the synonymous substitutions are accelerated must be clarified by future studies.

The molecular clock and Bayesian analyses using the nucleotide sequences estimated the date of duplication as 11.4-15.9, 8.8-12.6, and 3.3-7.8 Mya for trappin multigens of armadillo, cow, and pig, respectively. These results are consistent with previous experimental analyses demonstrating that the collared peccary that was separated from porcine 33 Mya [29,30], and sheep, which was separated from bovine 19.6 Mya [31], do not have trappin multigens [16,19]. The findings of recently-duplicated accelerated-evolved trappin multigens in three individual species demonstrate that mammalian genomes have the potential to form trappin multigens in several million years. The selective pressure that formed the trappin multigens may relate with some pathogens, and the variety of amino-acid sequences in the WAP-domain may contribute to the acquisition of antimicrobial activities for a large spectrum of pathogens. Tissue distribution of trappin paralogs in pig and cow has been shown to vary among genes: porcine trappin-2 is expressed in the trachea and the large intestine, porcine trappin-1 in the small intestine, bovine trappin-2 in the epidermis and the tongue, bovine trappin-4 in the trachea and the tongue, and bovine trappin-5 in the trachea [28]. Therefore, the selective pressures might also affect the regulation of the tissue-specific expression of trappin genes.

Our previous analyses revealed that guinea pig has a trappin-12 gene [22] and two derivative genes, SVP [32] and caltrin II [19]. SVP and caltrin II genes have significant homology with trappin including introns, non-coding region of exons, and flanking regions, but lack WAP and TGS domains, respectively. The molecular clock analysis estimated the date of the duplication of the guinea pig genes as 34.7-79.0 Mya. This date of duplication is much earlier than those of pig, cow, and armadillo.

In Afrotherians we found two trappin genes, trappin-2 and trappin-18, whose date of duplication was estimated as 91.9-244 Mya. This date is surprising, because it is earlier than the date of the periods of divergence of the major orders of eutherian mammals (70-10 Mya) [24,31], and suggests that the duplication of trappin-18 occurred in the ancestors of the eutherian mammals before the divergence of the species. In this context, most species lack trappin-18, however, only Afrotheria has retained the gene. The reason is still unknown, but it is conceivable that trappin-18 increases resistance to Afrotheria-specific pathogen. Another possible alternative explanation is that trappin-18 underwent substitutions at a faster rate per year than other trappin genes and that lead to the duplication time being overestimated.

Conclusions
- Typical trappin genes are only found in the genome sequences of eutherians but not in those of other vertebrate species.
- Trappin-2 is the most widely distributed and is the strongest candidate of the ancestral forms of trappin. Recently-duplicated species-specific trappin paralogs are present in the genomes of armadillo, pig, and cow, and the non-synonymous sites of those genes have undergone accelerated evolution as a result of positive Darwinian selection or relaxation of functional constraint.
- Synonymous sites of recently-duplicated trappin paralogs of armadillo and pig have also undergone accelerated evolution by unknown mechanisms.
- The anciently-duplicated trappin-18 gene is only retained in afrotherian species and is a fossil molecule of the trappin gene family.

Methods
Isolation of trappin genes from various animal species
The genome database of various species (URL: http://www.ensembl.org/index.html) [33] were screened using the amino-acid sequence of human trappin-2. The exon-intron organization was estimated by comparing it with that of the human trappin-2 gene. The nucleotide
sequences of the trappin genes were deposited in the DDBJ/EMBL/GenBank DNA databases as third party annotations (TPAs) under accession numbers BR000322 to BR000327 and BR000708 to BR000720.

Evaluation of the quality of genomic sequence with low coverage

To evaluate the quality of genomic sequences with low coverage, we compared known cDNA sequences with those of corresponding exons in the genome databases. We used cDNA sequences for SDHA, MDH2, ATP5B, GAPDH, SDHB, CS, and IDH1 which was determined by Kullberg et al. [34]. The corresponding exons of armadillo (2 × coverage), rabbit (2 ×), cat (1.87 ×), elephant (2 x), cow (7 ×), and human (Genome Reference Consortium GRCh37 assembly) were isolated and the numbers of nucleotide substitutions between the sequences of cDNA and the genome databases were calculated for each species using MEGA software [35]. The sequences used for the analysis are shown in Supplemental Table S1 (see Additional file 1).

Phylogenetic analyses

Nucleotide sequences of the WAP-coding regions of trappin, SLPI, and trappin-related genes were used to analyze their phylogenetical relationship. The introns, exon 3 (noncoding exon), and 3′-noncoding regions were used to analyze recent evolution of trappin genes in eutherian mammals. The nucleotide sequences were aligned using ClustalW software [36], and the best fit/gap placement was confirmed manually. Phylogenetic analysis was performed by the neighbor-joining (NJ) method [37,38] and maximum parsimony (MP) method [38] with 2,000 bootstrap replicates using MEGA software [35] or the maximum likelihood (ML) method with 200 bootstrap replicates using PHYML [39] plugin for Geneious software http://www.geneious.com. The sequences used are as follows with the accession numbers in parentheses: human (Homo sapiens) trappin-2 (D13156) and SLPI (X04502); chimpanzee (Pan troglodytes) trappin-2 (XM_514671) and SLPI (DP000037); macaque (Macaca mulatta) trappin-2 (XM_00110935) and SLPI (DP000043); bunnahy (Otolarnum garnettii) trappin-2 (BR000708) and SLPI (DP000040); mouse (Mus musculus) SLPI (AF002719); rat (Rattus norvegicus) SLPI (AHAX01026351); guinea pig (Cavia porcellus) trappin-12 (AB161363), caltrin II (AB161364) and SVP (U59711); European shrew (Sorex araneus) trappin-2 (BR000713) and SLPI (AA101303048); European hedgehog (Erinaceus europaeus) trappin-2 (BR000714) and SLPI (AAN01307740); microbat (Myotis lucifugus) trappin-2 (BR000712) and SLPI (AAPE01410948); megabat (Pteropus vampyrus) trappin-2 (ABRP01168531) and SLPI (ABRP01290205); horse (Equus caballus) trappin-2 (XM_001503186) and SLPI (XP_001503242); dog (Canis familiaris) trappin-2 (BR000710) and SLPI (AAEX02024101); cat (Felis catus) trappin-2 (BR000711) and SLPI (AANG01238466); bovine (Bos taurus) trappin-2 (AJ223216), trappin-4 (AJ223217), trappin-5 (AJ233218), trappin-6 (AB011010), trappin-19 (BR000718), trappin-20 (BR000719) and SLPI (AAFC0303522); sheep (Ovis aries) trappin-2 (NM_001035224) and SLPI (AY346135); porcine (Sus scrofa) trappin-1 (D50320), trappin-2 (D50319), trappin-3 (D50321), trappin-7 (D50323), trappin-8 (D50322), trappin-9 (AB003285) and SLPI (NM_213870); elephant (Loxodonta africana) trappin-2 (BR000716), trappin-18 (BR000717) and SLPI (AAUG01360578); hyrax (Procavia capensis) trappin-2 (ABRQ01439157), trappin-18 (ABRQ01360466), and trappin-19 (ABRQ01352342); tenrec (Echinops telfairi) trappin-2 (BR000715), trappin-18 (AAIY01696839), and SLPI (AAIY01696839); wart hog (Phacochoerus aethiopicus) trappin-1 (AB003282) and trappin-2 (AB003281); collared peccary (Pecari tajacu) trappin-10 (AB003283); hippocotomas (Hippopotamus amphibius) trappin-11 (AB003284); nine-banded armadillo (Dasypus novemcinctus) trappin-2 (BR000322), trappin-13 (BR000323), trappin-14 (BR000324), trappin-15 (BR000325), trappin-16 (BR000326), trappin-17 (BR000327) and SLPI; sloth (Choloepus hoffmanni) trappin-21 (ABVD01210669) and SLPI (ABVD01323747); platypus (Ornithorhynchus anatinus) SLPla (AAPN01348542), SLPlb (AAPN01336636), SLPlc (AAPN01050486), SLPlId (AAPN01048517) and SLPlP (AAPN01030446); chicken (Gallus gallus) trappin-related protein (NC_006107); finch (Taeniopygia guttata) trappin-related protein (ABQF01028586); and opossum (Monodelphis domestica) trappin-related protein (BR000720).

Molecular clock analysis and Bayesian divergence time estimation

The introns, exon 3 (noncoding exon), and 5′- and 3′-noncoding regions of pig, cow, armadillo, guinea pig, and Afrotheria (elephant and hyrax) trappins were aligned, and a phylogenetic tree was constructed by the NJ method. A linearized tree was constructed and the dates of the duplication events of trappin genes were calculated by MEGA 3.1 using the divergence time between Primate and Artiodactyla (96.2 Mya) [40] as a calibration point for dating.

As an additional method to investigate divergence times, we used the Bayesian method implemented in the software package BEAST 1.4.8 [25]. To generate divergence times, the following nine fossil calibration points were taken from the work by Benton and Donoghue [40] and implemented as priors in the analysis of both DNA sequence and amino acid sequence data: (1)
human-chimp: 6.5 Mya; (2) human-Macaque: 23.5 ± 0.5 Mya; (3) dog-cat: 43 ± 0.2 Mya; (4) cow-sheep: 18.3 ± 0.1 Mya; (5) cow-dog 96.2 ± 0.9 Mya; (6) human-cow: 96.2 ± 0.9 Mya; (7) human-armadillo: 96.2 ± 0.9 Mya; (8) tenrec-elephant: 48.6 ± 0.2 Mya; (9) human-opossum: 124.6 ± 0.1 Mya. The chains were run until convergence was reached (i.e., until the effective sample size for each parameter exceeded 200), which was 93 million states for the DNA sequence data and 10 million states for the amino acid sequence data. The HKY + gamma model was used for the analysis of the DNA sequence data, and the WAG model was used for the analysis of the amino acid sequence data. For both sequence data types, the birth-death speciation process was used as a tree prior.

**Calculation of nucleotide substitution rates**

Nucleotide sequences were separated into the following regions: the 5'-flanking region, exon 1, intron 1, exon 2, the WAP-coding region of exon 2, intron 2, exon 3, and the 3'-flanking region. These regions were aligned separately using ClustalW software. Jukes-Cantor (JC) distances [41] and Tamura-Nei (TN) distances [42] were calculated using MEGA software [35]. For the calculation of TN distances, we estimated the gamma shape parameter using MrBayes [43] plugin for Geneious software. Distance values for the synonymous substitutions per site (ds) and non-synonymous substitutions per site (dn) were calculated using the modified Nei-Gojobori (NG) method [44]. Standard errors were computed using the bootstrap method [45] with 2,000 replicates. Fisher’s exact test was used for the statistical analyses [46].

**Synteny and Harr plot analyses**

Synteny of neighboring genes of the *trappin* genes was investigated by surveying neighboring genes on horse genome cont2.26764 (AAWR02026765), megabat genome cont1.168530 (ABRP01168531), cow chromosome 13 (DAAA02036736) [47], hyrax genome cont1.336045 (ABRQ01336046), elephant SuperContig scaffold_19, human chromosome 20, mouse chromosome 2, rat chromosome 3, and dog chromosome 24. Harr plot analyses were performed at a 23/40 nucleotide stringency using Genetyx-win software (Genetyx Co., Tokyo).

**List of abbreviations**

TGS: transglutaminase substrate; WAP: whey acidic protein; SVP: seminal vesicle clotting protein; SLPI: secretory leukocyte proteinase inhibitor; Mya: million years ago; ds: distance values for synonymous substitutions per site; dn: distance values for non-synonymous substitutions per site; NJ: neighbor-joining; MP: maximum parsimony; JC: Jukes-Cantor; NG: Nei-Gojobori.

**Acknowledgements**

We thank Tomoko Okada for her secretarial assistance. This work was supported by the Ministry of Education, Culture, Sport, Science, and Technology of Japan (MEXT) 21st Century and Global Center of Excellence Program of MEXT.

**Author details**

1Department of Biological Sciences, Tokyo Institute of Technology, Yokohama, Japan. 2United States Department of Agriculture, National Center for Agricultural Research Utilization, Microbial Genomics Research Unit, USA. 3Current address: Laboratory for Neurobiology of Synapse, RIKEN Brain Science Institute, Saitama, Japan.

**Authors’ contributions**

AK and SH planned and designed the study. AK, YF, and AR performed the analyses. AK, AR, and SH wrote the manuscript. All authors read and approved the final manuscript.

**Received: 14 May 2009 Accepted: 29 January 2010 Published: 29 January 2010**

**References**

1. Schalkwijk J, Wiedow O, Hirose S: The trappin gene family: proteins defined by an N-terminal transglutaminase substrate domain and a C-terminal four-disulphide core. Biochem J 1999, 340(Pt 3):569-577.
2. Nara K, Ito S, Ito T, Suzuki Y, Ghoneim MA, Tachibana S, Hirose S: Elastase inhibitor elafin is a new type of proteinase inhibitor which has a transglutaminase-mediated anchoring sequence termed “cementoin”. J Biochem (Tokyo) 1994, 115(3):441-448.
3. Steinert PM, Marekov LN: Direct evidence that involucrin is a major early isoprotein cross-linked component of the keratinocyte commelin cell envelope. J Biol Chem 1997, 272(20):2021-2030.
4. Molhuizen HO, Alkemade HA, Zeewen PL, de Jongh GJ, Wieringa B, Schalkwijk J: SKALP/elafin: an elastase inhibitor from cultured human keratinocytes. Purification, cDNA sequence, and evidence for transglutaminase cross-linking. J Biol Chem 1993, 268(16):12028-12032.
5. Wiedow O, Luademann J, Utech S, Elafin is a potent inhibitor of proteinase 3. Biochem Biophys Res Commun 1991, 174(1):6-10.
6. Wiedow O, Schroder JM, Gregory H, Young JA, Christophers E: Elafin: an elastase-specific inhibitor of human skin. Purification, characterization, and complete amino acid sequence [published erratum appears in J Biol Chem 1997 Feb 15;266(5):3356]. J Biol Chem 1990, 265(23):14791-14795.
7. Sallenave JM: Antimicrobial activity of antiproteases. Biochem Soc Trans 2002, 30(2):111-115.
8. Simpson AJ, Maxwell AJ, Govan JR, Hadlett C, Sallenave JM: Elafin (elastase-specific inhibitor) has anti-microbial activity against gram-positive and gram-negative respiratory pathogens. FEBS Lett 1999, 452(3):309-313.
9. Baranger K, Zani ML, Chandenier J, Dallet-Choiry S, Moreau T: The antibacterial and antifungal properties of trappin-2 [pre-elafin] do not depend on its protease inhibitory function. FEBS J 2008, 275(9):2008-2020.
10. Molhuizen HO, Schalkwijk J: Structural, biochemical, and cell biological aspects of the serine proteinase inhibitor SKALP/elafin/ESL. Biochim Biophys Acta 2009, 1791(1):1-7.
11. Zaidi SH, You XM, Ciura S, O’Brien S, Husain M, Rabanovitch M: Suppressed smooth muscle proliferation and inflammatory cell invasion...
after arterial injury in elafin-overexpressing mice. J Clin Invest 2000, 105(12):1687-1695.
12. Williams SE, Brown TI, Roghanian A, Sallenave JM. SLPI and elafin: one glove, many fingers. Clin Sci (Lond) 2006, 110(1):21-35.
13. Fritz H. Human mucus proteinase inhibitor (human MPl). Human seminal inhibitor I (HSUI-I), antileukoprotease (ALP), secretory leukocyte protease inhibitor (SLPI). Biol Chem Hoppe Seyler 1988, 369(Suppl):78-82.
14. Sallenave JM, Shalmun J, Cressley J, Jordana M, Gauld J. Regulation of secretory leukocyte proteinase inhibitor (SLPI) and elastase-specific inhibitor (ESI/elafin) in human airway epithelial cells by cytokines and neutrophilic enzymes. Am J Respir Cell Mol Biol 1994, 11(6):733-741.
15. Bingle L, Trelley TD, Bingle CD. Cytokine-mediated induction of the human elafin gene in pulmonary epithelial cells is regulated by nuclear factor-kappaB. Am J Respir Cell Mol Biol 2001, 25(1):84-91.
16. Brown TI, Mistry R, Collie DD, Tate S, Sallenave JM. Trappin ovine molecule (TOM), the ovine ortholog of elafin, is an acute phase reactant in the lung. Physiol Genomics 2004, 19(1):11-21.
17. Claus A, Lilja H, Lundwall A. A locus on human chromosome 20 contains several genes expressing protease inhibitor domains with homology to whey acidic protein. Biochem J 2002, 368(Pt 1):233-242.
18. Tamechika I, Itakura M, Saruta Y, Furukawa M, Kato A, Tachibana S, Hirose S. Accelerated evolution in inhibitor domains of porcine elafin family homologs. J Biol Chem 1996, 271(12):702-718.
19. Furutani Y, Kato A, Yase Y, Alexander LJ, Beattie CW, Hirose S. Evolution of the trappin multigene family in the Suidae. J Biochem (Tokyo) 1998, 124(3):491-502.
20. Claus A, Lilja H, Lundwall A. The evolution of a genetic locus encoding small serine proteinase inhibitors. Biochim Biophys Acta 2005, 1733(2):383-389.
21. Furutani Y, Kato A, Kawai R, Fibriani A, Kojima S, Hirose S. Androgen-dependent expression, gene structure, and molecular evolution of guinea pig trappin II, a WAP-motif protein. Bio Reprod 2004, 71(5):1583-1590.
22. Furutani Y, Kato A, Fibriani A, Hirata T, Kawai R, Jeon JH, Fujii Y, Kim KS, Kojima S, Hirose S. Identification, evolution, and regulation of expression of Guinea pig trappin with an unusually long transglutaminase substrate domain. J Biochem 2005, 280(21):2004-2015.
23. Saheti T, Ito F, Hagiwara H, Saito Y, Kurori K, Tachibana S, Hirose S. Primary structure of the human elafin precursor proelafnin derived from the nucleotide sequence of its gene and the presence of unique repetitive sequences in the presequence. Biochim Biophys Acta 1992, 1185(2-4):240-245.
24. Nishihara H, Hasegawa M, Okada N. Pegasoferase, an unexpected mammalian clade revealed by tracking ancient retroposon insertions. Proc Natl Acad Sci USA 2006, 103(26):9929-9934.
25. Drummmond AJ, Rambaut A: BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol 2007, 7:214.
26. Hurle B, Swanwijk GV, Green ED. Comparative sequence analyses reveal rapid and divergent evolutionary changes of the WFDC locus in the primate lineage. Genome Res 2007, 17(3):276-286.
27. Lundwall A, Ulvback M. The gene of the protease inhibitor SKALP/elafin is a member of the REST gene family. Biochem Biophys Res Commun 1996, 221(2):323-327.
28. Zeszusiew PL, Hendriks W, de Jong WW, Schalkwijk J. Identification and sequence analysis of two new members of the SKALP/elafin and SPAI-2 gene family. Biochemical properties of the transglutaminase substrate motif and suggestions for a new nomenclature. J Biol Chem 1997, 272(33):20471-20478.
29. Pickford M. Old World suoid systematics, phylogeny, biogeography, and biostratigraphy. Palaeontol Evol 1993, 26:27-237-269.
30. Randi E, Lucchini V, Dong CH. Evolutionary genetics of the Suiformes as reconstructed using mtDNA sequencing. J Mamm Evol 1996, 3(2):163-194.
31. Kumar S, Hedges SB: A molecular timescale for vertebrate evolution. Nature 1998, 392(6679):917-920.
32. Hagstrom JE, Fautsch MF, Perdomoski V, Vrabel A, Wielben ED. Exons lost and found. Evolutionary reconstruction of a seminal vesicle transglutaminase substrate. J Biol Chem 1996, 271(35):21114-21119.
33. Hubbard TJ, Aken BL, Aytal S, Ballester B, Beal K, Braig E, Brett S, Chen Y, Chlapp M, Clarke L, Coates G, Fairley S, Fitzgerald S, Fernandez-Banet J, Gordon L, Graf S, Haider S, Hammond M, Hoque K, Jenkinson A, Johansson N, Kahan A, Kefee D, Keinan S, Kneissl R, Kokocinska E, Kuleba E, Lawson D, Longden I, Megy K, Meidi P, Overduin B, Parker A, Pritchard B, Rios D, Schuster M, Slaters G, Smledy D, Spooner W, Spudich G, Trevanion S, Vellela A, Vogel J, White S, Wilder S, Zadisa A, Birney E, Cunningham F, Curwen V, Duffin R, Fernandez-Suarez XM, Herrero J, Kasprzyk A, Proctor G, Smith J, Searle S, Sklice P. Ensemble 2009. Nucleic Acids Res 2009, 37 Database: D690-697.
34. Kullberg M, Nilsson MA, Arnason U, Harley EH, Janke A. Housekeeping genes for phylogenetic analysis of eutherian relationships. Mol Biol Evol 2006, 23(8):1493-1503.
35. Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 2007, 24(5):1596-1599.
36. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994, 22(22):4673-4680.
37. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 1987, 4(4):406-425.
38. Nei M, Kumar S. Molecular Evolution and Phylogenetics. New York, NY: Oxford Univ. Press 2000.
39. Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 2003, 52(5):696-704.
40. Benton MJ, Donoghue PC. Paleontological evidence to date the tree of life. Mol Biol Evol 2007, 24(1):26-33.
41. Jukes TH, Cantor CR. Evolution of protein molecules. Mammalian Protein Metabolism New York: Academic Press/Munro HN 1969, 21-132.
42. Tamura K, Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol 1993, 10(3):512-526.
43. Ronquist F, Huelsenbeck JP. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 2003, 19(12):1572-1574.
44. Zhang J, Rosenberg HF, Nei M. Positive Darwinian selection after gene duplication in primate ribonuclease genes. Proc Natl Acad Sci USA 1998, 95(7):308-313.
45. Dopazo J. Estimating errors and confidence intervals for branch lengths in phylogenetic trees by a bootstrap approach. J Mol Evol 1994, 38(3):300-304.
46. Zhang J, Kumar S, Nei M. Small-sample tests of episodic adaptive evolution: a case study of primate lysozymes [letter]. Mol Biol Evol 1997, 14(12):1335-1338.
47. Zimin AV, Delcher AL, Florea L, Krause J, Lander ES, Lomsadze A, Liu Z, Maddron H, Malahov A, Khosuev A,秘密2004, 214.20471-20478.
48. Hasegawa M, Torme JL, Kishino H. Time scale of eutherian evolution estimated without assuming a constant rate of molecular evolution. Genetics Genet Syst 2003, 78(4):267-283.
49. Hallstrom SM, Janke A. Resolution among major placental mammal interordinal relationships with genome data imply that speciation influenced their earliest radiations. BMC Evol Biol 2008, 8:162.
50. Arnason U, Adegolke JA, Bregalha A, Harley EH, Janke A, Kullberg M. Mitogenomic relationships of placental mammals and molecular estimates of their divergences. Gene 2008, 421(1-2):37-51.

Cite this article as: Kato et al. Evolution of trappin genes in mammals. BMC Evolutionary Biology 2010 10:31.

doi:10.1186/1471-2148-10-31