Identification of Novel Mammalian Caspases Reveals an Important Role of Gene Loss in Shaping the Human Caspase Repertoire

Leopold Eckhart,* Claudia Ballaun,* Marcela Hermann,† John L. VandeBerg,‡ Wolfgang Sipos,§ Aumaid Uthman,*|| Heinz Fischer,* and Erwin Tschachler*

*Department of Dermatology, Medical University of Vienna, Vienna, Austria; †Max F. Perutz Laboratories, Medical University of Vienna, Vienna, Austria; ‡Department of Genetics, Southwest Foundation for Biomedical Research, San Antonio, TX; §Department for Farm Animals and Herd Management, University of Veterinary Medicine Vienna, Vienna, Austria; and ||College of Veterinary Medicine, University of Sulaimani, Sulaimani, Iraq

Proteases of the caspase family play central roles in apoptosis and inflammation. Recently, we have described a new gene encoding caspase-15 that has been inactivated independently in different mammalian lineages. To determine the dynamics of gene duplication and loss in the entire caspase gene family, we performed a comprehensive evolutionary analysis of mammalian caspases. By comparative genomics and reverse transcriptase-polymerase chain reaction analyses, we identified 3 novel mammalian caspase genes, which we tentatively named caspases-16 through -18. Caspase-16, which is most similar in sequence to caspase-14, has been conserved in marsupials and placental mammals, including humans. Caspase-17, which is most similar to caspase-3, has been conserved among fish, frog, chicken, lizard, and the platypus but is absent from marsupials and placental mammals. Caspase-18, which is most similar to caspase-8, has been conserved among chicken, platypus, and opossum but is absent from placental mammals. These gene distribution patterns suggest that, in the evolutionary lineage leading to humans, caspase-17 was lost after the split of protherian and therian mammals and caspase-18 was lost after the split of marsupials and placental mammals. In the canine genome, the number of caspases has been reduced by the fusion of the neighboring genes caspases-1 and -4, resulting in a single coding region. Further lineage-specific gene inactivations were found for caspase-10 in murine rodents and caspase-12 in humans, rabbit, and cow. Lineage-specific gene duplications were found for caspases-1, -3, and -12 in opossum and caspase-4 in primates. Other caspases were generally conserved in all mammalian species investigated. Using the positions of introns as stable characters during recent vertebrate evolution, we define 3 phylogenetic clades of caspase genes: caspases-1/-2/4/5/9/12/14/15/16 (clade I), caspases-3/6/7/17 (clade II), and caspases-8/10/18/CFLAR (clade III). We conclude that gene inactivations have occurred in each of the 3 caspase clades and that gene loss has been as critical as gene duplication in the evolution of the human repertoire of caspases.

Introduction

Caspases are cysteine-dependent aspartate-directed proteases that have central roles in apoptosis, proinflammatory signaling, immune cell proliferation, epidermal barrier formation, and various other physiological processes (Degterev et al. 2003; Denecker et al. 2007; Lamkanfi et al. 2007). Because of their critical functions and their potential as therapeutic targets (Reed 2002; Hotchkiss and Nicholson 2006), caspases have attracted much interest and have been well characterized with respect to gene expression, structure, activation, and catalytic activity (Earnshaw et al. 1999; Degterev et al. 2003).

All caspases contain a peptidase C14 domain (caspase domain, pfam00656), which is approximately 230 amino acids long and comprises both alpha-helical segments and a central beta sheet that facilitates homodimerization (Fuentes-Prior and Salvesen 2004). Activation of pro-caspase molecules involves proteolytic processing, thereby generating 2 subunits from each caspase domain and removing the N-terminal prodomain. In humans, 14 proteins with a peptidase C14 domain have been described: caspases-1 (originally named interleukin-1 beta-converting enzyme), -2, -3, -4, -5, -6, -7, -8, -9, -10, -12, and -14; CFLAR (CASP8 and FADD-like apoptosis regulator, also known as c-FLIP); and the product of the mucosa-associated lymphoid tissue lymphoma translocation gene 1. The latter is not a true caspase but a paracaspase that deviates significantly from the canonical caspase structure (Uren et al. 2000). Proteolytic activity has been reported for all caspases except for caspase-12, which is inactivated by deleterious mutations in the majority of the human population (Fischer et al. 2002; Wang et al. 2006; Xue et al. 2006) and functions as an inhibitor of inflammation in the mouse (Saleh et al. 2006). Likewise, CFLAR inhibits caspase-8–dependent apoptosis (Irmler et al. 1997). Inconsistent with the rules of caspase nomenclature (Alnemri et al. 1996), caspases-11 and -13 represent the murine and bovine orthologs of human caspase-4, respectively (Koenig et al. 2001). We have recently cloned and characterized a new caspase, caspase-15, in a number of mammalian species (Eckhart et al. 2005; Eckhart et al. 2006).

Mammalian caspases have been grouped according to different criteria. Based on their functions, caspases-2, -8, -9, and -10 are apoptosis initiators, caspases-3, -6, and -7 are apoptosis executors, caspases-1, -4, and -5 are involved in inflammation, and caspase-14 has a role in terminal differentiation of epidermal keratinocytes (Fuentes-Prior and Salvesen 2004). Based on the length and structural organization of the N-terminal prodomain, caspases-3, -6, -7, and -14 are short-prodomain proteases, whereas all others have a long prodomain folding into a caspase activation and recruitment domain (CARD) (caspases-1, -2, -4, -5, -9, and -12), a tandem of 2 death effector domains (DEDs) (caspases-8 and -10 and CFLAR), or a pyrin domain (caspase-15). On the basis of different optimal cleavage substrate motifs, 3 caspase groups have been proposed: group I (caspases-1, -4, and -5), II (caspases-2, -3, and -7), and III (caspases-6, -8, and -9) (Thornberry et al. 1997), whereas a phylogenetic analysis based on amino acid

Key words: caspase, gene loss, gene duplication, comparative genomics, mammals.

E-mail: leopold.eckhart@meduniwien.ac.at.

doi:10.1093/molbev/msn012
Advance Access publication February 14, 2008

© The Author 2008. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution. All rights reserved.
For permissions, please e-mail: journals.permissions@oxfordjournals.org
sequence comparison suggested the existence of 3 “clusters”: cluster I (caspases-1, -2, -4, -5, -12, and -14), II (caspases-3, -6, and -7), and III (caspases-8, -9, and -10) (Lamkanfi et al. 2002).

Comparative genomics has revealed that genes similar to caspases are present in alpha-proteobacteria, which are relatives of the free-living ancestors of mitochondria (Koonin and Aravind 2002). This indicates that the ancestral caspase genes of euukaryotes have been derived from the mitochondrial endosymbiont. “Protocaspase” genes evolved into metacaspases in plants and fungi, paracaspases in slime molds and animals, and genuine caspases in animals. The number of caspase genes per genome differs among animal lineages with 4, 5, 7, and 31 genes encoding caspase domains being present in the nematode Caenorhabditis elegans, the cnidarian Nematostella vectensis, the fruit fly Drosophila melanogaster, and sea urchin Strongylocentrotus purpuratus (Robertson et al. 2006), respectively. As described above, mammals contain more than 10 caspase genes, most of which were cloned in the late 1990s. However, our recent identification of caspase-15, a novel ancient mammalian caspase that was lost in primates and other evolutionary lineages (Eckhart et al. 2006), indicated that the complexity of caspase evolution in mammals had previously been underestimated.

Here, we extended our comparative and evolutionary analysis to all mammalian caspase genes and to a large number of sequenced genomes. We report on the evolutionary history of novel mammalian caspases and propose a new classification of caspases into 3 phylogenetic clades. Our study demonstrates that evolutionary dynamics of the caspase gene content was strongly influenced by gene loss in mammals.

Materials and Methods

Sequence Queries, Alignments, and Phylogenetic Analyses

Blast searches for caspase homologs were performed using query sequences such as the human caspases, CFLAR, human caspase pseudogenes (Rocher et al. 1997; Centola et al. 1998), and caspases predicted by automated algorithms of the ENSEMBL and GenBank databases in the genome sequences of amniotes. Sequences were retrieved from the GenBank database and the ENSEMBL genome browser (http://www.ncbi.nlm.nih.gov/BLAST/; http://www.ensembl.org/Multi/blastview). Additional nucleotide sequence alignments were made with the University of California Santa Cruz genome browser (http://www.genome.ucsc.edu/) and with the program LALIGN (http://www.ch.embnet.org/software/LALIGN_form.html).

Gene orthology was evaluated according to the following criteria: 1) Reciprocal best hits (also known as symmetrical best hits) in Blast searches (Koonin 2005); 2) Features of gene organization that are stable in recent vertebrate evolution, namely the localization and phase of introns (Roy et al. 2003); 3) Syntenic position in the genome, that is, similarity in the arrangement of neighboring genes (Zheng et al. 2005).

Phylogenetic analyses were performed at the Phylomon Web server (Tárraga et al. 2007). Amino acid sequences of caspase catalytic domains were aligned using ClustalW. Phylogenetic trees were built with PHYLIP using the Jones–Taylor–Thornton matrix as distance algorithm and the Neighbor-Joining method for clustering. Trees were displayed using the TreeView program (Page 1996).

DNA Preparation, Polymerase Chain Reactions and Sequence Analysis

Genomic DNA was prepared from animal tissues according to a standard protocol (Ausubel 1998). Tissue RNAs were extracted with the Trizol reagent (Invitrogen, Vienna, Austria) according to the manufacturer’s instructions and reverse transcribed according to standard protocols (Eckhart et al. 1999). Polymerase chain reactions (PCRs) were performed according to a published protocol with a primer annealing temperature of 55 °C (Eckhart et al. 1999). PCR products were either purified and sequenced or cloned into the pCRII Topo vector (Invitrogen) according to the manufacturer’s instructions. Primer sequences are shown in supplementary table S1 (Supplementary Material online).

Results

Isolation of Caspase Genes from Databases and Nomenclature of Novel Mammalian Caspase Genes

Genome sequences of diverse mammals, mostly used as model species in biomedical research, as well as chicken and lizard (table 1) were searched for caspase homologs. In the course of these screenings, we identified novel caspase genes for which we propose the following taxonomy. 1) Caspases that evolved by gene duplication in evolutionary lineages not leading to humans are termed caspase-x-like y, where x is the number of the orthologous human caspase and y is a consecutive number ranging from 1 to the maximum number of homologs arisen by duplication (fig. 1A). 2) Caspases that are either still present in the human genome or which have been present in the genome of a mammalian ancestor of humans are assigned a new number because they are paralogous to all (other) human caspases (fig. 1B). We consider the criterion of presence in a mammalian ancestor of humans met if an ortholog of such a caspase is present in a mammalian sister group of humans and in a (mammalian or nonmammalian) sister group of the common ancestor of humans and the first, evolutionarily more close sister group (fig. 1). An example for this taxonomy is caspase-15, which is absent in humans but whose former presence in an ancestor of humans can be inferred from the detection of caspase-15 in pig (representing sister group 1) and in marsupial opossum (representing sister group 2) (Eckhart et al. 2006). The proposed caspase taxonomy extends the original guidelines of caspase nomenclature (Alnemri et al. 1996) by considering nonhuman caspases while ensuring phylogenetic relevance for human physiology.
Identification of Caspase-16, an Evolutionarily Conserved Mammalian Caspase Gene Most Similar to Caspase-14

Comparative genomics of all caspase-like genes in the human genome showed that a transcribed gene previously considered to be a pseudogene (Centola et al. 1998) was conserved among mammals. The chromosomal locus of this gene was syntenic in human (chromosome 16p13.3), opossum (fig. 2A), and other mammals (data not shown). Because the gene encoded a caspase domain and the cysteine and histidine residues forming the catalytic center of caspasas were conserved (fig. 2B), it appeared to represent a novel mammalian member of the caspase family. In accordance with the proposed guidelines of caspase nomenclature, we refer to the novel gene as caspase-16.

Caspase-16 orthologs were identified in all mammalian genome sequences (including platypus and opossum) available in September 2007, whereas no orthologs were found in nonmammalian genomes including the closest outgroup to mammals, that is, sauropsids, represented by chicken (Gallus gallus) and the lizard Anolis carolinensis. This suggested an evolutionary origin of caspase-16 after the divergence of the sauropsids (including reptiles and birds) and synapsids (including mammals) and prior to the split of protherians (monotremes) and therians (marsupials and placental mammals) approximately 220 MYA (van Rheede et al. 2006). Caspase-16 corresponds to a so-called caspase-14-like gene identified in the rat genome (Puente and Lopez-Otin 2004). Homology of human, chimpanzee, dog, mouse, and rat caspase-16 genes has also been detected by the automated system that feeds the National Center for Biotechnology Information HomoloGene database (Wheeler et al. 2007), and caspase-16 gene-related information is available at the HomoloGene database entry 72522 (http://www.ncbi.nlm.nih.gov/sites/entrez/query.fcgi?db=homologene). Analyses of the genome sequences listed in table 1 showed that the open reading frame of the caspase domain of caspase-16 was intact in all species except guinea pig, which contained in-frame stop codons. Caspase-16 contains a unique insertion of 25 amino acids at a position corresponding to the L3 loop of the caspase domain (Fuentes-Prior and Salvesen 2004). Excluding this segment from the alignment, the caspase-16 peptidase C14 domain displays the highest amino acid sequence similarity with caspase-14 (45% identity) and caspase-15 (29% identity). The prodomain of caspase-16 lacks significant similarity to other proteins. However, the structure of the N-terminus of the human caspase-16 protein requires further characterization because of differences in the corresponding region of the mRNA predicted by the GnomON algorithm of the GenBank and the caspase-16 cDNAs reported by Centola et al. (1998).

We analyzed the expression of caspase-16 by reverse transcriptase–polymerase chain reaction (RT-PCR) and nested PCR and found that caspase-16 was expressed in human spleen (supplementary fig. S1A, Supplementary Material online), confirming Northern blot data published by Centola et al. (1998). Unlike its closest genetic relative, that is, caspase-14 (Eckhart et al. 2000), caspase-16 was not expressed in differentiated epidermal keratinocytes (supplementary fig. S1A, Supplementary Material online). In a separate experiment, expression of caspase-16 was also detected in the liver of the opossum (GenBank accession number: EF397305) (supplementary fig. S1B, Supplementary Material online). Further studies of caspase-16 expression at the mRNA and protein levels as well as the investigation of its catalytic potential as a protease have been initiated in our laboratories.

Identification of Caspase-17, a Novel Caspase Similar to Caspase-3, in the Platypus and Nonmammalian Vertebrates

The gene sequence of the platypus, a protherian mammal, contained sequences homologous to 2 exons of an uncharacterized caspase gene present in the genome of chicken (supplementary fig. S2, Supplementary Material online). Expression of the novel gene, which we tentatively named caspase-17, was detected by RT-PCR in chicken oviduct and liver (supplementary fig. S2, Supplementary Material online).
whether the caspase-

highest number of a caspase identified previously. This is irrespective of

should be assigned a new name, that is, caspase-

n

gene has been duplicated in sister group 1 of

x

species (thick-lined frames) and gene (thin lines) tree shows an example

in which the caspase-x gene has been duplicated in sister group 1 of

humans. Both copies are orthologs of human caspase-x

.x

and placental mammals. Expressed sequence tags corre-

sponding to chicken caspase-18, but originally termed

initiator caspase-8/-10' (gi 67097571; submitted by

Janicke et al. 2006). Sur-

prisingly, a previously undescribed gene was found in both

opossum and chicken between the

CASP10 and

CASP8 genes (fig. 4A). This novel gene encoded a protein with

the same structure as that of caspase-8. It comprised 2

N-terminal DEDs and a caspase domain in which all amino

acid residues critical for catalytic activity were conserved

(fig. 4B). We have therefore tentatively termed the novel

gene caspase-18. In a phylogenetic analysis, chicken and

opossum caspase-18 genes clustered together to the exclu-

sion of CFLAR and caspases-8 and -10 of chicken, opos-

sum, and humans (fig. 4C). An uncharacterized cDNA corresponding to chicken caspase-18, but originally termed

initiator caspase-8/10" (gi 67097571; submitted by

Kazuhiro Sakamaki and Masami Nozaki, Kyoto University, Kyoto, Japan) was identified in the GenBank, which showed that this gene is expressed. We confirmed expres-

sion of caspase-18 in the liver of opossum and chicken, re-

spectively, by RT-PCR (GenBank accession numbers

EU165363–EU165365) (supplementary fig. S3, Supplementary Material online).

Identification of Caspase-18, a Caspase-8–Like Gene in

Chicken, Platypus, and Opossum

Caspases-8 and -10 and CFLAR are arranged in close

proximity at a chromosomal locus (human chromosome

2q33–q34), henceforth referred to as the caspase-8 subfam-

ily locus. Comparison of this locus in the genomes of hu-

man, mouse, opossum, and chicken showed that the outermost genes, CFLAR and CASP8, were conserved

and always represented the borders of this locus. In contrast to the other 3 species, the mouse lacked a caspase-10 gene, as has been reported previously (Janicke et al. 2006). Sur-

prisingly, a previously undescribed gene was found in both

opossum and chicken between the

CASP10 and

CASP8 genes (fig. 4A). This novel gene encoded a protein with

the same structure as that of caspase-8. It comprised 2

N-terminal DEDs and a caspase domain in which all amino

acid residues critical for catalytic activity were conserved

(fig. 4B). We have therefore tentatively termed the novel

gene caspase-18. In a phylogenetic analysis, chicken and

opossum caspase-18 genes clustered together to the exclu-

sion of CFLAR and caspases-8 and -10 of chicken, opos-

sum, and humans (fig. 4C). An uncharacterized cDNA corresponding to chicken caspase-18, but originally termed

initiator caspase-8/10" (gi 67097571; submitted by

Kazuhiro Sakamaki and Masami Nozaki, Kyoto University, Kyoto, Japan) was identified in the GenBank, which showed that this gene is expressed. We confirmed expres-

sion of caspase-18 in the liver of opossum and chicken, re-

spectively, by RT-PCR (GenBank accession numbers

EU165363–EU165365) (supplementary fig. S3, Supplementary Material online).
To determine the evolutionary history of caspase-18, we analyzed, in addition to the species mentioned above, Western clawed frog (X. tropicalis), platypus, armadillo, elephant, dog, and cow. A putative ortholog of caspase-18 was identified in the genome of the platypus but not in the genomes of the other species. This distribution suggested that caspase-18 originated, presumably by duplication of caspase-8 or -10, after the split of the amphibian lineage from amniotes, was further conserved in birds and protherian as well as metatherian mammals, but was deleted in a common ancestor of all placental (eutherian) mammals.

Exon–Intron Organization Defines 3 Phylogenetic Clades of Caspases

We reevaluated the phylogenetic relationships of mammalian caspases by defining phylogenetic clades of caspase genes using a character that is considered stable over the evolutionary period investigated, namely, the position of introns within the coding sequence (Roy et al. 2003). Amino acid sequence alignment and mapping of intron positions within the coding regions of mammalian caspases showed that there are 3 main, distinct patterns of exon–intron structure (fig. 5; supplementary fig. S4, Supplementary Material online). The genes encoding caspases-1, -2, -4, -5, -9, -12, -14, -15, and -16 (clade I) share 4 common intron positions within the region coding for the caspase domain, whereas caspases-3, -6, -7, and -17 (clade II) as well as caspases-8, -10, and -18 and CFLAR (clade III) share 3 common intron positions, respectively (fig. 5). One intron position is shared between clade I and II caspases. Caspases-2 and -9 share a homologous intron in the 5’ region of the caspase domain region and in the 3’ untranslated region. This commonality supports the notion that these caspases, which are also known as caspase-1 subfamily, form a subclade of clade I. Caspases-2 and -9 share a homologous intron in the 5’ region of the caspase domain region and in the 3’ untranslated region.
the region encoding the CARD motif of the prodomain (data not shown), supporting their close relationship, which is also suggested by common roles in the initiation phase of apoptosis.

The 3 phylogenetic clades of caspases, defined by shared intron positions within the caspase domain regions of the genes, have not been predicted by the phylogenetic analysis of amino acid sequences (Lamkanfi et al. 2002) but correlate well, though not perfectly, with the grouping of caspases according to their prodomain structure. All clade III caspases contain a tandem of 2 DEDs in their prodomain, and all clade II caspases have a short (caspases-3, -6, and -7) or no prodomain (caspase-17). The predominant prodomain motif among clade I caspases is CARD. Absence of a CARD in caspases-14, -15, and -16, which also cluster in phylogenetic analyses using the amino acid sequence of the caspase domain only, has probably been caused by loss of an ancestral CARD. It exceeded the scope of this study to assess which of the 3 caspase clades is phylogenetically closest to the primordial caspase of animals. However, a preliminary comparative genetics screening of more primitive metazoan species revealed that caspases with prodomains containing a CARD or 2 DEDs as well as caspases with short prodomains are encoded in the genome of the sea anemone *N. vectensis* (Eckhart L, unpublished data), which represents an evolutionary lineage that diverged from the bilaterian lineage approximately 600 MYA.

Evolutionary History of Caspase Genes in Mammals

Next we assessed the evolutionary history of each individual caspase in multiple mammalian lineages. Because of limitations in the availability of assembled genome sequences, which are required for an in-depth comparative gene analysis, and in order to facilitate follow-up studies in accessible animal models, we focused on a series of
mammalian model species (opossum, cow, dog, rabbit, guinea pig, rat, mouse, macaque, and chimpanzee). Blast searches in the respective databases and RT-PCR analyses on mRNAs from various animals were performed. Caspases-2, -6, -7, -8, -9, and -14 and CFLAR were conserved without deletion or duplication events in all mammalian species investigated, whereas interspecies differences were found for the other caspases. Besides the gene losses of caspases-17 and -18 described above, caspase-10 was lost in murine rodents but not in guinea pig and rabbit (GenBank accession number EF397306) and caspase-3 was duplicated in the opossum (supplementary fig. S5, Supplementary Material online).

The most striking interspecies differences were found at the locus of the so-called caspase-1 subfamily genes (caspases-1, -4, -5, and -12), also known as the subfamily of inflammatory caspases (Martinon and Tschopp 2004), which are clustered on human chromosome 11q23. Comparison of this chromosomal locus among mammals supported previous studies that have proposed a presumably prime-specific gene duplication leading to caspases-4 and -5 (Lamkanfi et al. 2002) and gene inactivation of caspase-12 in humans (Fischer et al. 2002). In addition, a gene duplication giving rise to a caspase-4-like transcribed pseudogene was identified in humans and other primates (fig. 6A). A phylogenetic tree of caspases-8 (C8), -10 (C10), and -18 (C18) and CFLAR of the species chicken, opossum, and human was built using the Neighbor-Joining method. The branch lengths in the tree are proportional to the number of substitutions per site (scale: 0.1 substitutions per site); chi, chicken; opo, opossum; and hum, human.
by the facultative inclusion of exon 4 (fig. 6B). The shorter variant was identical to a cDNA, which prior to our report that caspase-13 is the bovine ortholog of human caspase-4 (Koenig et al. 2001) had been named “hybrid caspase-1/caspase-13” (Taylor et al. 2000). From the presence of caspases-1 and -4 in humans, mouse, and cow, it can be inferred that both genes were present in the genome of the last common ancestor of Boreoeutheria (Murphy et al. 2006). However, the use of near-complete genome data from many species allows concomitant analyses of multiple genes that are likely to give phylogenetic results with high confidence.

The differences in caspase gene content in mammals may reflect adaptations to species-specific constraints or redundant functions of caspases. In the recent evolution of humans, caspase-12 (Fischer et al. 2002) has been under positive selection for mutual inactivation (Wang et al. 2006; Xue et al. 2006). Whether positive selection has also driven inactivation of other caspase genes, such as caspases-15, -17, and -18, or duplication of caspase genes, such as caspase-3 in the opossum and caspase-4 in primates, remains to be determined. In our view, any hypotheses regarding the evolutionary roles of these gene reorganization events require the understanding of the biological functions of the respective caspase in at least one species. All murine caspase genes except for caspase-16 have been investigated by targeted knockout. Mice deficient in caspases-1, -2, -6, -7, -11, -12, and -14 have relatively mild phenotypes under standard laboratory conditions, whereas inactivation of caspases-3, 8, or 9 is lethal (Kuida et al. 1996; Hakem et al. 1998; Varfolomeev et al. 1998). This indicates that most caspases have functions not critical for development and survival under nonstressed conditions.

### Discussion

The results of our comparative genomic analysis imply that the number of caspase gene deletions (caspases-17, -18, -15, and -12, ranked in the chronological order of inactivation) almost equaled the number of caspase gene innovations by duplication (caspases-12, -14, -16, -4, and -5) in the evolutionary lineage leading to humans. Furthermore, we found novel examples of caspase gene deletions that occurred in parallel in different lineages: loss of caspase-15 in chicken (this study) and some mammals (Eckhart et al. 2006) and deletions of caspase-12 in cow and rabbit (this study) and humans (Fischer et al. 2002) as well as caspase-18 in anole and the last common ancestor of placental mammals (this study). We conclude that gene deletion should not generally be considered a rare event during evolution (Rokas and Holland 2000), and great care should be taken when phylogenetic relationships among species are inferred from common presence or absence of individual genes. However, the use of near-complete genome data from many species allows concomitant analyses of multiple genes that may reflect adaptations to species-specific constraints or redundant functions of caspases.

### Fig. 5.—Exon–intron structure defines phylogenetic relationship of mammalian caspase genes. Schematic overview of the caspase domains of mammalian caspase genes. The position of introns is shown by triangles. Vertical lines indicate homologous intron positions. An intron, marked with an asterisk, of caspase-16 is interrupted by a gene-specific exon. Caspase-5 is not included in this scheme because it represents a primate-specific duplication product of the caspase-4 gene and because the exons encoding its caspase domain have the same organization as the caspase-4 exons. The exon–intron structure of caspase-4 is also identical to those of caspases-11 and -13, which are actually the murine and bovine orthologs of human caspase-4. An alignment of caspase sequences with exact positions and phases of exon borders is shown in the Supplementary Material online. CARD, caspase activation and recruitment domain; DED2, death effector domains in tandem arrangement; and PYD, pyrin domain.

<table>
<thead>
<tr>
<th>Casp-1</th>
<th>Casp-4</th>
<th>Casp-12</th>
<th>Casp-14</th>
<th>Casp-15</th>
<th>Casp-16</th>
<th>Casp-2</th>
<th>Casp-9</th>
<th>FLAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>CARD</td>
<td>CARD</td>
<td>CARD</td>
<td></td>
<td>PYD</td>
<td></td>
<td>CARD</td>
<td>CARD</td>
<td>DED2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### FIG. 5

- **Clade I**: Precursors of caspase-1 and caspase-4 domains are fused in bovine caspase-5.
- **Clade II**: Presence of caspase-15 and -12 domains in nonmammalian species.
- **Clade III**: Duplication of caspase-15, followed by gene copy–specific sequence modification, gave rise to caspase-14 and -16 in the evolutionary lineage leading to mammals.
but, most likely, are essential in distinct states of infection (Li et al. 1995) or exposure to noxious environmental challenges such as UV irradiation (Denecker et al. 2007). Because caspases-10, -15, -17, and -18 have been lost in the evolutionary lineage leading to the mouse, gene knockout or knockdown approaches must be established in alternative model species or cultured cells derived from species expressing these caspases.

Notably, we found the highest interspecies variability among members of the caspase-1 subfamily gene loci in mammals. Genes are represented by arrows. Note that cDNAs of the caspase-1 subfamily genes have not been cloned in the opossum and that orthology relationships of these genes should be considered preliminary. Asterisks indicate the presence of mutations that prevent the expression of a catalytically active caspase. (A) The caspases-1 and -4 genes have fused in the dog. Exons are represented by boxes in which the exon number is shown.

The results of the present study suggest that one of the central gene families involved in apoptosis, the caspases, has undergone extensive changes during mammalian evolution with approximately one-third of genes being deleted and a similar fraction of genes being duplicated. Our data show that the apoptotic machinery has not constantly increased in complexity during evolution in the vertebrate lineage as it was extrapolated from a comparative genomic analysis of model organisms such as nematode, fly, and mouse (Aravind et al. 2001) but, at least in mammals, has undergone both expansions and reductions in gene numbers during recent evolution.

**Supplementary Material**

Supplementary table S1 and figures S1–S6 are available at Molecular Biology and Evolution online (http://www.mbe.oxfordjournals.org/).

**Acknowledgments**

We thank Ulrich Koenig and Wolfgang J. Schneider for helpful discussions and Maria Buchberger, Sonja Gawlas, Christian Kittel (Medical University of Vienna, Austria), and Jeannie Chan (Southwest Foundation for
Biomedical Research, San Antonio, TX) for excellent technical support. The authors acknowledge all providers of public genome sequences, in particular the Broad Institute, Boston, MA, and the Genome Sequencing Center at the Washington University School of Medicine, St Louis, MO. We thank Heidemarie Rossiter for critically reading the manuscript.

**Literature Cited**


![Figure 7. Distribution of caspase genes in mammalian species.](https://academic.oup.com/mbe/article-abstract/25/5/831/1195498/FIG_7_Distribution_of_caspase_genes_in_mammalian_species)
Evolution of Mammalian Caspases  


Arndt von Haeseler, Associate Editor

Accepted January 8, 2008