

RESEARCH ARTICLE

# Evolution of Vertebrate *Adam* Genes; Duplication of Testicular *Adams* from Ancient *Adam9/9-like* Loci

Harinath Bahudhanapati<sup>☉</sup>, Shashwati Bhattacharya<sup>☉</sup>, Shuo Wei<sup>\*</sup>

Department of Biology, West Virginia University, Morgantown, WV, United States of America

☉ These authors contributed equally to this work.

\* [Shuo.Wei@mail.wvu.edu](mailto:Shuo.Wei@mail.wvu.edu)



**OPEN ACCESS**

**Citation:** Bahudhanapati H, Bhattacharya S, Wei S (2015) Evolution of Vertebrate *Adam* Genes; Duplication of Testicular *Adams* from Ancient *Adam9/9-like* Loci. PLoS ONE 10(8): e0136281. doi:10.1371/journal.pone.0136281

**Editor:** Hubert Vaudry, University of Rouen, France, FRANCE

**Received:** November 26, 2014

**Accepted:** August 2, 2015

**Published:** August 26, 2015

**Copyright:** © 2015 Bahudhanapati et al. 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 use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** This work was supported by the U.S. National Institute of Health (GM114105 and DE022813 to SW) and the March of Dimes Foundation (1-FY10-399 to SW). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

## Abstract

Members of the disintegrin metalloproteinase (ADAM) family have important functions in regulating cell-cell and cell-matrix interactions as well as cell signaling. There are two major types of ADAMs: the somatic ADAMs (sADAMs) that have a significant presence in somatic tissues, and the testicular ADAMs (tADAMs) that are expressed predominantly in the testis. Genes encoding tADAMs can be further divided into two groups: group I (intronless) and group II (intron-containing). To date, *tAdams* have only been reported in placental mammals, and their evolutionary origin and relationship to *sAdams* remain largely unknown. Using phylogenetic and syntenic tools, we analyzed the *Adam* genes in various vertebrates ranging from fishes to placental mammals. Our analyses reveal duplication and loss of some *sAdams* in certain vertebrate species. In particular, there exists an *Adam9-like* gene in non-mammalian vertebrates but not mammals. We also identified putative group I and group II *tAdams* in all amniote species that have been examined. These *tAdam* homologues are more closely related to *Adams 9* and *9-like* than to other *sAdams*. In all amniote species examined, group II *tAdams* lie in close vicinity to *Adam9* and hence likely arose from tandem duplication, whereas group I *tAdams* likely originated through retroposition because of their lack of introns. Clusters of multiple group I *tAdams* are also common, suggesting tandem duplication after retroposition. Therefore, *Adam9/9-like* and some of the derived *tAdam* loci are likely preferred targets for tandem duplication and/or retroposition. Consistent with this hypothesis, we identified a young retroposed gene that duplicated recently from *Adam9* in the opossum. As a result of gene duplication, some *tAdams* were pseudogenized in certain species, whereas others acquired new expression patterns and functions. The rapid duplication of *Adam* genes has a major contribution to the diversity of ADAMs in various vertebrate species.

## Introduction

The first two members of the ADAM family were originally identified as guinea pig sperm-surface antigens that were recognized by the monoclonal antibody PH-30, which strongly inhibits sperm-egg fusion [1]. Sequence analyses of the precursors of both proteins showed that they have a domain organization similar to that of snake venom metalloproteinases, containing a disintegrin domain and a metalloproteinase domain [2,3]. Genes encoding these two proteins, together with several homologues that were cloned from guinea pig and mouse testis, were therefore renamed “*Adams*” (a disintegrin and metalloproteinases) to reflect these important features [4]. It soon became clear that the expression and functions of ADAMs are not limited to the testis. Instead some ADAM family members (*i.e.*, the sADAMs) have a significant presence in somatic tissues and play important roles in embryonic development [5,6]. To date close to 40 *Adam* genes have been identified in vertebrate and invertebrate animals, and *Adam*-like sequences have also been found in the yeast species *Schizosaccharomyces pombe* [7,8]. Functions of ADAMs in normal development vary from neurogenesis and neural crest development to adipogenesis and myogenesis [6,9,10]. ADAMs have also been implicated in many diseases, including tumors, immune diseases and neurodegenerative diseases [7,8].

A typical ADAM is a type I transmembrane protein, although secreted soluble proteins can also be generated through alternative splicing. The unprocessed form of these transmembrane proteins contains a signal peptide, a pro-domain, a metalloproteinase domain, a disintegrin domain, a cysteine-rich domain, a transmembrane region and a cytoplasmic tail, from N- to C-terminus [7,8]. Only some ADAMs contain the consensus zinc-binding motif (HEXGHXXGXXHD) that is required for metalloproteinase activity, whereas the rest are considered proteolytically inactive. The pro-domain functions as a chaperone to facilitate the folding and secretion of the ADAMs; it also binds to the metalloproteinase domain and inhibits the autocleavage activity of proteolytically active ADAMs. In the processed (“mature”) form of ADAMs, the pro-domain is removed by pro-protein convertases. The mature form of the proteolytically active ADAMs can then act on its substrates, which are mainly cell-surface proteins, to regulate various cellular events including signaling and adhesion. The disintegrin and cysteine-rich domains have been shown to modulate cell-cell and cell-matrix interactions, as well as substrate recognition for the metalloproteinase domain. In particular, the disintegrin domain of several ADAMs can interact specifically with various integrins [7,8]. Finally, the cytoplasmic tail may regulate the functions of extracellular domains through responding to intercellular signals or controlling subcellular localization of ADAMs. It has also been reported that the cytoplasmic tail of ADAM13, an sADAM, is cleaved and translocates into the nucleus, where it regulates target gene expression [11].

The mouse genome contains at least 34 *Adam* genes, while 27 *Adam* loci have been identified to date in the human genome [12,13]. About half of these *Adam* genes are expressed exclusively or predominantly in the testis, and are referred to as *tAdams*. These *tAdams* are further classified into two groups: group I genes (*Adams* 1, 4, 6, 20, 21, 24–26, 29, 30, 34, and 36–39), which have no or few introns; and group II genes (*Adams* 2, 3, 5, 18 and 32), which have multiple introns. None of the group II tADAMs contains the consensus zinc-binding motif in the metalloproteinase domain, and are hence likely non-proteolytic. By contrast, group I tADAMs, except for ADAMs 4, 6 and 29, are predicted to be proteolytically active. It should be noted that some tADAMs, including the presumably proteolytic ADAM1, undergo further processing to remove the metalloproteinase domain during sperm maturation [12,14]. Another unique feature of the tADAMs is that they are often associated with each other and form oligomers on the cell surface [12]. Such complex formation seems to be important for the stability of some tADAMs, as knockout of one tADAM often leads to loss or reduction of another [12,15]. One

exception is mouse ADAM24, which has been shown to be present on the sperm surface as a monomer with an intact metalloproteinase domain. This ADAM and several other closely related paralogues that have been found in the mouse genome (ADAMs 25, 26, 34, and 36–39), all predicted to be proteolytically active, are also called “testases” [16,17].

Thus far, *tAdam* genes have only been identified in eutherians, and it is not clear how they originated during evolution and if there are any *tAdams* in other animals. Most of our current knowledge on the functions of tADAMs comes from studies in mice, which show that several *tAdam* genes are important for fertilization in this rodent species. For example, *Adam1a*<sup>-/-</sup>, *Adam2*<sup>-/-</sup> and *Adam3*<sup>-/-</sup> male mice are all infertile. Detailed analyses indicate that sperms lacking any of these tADAMs are defective in their migratory activities in the female reproductive tract, as well as in *in vitro* sperm-egg interaction [18–21]. A different phenotype was observed in *Adam24*<sup>-/-</sup> mice: these mice displayed reduced fertility with high level of polyspermic fertilization [22]. However, some of these genes, such as *Adams 1a* and 3, appear to have been pseudogenized in the human genome [23,24], suggesting that functions of these genes are species-specific. Therefore, a thorough investigation of the *tAdam* genes in various animal species will not only help us understand how these genes evolved, but also provide important clues on which of these genes may be expressed and functional in individual species. Given the importance of tADAMs in fertilization, the latter information may be of particular interest to veterinary and agricultural research.

In the current study we identified putative group I and group II *tAdams* in all amniote species that have been examined. Phylogenetic and syntenic (gene linkage) analyses suggest that these *tAdams* originally derived from ancient *Adam9*, a somatic *Adam* that is conserved in vertebrates, and/or *Adam9-like*, which we identified here as a close paralogue of *Adam9*. Two mechanisms, tandem duplication and retroposition, were found to underlie the expansion of the *tAdam* genes during vertebrate evolution. These results provide important insight into the evolution of *tAdams*, and form the basis for future functional studies of these important reproductive genes in various vertebrate species.

## Materials and Methods

### Identification of *Adam* genes in various vertebrate species

The sequences of annotated ADAM proteins from humans and mice were BLASTed against the EST databases and genome sequences of other vertebrate species to identify *Adam* genes. Specifically, we performed consistent mining of representative vertebrate species (opossum, chick, *Anolis* lizard, *Xenopus tropicalis* and zebrafish) to understand the broader evolution of the ADAM genes. This was augmented with data obtained from additional species, such as rat, rabbit, pig, dog, platypus, zebra finch, flycatcher, coelacanth, Japanese medaka, cave fish, spotted gar, elephant shark and lamprey, where a more specific understanding of synteny/phylogeny was needed to provide clarity on the data. Potential retroposed *Adam* genes were identified as *Adam* genes with no or few ( $\leq 2$ ) introns by UCSC Genome Browser BLAT search [25]. To identify the retroposon elements, the *Adam* genes, along with surrounding genome sequences that are 1 kb upstream and downstream, were analyzed with the CENSOR program [26].

### Establishment of orthology with existing *Adam* genes

Three main criteria, phylogeny, synteny, and conservation of exon/intron boundaries, were used to establish the orthology with existing human and mouse *Adams*. Phylogenetic and syntenic analyses were carried out as described below. To determine the exon/intron boundaries of *Adam* genes, deduced amino acid sequences were searched against the UCSC Genome Browser using BLAT [25].

## Phylogenetic analyses

For phylogenetic analyses, deduced amino acid sequences encoded by *Adam* genes from various species (S1 Table) were aligned using MUSCLE (Multiple Sequence Comparison by Log-Expectation; [27]) alignment, and all positions containing gaps were eliminated. Known pseudogenes, such as human *Adams 1, 3, 6* and *24*, were not included in the analyses. Phylogenetic trees, including Bayesian, neighbor joining and maximum likelihood trees, were built to assess the evolutionary relationship among ADAMs as described below.

The Bayesian analyses were conducted with MrBayes version 3.2.2 [28]. Each analysis was run for 3,000,000 iterations with six chains to achieve a mean standard deviation less than 0.01. Trees were sampled every 100 iterations. The final tree was visualized with FigTree v4.1.0 [29].

Neighbor joining and maximum likelihood trees were generated by using the corresponding algorithms included in the MEGA 6.0 software [30]. For neighbor joining trees, the bootstrap consensus trees inferred from 1000 replicates were used to represent the evolutionary history of the taxa analyzed. The evolutionary distances were computed using the numbers of amino acid substitutions per site with the Poisson correction method [31]. The Maximum Likelihood method is based on the JTT matrix-based model [32]. Initial trees were obtained by applying the neighbor joining and BioNJ algorithms to a matrix of pairwise distances using the JTT model, and then the topology was selected with superior log likelihood value.

## Syntenic analyses

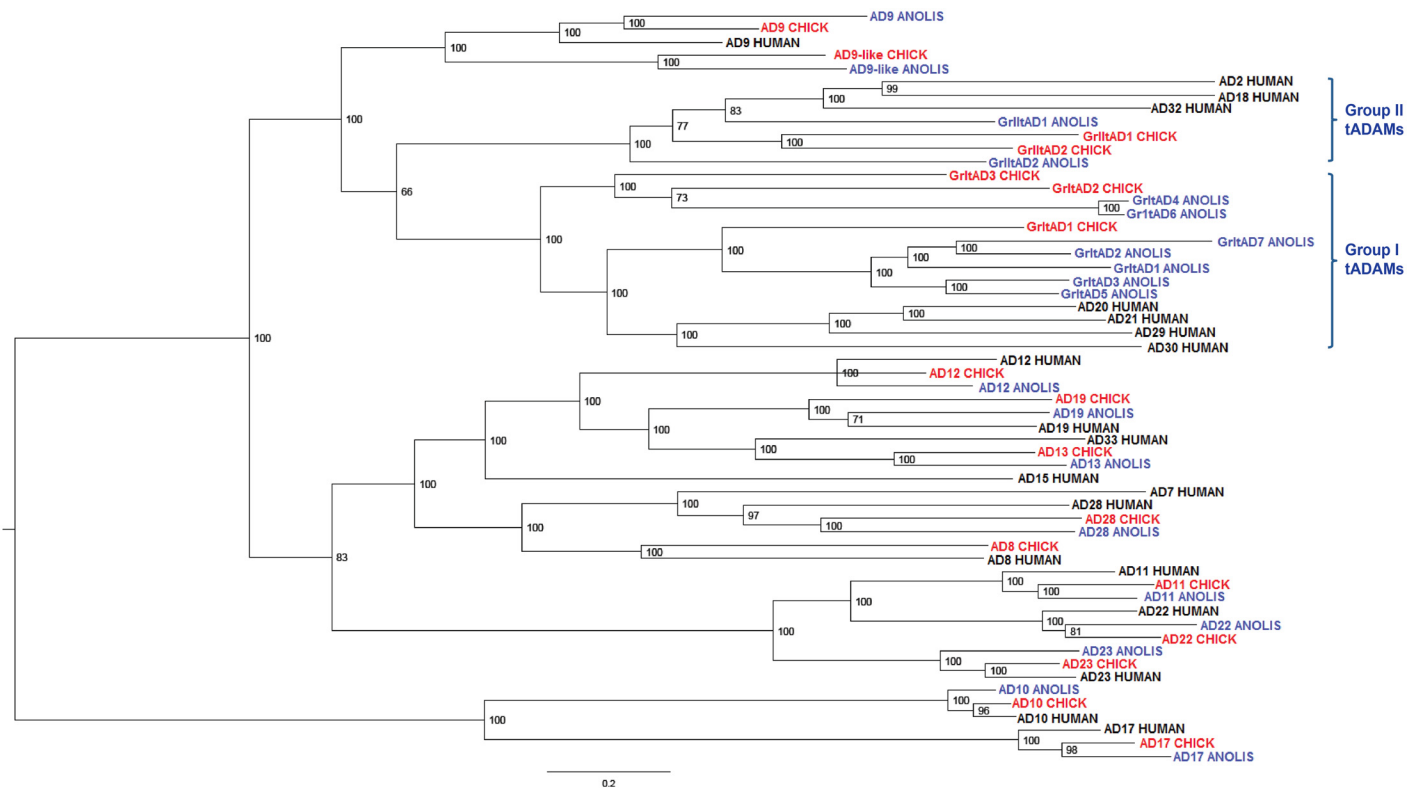
For syntenic analyses, an individual *Adam* gene was entered into the search box of Genomicus [33], Metazome [34], and NCBI Map Viewer to identify surrounding loci that are evolutionarily conserved and can serve as syntenic anchor genes. Data from these websites were combined to generate the syntenic maps. Loss of an *Adam* gene in a specific species was validated by searching the sequences of nearby genomic regions that include the syntenic anchor genes.

## Results and Discussion

### ADAMs in various animal species

A previous study by Huxley-Jones *et al.* identified 4 and 22 *Adam* genes in *Ciona* and zebrafish, respectively, and we have also reported the presence of 15 homologues of mammalian *sAdam* genes in *Xenopus tropicalis* [35,36]. In all published analyses on ADAMs from various invertebrate and vertebrate species, ADAMs 10 and 17 stand out phylogenetically as a separate clade from the rest of *sADAMs* [35–38]. These two ADAMs are highly conserved throughout animal evolution, with orthologues present in invertebrates such as *C. elegans*, *Drosophila* and *Ciona* [35,37]. One exception is sea urchin, where ADAM10 seems to be absent [38]. There is another ADAM in sea urchin and two in *Drosophila* and *Ciona*, which do not belong to the ADAM10/17 clade. These ADAMs cluster with the rest of mammalian *sADAMs*, but no distinct orthology could be established [35].

In contrast to the few *sADAMs* in invertebrates, orthologues of most mammalian *sADAMs*, including ADAMs 9, 10, 11, 12, 15, 17, 19, 22, 23, 28 and 33, have been identified in both zebrafish and *Xenopus* [35,36], suggesting an expansion of *sAdam* genes during early vertebrate evolution. However, these published studies failed to identify an orthologue of ADAM7 in zebrafish and *Xenopus*, and ADAM8 seems to be missing in *Xenopus* [35,36]. Moreover, no *tADAM* orthologue has been reported in any non-eutherian species [35–38]. To understand the evolution of *Adam* genes and, particularly, the origins of *tAdams*, we performed a search for *Adam* genes in various vertebrate species, including *X. tropicalis*, *Anolis*, chick, opossum, mouse and human. Phylogenetic trees of deduced protein sequences encoded by these genes



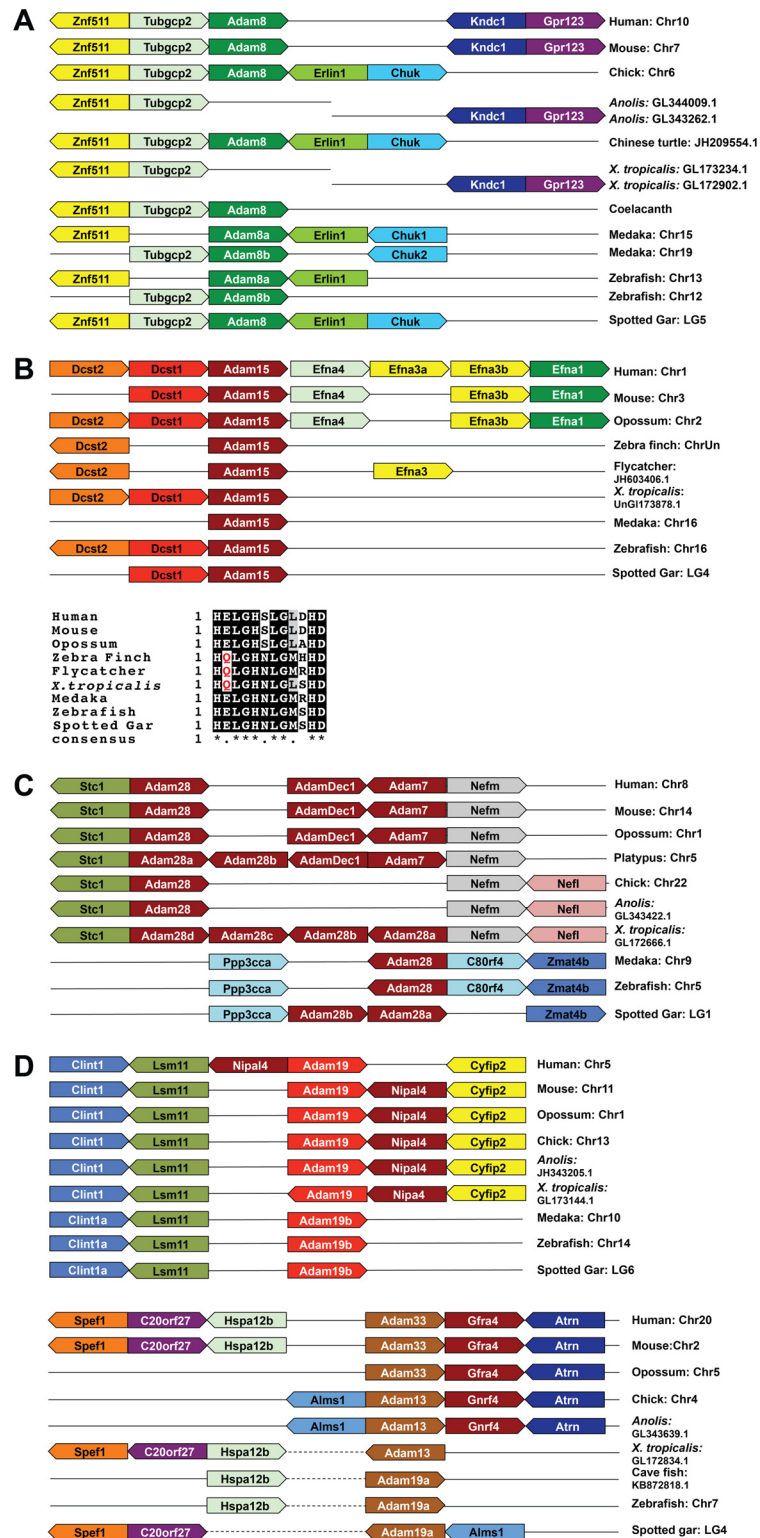
**Fig 1. Bayesian tree of ADAMs from *Anolis*, chick and human.** Sequences of ADAMs from *Anolis* (blue), chick (red) and human (black) were aligned, and Bayesian trees were generated as described in Materials and Methods. The numbers shown are Bayesian posterior probability values.

doi:10.1371/journal.pone.0136281.g001

were generated by three methods: Bayesian (Fig 1 and S1 Fig), neighbor joining (S2 Fig), and maximum likelihood (S3 Fig); the results obtained with different methods agree well with each other. We also carried out syntenic analyses for these *Adam* genes using the online comparative genomics tools such as Genomicus and Metazome [33]. The results are summarized below.

### Absence of several *sAdams* in certain vertebrate species

We showed previously that the genomic region surrounding *Adam8* underwent substantial reorganization during vertebrate evolution, likely resulting in the loss of *Adam8* in *X. tropicalis* [36]. Similarly, we could not identify *Adam8* in the *Anolis* genome, although an *Adam8* orthologue does exist in Chinese softshell turtle, another reptile species, and in coelacanth, one of the closest living fish relatives of tetrapods (Fig 2A; [39]). A single *Adam8* gene was found in all mammalian and avian species that have been examined, and there are two copies of *Adam8* in the genomes of several teleost fish species (Fig 2A and [35]), likely as a result of teleost genome duplication (TGD; [40,41]). This is supported by the presence of a single *Adam8* gene in the spotted gar (Fig 2A), which diverged from teleosts before the TGD and has a genome organization more similar to that of humans than teleosts [42]. Our syntenic analysis further shows that most non-mammalian vertebrates share a similar genome structure in this region, but *X. tropicalis* and *Anolis* have undergone chromosome reorganization that likely resulted in loss of *Adam8* (Fig 2A).



**Fig 2. Syntenic analyses of certain sADAMs.** A. Absence of *Adam8* in the *X. tropicalis* and *Anolis* genomes. B. Loss of ADAM15 gene or protease activity in non-mammalian tetrapods. Sequences of ADAM15 proteins from various vertebrate species were aligned with ClustalW and colored using BoxShade, and the glutamine found in certain species instead of the glutamic acid in the conserved zinc-binding motif is highlighted in red. C. Absence of *Adam7* and *Adamdec1* in non-mammalian vertebrates. D. Syntenic

analyses showing that fish *Adam19b* is the orthologue of tetrapod *Adam19*, whereas fish *Adam19a* is the orthologue of tetrapod *Adam13/33*. Syntenic maps were generated as described in Materials and Methods. Dashed lines indicate that genes are several loci away from each other (same below).

doi:10.1371/journal.pone.0136281.g002

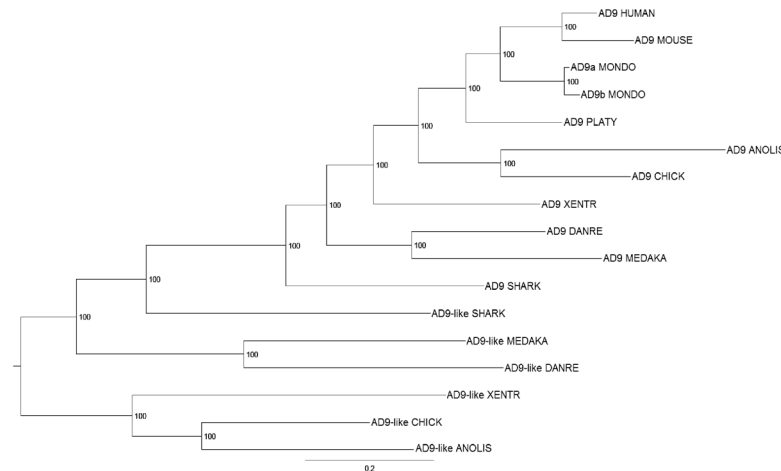
Another notable loss of gene or gene function lies in the *Adam15* locus. The mammalian *Adam15* encodes an active metalloproteinase, whose activity has been demonstrated using several substrates [43,44]. However, we have predicted that both *X. tropicalis* and *X. laevis* ADAM15 proteins are non-proteolytic, as they contain a glutamine instead of the glutamic acid that is found in the consensus zinc-binding motif [36]. In the current study we were unable to find an *Adam15* orthologue in the *Anolis* or chicken genome. *Adam15* does exist in the zebra finch and flycatcher, two avian species, but the deduced protein sequences show that both contain the E to Q substitution that was previously found in frogs (Fig 2B). A separate report also suggests that *Adam15* is lost in the chick but not zebra finch [45]. On the other hand, fishes seem to have a well-conserved *Adam15* gene encoding a proteolytic enzyme (Fig 2B), suggesting that the loss of ADAM15 or its proteolytic activity is probably specific for non-mammalian tetrapods.

On human chromosome 8p12, genes encoding ADAM28, ADAMDEC1 (a soluble ADAM-like protein) and ADAM7 form a cluster, likely as a result of tandem duplication [46]. Similar clustering of these three *Adams* was found in other mammalian species including marsupials and monotremes, except that in platypus *Adam28* was further duplicated to give rise to two genes in tandem (Fig 2C). Four and two tandem copies of *Adam28* exist in *X. tropicalis* and spotted gar, respectively, but no definitive orthologue of *Adam7* or *Adamdec1* could be identified (Fig 2C and [36]). We were also unable to identify either *Adam7* or *Adamdec1* in any other non-mammalian vertebrate genomes that have been investigated, including those of aves, reptiles and fishes. Instead, in most of these non-mammalian vertebrates, a single *Adam28* locus is present in this region (Fig 2C and data not shown). Therefore, the duplication of *Adam28* in the spotted gar, *X. tropicalis* and platypus seems to be independent events. These data also suggest that *Adam7* and *Adamdec1* were duplicated from *Adam28*, probably only in mammals. Subsequently, ADAM7 and ADAMDEC1 underwent neo- or subfunctionalization, as well as the loss of ADAM7 protease activity and truncation of ADAMDEC1 that removed part of the disintegrin domain and all the sequence C-terminal to it [36,46].

Our previous phylogenetic and syntenic analyses indicate that *Xenopus* ADAM13, which plays an important role in neural crest induction and migration as well as eye formation, is an orthologue of mammalian ADAM33 [11,36,47,48]. ADAM13/33 also exists in all avian and reptile species that have been examined (Fig 2D and data not shown). However, Huxley-Jones *et al.* reported the presence of two *Adam19* loci (*Adam19a* and *Adam19b*) but not *Adam13/33* in zebrafish [35], raising the question whether *Adam13/33* only exists in tetrapods. The two *Adam19* loci in zebrafish are unlikely the result of TGD, as both are also present in the spotted gar (Fig 2D). A syntenic comparison with other vertebrate species reveals that fish *Adam19b* is the orthologue of tetrapod *Adam19*, whereas fish *Adam19a* is in fact the orthologue of tetrapod *Adam13/33* (Fig 2D). This is consistent with our previous phylogenetic analysis, which showed that zebrafish ADAM19b clusters with tetrapod ADAM19 and zebrafish ADAM19a clusters with tetrapod ADAM13/33 with high confidence [36]. Thus ADAM13/33 is conserved from fishes to mammals.

### Presence of an *Adam9*-like gene in non-mammalian vertebrates

ADAM9 is an sADAM that plays important roles in tumor progression, including metastasis and angiogenesis [49]. Although *Adam9*<sup>-/-</sup> mice do not show any apparent developmental



**Fig 3. Bayesian trees of ADAMs 9 and 9-like from various vertebrate species.** MONDO, opossum; PLATY, platypus; XENTR, *X. tropicalis*; and DANRE, zebrafish. Note that opossum *Adam9b* is predicted to be a pseudogene with a frameshift.

doi:10.1371/journal.pone.0136281.g003

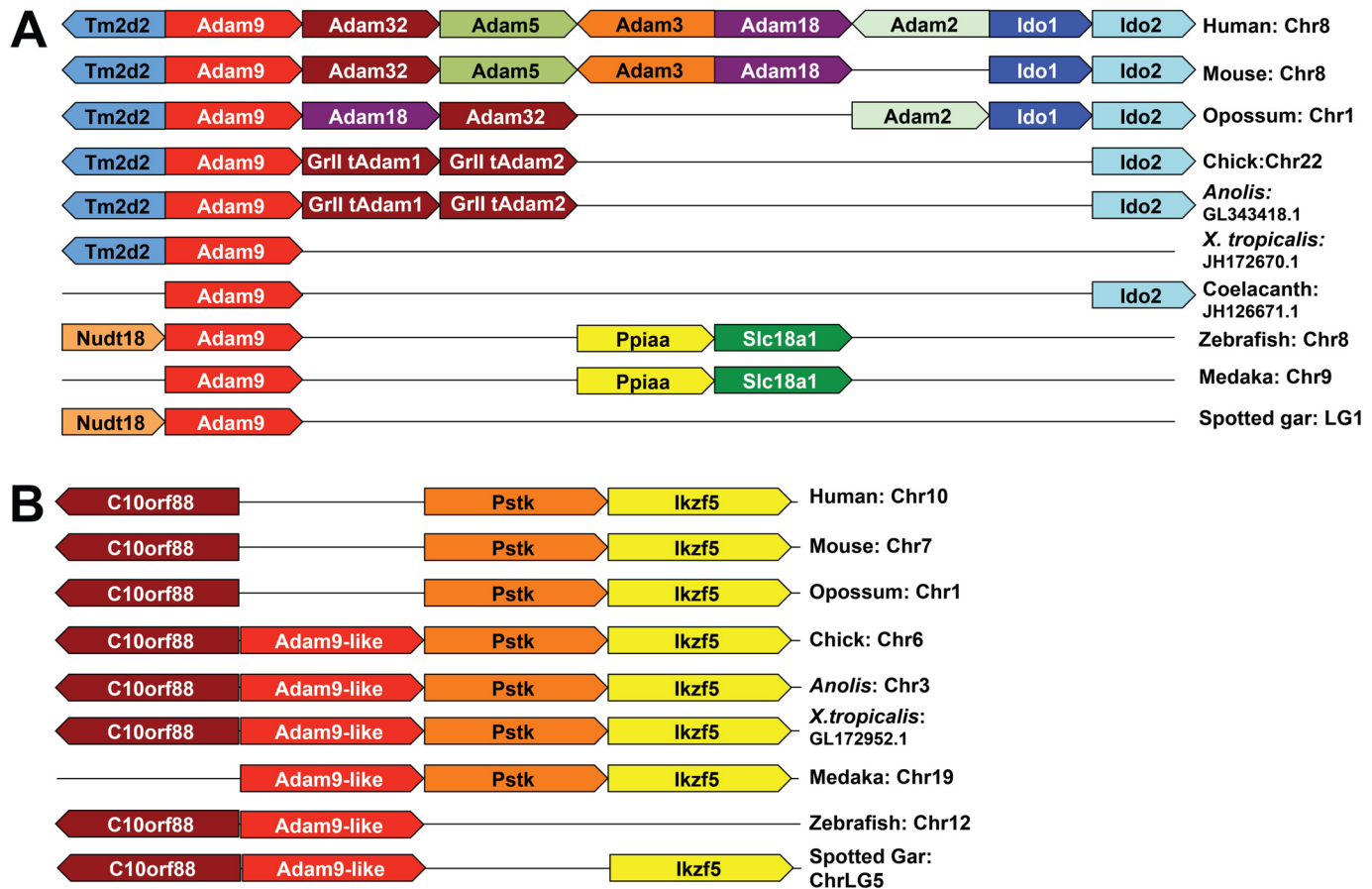
defects, they develop retinal degeneration during aging, and null mutations in *Adam9* have been linked to early onset cone-rod dystrophy in humans [50]. Previous studies did not identify an *Adam9* orthologue in invertebrates such as *Drosophila* and *Ciona*, but a partial sequence with high homology to mammalian *Adam9* was reported in zebrafish [35]. We searched the Ensembl database and found the full-length sequence of this gene; phylogenetic analysis indicates that it is the zebrafish orthologue of *Adam9* (Fig 3). An *Adam9* gene also exists in other fishes, including the cartilage fish elephant shark, as well as all other vertebrate species that have been examined (Figs 3 and 4A).

We also identified an *Adam9-like* gene in non-mammalian vertebrates, including fishes, amphibians, reptiles and aves; however, this gene seems to have been lost in mammals (Figs 3 and 4B). The predicted protein sequences encoded by the *Adam9-like* genes from various species are more similar to ADAM9 than to other sADAMs (Fig 1, S1A, S2A-S2C and S3A-S3C Figs). Similar to *Adam9*, the presence of *Adam9-like* can also be traced to the elephant shark (Fig 3), but neither *Adam9* nor *9-like* could be identified in the jawless fish lamprey. Thus both *Adams 9* and *9-like* likely emerged during early vertebrate evolution. However, we were unable to determine which gene emerged first, due to the lack of an available genome sequence between the lamprey and elephant shark.

### Group I and group II tADAMs exist in non-eutherian amniotes and form an ADAM subfamily with ADAMs 9 and 9-like

In an attempt to find potential *tAdam* genes in non-eutherian vertebrates, we identified close homologues of human and mouse group I and group II *tAdams* in *Anolis*, chick and opossum (Fig 1 and S1B Fig), but not in *X. tropicalis* or any of the fish species. Group I and group II tADAMs from various vertebrates form distinct clades, and both clades cluster with ADAMs 9 and 9-like (Fig 1, S1B, S1C, S2B-S2F and S3B-S3F Figs). Although no close homologue of any mammalian *tAdam* has been identified in the *X. tropicalis* genome [36], there is an *Adam16* (also called *xMdc*) gene that has been reported to be expressed exclusively in the testis of *Xenopus laevis* [51]. Interestingly, ADAM16 also clusters with ADAMs 9 and 9-like (S1A, S2A and S3A Figs), and hence may be a tADAM that evolved separately in frogs and related species.





**Fig 4. Syntenic analyses of ADAMs 9 (A) and 9-like (B).**

doi:10.1371/journal.pone.0136281.g004

Our findings that group I and group II tADAMs belong to the same ADAM subfamily that also includes ADAMs 9 and 9-like, and that ADAMs 9 and 9-like likely emerged earlier than both groups of tADAMs during vertebrate evolution, suggest that group I and group II tADAMs evolved from ancient ADAM9 and/or 9-like. In the literature, the clade of *Adams* 7 and 28 are sometimes classified together with the *tAdams* as “reproductive” *Adams*, as *Adams* 7 and 28 are predominantly expressed in the mouse epididymis [12]. However, information available in the UniGene database (UID 158438) shows that human *Adam28* is expressed in a wide variety of somatic tissues, and our phylogenetic analyses strongly indicate that ADAMs 7 and 28 do not belong to the subfamily that includes ADAMs 9, 9-like and tADAMs. Thus ADAMs 7 and 28 were not included in the subsequent analyses.

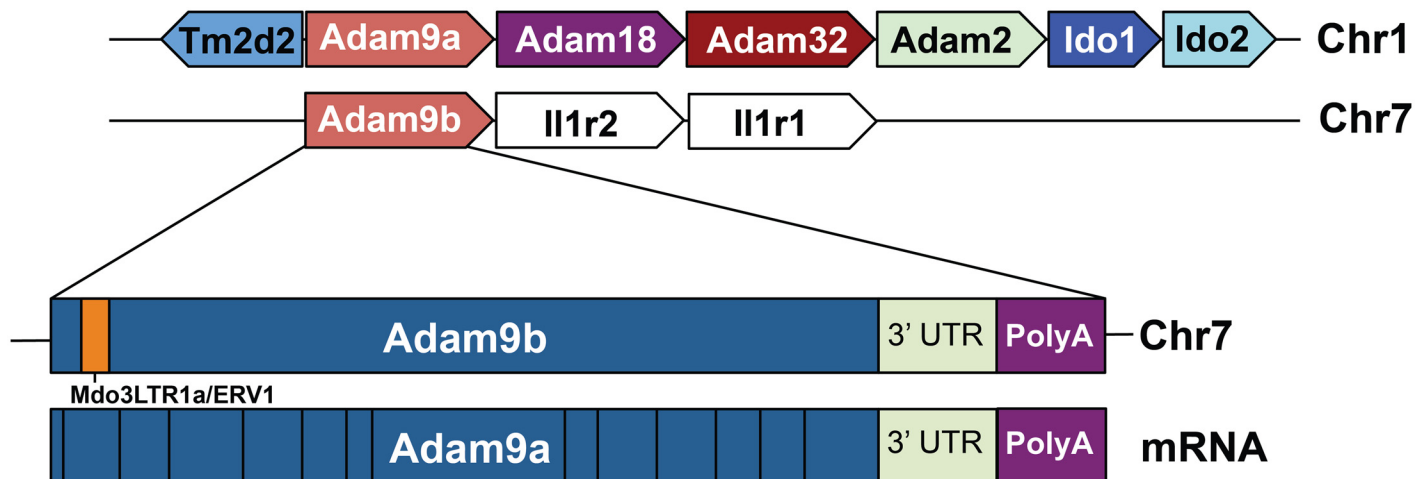
### Origins of group II and group I *tAdam* genes: evidence for tandem duplication and retroposition, respectively

Syntenic comparison of group II *tAdams* shows a unique feature: in all amniote species that were examined, group II *tAdams* cluster together in a genomic region that also includes *Adam9* (Fig 4A). The only exception is rodent *Adam2*, which is localized in a different region.

For example, in the mouse genome *Adam2* lies in chromosome 14 and *Adam9* in chromosome 8, whereas in the rat genome *Adam2* is localized in chromosome 15 and *Adam9* in chromosome 16. No other *Adam* gene was found in the close vicinity of *Adam9* in the *X. tropicalis* or zebrafish genome. By contrast, two group II *tAdam* genes lie next to *Adam9* in the *Anolis* and chicken genomes, whereas three, four and five were found in this region in the opossum, mouse and human genomes, respectively (Fig 4A). These results suggest that tandem duplication gave rise to group II *tAdam* genes, and that an ancient *Adam9* was likely the parental gene that generated the very first group II *tAdam* during evolution.

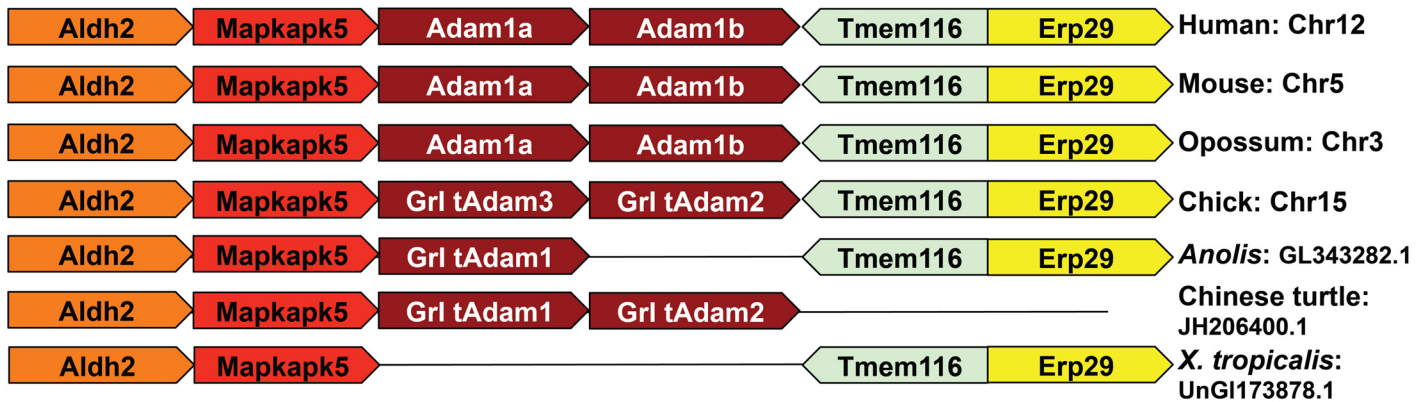
Unlike group II *tAdams*, which have multiple introns, group I *tAdams* have few or no introns. Because lack of introns is characteristic of retroposed genes, it has been speculated that group I *tAdams* derived from retroposition [12]. However, there is no published evidence that supports this speculation, and it was not clear from which gene(s) group I *tAdams* were retroposed. As discussed above, our phylogenetic and syntenic analyses suggest that group I *tAdams* share a common ancestor with group II *tAdams* as well as *Adams 9* and *9-like*. Thus the parental gene that gave rise to a specific group I *tAdam* could be *Adam9*, *9-like*, a group II *tAdam*, or another group I *tAdam*. Because nearly all of these group I *tAdams* emerged a long time ago, they have lost the signatures of retroposed genes (other than the lack of introns) that may serve as evidence for the retroposition events they underwent. We therefore started out to search for young retroposed genes that derived from the *Adam9/9-like/tAdam* subfamily, in which the characteristic features of retroposition may still be well preserved.

In the opossum genome, we identified an additional copy of *Adam9* that has only one intron (Fig 5). This gene, which we named *Adam9b*, is ~99% identical to the other copy (*Adam9a*) in nucleotide sequence. The *Adam9a* gene contains 24 introns and conserved exon/intron boundaries that are found in *Adam9* genes of other vertebrate species (data not shown), and is hence the opossum orthologue of *Adam9*. The lack of introns and very high sequence homology to *Adam9a* suggest that *Adam9b* was likely retroposed recently from *Adam9a*. This is further supported by the presence of a polyA tail in the *Adam9b* gene (Fig 5). Because the marsupial genomes contain a high percentage of retroposed elements [52], we tested if other *sAdams* have also been retroposed. No potential retroposed copy of any other *sAdams* could be



**Fig 5. A recent duplication of *Adam9* through retroposition.** Locations of *Adams 9a* and *9b* in the opossum genome, as well as the gene structure of *Adam9b*, are shown. Structure of the *Adam9a* mRNA is also shown for comparison, with black vertical lines representing splice sites (not to scale). The *Adam9b* pseudogene contains a polyA sequence, and the only intron was identified as an opossum retroposon (Mdo3LTR1a/ERV1), which was likely inserted after the original retroposition event that generated *Adam9b*. This insertion caused a frameshift.

doi:10.1371/journal.pone.0136281.g005



**Fig 6. Tandem duplication of *Adam1* in most amniote species.**

doi:10.1371/journal.pone.0136281.g006

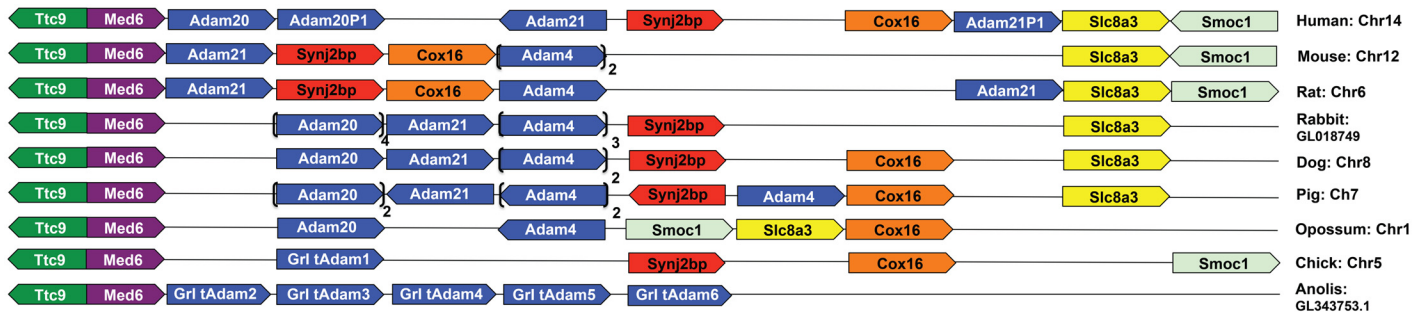
identified in the opossum genome, suggesting that *Adam9* may be a preferred target for retroposition. A recent study shows that the retroposon element L1 has preference for certain transcripts, presumably those that are expressed in the germ cells, where L1 activity is high [53]. In line with this finding, transcripts of *Adam9* are expressed in the testis of *Xenopus*, mouse and human ([51] and S2 Table).

### Further expansion of group I *tAdam* genes through tandem duplication

Clusters of group I *tAdam* genes were found in various amniote species, suggesting tandem duplication after retroposition. One example is the duplication of the *Adam1* locus (Fig 6). Although gene organization in the surrounding loci is conserved from frogs to mammals, no *Adam* gene was found in the *X. tropicalis* genome in this genomic region. By contrast, there is one copy of *Adam1* in the *Anolis* genome and two tandem copies in the chicken, opossum, mouse and human genomes (Fig 6). However, no strict orthology could be established between mammalian and chicken *Adam1* genes. One possible explanation for these observations is that *Adam1* originated before mammals diverged from reptiles and aves, and was duplicated separately in mammals and aves. Alternatively, *Adam1* may have been duplicated before the divergence of mammals, but one copy was lost in certain species such as the *Anolis*. In line with the latter model, two tandem copies of *Adam1* were also identified in the Chinese softshell turtle (Fig 6).

Tandem duplication was also found in the genomic region that includes *Adams 4, 20* and *21*. The total number of *Adam4/20/21* loci in this region varies from species to species among amniotes, ranging from one in the chick to eight in the rabbit (Fig 7). Accompanying gene duplication, occasional gene loss also occurred. For example, *Adam20* seems to be absent in rodents but not in most other eutherians (Fig 7).

The testases are a clade of proteolytically active group I tADAMs [16,17]. Nine testase genes have been cloned or predicted in the mouse genome, but it is not clear if orthologues of these genes exist elsewhere. We were unable to find any testase genes in any vertebrate species other than euarchontoglires; therefore the testases are probably specific for these mammals. On mouse chromosome 8, all nine testase genes cluster into two sub-regions that are separated by ~10 genes. Similar organizations were found on corresponding chromosomes of rat and rabbit, two other glires species, except that there are fewer testase genes than in the mouse genome (Fig 8). In the human genome, this region seems to have undergone chromosomal reorganization,



**Fig 7. Tandem duplication of *Adams* 4, 20 and 21 in mammals.** Multiple copies of *Adams* 4 and 20 are shown in brackets with copy numbers indicated.

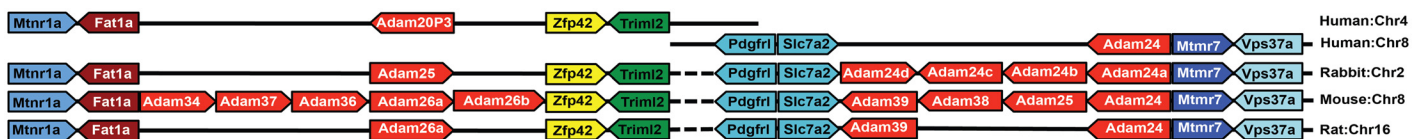
doi:10.1371/journal.pone.0136281.g007

and the two sub-regions were found in two separate chromosomes. There is one *tAdam* locus in each sub-region, but both loci appear to have been pseudogenized (Fig 8). Therefore, these two loci likely existed in early euarchontoglires, but the large-scale duplication of testase genes that occurred in the mouse, rat and rabbit likely happened after the divergence of glires.

### The fate of *tAdam* genes after duplication

A common consequence of gene duplication is the loss of function of duplicated genes. The HUGO Gene Nomenclature Committee website listed 11 human *Adam* pseudogenes, all of which are *tAdams*, including both group I and group II genes [13]. The pseudogenization of these *tAdams*, including group I *tAdams* that likely derived from retroposition, is mainly caused by premature stop codons rather than lack of transcription (data not shown). Retroposed genes were thought to be non-functional because they were not expected to carry any regulatory elements required for transcription [54]. However, it has been shown that many retroposed genes in the human genome are actively transcribed, probably by localizing closely to other genes and utilizing the regulatory elements and/or open chromatin structures of these neighboring genes [55]. Direct inheritance of basic promoters from parental genes has also been reported for retroposed genes [56]. Interestingly, a large fraction of retroposed genes, along with some other duplicated genes, is expressed initially in the testis [54,55]. This has been attributed to the hypertranscriptional state of the meiotic and post-meiotic spermatogenic cells, which facilitates transcription of otherwise silenced genes, such as those derived from retroposition or other means of gene duplication [57,58]. Therefore, it is not surprising that the group I *tAdam* genes, which likely derived from retroposition, are predominantly expressed in the testis.

As discussed above, recently duplicated genes are often expressed in the testis. These actively expressed genes may acquire new functions, and will be maintained if the new functions provide advantages during selection. These genes may also expand their expression/functions to



**Fig 8. Localization of testase genes in euarchontoglire genomes.** ADAMs 24, 25, 26a, 26b, 34, and 36–39 are testases 1, 2, 3a, 3b, 4, and 6–9, respectively. Both human *Adam20P3* and *Adam24* are pseudogenes, and *Adam20P3* is a close paralogue of *Adam20*.

doi:10.1371/journal.pone.0136281.g008

other tissues [54]. Consistent with this theory, the mouse *Adam2* promoter was reported to bind Mrf/Gm98, a transcriptional activator in the central nervous system [59], and neuroblasts from ADAM2 knockout mice show defects in their migration toward the olfactory bulb [60]. A search against the UniGene database revealed that other group II *tAdams* are also expressed in some somatic tissues in mice and humans (S2 Table). Therefore, the expression and functions of *tAdam* genes are not limited to the testis.

## Conclusions

In this study we show that certain *Adam* genes were frequently lost or duplicated during vertebrate evolution. We also report here the existence of homologues of mouse and human *tAdams* in various non-eutherian amniotes. Our phylogenetic and syntenic analyses suggest that *tAdams* form a subfamily with *Adam9* as well as an *Adam9-like* gene, which seems to be present in non-mammalian vertebrates only. Group II *tAdams* likely originated from an ancient *Adam9* gene through tandem duplication, whereas group I *tAdams* seem to have derived initially from retroposition and also undergone extensive tandem duplication in amniotes. One striking example is the clustering of the whole testase gene clade, which contains 9 group I *tAdams* in mice, in a single chromosomal region. Most of the *tAdam* loci maintain the distinct features of duplicated genes, such as tandem localization and lack of introns, providing excellent examples for studying gene duplication. It would be of interest in the future to determine if the non-eutherian *tAdam* genes that we identified in this study are indeed expressed predominantly in the testis. However, some of these genes may have expanded their expression to other tissues and acquired new functions.

## Supporting Information

**S1 Fig. Bayesian trees of ADAMs from *Xenopus*, opossum and mouse.**

(DOCX)

**S2 Fig. Neighbor joining trees of ADAMs from various vertebrate species.**

(DOCX)

**S3 Fig. Maximum likelihood trees of ADAMs from various vertebrate species.**

(DOCX)

**S1 Table. Accession numbers of ADAM proteins used in phylogenetic analyses.**

(XLSX)

**S2 Table. Expression patterns of *Adam9* and Group II *tAdams* in mice and humans.**

(DOCX)

## Acknowledgments

We thank Brittany Ott for help with generating Bayesian trees, Dr. Jennifer Hawkins for inspiring discussions, and Dr. Laura Christian for insightful comments on the manuscript.

## Author Contributions

Conceived and designed the experiments: SW. Performed the experiments: HB SB. Analyzed the data: SW HB SB. Contributed reagents/materials/analysis tools: SW HB SB. Wrote the paper: SW HB SB.

## References

1. Primakoff P, Hyatt H, Tredick-Kline J. Identification and purification of a sperm surface protein with a potential role in sperm-egg membrane fusion. *J Cell Biol.* 1987; 104: 141–149. PMID: [3793758](#)
2. Blobel CP, Wolfsberg TG, Turck CW, Myles DG, Primakoff P, White JM. A potential fusion peptide and an integrin ligand domain in a protein active in sperm-egg fusion. *Nature.* 1992; 356: 248–252. PMID: [1552944](#)
3. Wolfsberg TG, Bazan JF, Blobel CP, Myles DG, Primakoff P, White JM. The precursor region of a protein active in sperm-egg fusion contains a metalloprotease and a disintegrin domain: structural, functional, and evolutionary implications. *Proc Natl Acad Sci U S A.* 1993; 90: 10783–10787. PMID: [8248170](#)
4. Wolfsberg TG, Straight PD, Gerena RL, Huovila AP, Primakoff P, Myles DG, et al. ADAM, a widely distributed and developmentally regulated gene family encoding membrane proteins with a disintegrin and metalloprotease domain. *Dev Biol.* 1995; 169: 378–383. PMID: [7750654](#)
5. Yagami-Hiromasa T, Sato T, Kurisaki T, Kamijo K, Nabeshima Y, Fujisawa-Sehara A. A metalloprotease-disintegrin participating in myoblast fusion. *Nature.* 1995; 377: 652–656. PMID: [7566181](#)
6. Rooke J, Pan D, Xu T, Rubin GM. KUZ, a conserved metalloprotease-disintegrin protein with two roles in *Drosophila* neurogenesis. *Science.* 1996; 273: 1227–1231. PMID: [8703057](#)
7. Edwards DR, Handsley MM, Pennington CJ. The ADAM metalloproteinases. *Mol Aspects Med.* 2008; 29: 258–289. doi: [10.1016/j.mam.2008.08.001](#) PMID: [18762209](#)
8. Wei S. ADAM metalloproteinases. In: Rawlings ND, Guy S, editors. *The Handbook of Proteolytic Enzymes*: Academic Press; 2013. p. 1086–1094.
9. Christian L, Bahudhanapati H, Wei S. Extracellular metalloproteinases in neural crest development and craniofacial morphogenesis. *Crit Rev Biochem Mol Biol.* 2013; 48: 544–560. doi: [10.3109/10409238.2013.838203](#) PMID: [24066766](#)
10. Kurisaki T, Masuda A, Sudo K, Sakagami J, Higashiyama S, Matsuda Y, et al. Phenotypic analysis of Meltrin alpha (ADAM12)-deficient mice: involvement of Meltrin alpha in adipogenesis and myogenesis. *Mol Cell Biol.* 2003; 23: 55–61. PMID: [12482960](#)
11. Cousin H, Abbruzzese G, Kerdavid E, Gaultier A, Alfandari D. Translocation of the cytoplasmic domain of ADAM13 to the nucleus is essential for Calpain8-a expression and cranial neural crest cell migration. *Dev Cell.* 2011; 20: 256–263. doi: [10.1016/j.devcel.2010.12.009](#) PMID: [21316592](#)
12. Cho C. Testicular and epididymal ADAMs: expression and function during fertilization. *Nat Rev Urol.* 2012; 9: 550–560. doi: [10.1038/nrurol.2012.167](#) PMID: [22926424](#)
13. HUGO Gene Nomenclature Committee. Available: <http://www.genenames.org/genefamily/adam.php>.
14. Lum L, Blobel CP. Evidence for distinct serine protease activities with a potential role in processing the sperm protein fertilin. *Dev Biol.* 1997; 191: 131–145. PMID: [9356177](#)
15. Han C, Choi E, Park I, Lee B, Jin S, Kim do H, et al. Comprehensive analysis of reproductive ADAMs: relationship of ADAM4 and ADAM6 with an ADAM complex required for fertilization in mice. *Biol Reprod.* 2009; 80: 1001–1008. doi: [10.1095/biolreprod.108.073700](#) PMID: [19129510](#)
16. Zhu GZ, Myles DG, Primakoff P. Testase 1 (ADAM 24) a plasma membrane-anchored sperm protease implicated in sperm function during epididymal maturation or fertilization. *J Cell Sci.* 2001; 114: 1787–1794. PMID: [11309208](#)
17. Puente XS, Sanchez LM, Overall CM, Lopez-Otin C. Human and mouse proteases: a comparative genomic approach. *Nat Rev Genet.* 2003; 4: 544–558. PMID: [12838346](#)
18. Cho C, Bunch DO, Faure JE, Goulding EH, Eddy EM, Primakoff P, et al. Fertilization defects in sperm from mice lacking fertilin beta. *Science.* 1998; 281: 1857–1859. PMID: [9743500](#)
19. Shamsadin R, Adham IM, Nayernia K, Heinlein UA, Oberwinkler H, Engel W. Male mice deficient for germ-cell cyritestin are infertile. *Biol Reprod.* 1999; 61: 1445–1451. PMID: [10569988](#)
20. Yamaguchi R, Muro Y, Isotani A, Tokuhiko K, Takumi K, Adham I, et al. Disruption of ADAM3 impairs the migration of sperm into oviduct in mouse. *Biol Reprod.* 2009; 81: 142–146. doi: [10.1095/biolreprod.108.074021](#) PMID: [19339711](#)
21. Nishimura H, Kim E, Nakanishi T, Baba T. Possible function of the ADAM1a/ADAM2 Fertilin complex in the appearance of ADAM3 on the sperm surface. *J Biol Chem.* 2004; 279: 34957–34962. PMID: [15194697](#)
22. Zhu GZ, Gupta S, Myles DG, Primakoff P. Testase 1 (ADAM 24) a sperm surface metalloprotease is required for normal fertility in mice. *Mol Reprod Dev.* 2009; 76: 1106–1114. doi: [10.1002/mrd.21076](#) PMID: [19670298](#)
23. Jury JA, Frayne J, Hall L. The human fertilin alpha gene is non-functional: implications for its proposed role in fertilization. *Biochem J.* 1997; 321 (Pt 3): 577–581. PMID: [9032439](#)

24. Grzmil P, Kim Y, Shamsadin R, Neesen J, Adham IM, Heinlein UA, et al. Human cyritestin genes (CYRN1 and CYRN2) are non-functional. *Biochem J*. 2001; 357: 551–556. PMID: [11439107](#)
25. BLAT. Available: [https://genome.ucsc.edu/cgi-bin/hgBlat?command = start](https://genome.ucsc.edu/cgi-bin/hgBlat?command=start).
26. Kohany O, Gentles AJ, Hankus L, Jurka J. Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. *BMC Bioinformatics*. 2006; 7: 474. PMID: [17064419](#)
27. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. 2004; 32: 1792–1797. PMID: [15034147](#)
28. Ronquist F, Huelsenbeck JP. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*. 2003; 19: 1572–1574. PMID: [12912839](#)
29. FigTree. Available: <http://tree.bio.ed.ac.uk/software/figtree/>.
30. Tamura K, Stecher G, Peterson D, Filipksi A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol*. 2013; 30: 2725–2729. doi: [10.1093/molbev/mst197](#) PMID: [24132122](#)
31. Zuckerkandl E, Pauling L. Evolutionary divergence and convergence in proteins. In: Bryson V, Vogel HJ, editors. *Evolving Genes and Proteins*: Academic Press; 1965. p. 97–166.
32. Jones DT, Taylor WR, Thornton JM. The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci*. 1992; 8: 275–282. PMID: [1633570](#)
33. Louis A, Muffato M, Roest Crollius H. Genomicus: five genome browsers for comparative genomics in eukaryota. *Nucleic Acids Res*. 2013; 41: D700–705. doi: [10.1093/nar/gks1156](#) PMID: [23193262](#)
34. Metazome. Available: <http://www.metazome.net>.
35. Huxley-Jones J, Clarke TK, Beck C, Toubaris G, Robertson DL, Boot-Handford RP. The evolution of the vertebrate metzincins; insights from *Ciona intestinalis* and *Danio rerio*. *BMC Evol Biol*. 2007; 7: 63. PMID: [17439641](#)
36. Wei S, Whittaker CA, Xu G, Bridges LC, Shah A, White JM, et al. Conservation and divergence of ADAM family proteins in the *Xenopus* genome. *BMC Evol Biol*. 2010; 10: 211. doi: [10.1186/1471-2148-10-211](#) PMID: [20630080](#)
37. Sapir A, Assa-Kunik E, Tsruya R, Schejter E, Shilo BZ. Unidirectional Notch signaling depends on continuous cleavage of Delta. *Development*. 2005; 132: 123–132. PMID: [15576412](#)
38. Angerer L, Hussain S, Wei Z, Livingston BT. Sea urchin metalloproteases: a genomic survey of the BMP-1/tolloid-like, MMP and ADAM families. *Dev Biol*. 2006; 300: 267–281. PMID: [17059814](#)
39. Amemiya CT, Alfoldi J, Lee AP, Fan S, Philippe H, Maccallum I, et al. The African coelacanth genome provides insights into tetrapod evolution. *Nature*. 2013; 496: 311–316. doi: [10.1038/nature12027](#) PMID: [23598338](#)
40. Jaillon O, Aury JM, Brunet F, Petit JL, Stange-Thomann N, Mauceli E, et al. Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. *Nature*. 2004; 431: 946–957. PMID: [15496914](#)
41. Volf JN. Genome evolution and biodiversity in teleost fish. *Heredity (Edinb)*. 2005; 94: 280–294.
42. Amores A, Catchen J, Ferrara A, Fontenot Q, Postlethwait JH. Genome evolution and meiotic maps by massively parallel DNA sequencing: spotted gar, an outgroup for the teleost genome duplication. *Genetics*. 2011; 188: 799–808. doi: [10.1534/genetics.111.127324](#) PMID: [21828280](#)
43. Maretzky T, Yang G, Ouerfelli O, Overall CM, Worpenberg S, Hassiepen U, et al. Characterization of the catalytic activity of the membrane-anchored metalloproteinase ADAM15 in cell-based assays. *Biochem J*. 2009; 420: 105–113. doi: [10.1042/BJ20082127](#) PMID: [19207106](#)
44. Najy AJ, Day KC, Day ML. The ectodomain shedding of E-cadherin by ADAM15 supports ErbB receptor activation. *J Biol Chem*. 2008; 283: 18393–18401. doi: [10.1074/jbc.M801329200](#) PMID: [18434311](#)
45. Meslin C, Mugnier S, Callebaut I, Laurin M, Pascal G, Poupon A, et al. Evolution of genes involved in gamete interaction: evidence for positive selection, duplications and losses in vertebrates. *PLoS One*. 2012; 7: e44548. doi: [10.1371/journal.pone.0044548](#) PMID: [22957080](#)
46. Bates EE, Fridman WH, Mueller CG. The ADAMDEC1 (decysin) gene structure: evolution by duplication in a metalloprotease gene cluster on chromosome 8p12. *Immunogenetics*. 2002; 54: 96–105. PMID: [12037602](#)
47. Wei S, Xu G, Bridges LC, Williams P, White JM, DeSimone DW. ADAM13 induces cranial neural crest by cleaving class B Ephrins and regulating Wnt signaling. *Dev Cell*. 2010; 19: 345–352. doi: [10.1016/j.devcel.2010.07.012](#) PMID: [20708595](#)
48. Wei S, Xu G, Bridges LC, Williams P, Nakayama T, Shah A, et al. Roles of ADAM13-regulated Wnt activity in early *Xenopus* eye development. *Dev Biol*. 2012; 363: 147–154. doi: [10.1016/j.ydbio.2011.12.031](#) PMID: [22227340](#)

49. Peduto L. ADAM9 as a potential target molecule in cancer. *Curr Pharm Des.* 2009; 15: 2282–2287. PMID: [19601830](#)
50. Parry DA, Toomes C, Bida L, Danciger M, Towns KV, McKibbin M, et al. Loss of the metalloprotease ADAM9 leads to cone-rod dystrophy in humans and retinal degeneration in mice. *Am J Hum Genet.* 2009; 84: 683–691. doi: [10.1016/j.ajhg.2009.04.005](#) PMID: [19409519](#)
51. Shilling FM, Kratzschmar J, Cai H, Weskamp G, Gayko U, Leibow J, et al. Identification of metalloprotease/disintegrins in *Xenopus laevis* testis with a potential role in fertilization. *Dev Biol.* 1997; 186: 155–164. PMID: [9205136](#)
52. Mikkelsen TS, Wakefield MJ, Aken B, Amemiya CT, Chang JL, Duke S, et al. Genome of the marsupial *Monodelphis domestica* reveals innovation in non-coding sequences. *Nature.* 2007; 447: 167–177. PMID: [17495919](#)
53. Mandal PK, Ewing AD, Hancks DC, Kazazian HH Jr. Enrichment of processed pseudogene transcripts in L1-ribonucleoprotein particles. *Hum Mol Genet.* 2013; 22: 3730–3748. doi: [10.1093/hmg/ddt225](#) PMID: [23696454](#)
54. Kaessmann H, Vinckenbosch N, Long M. RNA-based gene duplication: mechanistic and evolutionary insights. *Nat Rev Genet.* 2009; 10: 19–31. doi: [10.1038/nrg2487](#) PMID: [19030023](#)
55. Vinckenbosch N, Dupanloup I, Kaessmann H. Evolutionary fate of retroposed gene copies in the human genome. *Proc Natl Acad Sci U S A.* 2006; 103: 3220–3225. PMID: [16492757](#)
56. Okamura K, Nakai K. Retrotransposition as a source of new promoters. *Mol Biol Evol.* 2008; 25: 1231–1238. doi: [10.1093/molbev/msn071](#) PMID: [18367464](#)
57. Marques AC, Dupanloup I, Vinckenbosch N, Reymond A, Kaessmann H. Emergence of young human genes after a burst of retroposition in primates. *PLoS Biol.* 2005; 3: e357. PMID: [16201836](#)
58. She X, Horvath JE, Jiang Z, Liu G, Furey TS, Christ L, et al. The structure and evolution of centromeric transition regions within the human genome. *Nature.* 2004; 430: 857–864. PMID: [15318213](#)
59. Choi H, Lee B, Jin S, Kwon JT, Kim J, Jeong J, et al. Identification and characterization of promoter and regulatory regions for mouse *Adam2* gene expression. *Mol Biol Rep.* 2013; 40: 787–796. doi: [10.1007/s11033-012-2116-8](#) PMID: [23065232](#)
60. Murase S, Cho C, White JM, Horwitz AF. ADAM2 promotes migration of neuroblasts in the rostral migratory stream to the olfactory bulb. *Eur J Neurosci.* 2008; 27: 1585–1595. doi: [10.1111/j.1460-9568.2008.06119.x](#) PMID: [18380661](#)