# Evolution of Mammalian Chitinase(-Like) Members of Family 18 Glycosyl Hydrolases

Anton P. Bussink, Dave Speijer, Johannes M. F. G. Aerts and Rolf G. Boot<sup>1</sup>

Department of Medical Biochemistry, Academic Medical Center, University of Amsterdam, 1105 AZ Amsterdam, The Netherlands

> Manuscript received May 11, 2007 Accepted for publication August 2, 2007

### ABSTRACT

Family 18 of glycosyl hydrolases encompasses chitinases and so-called chi-lectins lacking enzymatic activity due to amino acid substitutions in their active site. Both types of proteins widely occur in mammals although these organisms lack endogenous chitin. Their physiological function(s) as well as evolutionary relationships are still largely enigmatic. An overview of all family members is presented and their relationships are described. Molecular phylogenetic analyses suggest that both active chitinases (chitotriosidase and AMCase) result from an early gene duplication event. Further duplication events, followed by mutations leading to loss of chitinase activity, allowed evolution of the chi-lectins. The homologous genes encoding chitinase(-like) proteins are clustered in two distinct loci that display a high degree of synteny among mammals. Despite the shared chromosomal location and high homology, individual genes have evolved independently. Orthologs are more closely related than paralogues, and calculated substitution rate ratios indicate that protein-coding sequences underwent purifying selection. Substantial gene specialization has occurred in time, allowing for tissue-specific expression of pH optimized chitinases and chilectins. Finally, several family 18 chitinase-like proteins are present only in certain lineages of mammals, exemplifying recent evolutionary events in the chitinase protein family.

CHITIN, the linear polymer of *N*-acetylglucosamine, is the second most abundant polysaccharide in nature and serves as an indispensable structural component in a variety of organisms, including fungi and arthropods (THARANATHAN and KITTUR 2003). On the basis of sequence homologies, chitinases fall into two groups: families 18 and 19 of glycosyl hydrolases (HENRISSAT 1991). Members of family 18 employ a substrate-assisted reaction mechanism (TERWISSCHA VAN SCHELTINGA *et al.* 1995; VANAALTEN *et al.* 2001), whereas those of family 19 adopt a fold-and-reaction mechanism similar to that of lysozyme (MONZINGO *et al.* 1996), suggesting that these families evolved independently to deal with chitin.

Early reports on chitinolytic activity in vertebrates (JEUNIAUX 1961) were confirmed following investigations on Gaucher disease, the most common lysosomal storage disorder in humans caused by an inherited deficiency in glucocerebrosidase (BEUTLER and GRABOWSKI 1995). In the plasma of symptomatic patients with Gaucher disease, activity toward the artificial substrate 4-methylumbelliferyl-chitotriose is elevated several hundredfold (HOLLAK *et al.* 1994). The responsible enzyme, named chitotriosidase, was shown to be a true chitinase, hydrolyzing natural chitin and showing high sequence

<sup>1</sup>Corresponding author: Department of Medical Biochemistry, Academic Medical Center, University of Amsterdam, Meibergdreef 15, 1105 AZ Amsterdam, The Netherlands. E-mail: r.g.boot@amc.uva.nl

homology to chitinases from lower organisms (HOLLAK et al. 1994; Воот et al. 1995; RENKEMA et al. 1995). Other members of the mammalian chitinase family have been discovered since, including a second chitinase which, given its acidic pH optimum, was named acidic mammalian chitinase (AMCase) (BOOT et al. 2001). Since chitin is an important structural component of pathogens like fungi as well as a constituent of the mammalian diet, a dual function for mammalian chitinases in innate immunity and food digestion has been envisioned (SUZUKI et al. 2002; BOOT et al. 2005a). Indeed, for human chitotriosidase, an enzyme predominantly expressed by phagocytes, a fungistatic effect has been demonstrated recently (VAN EIJK et al. 2005). Several studies have tried to link a common chitotriosidase deficiency (BOOT et al. 1998) to susceptibility for infection by chitincontaining parasites (reviewed by BUSSINK et al. 2006). The physiological function of the second mammalian chitinase, AMCase, has recently attracted considerable attention due to a report linking the protein to the pathophysiology of asthma (ZHU et al. 2004).

In addition to active chitinases, highly homologous mammalian proteins lacking enzymatic activity due to substitution of active-site catalytic residues have been identified. Despite their lack of enzymatic activity, these proteins have retained active-site carbohydrate binding and hence have been named chi-lectins (RENKEMA *et al.* 1998; HOUSTON *et al.* 2003; BUSSINK *et al.* 2006). Like the active chitinases, chi-lectins belong to family 18 of glycosyl hydrolases, consisting of a 39-kDa catalytic domain having a TIM-barrel structure, one of the most versatile folds in nature (Sun et al. 2001; WIERENGA 2001; FUSETTI et al. 2002, 2003; HOUSTON et al. 2003). In contrast to both chitinases, chi-lectins lack the conserved additional chitinbinding domain (BOOT et al. 1995, 2001; RENKEMA et al. 1997). Despite the detailed knowledge regarding structure, insight into the exact physiological function of the various chi-lectins is limited (reviewed by BUSSINK et al. 2006). Similar to chitotriosidase and AMCase, chi-lectins are secreted locally or into the circulation and a role in inflammatory conditions is suggested. For example, human cartilage GP39 (Hcgp39/YKL-40/CHI3L1), a protein expressed by chondrocytes and phagocytes, has been implicated in arthritis, tissue remodeling, fibrosis, and cancer (HAKALA et al. 1993; JOHANSEN et al. 1993; VERHEIJDEN et al. 1997; RECKLIES et al. 2002; reviewed by JOHANSEN 2006). Similarly, the human chi-lectin YKL-39 (CHI3L2) and the murine Ym1 (Chi3L3/ECF-L) have been associated with the pathogenesis of arthritis (Hu et al. 1996; TSURUHA et al. 2002) and allergic airway inflammation, respectively (CHANG et al. 2001; WARD et al. 2001; HOMER et al. 2006).

The high-molecular-weight oviductins, consisting of the amino-terminal 39-kDa catalytic domain followed by a heavily glycosylated serine/threonine-rich domain, are secreted by nonciliated oviductal epithelial cells and have been shown to play a role in fertilization and early embryo development (reviewed by BUHI 2002).

To elucidate the evolution of members of family 18 of glycosyl hydrolases in mammals, we extensively surveyed available sequence information in terms of conserved protein features, molecular phylogeny, substitution rate ratios, and chromosomal distribution. This combined approach renders new insights into evolutionary origins, selective pressures, and genomic synteny. Moreover, our investigation has established the true orthologous relationships among mammalian chitinase(-like) proteins.

#### MATERIALS AND METHODS

**Database searches:** DNA and protein sequences for all of the members of glycosyl hydrolase family 18 were obtained by BLAST searches (blastn and blastp and translated blasts) of public databases, mainly the NCBI and ENSEMBL databases (http://www.ncbi.nlm.nih.gov and http://www.ensembl.org/ index.html). The sequences were checked against published sequences in literature, as well as genomic information. For mBclp2, only a partial sequence was available and the entire coding sequence was determined from the genomic sequence on the basis of consensus splice sites. All cDNA sequences obtained were checked against protein sequences and vice versa, by both BLAST searches and manual inspection.

**Protein—alignment and features:** Signal peptides and chitin-binding domain-coding sequences were omitted; hence only sequence information corresponding with the 39-kDa catalytic domain of the active chitinases was used for the alignment. ClustalX was implemented for both protein and DNA alignments (THOMPSON *et al.* 1997), which were checked man-

ually, and, where necessary, alignments were edited. ESPript 2.2 (GOUET *et al.* 1999) was used for visualization of the protein alignment (using RISLER similarity scoring) and automatic superimposition of secondary structures (GOUET *et al.* 2003) of human chitotriosidase on the basis of a published crystal structure (pdb accession code 1LQ0). The assigned helices, strands, and turns were checked manually against the published chitotriosidase structures.

Calculated isoelectric points were determined using the pI/ MW tool on the ExPASy server (http://www.expasy.org/tools/ pi\_tool.html) (GASTEIGER *et al.* 2003).

**Synteny analysis:** Genomic mapping was performed manually by linking the coding sequences retrieved (see above) to genomic location according to both the NCBI and Ensemble databases.

Phylogenetic analysis: Maximum-likelihood and parsimony analyses based on the cDNA alignment were performed by PHYLIP version 3.65 (available at http://evolution.genetics. washington.edu/phylip.html; FELSENSTEIN 2006), making use of the DnaML (after rejection of the molecular clock hypothesis by performing likelihood-ratio testing) and DnaPars programs. Support values were generated on a thousand bootstrapped replicate data sets. For the likelihood analyses, rates were considered equal over sites. To avoid bias caused by order of the sequence input, the order was randomized in all analyses. The consensus tree and bootstrap support values were determined using the Consense program, implemented in the PHYLIP package. The cladogram was generated by aligning the nucleotide sequences coding for the 39-kDa domain with the aid of ClustalX. The input comprised 34 taxa (excluding monkey sequences and including the outgroup sequence), with 1127 characters in each taxon. Inclusion and exclusion of gaps resulted in identical branching and near identical bootstrap values. The input for supplemental Figure 2 at http:// www.genetics.org/supplemental/ was extended with the two Xenopus tropicalis chitinase sequences.

**Calculation of substitution rates:** Rate ratios of nonsynonymous-to-synonymous substitutions  $d_N/d_S(\omega)$  were calculated by PAML, version 3.15 (YANG 1997), using a maximumlikelihood approach based on the consensus tree topology as determined by PHYLIP (see supplemental Figure 1 at http:// www.genetics.org/supplemental/). The program Codeml was used to calculate  $\omega$ -values for branches under the "free-ratio" model that allows a different  $\omega$  for specified branches (model 2). Codeml was also used to calculate  $\omega$ -values for sites (specific codons) under models that either exclude (M0 and M1a) or allow (M2a) positive selection. The different models were compared by performing likelihood-ratio testing.

## RESULTS

Identification of mammalian chitinase protein family members: Mining the literature and using NCBI or ENSEMBL BLAST searches led to the identification of 44 members of the chitinase protein family from 11 different mammalian species. Supplemental Table 1A at http://www.genetics.org/supplemental/ provides an overview of all sequences retrieved, including species specification, common protein aliases, existence of expression data, and NCBI accession numbers. Included are two chitinases from *X. tropicalis* and a chitinase from *Caenorhabditis elegans* that was used as an outgroup in the phylogenetic analyses. Supplemental Table 1B provides an overview of relevant genes identified in the mammalian genomes that are completely (*Homo sapiens*, Pan Troglodytes, and Mus musculus) or nearly completely sequenced (Rattus norvegicus and Bos taurus).

The true chitinase genes (coding for chitotriosidase and AMCase) are present in all mammalian species for which (near) complete genome data exist. The bovine chitotriosidase gene, mapped to chromosome 16, could be only partially aligned due to incompleteness of the genomic sequence.

The genes of the chi-lectins CHI3L1 (encoding Hcgp39 or GP39) and OVGP1 (oviductin) are also present in all analyzed species. Only in the case of rat oviductin could the complete coding sequence not be retrieved due to gaps in the available genomic sequence. The bovine genome contains a second GP39-like gene that is highly homologous (96% nucleotide identity), named BP40, that seems specific for the artiodactyls (even-toed ungulates or hoofed mammals).

Other chi-lectin genes are more specific for particular species. For example, the CHI3L2 (YKL-39) is present in the primate and cow genomes but not in the genomes of rodents. The opposite is the case for the Chi3l3 (Ym1), Chi3l4 (Ym2), and Bclp2 (brain chitinase-like protein 2) genes found only in rodent and not in primate genomes. Interestingly, BLAST searches on the rodent genomes revealed the presence of a previously unidentified paralogue, here referred to as BYm (basic Ym). BYm lacks the catalytic glutamic acid, suggesting that it is a chilectin and shows a high homology to Ym1 and Ym2. Expressed sequence tag data indicate that the gene is transcribed in murine olfactory epithelium and neonates. Previously, JIN et al. (1998) identified a cluster of four genes related to Ym1, naming the other three Ym2, Ym3, and Ym4. The newly identified BYm is not identical to any of these genes, suggesting that it is a novel enzymatically inactive member of the chitinase protein family. So far, there are no indications that the Ym3 gene is expressed. The Ym4 gene is not a complete gene since it lacks most upstream exons while those that are present contain a stop codon, suggesting that it is a pseudogene (JIN et al. 1998).

**Protein alignment and features:** To investigate the overall homology and shared features of the members of the chitinase protein family listed in supplemental Table 1 at http://www.genetics.org/supplemental/, we aligned the 39-kDa TIM-barrel domain. Left out of this alignment were primate species other than humans (chimpanzee and macaque), given the extreme homology. Figure 1 shows the high overall homology in the mammalian TIM-barrel domains, as indicated by the

shaded and solid backgrounds. The secondary structures of chitotriosidase, as observed in its crystal structure, generally correspond with regions of higher conservation, whereas loops separating two helices or strands correspond with gaps in the alignment. When secondary structures seen in crystal structures of two other family members, Hcgp39 and Ym1 (HOUSTON *et al.* 2003 and SUN *et al.* 2001, respectively) are superimposed on the alignment, the outcome is similar (not shown).

The cysteine residues involved in disulfide-bond formation, known to be essential for correct folding and stability of the TIM-barrel, are completely conserved (see Figure 1). Interestingly, both the AMCases and Ym proteins have two additional conserved cysteine residues, potentially allowing formation of a third disulfide bond at positions 28 and 371 (BOOT *et al.* 2001; SUN *et al.* 2001). The latter cysteine is located just outside the 39-kDa domain and is therefore not depicted in Figure 1.

The calculated isoelectric points (pl) of the chitinase (-like) proteins are also strongly conserved, as shown in Table 1. Both mouse and rat BYm have a basic pI. All AMCases have an acidic pI, being neutral in an acidic environment, in agreement with the observed expression in the gastrointestinal tract (SUZUKI *et al.* 2002; BOOT *et al.* 2005a). The oviductins, thought to be ubiquitously expressed in the slightly basic oviduct (HUGENTOBLER *et al.* 2004), have basic pI's. The largest variation in pI exists among chitotriosidases (see Table 1). Taken together, the data on overall homology, conserved cysteines, and isoelectric point indicate interspecies retention of protein structure among orthologs.

Phylogenetic analyses: The maximum-likelihood tree shown in Figure 2, generated as described in MATERIALS AND METHODS, is supported by high bootstrap support values. Importantly, a consensus maximum-parsimony analysis resulted in identical branching (see supplemental Figure 1 at http://www.genetics.org/supplemental/). The tree reveals clustering of all orthologs that are grouped in either the AMCase or the chitotriosidase clade. Both clades contain chi-lectins next to the chitinases, allowing a discrimination of AMCase-lectins and chitotriosidase-lectins. The chitotriosidase clade contains YKL39, GP39, and BP40 homologs. The latter two show complete conservation of the putative N-linked glycosylation sequence NIS near the N terminus (see Figure 1). The AMCase clade contains all oviductins and the rodent Ym proteins.

Substitution rate ratios: The phylogenetic tree provides information on evolutionary relationships of

FIGURE 1.—Protein alignment of the mature 39-kDa catalytic domain of mammalian chitinase(-like) proteins. Completely conserved amino acids are indicated a with solid background and those with a high similarity score (according to RISLER) are in a shaded background. The secondary structures (strands, arrows; helices and turns, T's) of human chitotriosidase, as published, are superimposed on the alignment. Cysteines involved in disulfide bridging are indicated by the numbers below the alignment. Chito, chitotriosidase; hcGP39, human cartilage glycoprotein 39; Ovi, oviductin; h, *H. sapiens*; m, *M. musculus*; r, *R. norvegicus*; b, *B. taurus*; c, *Capra hircus*; o, *Ovis aries*; s, *Sus scrofa*; cf, *Canis familiaris*; ma, *Mesocricetus auratus*. For further details, see text.

A. P. Bussink et al.

	ee	٩		тт—►	وووووووو	2000
	1 10	20	30 40		6 C	
hChito	AKLVCYFTNWAQ	YRQGEARF LPKDLDP YDTEAVDEEDDDVDD	SLCTHLIYAFAGMT NLCTHVIFAFAGMD	.NHQLSTTE.	WNDETLYQEF	IGLKKMNPK
rChito	AKLECYETNWAO	YRSGAARFLPRDVDP	NLCTHVIYAFAGIN	.NHOVSTVE	PNDELEYOELN	SLKKRNPK
bBP40	YKLICYYTSWSQ	YREGDGSCFPDAIDP	FLCTHVIYSFANIS	.NNEIDTWE.	WNDVTLYDTL	TLKNRNPN
cBP40	YKLICYYTSWSQ	Y <b>R</b> EGD <b>GS</b> CF <b>P</b> DA <b>IDP</b>	FLCTHIIYSFANIS	.NNEIDTWE.	WNDVTLYDTL	IT <b>lknrn</b> pk
oBP40	YKLICYYTSWSQ	YREGDGSCFPDAIDP	FLCTHVIYSFANIS	.NNEIDTWE.	WNDVTLYDTL	TLKNRNPK
sBP40	YKLICYYTSWSQ	YREGDGSCFPDAIDP YREGDCSCFPDAIDP	FLCTHVIYSFANIS	NNEIDTWE.	WNDVTLYDTL	NTLKNRNPN
mGP39	YKIVCYFTSWSO	Y BE G V G S E L P D A L O P	FLOTHIIYSFANIS	SDNML STWE	WNDESNYDKL	I L K T R N T N
rGP39	YKLVCYYTNWSQ	YREGNGSCFPDALDH	SLCTHIIYSFANIS	.NNKLSTSE.	WNDVTLYGML	TLKTRNPR
bGP39	YKLICYYTSWSQ	Y <b>R</b> EGD <b>GS</b> CF <b>P</b> DA <b>IDP</b>	FLCTHVIYSFANIS	. NNEIDTWE.	WNDVTLYDTL	IT <b>lknrn</b> pn
sGP39	YKLVCYYTSWSQ	Y R E G D G S C F P D A I D P	FLCTHIIYSFANIS	.NNEIDTLE.	WNDVTLYDTL	NTLKNRNPN
hYKL39	YKLVCYFTNWSQ	DROEPGKETPENIDP	FLCSHLIYSFASIE	.NNKVIIKD.	KSEVMLYQTIN	IS LKTKNPK
hAMCase	YOLTCYFTNWAO	YRPGLGRFMPDNIDP	CLCTHLIYAFAGRO	NNEITTIE.	WNDVTLYOAFN	IG <b>LKNKN</b> SO
mAMCase	YNLICYFTNWAQ	Y <b>R</b> P G L <b>G S</b> F K <b>P</b> D D <b>I N P</b>	CLCTHLIYAFAGMQ	.NNEITTIE.	WNDVTLYKAFN	DLKNRNSK
rAMCase	Y N L V C Y F T N W A Q	Y <b>r</b> pgl <b>gs</b> fk <b>p</b> dd <b>inp</b>	C <b>lCTHLIYAFA</b> GMQ	. NNQITTIE.	WNDVTLYKAFI	D <b>lknrn</b> S <b>k</b>
bAMCase	YQLVCYFSNWAQ	Y R P G L G S F K P D N I D P	CLCTHLIYAFAGMS	.NSEITTIE.	WNDVALYSSF	DLKKKNSQ
ciAMCase hOvi	HKIVCYFTNWAQ	SBPGPASTIPHDIDP	CICTHLIYAFAGMK	NNCTVAKDI	ODEKTIYDEEN	IKIKERNE IKIKERNE
mOvi	YKLVCYFTNWAH	SRPGPASIMPHDLDP	FLCTHLIFAFASMS	NNOIVAKNI	ODENVLYPEFN	KLKERNRE
bOvi	HKLVCYFTNWAF	SRPGPASILPRDLDP	FLCTHLVFAFASMS	. NNQIVPKDE	<b>DEKILY</b> P <b>EF</b>	K <b>lkern</b> rg
oOvi	H K L V C Y F T N W A F	SRPGSASILPRDLDP	F <b>LCTHLVFAFA</b> SMN	. NNQI VPKDE	LDEKILYPEFN	K <b>lkern</b> r <b>g</b>
sOvi	HKLVCYFANWAF	SRPGPASILPRDLDP	FLCTHLVFAFASMN	. DSQIVAKDA	RDESIFYPEF	QLKERNEK
maOvi	YKLVCYFASWAQ	SRPUPASTIPRDIDP	F LCTHLIFAFATME	NNQIVAMIS NNOIVANNI	ODEKIVYPEEN	KIKERNRE
mYm1	YOLMCYYTSWAK	DRPIEGSFKPGNIDP	CLCTHLIYAFAGMO	.NNEITYTH.	EODLRDYEAL	GLKDKNTE
mYm2	Y Q <b>lmcy</b> ytswak	D <b>R</b> PTE <b>GS</b> FK <b>P</b> GN <b>IDP</b>	CLCTHLIYAFAGMK	.NNEITYLS.	EQDLRDYEAL	GLKDRNTE
mYm3	Y Q <b>L M C Y</b> Y T S W A K	D <b>R</b> PRQ <b>GS</b> FK <b>P</b> GT <b>IDP</b>	C <b>LCTHLIYAFA</b> GMQ	. NNEITYTH.	EQDLRDYETL	IG <b>lkdrn</b> te
rYml	YQLMCYYTSWAK	DRPTVGSFKIGNIDP NRPKICSENRADIDP	CLCTHLIYAFAGMR	. NNEIINTS.	EQDLIDYEAIN	NYLKDRNTE
mBYm	YOTMCYFNNWPO	HOPDVBDIKHEDIDP	CLCTHLIYSFAGIW	ENNETMERS.	RKELDDYKGEN	IDI.KKRNNK
rBYm	YQLMCYFNNWPQ	YQPDVRGMKLDDIDP	CLCTHLIYSFAGIW	.ENNNTMTK.	RKDLDDYKEF	DLKKRNNK
	7 <u>0</u> 80	ooooo ومومو مومومو	22222222222222222 100 11	0 12	TT TT ( 20 130	200000000
hChito	7 0 8 0 LKTLLAIGGWNF	202020 90 GIQKETDMVATANNR	00000000000000000000000000000000000000		TT TT 20 130 YPGSQGSPAVI	KER <b>FTTLV</b>
hChito mChito	70 80 LKTLLAIGGWNF LKTLLAVGGWTF	<u>200000</u> 90 GIOKETDMVATANNR GIOKETDMVATASNR GIOKETDMVATASTR	00000000000000000000000000000000000000		TT TT 50 20 130 YPGSOGSPAVI FPGCRGSPTVI	O O O O O O O O O O O O O O O O O O O
hChito mChito rChito bBP40	70, 80 LKTLLAIGGWNF LKTLLAVGGWF LKTLLAVGGWSF LKTLLAVGGWNF	<u>200000</u> 90 GIQKETDMVATANNR GIQKETDMVATASNR GIQKETDMVATASTR GSERFSKIASKTOSR	Q 100 100 11 Q T F V S A L S A A A S A A S A A A A A A A A A A A A A	9 12 SFDGLDLDWE GFDGLDLDWE GFDGLDLDWE GFDGLDLAWI	TT TT 50 TPGSQGSPAVI FPGCRGSPTVI YPGSRGSPAVI YPGWR	XERFTTLV KERFTTLV KERFTTLI KERFTALI KERFTALI
hChito mChito rChito bBP40 cBP40	70, 80 LKTLLAIGGWNF LKTLLAVGGWTF LKTLLAVGGWSF LKTLLSVGGWNF LKTLLSVGGWNF	<u>000000</u> 90 GTQKETDMVATANNR GTQKETDMVATASNR GTQKETDMVATASTR GSERFSKIASKTQSR GPERFSKIASKTQSR	00000000000000000000000000000000000000	9 12 SFDGLDLDWE GFDGLDLDWE GFDGLDLDWE GFDGLDLAWI GFDGLDLAWI	TT TT 5 20 130 TPG SQGSPAVI FPG GRGSPTVI YPG SRGSPAVI YPG RRI	KERFTTLV KERFTTLV KERFTTLI KERFTALI KRHLTTLV
hChito mChito rChito bBP40 cBP40 oBP40 oBP40	70 LKTLLAIGGWNF LKTLLAVGGWTF LKTLLAVGGWNF LKTLLSVGGWNF LKTLLSVGGWNF	<u>000000</u> 90 GTQKETDMVATANNR GTQKETDMVATASNR GTQKETDMVATASTR GSERESKIASKTQSR GPERESKIASKTQSR GPERESAIASKTQSR	00000000000000000000000000000000000000	0     12       SFDGLDLDWE       GFDGLDLDWE       GFDGLDLWE       GFDGLDLAWE       GFDGLDLAWE       GFDGLDLAWE	TT     TT       20     130       YPG     200       FPG     200       YPG     200	KERFTTLV KERFTALI KERFTALI KRHLTTLV KRHLTTLV
hChito mChito bBP40 cBP40 oBP40 sBP40 sBP40 bbcCP39	70 LKTLLAIGGWNF LKTLLAVGGWF LKTLLAVGGWF LKTLLSVGGWNF LKTLLSVGGWNF LKTLSVGGWNF LKTLSVGGWNF	OCOUCO 90 GTQKETDMVATANNR GTQKETDMVATASNR GTQKETDMVATASNR GTQKETDMVATASTR GSERESKIASKTQSR GPERESKIASKTQSR GPERESKIASKTQSR GPERESKIASKTQSR	00000000000000000000000000000000000000	0     12       SFDGLDLDWE       GFDGLDLDWE       GFDGLDLWE       GFDGLDLAWE       GFDGLDLAWE       GFDGLDLAWE       GFDGLDLAWE       GFDGLDLAWE	TT     TT     J       20     130       YPG     SQGSPAVI       FPGGRGSPTVI       YPG     R       YPGRR     I       YPGRR     I       YPGRR     I	KERFTTLV KERFTALI KERFTALI KRHLTTLV KRHLTTLV KRHLTTLV
hChito mChito rChito bBP40 cBP40 oBP40 sBP40 hcGP39 mGP39	70 LKTLLAIGGWNF LKTLLAVGGWTF LKTLLSVGGWNF LKTLLSVGGWNF LKTLSVGGWNF LKTLLSVGGWNF LKTLLSVGGWNF LKTLLSVGGWNF	00000 90 GTQKETDMVATANNR GTQKETDMVATASNR GTQKETDMVATASNR GTQKETDMVATASTR GSERESKIASKTQSR GPERESKIASKTQSR GPQRESKIASKTQSR GSQRESKIASNTQSR GEKRESEIASNTERR	Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q	0     12       SFDGLDLDWE       GFDGLDLDWE       GFDGLDLWE       GFDGLDLAWE	TT     TT       20     130       YPG     SQGSPAVI       FPGGRGSPTVI       YPGSRGSPAVI       YPGRR     I	XERFTTLV KERFTALI KERFTALI KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KQHFTTLI
hChito mChito rChito bBP40 cBP40 oBP40 sBP40 hcGP39 mGP39 rGP39	70 LKTLLAIGGWNF LKTLLAVGGWF LKTLLSVGGWNF LKTLLSVGGWNF LKTLSVGGWNF LKTLSVGGWNF LKTLSVGGWNF LKTLSVGGWNF LKTLSVGGWNF LKTLSVGGWNF	OCOCO 90 GTQKETDMVATANNR GTQKETDMVATASNR GTQKETDMVATASNR GTQKETDMVATASTR GSERFSKIASKTQSR GPERFSKIASKTQSR GPERFSKIASKTQSR GSQRFSKIASKTQSR GEKRFSEIASNTERR GSERFSRIVSNAKSR	Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q	0     12       SFDGLDLDWE       GFDGLDLDWE       GFDGLDLWE       GFDGLDLAWE	TT       TT         20       130         YPG       SQGSPAVE         FPGGRGRSPTVE       YPGSRGSPAVE         YPGRR       I         YPR       I         YPGRR       I         YPG	XERFTTLV KERFTALI KERFTALI KERFTALI KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KQHFTTLI KQHFTTLI
hChito mChito rChito bBP40 cBP40 oBP40 sBP40 hcGP39 mGP39 rGP39 rGP39 bGP39	70 LKTLLAIGGWNF LKTLLAVGGWSF LKTLLSVGGWNF LKTLLSVGGWNF LKTLLSVGGWNF LKTLLSVGGWNF LKTLLSVGGWNF LKTLLSVGGWNF LKTLLSVGGWNF	90 90 GTQK FTD MVAT ANNR GTQK FTD MVAT ANNR GTQK FTD MVAT ASNR GTQK FTD MVAT ASTR GS ERFSK IASK TQSR GP ERFSK IASK TQSR GP QRFSK IASK TQSR GSQRFSK IASN TERR GSQRFSK IASN TERR GSQRFSK IASK TQSR	Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q	0     12       SFDGLDLDWE       GFDGLDLDWE       GFDGLDLWE       GFDGLDLAWI	TT       TT         20       130         YPG       SQGSPAVE         FPGGRGSPTVE       YPGRR	OKERFTTLV KERFTALI KERFTALI KERFTALI KERFTALI KERFTALI KERLTTLV KERLTTLV KERLTTLV KERLTTLV KQHFTTLI KQHFTTLI KQHFTTLI KRHLTTLV
hChito mChito rChito bBP40 cBP40 oBP40 sBP40 hcGP39 mGP39 rGP39 bGP39 sGP39 bQP39 bQP39 bQP39	70 LKTLLAIGGWNF LKTLLAVGGWF LKTLLSVGGWNF LKTLLSVGGWNF LKTLLSVGGWNF LKTLLSVGGWNF LKTLLSVGGWNF LKTLLSVGGWNF LKTLLSVGGWNF LKTLLSVGGWNF LKTLLSVGGWNF LKTLLSVGGWNF LKTLLSVGGWNF	90 90 90 GTQK FTDMVATANNR GTQK FTDMVATASNR GTQK FTDMVATASNR GTQK FTDMVATASTR GSERFSKIASKTQSR GPQRFSKIASKTQSR GSQRFSKIASNTQSR GSQRFSKIASNTERR GSQRFSKIASNTQSR GSQRFSKIASNTQSR GSQRFSKIASNTQSR GSQRFSKIASNTQSR	Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q	0     12       SFDGLDLDWE       GFDGLDLDWE       GFDGLDLAWI	TT       TT         20       130         YPG       SQGSPAVI         YPGSRGSPAVI       I         YPGRR	COOCOCOCO KERFTTLV KERFTALI KERFTALI KERFTALI KERLTTLV KERLTTLV KERLTTLV KERLTTLV KERLTTLV KERLTTLV KERLTTLV
hChito mChito rChito bBP40 cBP40 oBP40 sBP40 hcGP39 mGP39 rGP39 bGP39 sGP39 hYKL39 bYKL39	70 LKTLLAIGGWNF LKTLLAVGGWF LKTLLSVGGWNF LKTLLSVGGWNF LKTLLSVGGWNF LKTLLSVGGWNF LKTLLSVGGWNF LKTLLSVGGWNF LKTLLSVGGWNF LKTLLSVGGWF LKTLLSVGGWF LKTLLSIGGYLF	90 90 GTQK FTDMVATANNR GTQK FTDMVATASNR GTQK FTDMVATASNR GTQK FTDMVATASNR GSERFSKIASKTQSR GPERFSKIASKTQSR GPQRFSKIASKTQSR GSQRFSKIASNTQSR GSQRFSKIASNTQSR GSQRFSKIASNTQSR GSQRFSKIASNTQSR GSQRFSKIASNTQSR GSKGFHPMVDSSTSS	Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q	0     12       SFDGLDLDWE       GFDGLDLDWE       GFDGLDLWE       GFDGLDLAWE       MFDGLDE       WFDGLDE	TT       TT         20       130         YPG       SQGSPAVI         YPGSRGSPAVI       I         YPGRR	COOCOCOCO KERFTTLV KERFTALI KERFTALI KERFTALI KERLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KQHFTTLI KQHFTTLI KQHFTTLI KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV
hChito mChito rChito bBP40 cBP40 sBP40 hcGP39 mGP39 rGP39 bGP39 sGP39 hYKL39 bYKL39 hAMCase	70 LKTLLAIGGWNF LKTLLAVGGWF LKTLLSVGGWNF LKTLLSVGGWNF LKTLLSVGGWNF LKTLLSVGGWNF LKTLLSVGGWNF LKTLLSVGGWNF LKTLLSVGGWNF LKTLLSVGGWNF LKTLLSVGGWNF LKTLLSIGGYLF LKTLLSIGGYLF LKTLLSIGGWF	90 90 G T QK F T D MVAT ANNR G T QK F T D MVAT ANNR G T QK F T D MVAT AS T R G S E R F SK I AS K T Q S R G P E R F SK I AS K T Q S R G P Q R F SK I AS K T Q S R G S Q R F SK I AS K T Q S R G S Q R F SK I AS N T E R R G S Q R F SK I AS N T E R R G S Q R F SK I AS N T Z S R G S Q R F SK I AS N T Q S R G S Q R F SK I AS N T Q S R G S Q R F SK I AS N T Q S R G S Q R F SK I AS N T Q S R G S Q R F SK I AS N T Q S R G S Q R F SK I AS N T Q S R G S Q R F SK I AS N T Q S R G S Q R F SK I AS N T Q S R G S Q R F H P MVD S S S S S T G T A P F T A MVS T P E N R	0       0       0       0       0       1         0       10       1       1       1         0       TFVNSAIRFIRKY       0       0       1         0       TFVNSAISFIRTQ       0       1       1         0       TFIKSVPPFIRTH       1       1       1       1       1         0       TFIKSVPPFIRTH       TAFVRSVAPFIRTY       1	0       12         SFDGLDLDWE       GFDGLDLDWE         GFDGLDLDWE       GFDGLDLAWI         GFDGLDLAWI       SFDGLDLAWI         GFDGLDLAWI       GFDGLDLAWI         GFDGLDLAWI       SFDGLDLAWI         GFDGLDLAWI       SFDGLDV         SFDGLDLAWI       SFDGLDV         SFDGLDV       SWI         SFDGLDV       SWI         SFDGLDV       SWI         SFDGLDV       SWI         SFDGLDV       SWI         SFDGLDV       SWI	TT       TT       J         20       130         YPG       SQSPAVE         YPGSRGSPAVE         YPGRR	COOCOCOCO KERFTTLV KERFTALI KERFTALI KERFTALI KERFTALI KERLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KQHFTTLI KQHFTTLI KQHFTTLI KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTVI
hChito mChito rChito bBP40 cBP40 oBP40 sBP40 hcGP39 mGP39 mGP39 bGP39 sGP39 hYKL39 bYKL39 hAMCase mAMCase	70       80         LKT       LAIGGWNF         LKT       LAVGGWF         LKT       LSVGGWF         LKT       LSGGYLF         LKT       LSGGWF         LKT       LSGGWF         LKT       LSGGYLF         LKT       LAIGGWF         LKT       LAIGGWF         LKT       LAIGGWF	90 90 G T Q K F T D MVAT ANNR G T Q K F T D MVAT ANNR G T Q K F T D MVAT AS NR G T Q K F T D MVAT AS T R G S E R F S K I AS K T Q S R G P Q R F S K I AS K T Q S R G P Q R F S K I AS K T Q S R G S Q R F S K I AS N T E R R G S Q R F S K I AS N T E R R G S Q R F S K I AS N T Z S R G S Q R F S K I AS N T Q S R G S Q R F S K I AS N T Q S R G S K G F H P MVD S S T S S G S K G F H P VVE S S S S T G T A P F T A MVS T P E N R G T A P F T T MVS T S Q N R	0       0       0       0       0       0       1         0       10       1	0       12         S F D G L       D L D W E         G F D G L D L D W E         G F D G L D L D W E         G F D G L       D L A W I         G F D G L D L W E         G F D G L D U W E         G F D G L D U W E	TT       TT       J         20       130         YPG       SQGSPAVI         YPGSRGSPAVI         YPGRR	COOCOCOCO KERFTTLV KERFTALI KERFTALI KERFTALI KERLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTVI KRHLTTVI KRHLFTVI
hChito mChito rChito bBP40 cBP40 oBP40 sBP40 hcGP39 mGP39 rGP39 bGP39 sGP39 hYKL39 bYKL39 hAMCase mAMCase rAMCase	70       80         LKT       LLAIGGWNF         LKT       LLAVGGWF         LKT       LLVGGWF         LKT       LSVGGWF         LKT       LSGYLF         LKT       LSGYLF         LKT       LAIGGWF	90 90 GTQK FTD MVATANNR GTQK FTD MVATANNR GTQK FTD MVATASNR GTQK FTD MVATASNR GS ERFSKIASKTQSR GP ERFSKIASKTQSR GP QRFSKIASKTQSR GS QRFSKIASKTQSR GS QRFSKIASNTQSR GS QRFSKIASNTQSR GS QRFSKIASNTQSR GS KGFHP VVSSSSST GTAPFTAMVSTSQNR GTAPFTTMVSTSQNR	0       0       0       0       0       0       0       1         0       10       11       1 <th>0       12         S F DGL DL DWF       GF DGL DL DWF         G F DGL DL DWF       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI</th> <th>TT       TT         20       130         YPG       SQGSPAVI         YPGSRGSPAVI       YPGRR</th> <th>COOCOCOCO KERFTTLV KERFTALI KERFTALI KERFTALI KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTVLI KRHLTTVLI KHLFTVLI KHLFTVLV</th>	0       12         S F DGL DL DWF       GF DGL DL DWF         G F DGL DL DWF       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI	TT       TT         20       130         YPG       SQGSPAVI         YPGSRGSPAVI       YPGRR	COOCOCOCO KERFTTLV KERFTALI KERFTALI KERFTALI KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTVLI KRHLTTVLI KHLFTVLI KHLFTVLV
hChito mChito rChito bBP40 cBP40 sBP40 hcGP39 rGP39 rGP39 bGP39 sGP39 hYKL39 bYKL39 hAMCase mAMCase rAMCase cfAMCase	70 LKT LKT LLAIGGWNF LKT LKT LSVGGWNF LKT LSVGGWNF LKT LSVGGWNF LKT LSVGGWNF LKT LSVGGWNF LKT LSVGGWNF LKT LSVGGWNF LKT LSSVGGWNF LSSVG LSSVG LSSVG LSSVGWF LSSVG LS	90 90 GTQK FTD MVATANNR GTQK FTD MVATANNR GTQK FTD MVATASNR GTQK FTD MVATASNR GS ERFSKIASKTQSR GP ERFSKIASKTQSR GP QRFSKIASKTQSR GS QRFSKIASKTQSR GS QRFSKIASNTQSR GS QRFSKIASNTQSR GS QRFSKIASNTQSR GS KGF HP MVSSSSST GTAPFTAMVSTSQNR GTAPFTAMVSSPENR	Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q	0       12         S F DGL DL DWE       GF DGL DL DWE         G F DGL DL DWE       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL AWI       GF DGL DL AWI         G F DGL DL WI       GF DGL DL WI         G F DGL DL WI       GF DGL DL WI         G F DGL DL WI       GF DGL DL WI         G F DGL DF WI       GF DGL DF WI	TT       TT         20       130         YPG       SQGSPAVI         YPGSRGSPAVI       YPGRR	COOCOCOCO KERFTTLV KERFTALI KERFTALI KERFTALI KERFTALI KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV
hChito mChito rChito bBP40 cBP40 sBP40 hcGP39 rGP39 rGP39 bGP39 sGP39 hYKL39 hAWCase rAMCase rAMCase cfAMCase hOvi	70       80         LKT       LLAIGGWNF         LKT       LLAVGGWF         LKT       LLVGGWNF         LKT       LSVGGWNF         LKT       LAIGGWNF	90 90 G TQK F TD MVAT ANNR G TQK F TD MVAT ANNR G TQK F TD MVAT ASNR G TQK F TD MVAT ASNR G S E R F SK IASK TQSR G P E R F SK IASK TQSR G P Q R F SK IASK TQSR G SQR F SK IASK TQSR G SKG F HP MVS TSS G SKG F HP VVES SSST G TAP F TAMVS TPENR G TAP F TAMVS TPENR G TAP F TAMVS TPENR G TAP F TAMVS SPENR G TAP F TAMVS SPENR	0       0       0       0       0       0       0       1         1       0       1       1       1       1       1         0       1       0       1 <th>9 SFDGLDLDWE GFDGLDLDWE GFDGLDLAWI GFDGLDF</th> <th>TT       TT         20       130         21       130         29       130         29       130         29       130         29       130         29       130         29       130         29       130         29       130         29       130         29       130         29       130         29       130         29       130         29       100</th> <th>COOCOCOCO KERFTTLV KERFTALI KERFTALI KERFTALI KERFTALI KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV</th>	9 SFDGLDLDWE GFDGLDLDWE GFDGLDLAWI GFDGLDF	TT       TT         20       130         21       130         29       130         29       130         29       130         29       130         29       130         29       130         29       130         29       130         29       130         29       130         29       130         29       130         29       130         29       100	COOCOCOCO KERFTTLV KERFTALI KERFTALI KERFTALI KERFTALI KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV
hChito mChito rChito bBP40 cBP40 sBP40 hcGP39 rGP39 rGP39 bGP39 sGP39 hYKL39 hAMCase rAMCase rAMCase bAMCase hOvi mOvi	70       80         LKT       LLAIGGWNF         LKT       LLAVGGWF         LKT       LLVGGWNF         LKT       LSVGGWNF         LKT       LAIGGWNF         LKT       LAIGGWNF         LKT       LAIGGWNF         LKT       LAIGGWNF         LKT       LAIGGWNF         LKT       LAIGGWNF         LKT       LSIGGWNF         LKT       LSIGGWNF	90 90 G TQK F TD MVAT ANNR G TQK F TD MVAT ANNR G TQK F TD MVAT ASNR G TQK F TD MVAT ASNR G S E R F SK IASK TQSR G P E R F SK IASK TQSR G P Q R F SK IASK TQSR G SQR F SK IASK TQSR G SKG F HP MVS TSS G SKG F HP VVES SSST G TAP F TAMVS TPENR G TAP F TAMVS TPENR G TAP F TAMVS TPENR G TAP F TAMVS SPENR G TSR F TT MLS T LANR	0       0       0       0       0       0       0       0       1         Q       T       FVNSAIRFIRKY       0       0       0       0       0       1       1         Q       T       FVNSAIRFIRKY       0       0       1       1       1       1         Q       T       FVNSAIRFIRKY       0       1 <th>9 SFDGLDLDWE GFDGLDLDWE GFDGLDLAWI GFDGLDL</th> <th>TT       TT         130       130         YPG       QGSPAVI         FPG       RGSPAVI         YPG       RGSPAVI         YPG       RGSPAVI         YPG       R         YPG</th> <th>COOCOCOCO KERFTTLV KERFTALI KERFTALI KERFTALI KERFTALI KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTVLI KRHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV</th>	9 SFDGLDLDWE GFDGLDLDWE GFDGLDLAWI GFDGLDL	TT       TT         130       130         YPG       QGSPAVI         FPG       RGSPAVI         YPG       RGSPAVI         YPG       RGSPAVI         YPG       R         YPG	COOCOCOCO KERFTTLV KERFTALI KERFTALI KERFTALI KERFTALI KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTLV KRHLTTVLI KRHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV KHLFTVLV
hChito mChito rChito bBP40 cBP40 oBP40 sBP40 hcGP39 rGP39 bGP39 sGP39 hYKL39 bYKL39 hAMCase rAMCase rAMCase cfAMCase hOvi mOvi bOvi	70       80         LKT       LLAIGGWNF         LKT       LLAVGGWF         LKT       LLVGGWNF         LKT       LSVGGWNF         LKT       LAIGGWNF         LKT       LSIGGWNF	90 90 G TQK F TD MVAT ANNR G TQK F TD MVAT ANNR G TQK F TD MVAT ASNR G TQK F TD MVAT ASNR G S E R F SK IASK TQSR G P E R F SK IASK TQSR G P E R F SK IASK TQSR G P Q R F SK IASK TQSR G S G C F HP MVS TSS G S G C F HP VVESSSST G TAP F TAMVS TPENR G TAP F TAMVS TPENR G TAP F TAMVS TPENR G TAP F TAMVS SPENR G TAP F TAMVS SPENR G TSR F TT MLST FANR G T SR F TAMLST LANR G T VR F TTMLST FSNR	Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q	9 SFDGLDLDWE GFDGLDLDWE GFDGLDLAWI GFDGLDL	TT       TT         130         130         130         141         141         142         142         143         143         144         145         145         146         147         147         148         149         149         140         141         142         143         144         144         145         146         147         147         148         148	COOCOCOCO KERFTTLV KERFTALI KERFTALI KERFTALI KERFTALI KERFTALI KERFTALI KERFTALI KERFTTLV KERLTTLV KERLTTLV KERLTTLV KERFTVLV KERFTVLV KERFTVLV KERFTVLV KERFTVLV KERFTVLV KERFTVLV KERFTVLV KERFTVLV KERFTVLV KERFTTLL KERFTALI KERFTTLV KERFTTLV KERFTTLV KERFTTLV KERFTTLV KERFTTLI KERFT
hChito mChito rChito bBP40 cBP40 oBP40 sBP40 hcGP39 rGP39 bGP39 sGP39 hYKL39 bYKL39 hAMCase rAMCase cfAMCase cfAMCase hOvi mOvi bOVi oOvi sOvi	70       80         LKT       LLAIGGWNF         LKT       LLAVGGWF         LKT       LLSVGGWNF         LKT       LSVGGWNF         LKT       LSVGWNF	90 90 G TQK F TD MVAT ANNR G TQK F TD MVAT ANNR G TQK F TD MVAT ASNR G TQK F TD MVAT ASNR G S E R F SK IASK TQSR G P E R F SK IASK TQSR G P E R F SK IASK TQSR G P Q R F SK IASK TQSR G S G C F HP MVS TSS G S G C F HP VVE SSS TS G TAP F TAMVS TPENR G TAP F TAMVS TPENR G TAP F TAMVS TPENR G TAP F TAMVS TPENR G TAP F TAMVS SPENR G TSR F TT MLS TFANR G T SR F TT MLS TFSNR G T SR F TT MLS TFSNR	Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q	9 SFDGLDLDWE GFDGLDLDWE GFDGLDLAWI GFDGLDLFFI GFDGLDLFFI GFDGLDLFFI GFDGLDLFFI	TT       TT         130         130         141         141         142         142         143         143         144         145         145         146         147         147         148         149         149         140         141         142         142         143         144         144         145         146         147	D C C C C C C C C C C C C C
hChito mChito rChito bBP40 cBP40 oBP40 sBP40 hcGP39 rGP39 bGP39 sGP39 hYKL39 bYKL39 bYKL39 bYKL39 hAMCase rAMCase cfAMCase cfAMCase hOvi mOvi bOVi oOvi sOvi cfOvi	70       80         LKT       LLAIGGWNF         LKT       LLAVGGWNF         LKT       LLSVGGWNF         LKT       LSVGGWNF	90 90 G TQK F TD MVAT ANNR G TQK F TD MVAT ANNR G TQK F TD MVAT ASNR G TQK F TD MVAT ASNR G S E R F SK IASK TQSR G P E R F SK IASK TQSR G P E R F SK IASK TQSR G P Q R F SK IASK TQSR G S G C F HP MVS TSS G S G C F HP VVE SSS TS G TAP F TAMVS TPENR G TAP F TAMVS SPENR G TSR F TTMLS TFSNR G TSR F TKMLS TFSNR G TSR F TKMLS TFSNR G TSR F TTMLS TFTNR	Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q	9       12         S F D G L D L D WE       G F D G L D L D WE         G F D G L D L D WE       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L D WE       G F D G L D L D WE         G F D G L D L D WI       G F D G L D L F F I         G F D G L D L D WE       G F D G L D L F F I         G F D G L D L D WE       G F D G L D L F F I         G F D G L D L D WE       G F D G L D L F F I         G F D G L D L F F I       G F D G L D L F F I         G F D G L D L F F I       G F D G L D L F F I         G F D G L D L F F I       G F	TT       TT         130         130         140	QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ
hChito mChito rChito bBP40 cBP40 oBP40 sBP40 hcGP39 mGP39 rGP39 bGP39 bGP39 bGP39 bGP39 bYKL39 bYKL39 bYKL39 bYKL39 bYKL39 bAMCase rAMCase cfAMCase cfAMCase cfAMCase hOvi mOvi bOvi oOvi sOvi cfOvi maOvi	70       80         LKT       LLAIGGWNF         LKT       LLAVGGWNF         LKT       LLSVGGWNF         LKT       LSVGGWNF         LKT       LAIGGWNF         LKT       LAIGGWNF         LKT       LSIGGWNF         LKT       SVGGWNF         LKT       SVGGWNF         LKT       SVGGWNF         LKT       LSVGWNF <th>90 90 G TQK F TD MVAT ANNR G TQK F TD MVAT ANNR G TQK F TD MVAT ASNR G TQK F TD MVAT ASNR G S E R F SK IASK TQSR G P E R F SK IASK TQSR G P E R F SK IASK TQSR G P Q R F SK IASK TQSR G Q R F SK IASK TQSR G SQR F SK IASK TQSR G SQR F SK IASK TQSR G SQR F SK IASK TQSR G SKG F HP MVS TSS G SKG F HP VVE SSS TS G TAP F TAMVS TPENR G TSR F TTMLS TFSNR G TSR F TTMLS TFSNR G TSR F TTMLS TFSNR G TSR F TTMLS TLASR G TSR F TTMLS TLASR</th> <th>0       0</th> <th>9       12         S F D G L D L D WE       G F D G L D L D WE         G F D G L D L D WE       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L D WE       G F D G L D L MI         G F D G L D L D WE       G F D G L D L F F I         G F D G L D L F F I       G F D G L D L F F I         G F D G L D L F F I       G F D G L D L F F I         G F D G L D L F F I       G F D G L D L F F I         G F D G L D L F F I       G F D G L D L F F I         G F D G L D L F F I       G F D G L D L F F I         G F D G L D L F F I       G F D G L D L F F I         G F D G L D L F F I       G</th> <th>TT       TT         13         14         15         17       13         17       13         17       13         17       13         17       13         17       13         17       13         17       13         17       15         17       16         17       17        17         &lt;</th> <th>QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ</th>	90 90 G TQK F TD MVAT ANNR G TQK F TD MVAT ANNR G TQK F TD MVAT ASNR G TQK F TD MVAT ASNR G S E R F SK IASK TQSR G P E R F SK IASK TQSR G P E R F SK IASK TQSR G P Q R F SK IASK TQSR G Q R F SK IASK TQSR G SQR F SK IASK TQSR G SQR F SK IASK TQSR G SQR F SK IASK TQSR G SKG F HP MVS TSS G SKG F HP VVE SSS TS G TAP F TAMVS TPENR G TSR F TTMLS TFSNR G TSR F TTMLS TFSNR G TSR F TTMLS TFSNR G TSR F TTMLS TLASR G TSR F TTMLS TLASR	0       0	9       12         S F D G L D L D WE       G F D G L D L D WE         G F D G L D L D WE       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L D WE       G F D G L D L MI         G F D G L D L D WE       G F D G L D L F F I         G F D G L D L F F I       G F D G L D L F F I         G F D G L D L F F I       G F D G L D L F F I         G F D G L D L F F I       G F D G L D L F F I         G F D G L D L F F I       G F D G L D L F F I         G F D G L D L F F I       G F D G L D L F F I         G F D G L D L F F I       G F D G L D L F F I         G F D G L D L F F I       G	TT       TT         13         14         15         17       13         17       13         17       13         17       13         17       13         17       13         17       13         17       13         17       15         17       16         17       17        17         <	QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ
hChito mChito rChito bBP40 cBP40 oBP40 sBP40 hcGP39 rGP39 rGP39 bGP39 bGP39 bGP39 bYKL39 hXKL39 bAWCase rAMCase cfAMCase cfAMCase cfAMCase cfAMCase cfAMCase cfAMCase cfAMCase cfAMCase cfAMCase cfAMCase cfAMCase cfAMCase	70       80         LKT       LLAIGGWNF         LKT       LLAVGGWNF         LKT       LLSVGGWNF         LKT       LSVGGWNF         LKT       LAIGGWNF         LKT       LAIGGWNF         LKT       LAIGGWNF         LKT       LSIGGWNF	90 90 G TQK F TD MVATANNR G TQK F TD MVATASNR G TQK F TD MVATASNR G TQK F TD MVATASNR G S E R F SK IASK TQSR G P E R F SK IASK TQSR G P E R F SK IASK TQSR G P Q R F SK IASK TQSR G Q R F SK IASK TQSR G Q R F SK IASK TQSR G S G F HP VVS SSS ST G TAP F TAMVS TPENR G TAP F TAMVS TPENR G TAP F TAMVS SPENR G TSR F TTMLST FANR G TSR F TTMLST FSNR G TSR F TTMLST FSNR G TSR F TTMLST FSNR G TSR F TTMLST FSNR G TSR F TTMLST FNR G TSR F TTMLST FNR G TSR F TTMLST STNR G TSR F TTMLST STNR F TTMLST STNR G TSR F TTMLST STNR F TTMLST STNR F TTMLST STNR G TSR F TTMLST STNR	Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q	9       12         S F D G L       D L D W F         G F D G L       D L D W F         G F D G L       D L D W F         G F D G L       D L D W F         G F D G L       D L A W I         G F D G L       D L A W I         G F D G L       D L A W I         G F D G L       D L A W I         G F D G L       D L A W I         G F D G L       D L A W I         G F D G L       D L A W I         G F D G L       D L A W I         G F D G L       D L A W I         G F D G L       D L A W I         G F D G L       D L A W I         G F D G L       D L A W I         G F D G L       D L A W I         G F D G L       D L A W I         G F D G L       D L A W I         G F D G L       D L M W I         G F D G L       D L F D W F         G F D G L       D L F F I         G F D G L       D L F F I         G F D G L       D L F F I         G F D G L       D L F F I         G F D G L       D L F F I         G F D G L       D L F F I         G F D G L       D L F F I         G F D G L       D	TT       TT         130         130         130         130         130         130         130         141         141         142         142         143         143         144         145         145         146         147         147         148         149         149         140         140         141         142         143         144         144         145         146         147	QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ
hChito mChito rChito bBP40 cBP40 oBP40 sBP40 hcGP39 rGP39 rGP39 bGP39 bGP39 bYKL39 hXKL39 hXKL39 hXKL39 hXKL39 hXKL39 bAWCase cfAMCA cfAMCA c	70 LKT LLAIGGWNF LKT LLAVGGWNF LKT LKT LSVGGWNF LKT LSVGGWNF LKT LSVGGWNF LKT LSVGGWNF LKT LSVGGWNF LSVGGWNF LKT LSVGGWNF LSVGWNF LSVGGWNF LSVGGWNF LSVGGWNF LSVGGWNF LSVGGWNF LSVGGWNF LSVGWNF LSVGWNF LSVGWNF LSVGWNF LSVGWNF L	90 90 GTQK FTD MVATANNR GTQK FTD MVATANNR GTQK FTD MVATASNR GTQK FTD MVATASNR GS ERFSKIASKTQSR GPERFSKIASKTQSR GPQRFSKIASKTQSR GSQRFSKIASKTQSR GSQRFSKIASKTQSR GSQRFSKIASKTQSR GSQRFSKIASKTQSR GSKGFHPWVSSSSST GTAPFTAMVSTPENR GTAPFTAMVSTSQNR GTAPFTAMVSTSQNR GTAPFTAMVSTSQNR GTAPFTAMVSTSQNR GTAPFTAMVSTSQNR GTAPFTAMVSTSQNR GTAPFTAMVSTSQNR GTAPFTAMVSTSQNR GTAPFTAMVSTSQNR GTAPFTAMVSTSQNR GTSRFTTMLSTFANR GTSRFTTMLSTFSNR GTSRFTTMLSTFSNR GTSRFTTMLSTFSNR GTSRFTTMLSTFSNR GTSRFTTMLSTFNN GTSRFTTMLSTFNNR GTSRFTTMLSTFNNR GTSRFTTMLSTFNNR	Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q	9       12         S F D G L D L D WE       G F D G L D L D WE         G F D G L D L D WE       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L A WI       G F D G L D L A WI         G F D G L D L D WE       G F D G L D L F F I         G F D G L D L D WE       G F D G L D L F F I         G F D G L D L F F I       G F D G L D L F F I         G F D G L D L F F I       G F D G L D L F F I         G F D G L D L F F I       G F D G L D L F F I         G F D G L D L F F I       G F D G L D L F F I         G F D G L D L F F I       G F D G L D L F F I         G F D G L D L F F I       G F D G L D L F F I         G F D G L D L F F I       G F D G L D L F F I	TT       TT         130         130         130         141         141         142         142         143         143         144         145         145         146         147         147         148         148         149         149         149         140         141         142         143         144         144         145         145	000000000000000000000000000000000000
hChito mChito rChito bBP40 cBP40 oBP40 sBP40 hcGP39 rGP39 bGP39 sGP39 hYKL39 hAMCase rAMCase rAMCase cfAMCase cfAMCase hOvi bOvi oOvi sOvi cfOvi maOvi maOvi mYm1 mYm2 mYm3 rYm1	70 LKT LLAIGGWNF LKT LLAVGGWNF LKT LKT LSVGGWNF LKT LSVGGWNF LKT LSVGGWNF LKT LSVGGWNF LKT LSVGGWNF LSVGWNF LSVGGWNF LSVGGWNF LSVGGWNF LSVGGWNF LSVGGWNF LSVGGWNF LSVGGWNF LSVGGWNF LSVGGWNF LSVGWNF LSVGWNF LSVGWNF LSVG	90 90 GTQK FTD MVATANNR GTQK FTD MVATANNR GTQK FTD MVATASNR GTQK FTD MVATASNR GSERFSKIASKTQSR GPERFSKIASKTQSR GPQRFSKIASKTQSR GSQRFSKIASKTQSR GSQRFSKIASKTQSR GSQRFSKIASKTQSR GSQRFSKIASKTQSR GSKGFHPWVDSSTSR GSKGFHPWVDSSTSR GSKGFHPWVDSSTSR GSKGFHPWVDSSTSR GSKGFHPWVDSSTSR GTAPFTTMVSTSQNR GTAPFTTMVSTSQNR GTAPFTTMVSTSQNR GTAPFTTMVSTSQNR GTAPFTTMVSTSQNR GTAPFTTMVSTSQNR GTAPFTTMVSTSQNR GTSRFTTMLSTFANR GTSRFTTMLSTFSNR GTSRFTTMLSTFSNR GTSRFTTMLSTFSNR GTSRFTTMLSTFSNR GTSRFTTMLSTFSNR GTSRFTTMLSTFNN GTSRFTTMLSTFSNR GTSRFTTMLSTFNN GTSRFTTMLSTFNN GTSRFTTMLSTFNN GTSRFTTMLSTPQNR GSAPFSAMVSTPQNR GPAPFSAMVSTPQNN	QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ	0       12         S F D G L D L D WE         G F D G L D L D WE         G F D G L D L D WE         G F D G L D L A WI         G F D G L D L A WI         G F D G L D L A WI         G F D G L D L A WI         G F D G L D L A WI         G F D G L D L A WI         G F D G L D L A WI         G F D G L D L A WI         G F D G L D L A WI         G F D G L D L A WI         G F D G L D L A WI         G F D G L D L A WI         G F D G L D L A WI         G F D G L D L A WI         G F D G L D L A WI         G F D G L D L A WI         G F D G L D L A WI         G F D G L D L A WI         G F D G L D L M WI         G F D G L D L M WI         G F D G L D L M WI         G F D G L D L M WI         G F D G L D L F F I         G F D G L D L F F I         G F D G L D L F F I         G F D G L D L F F I         G F D G L D L F F I         G F D G L D L F F I         G F D G L D L F F I         G F D G L D L F F I         G F D G L D L M WI         M F D G L N L D WI         K F D G L N L D WI	TT       TT         130       130         140       140	000000000000000000000000000000000000
hChito mChito rChito bBP40 cBP40 oBP40 sBP40 hcGP39 rGP39 bGP39 sGP39 hYKL39 hAMCase rAMCase rAMCase cfAMCase cfAMCase hOvi bOvi oOvi sOvi cfOvi maOvi sOvi cfOvi maOvi sOvi rI m2 m2 m3 rYm1 mYm2 mYm3 rYm1 mBc1p2	70       80         LKT       LLAIGGWNFF         LKT       LLAVGGWNFF         LKT       LLSVGGWNFF         LKT       LLSIGGWNFF         LKT       LLSIGGWNFF         LKT       LLSIGGWNFF         LKT       LLSIGGWNFF         LKT       LSIGGWNFF         LKT       LSIGGWNFF         LKT       LSIGGWNFF         LKT       LSIGGWNFF         LKT       LAIGGWNFF         LKT       LAIGGWNFF         LKT       LAIGGWNFF         LKT       LAIGGWNFF         LKT       LAIGGWNFF	90 90 GTQK FTD MVATANNR GTQK FTD MVATANNR GTQK FTD MVATASNR GTQK FTD MVATASNR GSERFSKIASKTQSR GPERFSKIASKTQSR GPQRFSKIASKTQSR GPQRFSKIASKTQSR GSQRFSKIASKTQSR GSQRFSKIASKTQSR GSQRFSKIASKTQSR GSKGFHPWVDSSTSR GSKGFHPWVDSSTSR GSKGFHPWVDSSTSR GSKGFHPWVDSSTSR GSKGFHPWVDSSTSR GTAPFTTMVSTSQNR GTAPFTTMVSTSQNR GTAPFTTMVSTSQNR GTAPFTTMVSTSQNR GTAPFTTMVSTSQNR GTAPFTTMVSTSQNR GTSRFTTMLSTFANR GTSRFTTMLSTFSNR GTSRFTTMLSTFSNR GTSRFTTMLSTFSNR GTSRFTTMLSTFSNR GTSRFTTMLSTFSNR GTSRFTTMLSTFSNR GTSRFTTMLSTFNN GTSRFTTMLSTLANR GTSRFTTMLSTLANR GTSRFTTMLSTLANR GTSRFTTMLSTLANR GTSRFTTMLSTLANR GTSRFTTMLSTLANR GTSRFTTMLSTLANR GTSRFTTMLSTLANR GPAPFSAMVSTPQNR GPAPFSAMVSTPQNR	Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q	0       12         S F D G L D L D WE         G F D G L D L D WE         G F D G L D L D WE         G F D G L D L A WI         G F D G L D L F F I         G F D G L D L F F I         G F D G L D L F F I         G F D G L D L F F I         G F D G L D L F F I         G F D G L D L F F I         G F D G L D L F F I         G F D G L D L F F I         G F D G L N L D W C         K F D G L N L D W C         K F D G L N L D W C         G F D G L N L D W C         G F D G L N L D W C	TT       TT         1       1 <td< th=""><th>000000000000000000000000000000000000</th></td<>	000000000000000000000000000000000000
hChito mChito rChito bBP40 cBP40 sBP40 hcGP39 rGP39 rGP39 bGP39 sGP39 hXKL39 hAMCase rAMCase rAMCase cfAMCase cfAMCase hOvi bOvi oOvi sOvi cfOvi maOvi sOvi cfOvi maOvi sOvi rI m2 m2 m3 rYm1 mYm3 rYm1 mBclp2 mBYm	70       80         LKT       LLAIGGWNF         LKT       LLAVGGWNF         LKT       LLVGGWNF         LKT       LLSVGGWNF         LKT       LLSVGGWNF         LKT       LLSVGGWNF         LKT       LLSVGGWNF         LKT       LLSVGGWNF         LKT       LSVGGWNF         LKT       LAIGGWNF         LKT       LAIGGWNF         LKT       LAIGGWNF         LKT       LAIGGWNF <th>90 90 G TQK F TD MVAT ANNR G TQK F TD MVAT ANNR G TQK F TD MVAT ASNR G TQK F TD MVAT ASNR G S E R F SK IASK TQSR G P E R F SK IASK TQSR G P E R F SK IASK TQSR G P E R F SK IASK TQSR G S Q R F SK IASK TQSR G S G G F HP WVS SS SS ST G TAP F TAMVS TPENR G TAP F TAMVS TPENR G TAP F TAMVS TPENR G TAP F TAMVS TPENR G TSR F TT MLS TFANR G TSR F TT MLS TFSNR G TSR F TT MLS TFNN G TSR F T T MLS T N G TSR F T T MLS T N G TSR F T T MLS T N G T SR F T T M S T N G T SR F T T M S T N G T SR F T T M S T N G T SR F T T M S T N G T SR F T T M S</th> <th>Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q</th> <th>0       12         S F D G L D L D WE         G F D G L D L D WE         G F D G L D L D WE         G F D G L D L A WI         G F D G L D L F F I         G F D G L D L F F I         G F D G L D L F F I         G F D G L D L F F I         G F D G L D L F F I         G F D G L D L F F I         G F D G L D L F F I         G F D G L N L D WC         K F D G L N L D WC         K F D G L N L D WC         G F D G L N L D WC         G F D G L N L D WC         G F D G L N L D WC</th> <th>TT       TT         1       1         <td< th=""><th>000000000000000000000000000000000000</th></td<></th>	90 90 G TQK F TD MVAT ANNR G TQK F TD MVAT ANNR G TQK F TD MVAT ASNR G TQK F TD MVAT ASNR G S E R F SK IASK TQSR G P E R F SK IASK TQSR G P E R F SK IASK TQSR G P E R F SK IASK TQSR G S Q R F SK IASK TQSR G S G G F HP WVS SS SS ST G TAP F TAMVS TPENR G TAP F TAMVS TPENR G TAP F TAMVS TPENR G TAP F TAMVS TPENR G TSR F TT MLS TFANR G TSR F TT MLS TFSNR G TSR F TT MLS TFNN G TSR F T T MLS T N G TSR F T T MLS T N G TSR F T T MLS T N G T SR F T T M S T N G T SR F T T M S T N G T SR F T T M S T N G T SR F T T M S T N G T SR F T T M S	Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q	0       12         S F D G L D L D WE         G F D G L D L D WE         G F D G L D L D WE         G F D G L D L A WI         G F D G L D L F F I         G F D G L D L F F I         G F D G L D L F F I         G F D G L D L F F I         G F D G L D L F F I         G F D G L D L F F I         G F D G L D L F F I         G F D G L N L D WC         K F D G L N L D WC         K F D G L N L D WC         G F D G L N L D WC         G F D G L N L D WC         G F D G L N L D WC	TT       TT         1       1 <td< th=""><th>000000000000000000000000000000000000</th></td<>	000000000000000000000000000000000000

Evolution of Mammalian Chitinases

	موموموموموم		000000	2222222	<b>→</b>	тт
:	140 15	0 160	170	180	190	200
hChito	QDLANAFQQEA	QTSGK.ERLLLSAA	VPAGQTYVDAGY	EVDKIAQNLI	<b>FVNLMAYDF</b>	HGSWEKVTGHNS
rChito	ODLAKAFOEEA	RASGK.SRLLLTAA	VPTGRGHVDAGY	EVDKIVOSL	FINLMAYDF	HSSWDKTTGHNS
bBP40	<b><i>Ř</i>EMKAEFVREA</b>	Q.AGT.EQLLLSAA	VTAGKIAIDRGY	DIAQISRHLI	DFISLLTYDF	HGGWRGTVGHHS
cBP40	KEMKAE <b>FAREA</b>	Q.AGT.ERLLLSAA	VSAGKIAIDRGY	DIAQISRHLI	DFISLLTYDF	HGAWRQTVGHHS
oBP40 sBP40	KEMKAEFIREA(	Q.AGT.EQULLSAA	VSAGKIAIDRGY VSAGKIAIDRGY	DIAQISRHI	DFISLLTYDF FISLLTYDF	HGAWRQTVGHHS HGAWRO TVGHHS
hcGP39	KEMKAEFIKEA	O.PGK.KOLLISAA	LSAGKVTIDSSY	DIAKISQHLI	FISIMTYDF	HGAWRGTTGHHS
mGP39	KELNAE <mark>F</mark> TKEV	Q.PGR.EKLLLSAA	LSAGKVAIDTG <mark>Y</mark>	DIAQIAQH	DFINLMTYDF	HGVWRQITGHHS
rGP39	KELKAEFTKEV	Q.PGT.EKLLLSAA	VSAGKVTLDSGY	DVAQIAQHLI DIAQIAQHLI	DFINLMTYDF	HGTWRHTTGHHS
sGP39	KEMKAEFVREA	L.PGT.ERLLLSGA	VSAGKVAIDRGY	DIAOISCHU	DFISLLTYDF	HGAWROTTGHHS
hYKL39	HELAEAFQKDF	TKSTK.ERLLLTAG	V S A G R Q M I D N S Y	QVEKLAKDLI	FINLLSFDF	HGSWEKPLITGHNS
bykl39	HKLAEAFQQDF	VKSTK.ERLLLTAG	<b>VSA</b> GRQM <b>I</b> D <b>NSY</b>	QIKELAKDLI	DFINLLSFDF	HGSWEKPLVTGHNS
hAMCase	QEMREAFEQEA KEMPEAFEOEA	K.QINKPRUMVTAA T FSNDD <b>DIMVTAA</b>	VAAGISNIQSGY VACCISNIQACY	EIPQLSQYLI	Y I HVMTYDL F THVMTYDI	HGSWEGYTGENS HCSWEG. VTCENS
rAMCase	KELREAFEQEA	I.ESNRPRLMVTAA	VAAGISNIQAGY	EIPELSQYLI	FIHVMTYDL	HGSWDGYTGENS
bAMCase	QETREA <mark>FEQEA</mark> I	K.QTNKP <b>RLLVTAA</b>	VAAGISNIQAGY	EIPQLSQY	DFIHVMTYDF	HGSWEGYTGEN <mark>S</mark>
cfAMCase	QEMREAFEQEA	Q.QINKPRLMITAA	VAAGISNIQSGY	DIPQLSQYLI	YIHVMTYDF FINNISYDI	HASWEGYTGENS
mOvi	EELOFAFEREA	LLTOH.PRLLISAA	VSGIPSIIHTSY	DALLIGREL	FINVLSIDL	HGSWEK. FTGHNS
bOvi	EELLQAFKNEA	QLTMR.PRLLLSAA	VSGDPHVVQKAY	EARLLGRLL	FISVLSYDL	HGSWEKVTGHNS
oOvi	EELLQAFKNEA	QLTMR.P <b>RLLLSAA</b>	VSGDPHVIQKAY	D <b>a</b> r <b>llgr</b> l <b>l</b> i	DFISVLSYDL	HGSWEKVTGHNS
sOvi	EELLLAFRREA	QLTMR.P <b>RULLSAA</b> KLDTR P <b>RULLSAA</b>	VSADPHVIQKAY VSCDPVIIOTAY	DVHLLGKL	DFINVLSYDL FINVLSYDE	HGSWEKVTGHNS
maOvi	EELQFAFEKEA	LLTQR.PRLLLSAA	VSGIPYIIQTSY	DVHLLGRRLI	FINVLSYDL	HGSWEKSTGHNS
mYm1	KEMRKA <mark>F</mark> EEES'	V.EKDIP <b>RLLLTST</b>	GAGIIDVIKSGY	KIPELSQS <mark>L</mark> I	DYIQVMTYDL	HDPKDGY <b>TG</b> EN <mark>S</mark>
mYm2	QEMRKAFEEES'	T.LNHIPRLLLTST	GAGFIDVIKSGY CACLEDVIKSCY	KIPELSOSLI KIPELSOSLI	DYIQVMTYDL	HDPKNGYTGENS
rYm1	OEMRKAFEKES'	T.EOEIPRLLLTAT	VAGVIDTIOSGY	KIPELSOSL	YF OVMTYNL	HDFONG. YTRENS
mBclp2	QKIREA <b>FELEA</b> :	I.ENKSP <b>rlmvtat</b>	<b>VAG</b> VIST <b>I</b> Q <b>SGY</b>	EIPQLSĤFLI	<b>DYI</b> Q <b>VMTYNL</b>	HGSQ́DGY <b>TG</b> EN <mark>S</mark>
mBYm	HEIRKAFEKEV	S. KNKKPRLMVTAA	VAGVISTIQFGY	EIPQLSQSLI	DYIQVMTYDL	HGSWDGYTGENS
	TT 22 210	220 <u>220</u> 230	240	250	TT TT 260	→ TTT 270
hChito	TT <u>20</u> 210 <b>PIYK</b> RQE <b>E</b> SG <b>A</b>	220 220 AASLN <b>V</b> D <b>AAVQQ</b> W <b>L</b>	Q <b>240</b> Q K <b>G T</b> P A <b>S K L I L G</b>	250 MPTYGRSFT	► TT TT 260 ASSSDTRVG	TTT 270 APATGSGTPGPFTK
hChito mChito	TT 22 210 PIYKRQEESGA PIYKRQGESGA	220 220 230 AASLNVD <b>AAVQQ</b> W <b>L</b> AAEQNVD <b>AAVTLWL</b>	QKGTPASKLILG QKGTPASKLILG	250 MPTYGRSFT MPTYGRSFT	TT TT 260 ASSSDTRVG ASSSDNGVG	TTT 270 APATGSGTPGPFTK APATGPGAPGPYTK
hChito mChito rChito bBP40	TT QQ 210 PIYKRQEESGA PIYKRQGESGA PIYKRQGETGK PIFRGNSDGSS	QQQ         QQQQQQQQ           220         230           AASLNVDAAVQQWL         AAEQNVDAAVTLWL           AAEQNVDAAVTLWL         AAEXTLWL           AAEXNVDAAVTLWL         AAEXTLWL           AAEXNVDAAVTLWL         AAEXTLWL	Q QK <mark>GT</mark> PA <mark>SKLILG</mark> QK <b>GT</b> PASKLILG QK <b>GT</b> PASKLMLG RL <b>GA</b> PANKLVMG	250 MPTYGRSFT MPTYGRSFT MPAYGRSFT IPTFGRSYT	TT 260 LASSSDTRVG LASSSDNGVG LASSSDSGVG LASSSSTRVG	TTT 270 APATGSGTPGPFTK APATGPGAPGPYTK APATGPGAPGPYTK APISGPGIPGOFTK
hChito mChito rChito bBP40 cBP40	TT QQ 210 PLYKRQEESGA PLYKRQGESGA PLYKRQGETGK PLFRGNSDGSS PLFRGQEDASS	Q000000000000000000000000000000000000	Q QKGTPASKLILG QKGTPASKLILG QKGTPASKLMLG RLGAPANKLVMG RLGAPANKLVMG	250 MPTYGRSFT MPTYGRSFT MPAYGRSFT IPTFGRSYT IPTFGRSFT	TT 260 ASSSDTRVG ASSSDNGVG ASSSDSGVG ASSSSSSGVG ASSSSTRVG ASSSTRVG	TTT 270 APATGSGTPGPFTK APATGPGAPGPYTK APATGPGAPGPYTK APISGPGIPGQFTK APISGPGIPGRFTK
hChito mChito rChito bBP40 cBP40 oBP40	TT QQ 210 PLYKRQEESGA PLYKRQGESGA PLYKRQGETGK PLFRGNSDGSS PLFRGQEDASS PLFRGQEDASS PLFRGQEDASS	Q000         Q00000000           220         230           AASLNVDAAVTLWL           AAEQNVDAAVTLWL           DAEKNVDAAVTLWL	Q QKGTPASKLILG QKGTPASKLILG QKGTPASKLNLG RLGAPANKLVMG RLGAPANKLVMG RLGAPANKLVMG	250 MPTYGRSFT MPTYGRSFT MPAYGRSFT IPTFGRSYT IPTFGRSFT IPTFCRSFT	TT TT 260 ASSSDTRVG ASSSDNGVG ASSSDSGVG ASSSSDSGVG ASS.KTDVG ASS.KTDVG	TTT 270 APATGSGTPGPFTK APATGPGAPGPYTK APATGPGAPGPYTK APISGPGIPGQFTK APISGPGIPGRFTK APVSGPGVPGRFTK
hChito mChito rChito bBP40 cBP40 oBP40 sBP40 hcGP39	TT QQ 210 PLYKRQEESGA PLYKRQGESGA PLYKRQGETGK PLFRGNSDGSS PLFRGQEDASS PLFRGQEDASS PLFRGQEDASS PLFRGQEDASS	Q000000000000000000000000000000000000	Q QKGTPASKLILG QKGTPASKLILG QKGTPASKLNLG RLGAPANKLVMG RLGAPANKLVMG RLGAPANKLVMG RLGAPANKLVMG RLGAPASKLVMG	250 MPTYGRSFT MPTYGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT	TT TT 260 ASSSDTRVG ASSSDNGVG ASSSDSGVG ASSSTRVG ASS.KTDVG ASS.KTDVG ASS.KTDVG ASS.ETGVG	TTT 270 APATGSGTPGPFTK APATGPGAPGPYTK APATGPGAPGPYTK APISGPGIPGQFTK APVSGPGIPGRFTK APVSGPGIPGRFTK APSGPGIPGRFTK
hChito mChito bBP40 cBP40 oBP40 sBP40 sBP40 hcGP39 mGP39	TT QQ 210 PIYKRQEESGA PIYKRQGESGA PIYKRQGETGK PIFRGNSDGSS PIFRGQEDASS PIFRQCEDASS PIFRQCEDASS PIFRQCEDASS PIFRQCEDASS PIFRQCEDASS	QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ	Q Q K G T P A S K C T P A S K L C A P A N K L M C A P A N K L M G A P A N K L M G A P A N K L M G C C C C C C C C C C C C C	250 MPTYGRSFT MPTYGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGKSFT	TT 260 260 260 260 260 260 260 260	TTT 270 APATGSGTPGPFTK APATGPGAPGPYTK APATGPGAPGPYTK APISGPGIPGQFTK APVSGPGVPGRFTK APVSGPGIPGRFTK APSGPGIPGRFTK APSGPGIPGRFTK APSGPGIPGRFTK
hChito mChito bBP40 cBP40 oBP40 sBP40 sBP40 hcGP39 mGP39 rGP39	TT QQ 210 PIYKRQEESGA PIYKRQGESGA PIFRGQEDGSS PIFRGQEDASS PIFRGQEDASS PIFRQQEDASS PIFRQQEDASS PIFRQQEDASS PIFRQQEDASS PIFRQQEDASS	QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ	Q Q K G T P A S K L L C A P A N K L C A P A N K L V M G A P A N K L V M G A P A N K L V M G C A P A N K L V M G A N K L V M G A N K L V M G A N K L V M G A N K L V M G A N K L V M G A N K L V M G A N K L V M G A N K L V M G A N K L V M G A N K L V M G A N K L V M G N N N N N N N N N N N N N	250 MPTYGRSFT MPTYGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGKSFT IPTFGKSFT IPTFGKSFT IPTFGKSFT	TT T 260 260 260 260 260 260 260 260	TTT 270 APATGSGTPGPFTK APATGPGAPGPYTK APATGPGAPGPYTK APATGPGIPGQFTK APISGPGIPGQFTK APISGPGIPGQFTK APVSGPGIPGQFTK APISGEGLPGRFTK APISGEGLPGRFTK
hChito mChito bBP40 cBP40 oBP40 sBP40 hcGP39 mGP39 rGP39 bGP39 sGP39	TT QQ 210 PIYKRQGESGA PIYKRQGETGK PIFRGNEDGSS PIFRGQEDASS PIFRGQEDASS PIFRGQEDASS PIFRGQEDASS PIFRGQQDTGP PIFRGQQDASS	220 220 230 AASLNVDAAVTLWL AAEQNVDAAVTLWL RFSNADYAVSYML RFSNADYAVSYML RFSNADYAVSYML RFSNADYAVSYML RFSNADYAVSYML DRFSNYDYAVQYML DRFSNYDYGVGYML DRFSNADYAVSYML DRFSNADYAVSYML DRFSNADYAVSYML DRFSNADYAVSYML	Q Q Q K G T P A S K L L L C L C C C C C C C C C C C C C	250 MPTYGRSFT MPTYGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT	TT 260 260 260 260 260 260 260 260 260 260	TTT 270 APATGSGTPGPFTK APATGPGAPGPYTK APATGPGAPGPYTK APATGPGAPGQFTK APISGPGIPGQFTK APISGPGIPGQFTK APISGEGLPGRFTK APISGEGLPGQFTK APISGEGLPGQFTK APISGPGIPGQFTK APASGPGIPGRFTK
hChito mChito bBP40 cBP40 oBP40 sBP40 hcGP39 mGP39 rGP39 bGP39 sGP39 hYKL39	TT 210 PIYK RQEESGA PIYK RQGESGA PIFR QGETGK PIFR GQEDASS PIFR GQEDASS PIFR GQEDASS PIFR GQEDASS PIFR GQEDASS PIFR GQEDASS PIFR GQEDASS PIFR GQEDASS PIFR GQEDASS PIFR GQEDASS	220         230           220         230           AASLNVDAAVTLWL           AAEQNVDAAVTLWL           AFSNADYAVSYML           DRFSNADYAVSYML           .RFSNADYAVSYML	Q Q K G T T C C C C C C C C C C C C C	250 MPTYGRSFT MPTYGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT	TT 260 260 260 260 260 260 260 260 260 260	TTT 270 APATGSGTPGPFTK APATGPGAPGPYTK APATGPGAPGPYTK APATGPGAPGPYTK APATGPGAPGPTF APTSGPGIPGGFTK APVSGPGIPGGFTK APISGEGLPGRFTK APISGEGLPGRFTK APISGPGIPGGFTK APISGPGIPGRFTK APASGPGAAGPITE
hChito mChito rChito bBP40 cBP40 oBP40 sBP40 hcGP39 mGP39 rGP39 bGP39 sGP39 bGP39 sGP39 hYKL39 bYKL39	TT QQ 210 PIYK RQEESGA PIYK RQGESGA PIFR GQEDASS PIFR GQEDASS	220         230           220         230           AASLNVDAAVTIWL         AAEQNVDAAVTIWL           AAEQNVDAAVTIWL         AAEQNVDAAVTIWL           DAEKNVDAAVTIWL         BRSNADYAVSYML           RFSNADYAVSYML         BRFSNADYAVSYML           RFSNADYAVSYML         BRFSNADYAVSYML           DRFSNADYAVSYML         DRFSNADYAVSYML           DRFSNADYAVSYML         DRFSNADYAVSYML           DRFSNADYAVSYML         DRFSNADYAVSYML           DRFSNADYAVSYML         DRFSNADYAVSYML           DRFSNADYAVSYML         DRFSNADYAVSYML           DRFSNADYAVSYML         SYNVEYAVGYMI           DRFSNADYAVSYML         SYNVEYAVGYMI           DRFSNADYAVSYML         SYNVEYAVGYWI	2 2 4 0 2 4 0	250 MPTYGRSFT MPTYGRSFT IPTYGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTYGRSFT	TT 260 260 260 260 260 260 260 260 260 260	TTT 270 APATGSGTPGPFTK APATGPGAPGPYTK APATGPGAPGPYTK APATGPGAPGPYTK APATGPGAPGPYTK APSGPGIPGQFTK APISGPGIPGRFTK APISGPGIPGRFTK APISGPGIPGRFTK APISGPGIPGRFTK APISGPGIPGRFTK APASGPGAAGPITK APASGPGAAGPITK
hChito mChito bBP40 cBP40 oBP40 sBP40 hcGP39 mGP39 rGP39 bGP39 sGP39 bGP39 bGP39 bGP39 hYKL39 hAMCase mAMCase	TT 210 210 PIYKRQEESGA PIYKRQGESGA PIYKRQGETGK PIFRGNSDGSS PIFRGQEDASS PIFRGQESS PIFRGQESS PIFRGQESS PIFRGQESS PIFRGQESS PIFRGQESS PIFRGQESS PIFRGQESS PIFRGQESS PIFRGQESS PIFRGQESS PIFRGQESS PIFRGQESS PIFRGSS PIFRSS PIFRGSS PIFRGSS PIFRGSS PIFRGSS PIFRS	220 220 230 AASLNVDAAVTLWL AAEQNVDAAVTLWL DAEKNVDAAVTLWL RFSNADYAVSYML RFSNADYAVSYML RFSNADYAVSYML RFSNADYAVSYML DRFSNTDYAVGYMI DRFSNVDYAVGYMI DRFSNADYAVSYML DRFSNADYAVSYML DRFSNADYAVSYML DRFSNADYAVSYML DRFSNADYAVSYML SSYYNVEYAVGYWI SSYYNVEYAVGYWI NAYLNVDYVMNYWK	2 2 4 0 2 4 0	250 MPTYGRSFT MPTYGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTYGRSFT IPTYGRSFT FPTYGRNFI	TT 260 260 260 260 260 260 260 260	TTT 270 APATGSGTPGPFTK APATGPGAPGPYTK APATGPGAPGPYTK APATGPGAPGPYTK APATGPGAPGPYTK APISGPGIPGQFTK APISGPGIPGRFTK APISGEGLPGRFTK APISGPGIPGQFTK APISGPGIPGQFTK APASGPGAPITE APASGPGAAGPITK APASGPGAAGPITK APASGPGAAGPYXK APTSGAGPAGPYXK
hChito mChito bBP40 cBP40 oBP40 sBP40 hcGP39 mGP39 rGP39 bGP39 sGP39 bGP39 sGP39 hYKL39 bYKL39 hAMCase rAMCase	TT       QQ         210       Q         PIYK RQEESGA       Q         PIYK RQETGK       Q         PIFRGQEDASS       Q         PIFRGQEDASS       Q         PIFRQEDASS       Q         PIFRQEQEDASS       Q         PIFRQEQEDASS       Q         PIFRQEQEDASS       Q         PIFRQEQEDASS       Q         PIFRQEQEDRS       Q         PIFRQEQEDRS       Q         PIFRQEQEDRS       Q         PIFRQE       Q	QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ	2 2 4 0 2 4 0 2 K GTP A SK LILG 2 K GTP A SK LILG 2 K GTP A SK LING R LGAP A NK LVMG R LGAP A NK LVMG R LGAP A NK LVMG R LGAP A NK LVMG R LGAP A SK LVMG R LGAP A SK LVMG R LGAP A SK LVMG R LGAP A K LVMG N LGAP A K LVMG N NGAP A EK LIVG N NGAP A EK LIVG N NGAP A EK LIVG	250 MPTYGRSFT MPTYGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTYGRSFT IPTYGRSFT FPEYGHTFI FPEYGHTYI	TT 260 ASSSDTRVG ASSSDNGVG ASSSDSGVG ASSSDSGVG ASSS.TRVG ASS.KTDVG ASS.KTDVG ASS.ENQVG ASS.ENQVG ASS.ENQVG ASS.ENQVG ASS.ENQVG SS.ENQVG ASS.ENQVG SS.ENQVG ASS.ENTDVG ASS.ENQVG	TTT 270 APATGSGTPGPFTK APATGPGAPGPYTK APATGPGAPGPYTK APATGPGAPGPYTK APATGPGIPGQFTK APISGPGIPGRFTK APISGPGIPGRFTK APISGPGIPGRFTK APISGPGIPGRFTK APISGACPIPGRFTK APISGACPIPGRFTK APISGACPIPGRFTK APISGACPIPGRFTK APISGACPIPGRFTK APASGPGACPITE APASGPGACPITE APTSGACPACPITAK APTSGACPACPITAK
hChito mChito rChito bBP40 cBP40 oBP40 sBP40 hcGP39 mGP39 rGP39 rGP39 bGP39 sGP39 bGP39 sGP39 bYKL39 hAMCase mAMCase rAMCase	TT       QQ         210       QE         PIYK       RQE       SGA         PIYK       RQE       SGA         PIYK       RQE       SGA         PIYK       RQE       SGA         PIFK       GVE       DSS         PIFR       GVE       DASS         PIFR       GQE       DASS         PIFR       GUE       COS         PIFR       GUE <td< th=""><th>220 220 230 AASLNVDAAVTLWL AAEQNVDAAVTLWL DAEKNVDAAVTLWL DAEKNVDAAVTLWL DRFSNADYAVSYML RFSNADYAVSYML RFSNADYAVSYML DRFSNADYAVSYML DRFSNADYAVSYML DRFSNADYAVSYML DRFSNADYAVSYML DRFSNADYAVSYML DRFSNADYAVSYML DRFSNADYAVSYML DRFSNADYAVSYML NAYLNVDYVMYWK NAYLNVDYVMNYWK NAYLNVDYVMNYWK</th><th>2 4 0 QKGTPASKLILG QKGTPASKLILG QKGTPASKLIMG RLGAPANKLVMG RLGAPANKLVMG RLGAPANKLVMG RLGAPANKLVMG RLGAPASKLVMG RLGAPASKLVMG RLGAPASKLVMG RLGAPASKLVMG RLGAPASKLVMG NKGMPAEKVVMG NKGAPAEKLIVG NNGAPAEKLIVG NNGAPAEKLIVG NNGAPAEKLIVG NNGAPAEKLIVG</th><th>250 MPTYGRSFT MPTYGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTYGHSFT FPTYGHNFI FPEYGHTFI FPEYGHTFI</th><th>TT T TT 260 ASSSDTRVG ASSSDNGVG ASSSDSGVG ASSSSDSGVG ASSS.TRVG ASS.KTDVG ASS.KTDVG ASS.ENQUG ASS.ENQUG ASS.ENQUG ASS.ENQUG ASS.KTDVG SASS.KTDVG ASS.KTDVG ASS.ETAVG ASS.ETAVG</th><th>TTT 270 APATGSGTPGPFTK APATGPGAPGPYTK APATGPGAPGPYTK APATGPGAPGPYTK APISGPGIPGCFTK APISGPGIPGCFTK APSGPGIPGCFTK APISGCGLPGCFTK APISGCGLPGCFTK APISGPGIPGCFTK APISGPGIPGCFTK APASGPGACPITK APASGPGACPYTK APASGPGACPYTR APASGPGACPYTR APTSGACPACPYTR</th></td<>	220 220 230 AASLNVDAAVTLWL AAEQNVDAAVTLWL DAEKNVDAAVTLWL DAEKNVDAAVTLWL DRFSNADYAVSYML RFSNADYAVSYML RFSNADYAVSYML DRFSNADYAVSYML DRFSNADYAVSYML DRFSNADYAVSYML DRFSNADYAVSYML DRFSNADYAVSYML DRFSNADYAVSYML DRFSNADYAVSYML DRFSNADYAVSYML NAYLNVDYVMYWK NAYLNVDYVMNYWK NAYLNVDYVMNYWK	2 4 0 QKGTPASKLILG QKGTPASKLILG QKGTPASKLIMG RLGAPANKLVMG RLGAPANKLVMG RLGAPANKLVMG RLGAPANKLVMG RLGAPASKLVMG RLGAPASKLVMG RLGAPASKLVMG RLGAPASKLVMG RLGAPASKLVMG NKGMPAEKVVMG NKGAPAEKLIVG NNGAPAEKLIVG NNGAPAEKLIVG NNGAPAEKLIVG NNGAPAEKLIVG	250 MPTYGRSFT MPTYGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTYGHSFT FPTYGHNFI FPEYGHTFI FPEYGHTFI	TT T TT 260 ASSSDTRVG ASSSDNGVG ASSSDSGVG ASSSSDSGVG ASSS.TRVG ASS.KTDVG ASS.KTDVG ASS.ENQUG ASS.ENQUG ASS.ENQUG ASS.ENQUG ASS.KTDVG SASS.KTDVG ASS.KTDVG ASS.ETAVG ASS.ETAVG	TTT 270 APATGSGTPGPFTK APATGPGAPGPYTK APATGPGAPGPYTK APATGPGAPGPYTK APISGPGIPGCFTK APISGPGIPGCFTK APSGPGIPGCFTK APISGCGLPGCFTK APISGCGLPGCFTK APISGPGIPGCFTK APISGPGIPGCFTK APASGPGACPITK APASGPGACPYTK APASGPGACPYTR APASGPGACPYTR APTSGACPACPYTR
hChito mChito bBP40 cBP40 oBP40 sBP40 hcGP39 mGP39 rGP39 sGP39 bGP39 sGP39 bYKL39 bYKL39 bYKL39 hAMCase mAMCase rAMCase cfAMCase hOvi	TT       0.0         210       0         PIYK RQEESGA       PIYK RQEESGA         PIYK RQEESGA       PIYK RQEESGA         PIFRGQEDASS       PIFRGQEDASS         PIFR QEDASS       PIFR QEDASS         PIFR VY TOTGS       PIFR QEDASS         PIFR SLPEDPKS       PIFR QEDASS	220 220 230 AASLNVDAAVTLWL AAEQNVDAAVTLWL DAEKNVDAAVTLWL DAEKNVDAAVTLWL CRFSNADYAVSYML RFSNADYAVSYML RFSNADYAVSYML RFSNADYAVSYML DRFSNADYAVSYML DRFSNADYAVSYML DRFSNADYAVSYML DRFSNADYAVSYML DRFSNADYAVSYML DRFSNADYAVSYML SSYYNVEYAVGYMI SSYYNVEYAVGYMI SSYYNVEYAVGYMI SSYYNVEYAVGYMI NAYLNVDYVMNYWK NAYLNVDYVMNYWK NAYLNVDYVMNYWK NAYLNVDYVMNYWK NAYLNVDYVMNYWK NAYLNVDYVMNYWK	2 2 4 0 QK GTP A SK LILG QK GTP A SK LILG QK GTP A SK LILG QK GTP A SK LIMLG R LGAP ANK LVMG R LGAP A SK LVMG R LGAP A SK LVMG N GAP A SK LVMG D N GAP A SK LIVG D N GAP A SK LIVG N GAP A SK LIVG D N GAP A SK LIVG C M P A SK LIVG D N GAP A SK LIVG C M P A SK	250 MPTYGRSFT MPTYGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTYGRSFT FPTYGHNFI FPEYGHTFI FPAYGHTFI IPTYGRTF	TT T 260 260 260 260 260 260 260 260	TTT 270 APATGSSGTPGPFTK APATGPGAPGPYTK APATGPGAPGPYTK APATGPGAPGPYTK APISGPGIPGCPFTK APISGPGIPGCPFTK APSGPGIPGCPFTK APSGPGIPGCPFTK APSGPGIPGCPFTK APSGPGIPGCPFTK APSGPGAAGPYTK APSGPGAAGPYTK APASGPGAAGPYTK APASGPGAAGPYTK APASGPGAAGPYTK APASGPGAAGPYTK APASGPGAAGPYTK APASGPGAAGPYTK APASGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK
hChito mChito rChito bBP40 cBP40 oBP40 sBP40 hcGP39 mGP39 rGP39 bGP39 bGP39 bGP39 bYKL39 bYKL39 bYKL39 hAMCase mAMCase rAMCase cfAMCase hOvi mOvi	TT       0.0         210       0         PIYK RQEESGA       0.0         PIYK RQEESGA       0.0         PIYK RQEESGA       0.0         PIYK RQEESGA       0.0         PIFRGQEDASS       0.0         PIFRGQEDASS       0.0         PIFRQEDASS       0.0         PIFRUS       0.0         PIFRUS       0.0	220         230           220         230           AASLNVDAAVTLWL         AACQNVDAAVTLWL           AAEQNVDAAVTLWL         AAEQNVDAAVTLWL           DAFSNADYAVSYML         RFSNADYAVSYML           RFSNADYAVSYML         RFSNADYAVSYML           RFSNADYAVSYML         RFSNADYAVSYML           DRFSNADYAVSYML         RFSNADYAVSYML           DRFSNADYAVSYML         RFSNADYAVSYML           DRFSNADYAVSYML         SSYNVEYAVGYMI           DRFSNADYAVSYML         SSYNVEYAVGYMI           DRFSNADYAVSYML         SSYNVEYAVGYMI           SSYNVEYAVGYWI         SSYNVEYAVGYWI           SSYNVEYAVGYWI         SSYNVEYAVGYWI           NAYLNVDYVMNYWK         NAYLNVDYVMNYWK           NAYLNVDYVMNYWK         NAYLNVDYVMNYWK           NAYLNVDYVMNYWK         NAYLNVDYVMNYWK           NAYLNVDYVMNYWK         NAYLNVDYVMNYWK           NAYLNVDYVMNYWK         NAYLNVDYVMNYWK           NAYLNVDYVMNYWK         NAYLNVDYVMNYWK	2 2 4 0 QK GTP A SK LILG QK GTP A SK LILG QK GTP A SK LILG QK GTP A SK LIMLG R LGAP ANK LVMG R LGAP ANK LVMG N GAP A K LVMG D N GAP A K LVMG D N GAP A K LVMG D N GAP A K LIVG D N GAP A K LIVG N N GAP A K LIVG C A P A K K K K K K K K K K K K K K K K K	250 MPTYGRSFT MPTYGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTFGRSFT IPTYGRSFT IPTYGRSFT FPYGHNFI FPYGHTFI FPYGHTFI FPAYGHTFI FPAYGHTFI FPAYGRTFR FPTYGRFT	TT 260 260 260 260 260 260 260 260	TTT 270 APATGSGTPGPYTK APATGPGAPGPYTK APATGPGAPGPYTK APATGPGAPGPYTK APATGPGAPGPYTK APISGPGIPGCPFTK APISGPGIPGCPFTK APVSGPGIPGCPFTK APSGPGIPGCPFTK APSGPGIPGCPFTK APISGCCPGPCPTK APSGPGAAGPYTK APSGPGAAGPYTK APASGPGAAGPYTK APASGPGAAGPYTK APASGPGAAGPYTK APASGPGAAGPYTK APASGPGAAGPYTK APASGPGAAGPYTK APASGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK
hChito mChito rChito bBP40 cBP40 oBP40 sBP40 hcGP39 rGP39 bGP39 sGP39 hYKL39 hYKL39 hYKL39 hYKL39 hYKL39 hAMCase rAMCase cfAMCase hOvi mOvi bOvi accident	TT 210 PIYKRQGESGA PIYKRQGESGA PIYKRQGETGK PIFRGQEDASS PIFFRGQEDASS PIFRGQEDASS PIFF	000000000000000000000000000000000000	Q Q Q Q C C C C C C C C C C C C C	2 5 0 MP TYG RSFT MP TYG RSFT MP AYG RSFT IP TFG RSFT IP TYG RSFT IP TYG RSFT IP TYG RSFT IP TYG RSFT IP TYG RSFT IP TYG RTFR FP AYG HTFI IP TYG RTFR FP TYG RTFR FP TYG RTFR	TT T 260 260 260 260 260 260 260 260 260 260	TTT 270 APATGSGTPGPFTK APATGPGAPGPYTK APATGPGAPGPYTK APATGPGAPGPYTK APISGPGIPGGRFTK APISGPGIPGGRFTK APSGPGIPGGRFTK APISGEGLPGRFTK APISGEGLPGRFTK APISGEGLPGRFTK APISGEGLPGRFTK APASGPGAFFTK APASGPGFTK APA
hChito mChito rChito bBP40 cBP40 oBP40 sBP40 hcGP39 mGP39 rGP39 bGP39 bGP39 bGP39 bGP39 bYKL39 hYKL39 hYKL39 hYKL39 hYKL39 bYKL39 hAMCase rAMCase rAMCase cfAMCase cfAMCase hOvi mOvi bOvi oOvi sOvi	TT 210 PIYK RQEESGA PIYK RQGESGA PIYK RQGETGK PIFRGQEDASS PIFRGQEDASS PIFRGQEDASS PIFRGQEDASS PIFRGQQDTGP FRGQCDASS PIFRGQCDASS PIFRGQCDASS PIFRGQCDASS PIFRGQCDASS PIFRGQCDASS PIFRGQCDASS PIFRGQCDASS PIFRGQCDASS PIFRGQCDASS PIFRGQCDASS PIFRGQCDASS PIFRGQCDASS PIFRGQCDASS PIFRGQCDASS PIFRGQCDASS PIFRGQCDASS PIFSLPCDASS PIFSLPEDSKS PIFSLPEDSKS PIFSLPCDPKS PIFSLPCDPKS	000000000000000000000000000000000000	Q Q Q Q Q Q Q Q Q Q Q Q Q Q	2 5 0 MP TYG RSFT MP TYG RSFT MP TYG RSFT IP TFG RSFT IP TYG RSFT IP TYG RSFT IP TYG RTFR FP AYG HTFI IP TYG RTFR FP TYG RTFF IP TYG RTFF IP TYG RTFF	TT T 2 6 0 2 7 6 0 2 6 0 2 7 7 6 0 2 7 6 0 2 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7	TTT 270 APATGSGTPGPFTK APATGPGAPGPYTK APATGPGAPGPYTK APATGPGAPGPYTK APATGPGAPGPYTK APISGPGIPGGRFTK APISGPGIPGGRFTK APISGCLPGRFTK APISGCLPGRFTK APISGCLPGRFTK APISGCLPGPTF APASGPGAPCFTK APASGPGAPT APASGPGAPYTR APASGPGAPYTR APASGPGAPYTR APASGPGAPYTR APASGPGPAGPYTR APTSGAGPAGPYTR APTSGAGPAGPYTR APTSGAGPAGPYTR APTSGAGPAGPYTR APTSGAGPAGPYTR APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPASPGKYTK ACAVGPASPGKYTK
hChito mChito rChito bBP40 cBP40 oBP40 sBP40 hcGP39 mGP39 rGP39 bGP39 sGP39 hYKL39 bYKL39 hAMCase rAMCase cfAMCase cfAMCase cfAMCase hOvi mOvi bOvi oOvi sOvi cfOvi	TT       0.0         210       0         PIYK RQGE SGA       0.0         PIYK RQGE SGA       0.0         PIYK RQGE SGA       0.0         PIFRGQEDASS       0.0         PIFRCQCDASS       0.0         PIFSLECCDASS<	000000000000000000000000000000000000	Q Q Q Q Q Q Q Q Q Q Q Q Q Q	2 5 0 MP TYG RSFT MP TYG RSFT IP TYG RSFT IP TFG RSFT IP TYG RTFH IP TYG RTFH IP TYG RTFH IP TYG RTFH IP TYG RTFH IP TYG RTFH	TT 260 260 260 260 260 260 260 260 260 260	TTT 270 APATGSGTPGPFTK APATGPGAPGPYTK APATGPGAPGPYTK APATGPGAPGPYTK APATGPGAPGPYTK APATGPGAPGPYTK APASGPGIPGGRFTK APISGPGIPGGRFTK APISGCLPGRFTK APISGCLPGRFTK APISGCLPGRFTK APISGCLPGRFTK APISGCLPGRFTK APISGCLPGRFTK APISGCLPGRFTK APISGCLPGRFTK APISGCLPGRFTK APISGCLPGRFTK APISGCLPGRFTK APISGCLPGRFTK APISGCLPGRFTK APISGCLPGRFTK APISGCLPGRFTK APISGCLPGRFTK APISGCCPACPYTR APTSGCCPACPYTR APTSGCCPACPYTR APTSGCCPACPYTK AQAVGPASPGKYTK ACAGPASPGKYTK ACAVGPASPGKYTK ACAVGPASPGKYTK ACAVGPASPGKYTK ACAVGPASPGKYTK
hChito mChito rChito bBP40 cBP40 sBP40 hCGP39 mGP39 rGP39 bGP39 bGP39 bGP39 bGP39 bGP39 bYKL39 hAMCase rAMCase cfAMCase cfAMCase cfAMCase hOvi mOvi bOvi oOvi sOvi	TT 210 PIYK RQEESGA PIYK RQGESGA PIYK RQGETGK PIFRGQEDASS PIFRGQEDASS PIFRGQEDASS PIFRGQEDASS PIFRGQCDASS PIFRGQCDASS PIFRGQCDASS PIFRGQCDASS PIFRGQCDASS PIFRGQCDASS PIFRGQCDASS PIFRGQCDASS PIFRGQCDASS PIFRGQCDASS PIFRGQCDASS PIFRGQCDASS PIFRGQCDASS PIFRGQCDASS PIFRGQCDASS PIFSLPCDASS PIFSLPCDSS PIFSLPCDSS PIFSLPCDPKS PIFSLPCDPKS PIFSLPCDPKS PIFSLPCDPKS PIFSLPCDPKS	000000000000000000000000000000000000	2 2 4 0 QK GTPASKLILG QK GTPASKLILG QK GTPASKLIMEG QK GTPASKLIMEG RLGAPANKLVMG RLGAPANKLVMG RLGAPANKLVMG RLGAPANKLVMG RLGAPASKLIMG RLGAPASKLIMG RLGAPAKLVMG RLGAPAKLVMG CMPSEKVVMG CMPSEKVVMG CMPSEKLIMG QLGAPAEKLIVG CMPAEKLIMG QLGAPAEKLIMG QLGAPPEKLIMG QLGAPPEKLIMG QLGAPPEKLIMG CMPERC CMPPEKLIMG CMPPEKLIMG CMPEC CMPPEKLIMG CMPEC	2 5 0 MP TYC RSFT MP TYC RSFT IP TYC RSFT IP TFC RSFT IP TYC RSFT IP TYC RSFT IP TYC RTF FP AYC RTFH IP TYC RTFH IP TYC RTFH FP TYC RTFH FP TYC RTFH FP TYC RTFH	TTTT260ASASSDTRVGASSDTRVGASSSDSASSSDSASSSTRASSSTRASSSTRASSSTRASSSTRASSSTRASSSTRASSSTRASSTTASSTTASSTTSTTSTTSTTST<	TTT 270 APATGSGTPGPFTK APATGPGAPGPYTK APATGPGAPGPYTK APATGPGAPGPYTK APATGPGAPGPYTK APATGPGAPGPYTK APISGPGIPGQFTK APISGPGIPGCRFTK APISGCLPGRFTK APISGCLPGRFTK APISGCLPGRFTK APISGCLPGRFTK APISGCLPGPTF APISGCPGIPGQFTK APISGCPGIPGQFTK APASGPGACPYTK APASGPGACPYTK APTSGCPACPYTR APTSGCPACPYTR APTSGCPACPYTR APTSGCPACPYTR APTSGCPACPYTR APTSGCPACPYTR APTSGCPACPYTR APTSGCPACPYTR APTSGCPACPYTR APTSCCCPACPYTR APTSCCCCPACPYTR APTSCCCCPACPYTR APTSCCCCPACPYTR APTSCCCCPACPYTR APTSCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
hChito mChito rChito bBP40 cBP40 sBP40 sBP40 hCGP39 rGP39 rGP39 bGP39 sGP39 hYKL39 hAKL39 hAKCase rAMCase cfAMCase cfAMCase cfAMCase hOvi bOvi bOvi sOvi cfOvi maOvi sOvi cfOvi maOvi mYm1 mYm2	TT 210 PIYK RQE ESGA PIYK RQGE SGA PIYK RQGE SGA PIYK RQGE TGK PIFR GQE DASS PIFR GQE CASS PIFR GQE DASS PIFR GQE CASS PIFR GQE CAS	220         230           220         230           AASLNVDAAVTIWL         AAEQNVDAAVTIWL           AAEQNVDAAVTIWL         AAEQNVDAAVTIWL           AFSNADYAVSYML         AAFSNADYAVSYML           RFSNADYAVSYML         BRFSNADYAVSYML           RFSNADYAVSYML         BRFSNADYAVSYML           DRFSNADYAVSYML         SYNVEYAVGYWI           DRFSNADYAVSYML         SYNVEYAVGYWI           DRFSNADYLNYW         SYNW           SYNVEYAVGYWI         SSYNW           SYNVEYAVGYWI         SSYNW           SYNVEYAVGYWI         SSYNW           SYNVEYAVGYWI         SSYNW           SYNVEYAVGYWI         SSYNW           SYNVEYAVGYWI         SSYNW           SAYAMNYWR         SAYAMNYWR           SAYAMNYWR         SAYAMNYWR           SAYAMNYWR         SAYAMNYWR           SAYAMNYWR         SAYAMNYWR           SAYAMNYWR         SAYAMNY	Q Q Q Q Q Q Q Q Q Q Q Q Q Q	250 MPTYGRSFT MPTYGRSFT IPTYGRSFT IPTFGR	TT 260 260 260 260 260 260 260 260	TTT 270 APATGSGTPGPFTK APATGPGAPGPYTK APATGPGAPGPYTK APATGPGAPGPYTK APATGPGIPGQFTK APTSGPGIPGQFTK APISGPGIPGQFTK APISGPGIPGQFTK APISGCPGIPGQFTK APISGPGIPGQFTK APISGPGIPGQFTK APISGPGIPGQFTK APISGAGPAGPYTK APASGPGAAGPYTK APASGPGAAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPASPGKYTK AQAXGPASPGKYTK AQAXGPASPGKYTK AQAYGPASPGKYTK AQAYGPASPGKYTK AQAYGPASPGKYTK AQAYGPASPGKYTK AQAYGPASPGKYTK AQAYGPASPGKYTK
hChito mChito rChito bBP40 cBP40 sBP40 sBP40 hCGP39 mGP39 rGP39 bGP39 sGP39 hYKL39 hAMCase rAMCase rAMCase cfAMCase cfAMCase hOvi bOvi bOvi sOvi cfOvi maOvi sOvi cfOvi mYm1 mYm2 mYm3	TT 210 PIYK RQE ESGA PIYK RQGE SGA PIYK RQGE TGK PIFR GQEDASS PIFR	220         230           220         230           AASLNVDAAVTIWL         AAEQNVDAAVTIWL           AAEQNVDAAVTIWL         AAEQNVDAAVTIWL           AFSNADYAVSYML         DAFSNADYAVSYML           RFSNADYAVSYML         DRFSNADYAVSYML           DRFSNADYAVSYML         NAYLNVDYWWWWWWWW           NAYLNVDYVMNYWK         NAYLNVDYVMNYWK           NAYLNVDYVMNYWK         NAYLNVDYVMNYWK           NAYLNVDYVMNYWK         SAYAMNYWR          SAYAMNYWR        SAYAMNYWR          SAYAMNYWR        SAYAMNYWR          SAYAMNYWR        SAFAMNYWR          SAFAMNYWR        SAFAMNYWR          SAFAMNYWR        SAFAMNYWR          SAFAMNYWR        SAFAMNYWR          SAFAMNYWR        SAFAMNYWR          SAFAMNYWR        SAFAMNYWR          SAFAMNYWR        SADLNVDSIISYWK		250 MPTYGRSFT MPTYGRSFT IPTFGR	TT 260 260 250 250 250 250 250 250 250 25	TTT 270 APATGSGTPGPYTK APATGPGAPGPYTK APATGPGAPGPYTK APATGPGAPGPYTK APATGPGIPGCFTK APATGPGIPGCFTK APISGPGIPGCFTK APISGPGIPGCFTK APISGCLPGCFTK APISGCLPGCFTK APISGCLPGCFTK APISGCLPGCFTK APISGCLPGCFTK APISGAGPAGPYTK APASGPGAAGPYTK APASGPGAAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK APTSGAGPAGPYTK ACANGPASPGKYTK ACANGPASPGKYTK ACANGPASPGKYTK ACANGPASPGKYTK ACANGPASPGKYTK ACANGPASPGKYTK ACANGPASPGKYTK ACANGPASPGCYTT
hChito mChito rChito bBP40 cBP40 sBP40 sBP40 hCGP39 mGP39 rGP39 bGP39 sGP39 hYKL39 hAMCase rAMCase rAMCase rAMCase cfAMCase hOvi bOvi bOvi sOvi cfOvi mOvi sOvi cfOvi mAVI sOvi rMN1 mYm1 mYm3 rYm1	TT       0.0         210       0         PLYK RQGE SGA       0.0         PLYK RQGE SGA       0.0         PLYK RQGE SGA       0.0         PLYK RQGE SGA       0.0         PLFR GQE DASS       0.0         PLFS LPE DASS <td< th=""><th>220       230         220       230         AASLNVDAAVTIWL         AAEQNVDAAVTIWL         AAEQNVDAAVTIWL         DAEKNVDAAVTIWL         RFSNADYAVSYML         DRFSNADYAVSYML         NAYLNVDYWNYWK         NAYLNVDYVMNYWK         NAYLNVDYVMNYWK         NAYLNVDYVMNYWK         NAYLNVDYVMNYWK         NAYLNVDYVMNYWK        SAYAMNYWR        SAYAMNYWR        SAYAMNYWR        SAYAMNYWR        SAYAMNYWR        SAYAMNYWR        SAFAMNYWR        SAFAMNYWR        SAFAMNYWR        SAFAMNYWR        </th><th></th><th>250 MPTYGRSFT MPTYGRSFT IPTFGR</th><th>TT 260 260 250 250 250 250 250 250 250 25</th><th>TTT         270         AP ATG SGTPCPYTK         APATGPGAPCPYTK         APATGPCPGAPCPYTK         APATGPCAPCPYTK         APATGPCAPCPYTK         APATGPCAPCPYTK         APATGPCAPCPYTK         APATGPCAPCPYTK         APATGPCAPCPYTK         APATGPCAPCPYTK         APATGPCAPCPS         APTSGPCAPCPS         APASCPCAPCPS         APASCPCAPS         APASCPCASPCS         APASCPCASPCS         APASCPCASPCS         APASCPCASPCS         APASCPCASPCS</th></td<>	220       230         220       230         AASLNVDAAVTIWL         AAEQNVDAAVTIWL         AAEQNVDAAVTIWL         DAEKNVDAAVTIWL         RFSNADYAVSYML         DRFSNADYAVSYML         NAYLNVDYWNYWK         NAYLNVDYVMNYWK         NAYLNVDYVMNYWK         NAYLNVDYVMNYWK         NAYLNVDYVMNYWK         NAYLNVDYVMNYWK        SAYAMNYWR        SAYAMNYWR        SAYAMNYWR        SAYAMNYWR        SAYAMNYWR        SAYAMNYWR        SAFAMNYWR        SAFAMNYWR        SAFAMNYWR        SAFAMNYWR		250 MPTYGRSFT MPTYGRSFT IPTFGR	TT 260 260 250 250 250 250 250 250 250 25	TTT         270         AP ATG SGTPCPYTK         APATGPGAPCPYTK         APATGPCPGAPCPYTK         APATGPCAPCPYTK         APATGPCAPCPYTK         APATGPCAPCPYTK         APATGPCAPCPYTK         APATGPCAPCPYTK         APATGPCAPCPYTK         APATGPCAPCPYTK         APATGPCAPCPS         APTSGPCAPCPS         APASCPCAPCPS         APASCPCAPS         APASCPCASPCS         APASCPCASPCS         APASCPCASPCS         APASCPCASPCS         APASCPCASPCS
hChito mChito rChito bBP40 cBP40 sBP40 sBP40 hCGP39 mGP39 rGP39 bGP39 sGP39 bYKL39 bYKL39 bYKL39 bYKL39 hAMCase rAMCase rAMCase cfAMCase cfAMCase cfAMCase cfAMCase cfAMCase cfAMCase cfAMCase cfAMCase cfAMCase rAMCA rAMC rAMC rAMC rAMC rAMC rAMC rAM	TT 210 PIYK RQEESGA PIYK RQGESGA PIYK RQGETGK PIFRGQEDASS PIFRGQES PIFRGQEDASS PIFRGQEDASS PIFRGQEDASS PIFRGQEDASS PIFRGQEDASS PIFRGQEDASS PIFRGQEDASS PIFRGQEDASS PIFRGQEDASS PIFRGQES	220 220 230 AASLNVDAAVTLWL AAEQNVDAAVTLWL DAEKNVDAAVTLWL DAEKNVDAAVTLWL DRFSNADYAVSYML CRFSNADYAVSYML		250 MPTYGRSFT MPTYGRSFT IPTFGR	TT 260 260 250 250 250 250 250 250 250 25	TTT 270 AP ATG SGTP GP FTK AP ATG PGAP GP YTK AP ATG PGAP GP YTK AP ATG PGAP GP YTK AP ATG PGAP GP YTK AP ISG PGIP GR FTK AP ISG PGIP GR FTK AP ISG PGIP GR FTK AP ISG PGIP GR FTK AP ISG CG PG CR FTK AP ISG CG PG PG CR FTK AP ISG CG PG PG CR FTK AP ISG CG PG PG CR FTK AP SG PG PG PG CR FTK AP SG CG PG PG CR FTK AP SG CG PG PG CR FTK AP SG CG PG CR FTK AP SG CG PG CR FTK AP SG CG CR FTK AP SG CG CR CR TTK AP SG CG CR

FIGURE 1.—Continued.

# 964

# A. P. Bussink et al.

	TT-→QQ	e T	. T	<b>→</b> TT.	т —	>TT-		<u>ووو</u>	عووو	معمعهم		•	
	280	2	90		.00	3.	ιċ	3.	20	330		340	?
hChito	EGGMLAYY	OVCSWK	.GA	TKQRIQD.	QKVP	YIFRDNO	QWVGI	FDDVESI	KTKV	SYLKQKG	LG <mark>GA</mark> MV	ÀALI	)LDD
mChito	DKGVLAYY	DACSWK		ERHRIED.	QKVP	YAFQDNO	QWVSB	TDDVESI	KA <b>KA</b> .	AYLKQKG	LG <mark>GA</mark> MV	<b>VL</b> I	) L D D
rChito	EKGILVYF	DVCSWK		GKQRIED.	QKVP	YVSQGN	QWVGI	DDRES	RAKA.	AYVKQKG	LG <b>GA</b> MV	ŴILI	GDD
bBP40	EKGILAYY	JICDFL	HGA	TTHRFRD.	QQVP	YATKGN	QWVAY	YDDQE <mark>S</mark> Y	KNKA	<b>R</b> Y <b>LK</b> N <b>R</b> Q	LAGAMV	ÀALI	)LDD
cBP40	EKGILAYY	ICDFL	HGA	TTHRFRD.	QQVP	YATKGN	QŴVAY	YDDQE <mark>S</mark> Y	KNKA	<b>R</b> Y <b>LK</b> N <b>R</b> Q	LA <b>GA</b> MV	ÀALI	)LDD
oBP40	EKGILAYY	JICDFL	HGA	TTHRFRD.	QQVP	YATKGN	QWVAY	YDDQES	KNKA	<b>R</b> Y <b>LK</b> N <b>R</b> Q	LA <b>GA</b> MV	ÀALI	)LDD
sBP40	EKGILAYY	ICDFL	QGA	TTHRFRD.	QQVP	YATKGN	QWVAY	YDDQE <mark>S</mark> Y	KNKA	<b>R</b> Y <b>LK</b> N <b>R</b> Q	LAGAMV	<b>ALI</b>	)LDD
hcGP39	EAGTLAYY	ICDFL	RGA	TVHRTLG.	QQVP	YATKGN	QWVGY	YDDQES	/KS <mark>KV</mark>	QYLKDRQ	LA <mark>GA</mark> MV	ÀALI	DDJC
mGP39	EAGTLAYY	JICDFL	KGA	EVHRLSN.	EKVP	FATKGN	QWVGX	YEDKES	KNKV	GFLKEKK	LA <b>GA</b> MV	ÀALI	)LDD
rGP39	EKGTLAYY	JICDFL	RGA	EVHRILG.	QQVP	FATKGN	QWVGX	YDDPES	KNKV	<b>K</b> Y <b>LK</b> N <b>K</b> Q	LAGAMV	ÂVI	DDLCD
bGP39	EKGILAYY	ICDFL	HGA	ITHRFRD.	QQVP	YATKGN	QWVAY	YDDQES	KNKA	RYLKNRQ	LA <mark>GA</mark> MV	ÀALI	DDJC
sGP39	EKGILAYY	JICDFL	QGA	TVRRPLG.	QQVP	YATKGN	QWVGX	YDDQES	KNKA	KYLKSRQ	LA <mark>GA</mark> MV	ŇΤLΓ	)LDD
hYKL39	SSGFLAYY	ICQFL	KGA	KITRLQD.	QQVP	YAVKGN	QWVGY	YDDVK <mark>S</mark> i	4 E T K V	QF <b>LK</b> N <b>L</b> N	LG <mark>GA</mark> MI	ŴSII	) MDD
bYKL39	SSGFLAYY	ICQFL	QGA	KITRLQD.	QQVP	YAVKGN	QWVGY	YDDVES	/ETKV	QF <b>LK</b> N <b>L</b> N	LG <mark>GA</mark> MI	ŴSII	) M D D
hAMCase	ESGIWAYY	JICTFL	KNG	ATQGWDAP	QEVP	YAYQGN	VWVGY	YDNIKSI	DIKA	QWLKHNK	F G <mark>G A</mark> M V	ΩAI Γ	)LDD
mAMCase	QAGFWAYY	JICTFL	RSG	ATEVWDAS	QEVP	YAYKANI	EWLGY	YDNIK <mark>S</mark> I	SVKA	QWLKQNN	FG <mark>GA</mark> MI	ŶΑΙΓ	)LDD
rAMCase	QAGFWAYY	ICTFL	RNG	ATQDWDAP	QEVP	YAYKGNI	EWVGY	YDNIKSI	SVKA	QWLKQNN	FG <mark>GA</mark> MI	ŶΑΙΓ	)LDD
bAMCase	EAGFWAYY	ICAFL	KDG	ATEAWDDS	QNVP	YAYKGTI	EWVGY	Y D N V N S I	RIKA	QWLKENN	F G <mark>G A</mark> M V	I I	)LDD
cfAMCase	QAGFWAYY	JICTFL	KNG	ATQAWDAP	QDVP	YAYQGNI	EWVGY	YDDVK <mark>S</mark> I	GIKA	QWLKENN	F G <mark>G A</mark> M V	I I	)LDD
hOvi	QEGFLAYF	ICSFV	WGA	KKHWIDY.	QYVP	YANKGKI	EWVGY	YENAISI	SYKA	WFIRRH	F G <mark>G A</mark> M V	ŴΤLΓ	) M D D
mOvi	QAGFLAYY	<b>VC</b> SFV	QRA	KKHWIDY.	QYVP	YAFKGKI	EWLGY	YDDTISI	SYKA	MY <b>VK</b> R <b>E</b> H	FG <b>GA</b> MV	ATLI	) MDD
bOvi	QAGFLAYY	JCCFV	RRA	KKRWIND.	QYVP	YAFKGKI	EWVGY	YDDAISI	GYKA	FFIKREH	F G <mark>G A</mark> M V	NTLI	)LDD
oOvi	QAGFLAYY	<b>VC</b> SFV	QRA	KKRWIND.	QYVP	YAFKGKI	EWVGY	YDDAI <mark>S</mark> I	GYKA	FFIKREH	F G <mark>G A</mark> M V	ØΤLΓ	) L D D
sOvi	QAGFLAYY	VCSFV	QRA	KKRWIDH.	QYVP	YAYRGKI	EWVGY	YDDDI <mark>S</mark> I	SYKA	FFIKKEH	F G <mark>G A</mark> M V	ATLI	)LDD
cfOvi	HPGFLAYY	ICSFL	QRA	TKRWIDF.	QQVP	YAYKGK	VWVGY	YDDANSI	SSKA	MFIKEEH	F G <mark>G A</mark> M V	ΠLΓ	)LDD
maOvi	QAGFLAYY	VCSFI	QRA	EKHWIDH.	QYVP	YAYKGKI	EWVGY	YDDAV <mark>S</mark> I	SYKA	MF <b>VK</b> K <b>E</b> H	F G <mark>G A</mark> M V	ΠLΓ	) M D D
mYm1	ESGLLAYY	JVCTFL	NEG	ATEVWDAP	QEVP	YAYQGNI	EWVGY	Y D N V R SI	KLKA	QWLKDNN	LG <mark>GA</mark> VV	PLI	) M D D
mYm2	EQGLLAYF	JICTFL	NEG	ATEIFDAI	QEVP	YAYLGNI	EWVGY	Y D N V R S I	KLKA	QWLKDNN	LG <mark>GA</mark> VV	ΫΡLΓ	) M D D
mYm3	VPGLLAYY	JICTFL	NEG	ATEVWDAP	QEVP	YAYQGNI	EWVGY	Y D N V R S I	KLKA	QWVKDNN	LG <mark>GA</mark> VV	ΫPLΓ	) M D D
rYm1	EAGLWAYY	JICTFL	NDG	ATDLWDGP	QEVP	YAVQGN	VWVG	YDNVKSI	KIKA	QWLKDNK	LG <b>GA</b> MV	PLI	) M D D
mBclp2	ESGTWAYY	ICSFL	NDG	ATEAWDSA	QEVP	YAYQGNI	KŴVGY	YDNVKSI	RIKA	EWLKQNN	LG <mark>GA</mark> ML	ŴΤLΓ	) M D D
mBYm	QTGFWAYY	JCTFL	KNG	AIQVWNAA	QQVP	YAFHGNI	EWVGY	YDNIK <mark>S</mark> I	HIKA	QWLKRNN	YG <b>GA</b> MI	ÄΤΙΓ	) M D D
rBYm	KTGFWAYY	JICAFL	KNG	AIQVWNAA	QQVP	YAFHGNI	EWVGY	YDNVKSI	HIKA	QWLKNNN	FG <b>GA</b> MI	AIC	GMDD
		2											

	тт										Q	Q	Q	Q	Q	Q	Q	Q	Q	Q
			Э	5	ò									3	6	o				
hChito	FA	G	FS	C	Ν	Q	G		R	Y	₽	L	I	Q	т	L	R	Q	Е	L
mChito	FK	G	SF	C	Ν	Q	G		Ρ	Y	P	L	Ι	R	т	L	R	Q	Е	L
rChito	FK	G	SF	C	Ν	Е	G		Q	Y	₽	L	Ι	R	т	L	н	Q	Е	L
bBP40	FR	G	ТF	C	G	Q	Ν	L	Т	F	P	L	т	S	А	I	ĸ	D	v	L
cBP40	FR	G	ТF	C	G	Q	Ν	L	Т	F	P	L	т	S	А	v	ĸ	D	v	L
oBP40	FR	G	ТF	C	G	Q	Ν	L	Т	F	P	L	т	S	A	v	ĸ	D	v	L
sBP40	FR	G	ТF	C	G	Q	Ν	L	Т	F	P	L	т	S	A	v	ĸ	D	v	L
hcGP39	FQ	G	SE	C	G	Q	D	L	R	F	P	L	т	Ν	A	Ι	ĸ	D	A	L
mGP39	FQ	G	. Т	C	Q	Ρ	K	E	F	F	P	L	т	N	A	I	ĸ	D	A	L
rGP39	FR	G	SF	C	G	H	Ν	V	Η	F	P	L	т	N	А	Ι	ĸ	Е	Α	L
bGP39	FR	G	ТF	C	G	Q	Ν	L	A	F	P	L	т	S	A	Ι	ĸ	D	v	L
sGP39	FR	G	N F	C	G	Q	Ν	L	R	F	P	L	т	S	A	I	ĸ	D	v	L
hYKL39	FΤ	G	KS	С	Ν	Q	G	•	Ρ	Y	₽	L	v	Q	A	v	ĸ	R	S	L
bYKL39	FΤ	G	ΚI	C	S	Q	G	•	Ρ	Y	₽	L	v	Q	A	v	ĸ	R	S	L
hAMCase	FΤ	G	ТF	C	Ν	Q	G	•	Κ	F	₽	L	Ι	S	т	L	ĸ	K	A	L
mAMCase	FΤ	G	SF	C	D	Q	G	•	Κ	F	P	L	т	S	т	L	N	K	A	L
rAMCase	FΤ	G	SF	C	D	Q	G	•	Κ	F	₽	L	т	S	т	L	N	K	A	L
bAMCase	FΤ	G	ТF	C	Ν	Q	G	•	Κ	F	P	L	Ι	N	т	L	ĸ	D	Α	L
cfAMCase	FΤ	G	ТF	C	Ν	Q	G	•	Κ	F	₽	L	v	Ν	т	L	ĸ	K	A	L
hOvi	VR	G	ТF	C	G	т	G	•	Ρ	F	₽	L	v	Y	v	L	N	D	Ι	L
mOvi	VR	G	ΤF	С	G	N	G	•	Ρ	F	₽	L	V	Η	Ι	L	N	Е	L	L
bOvi	FR	G	ΥF	C	G	т	G	•	Ρ	F	₽	L	v	Η	т	L	N	Ν	L	L
oOvi	FR	G	N F	С	G	т	G	•	Ρ	F	P	L	A	Η	т	L	N	Ν	L	L
sOvi	VR	G	ΤF	C	G	т	G	•	Ρ	F	P	L	v	Y	М	L	N	D	L	L
cfOvi	AK	G	ΤF	C	R	т	G	•	Ρ	F	P	L	v	Η	ĸ	L	н	S	L	L
maOvi	VR	G	ΤF	C	G	N	G	•	Ρ	F	P	L	v	Η	Ι	L	N	Ε	L	L
mYm1	FS	G	SF	C	Η	Q	R	•	Η	F	P	L	т	S	т	L	ĸ	G	D	L
mYm2	FS	G	SF	C	Η	Q	G	•	R	F	P	L	т	Τ	т	L	ĸ	R	D	L
mYm3	FS	G	SF	C	Η	Q	G	•	R	F	P	L	т	S	т	L	ĸ	R	D	L
rYm1	FΤ	G	SF	C	Q	Q	G	•	R	F	P	L	т	S	т	L	ĸ	Η	Y	L
mBclp2	FΤ	G	SE	C	Ν	Q	G	•	Q	F	P	L	Т	S	Т	L	K	Ν	A	L
mBYm	ΥT	G	SF	C	G	Q	G	•	Т	F	P	L	т	S	Ι	L	ĸ	K	т	L
rBYm	ΥT	G	SF	C	D	Q	G	•	Ρ	F	P	L	т	S	Т	L	ĸ	Ν	Α	L
				2																

FIGURE 1.—Continued.

TABLE 1	
Theoretically determined isoelectric point	nts

	Chito	AMCase	GP39/BP40	YKL39	Oviductin	Ym1/2/3 and Bclp2	BYm
H. sapiens	6.52	5.42	8.65	7.24	9.17	_	_
P. troglodytes	7.10		8.65	7.24	9.17	_	
Macaca mulatta	7.76	4.90	8.19	6.83	9.17	_	
M. musculus	5.94	5.14	8.65		8.49	5.16, 5.45, 5.62, 5.16	8.92
R. norvegicus	8.49	5.02	8.98			5.67	8.82
B. taurus		5.12	8.65, 8.80	8.89	9.41	_	
Capra hircus			8.93	_		_	
Ovis aries			8.66	_	9.44	_	_
Sus scrofa			9.14, 8.92	_	9.22	_	
Canis familiaris		4.68		_	9.33	_	
Mesocricetus auratus		—	—	—	8.71	—	—

chitinase(-like) protein-coding genes. However, it does not provide information on selective forces. Since the AMCases have evolved to function in an acid environment, it might be hypothesized that there have been episodes of positive selection after the initial gene duplication to allow for rapid adaptation, in analogy to mammalian stomach lysozymes (MESSIER and STEWART 1997; YANG 1998). In addition, the repeated occurrence of the chi-lectin (loss of enzymatic function) mutations suggests site-specific positive selection. To evaluate more closely such selective forces, substitution rate ratios of nonsynonymous vs. synonymous mutations  $(d_N/d_S, \omega)$ have been calculated with PAML. Values <1 indicate the occurrence of purifying (negative) selection, i.e., elimination of mutations that would result in a change in protein composition, whereas values >1 indicate a selective pressure to maintain the changes in the protein (positive or adaptive selection).

 $\omega$ -Values were first calculated using the branch model to assess selection within clades. As a positive control, a data set consisting of seven primate lysozyme sequences was used, in which the occurrence of positive selection had previously been determined (MESSIER and STEWART 1997; YANG 1998). Although the results obtained by YANG (1998) could be exactly reproduced, applying various similar models to the chitinase family data set did not reveal episodes of adaptive selection. For the most parameter-rich model used, in which, in addition to a single  $\omega$  for each ortholog, every ancestral branch was allowed a different  $\omega$ , all  $\omega$ -values were found to be substantially <1, thus giving no direct support for the occurrence of positive selection. Table 2 shows free ingroup ω-values for all orthologs using this model. Values for ancestral branches ranged from 0.11 to 0.44 in all models.

It is conceivable that the chitinase(-like) proteins contain constrained amino acid sites subjected to purifying selection with  $\omega$  close to zero as well as sites that could be subjected to positive selection. A large number of constrained sites would mask a signal of positive selection when the  $\omega$ -values are averaged over all sites.  $\omega$ -Values were therefore also calculated using the so-called site model. SWANSON *et al.* (2001) had to employ a similar approach to identify a few amino acid positions in oviductins that were subjected to positive selection. An analysis using a data set consisting only of oviductin sequences confirmed the reported findings by SWANSON *et al.* (2001) but expansion of the data set with other family members did not identify amino acid positions in other chitinase(-like) proteins subjected to positive selection.

Genomic synteny of chitotriosidase and AMCase loci in humans and rodents: The chromosomal location of all chitinase(-like) protein-coding genes in mice and humans has been mapped. The chitotriosidase locus (1q32 in humans and 1F4 in mice) encoding chitotriosidase and GP39 and flanked by the genes coding for adenosine A1 receptor (ADORA1) and fibromodulin is syntenic (BOOT et al. 2005a). In Figure 3 the synteny analysis is extended to the AMCase loci, corresponding to locus 1p13 in humans and 3F3 in mice. Again, the human and murine regions are flanked by the same genes, in this case coding for adenosine A3 receptor and transmembrane protein 77. The presence of an ADORA paralogue on chitotriosidase and AMCase loci suggests that the genes encoding the two active chitinases result from a large-scale duplication.

The AMCase locus reveals major differences between mice and humans. First, additional open reading frames exist in the mouse genome encoding Ym1, Ym2, Bclp2, and BYm (encoded by Chi3l3, Chi3l4, BCLP2, and LOC229688, respectively), whereas these genes are absent from human chromosome 1p13. The opposite holds true for Chi3L2 encoding YKL39. Second, additional AMCase-like pseudogenes, LOC728204 and LOC149620, are present on human chromosome 1. Finally, the orientation of many genes in the human and mouse AMCase loci differs (see Figure 3), suggesting the occurrence of multiple and diverse recombination events.



\_\_\_\_0.1\_\_\_10 expected substitutions / 100 nucleotides

It is of interest to note that the chi-lectin gene YKL39 is part of the AMCase locus in humans and cows, but, on the basis of both phylogenetic analyses and protein features, YKL39 likely results from a gene duplication event in the chitotriosidase locus. Apparently, in the case of the YKL39 gene, an additional rearrangement occurred.

FIGURE 2.—Maximum-likelihood tree of mammalian chitinase(-like) genes. The number on every node indicates the support bootstrap value for the likelihood analyses of a thousand replicate data sets. Only values other than 1000 are shown. The tree was rerooted by an outgroup (ceCht-1, *Caenorhabditis elegans* chitinase-1). Abbreviations as in Figure 1. For further details, see text.

## DISCUSSION

Our investigation rendered new insights into the evolutionary relationships of mammalian chitinase (-like) proteins. Both phylogenetic analyses and genomic synteny point to the same evolution of mammalian family 18 chitinase proteins (Figure 4). First, a gene

TABLE 2

Substitution rates $d_N$ , $d_S$	s, and their ratio $\omega$ (	$(d_{\rm N}/d_{\rm S})$ for eac	h orthologous group
----------------------------------	-------------------------------	---------------------------------	---------------------

	Chito	AMCase	GP39/BP40	YKL39	Oviductin	Ym1/2/3 and Bclp2	BYm
$d_{ m N}$	0.11	0.11	0.17	0.17	0.20	0.04	0.11
$d_{\rm S}$	0.48	0.61	0.60	0.83	0.82	0.10	0.29
$\omega (d_{\rm N}/d_{\rm S})$	0.23	0.18	0.28	0.21	0.25	0.37	0.38



FIGURE 3.—Synteny of loci encoding chitinase(-like) proteins between mice and humans. Schematic overview of the synteny of mouse locus 1F4 with human 1q32 and mouse locus 3F3 with human 1p13. The orientation and position of the genes are indicated with arrows. The genes of members of the chitinase protein family are depicted by solid arrows, whereas the other genes in the loci are depicted by shaded arrows.



FIGURE 4.—Overview of the evolution of chitinase(-like) genes. •, the "ancestral" gene duplications;  $\bigcirc$ , rodent-specific gene duplication;  $\square$ , artiodactyle-specific gene duplication; a cross indicates the loss of catalytic activity mutations. "Chitolectins" are chi-lectins evolved from the chitotriosidase gene (duplication).

duplication event allowed the specialization of two active chitinases, chitotriosidase and AMCase. Duplications of both chitotriosidase and AMCase genes, followed by loss-of-enzymatic-function mutations, led to the subsequent evolution of chi-lectins.

The duplication of the active chitinase most likely has an ancient origin as Xenopus already has two active chitinases, one of which, like AMCase, is expressed in the stomach (Fujiмото et al. 2002). Indeed, the Xenopus AMCase homolog clusters with the mammalian AMCases (see supplemental Figure 2 at http://www.genetics.org/ supplemental/). This suggests that the gene duplication allowing evolution of chitotriosidases and AMCases occurred very early in tetrapod evolution in the wake of the development of the acidic stomach. The evolution of the various mammalian chi-lectins is most likely a more recent event. The existence of a molluscan chilectin remarkably homologous to Hcgp39 has been reported (BADARIOTTI et al. 2006). However, the molluscan sequence does not clade within the GP39 group (results not shown), suggesting that it is a product of an independent gene duplication and loss-of-function mutation. BLAST searches on the genome of the nematode C. elegans reveal several genes likely to encode chi-lectins, yet phylogenetic analyses again show none of them to group within the tree (results not shown). Recently, the occurrence in several plant species of chi-lectins homologous to chitinases has also been reported (VAN DAMME et al. 2006). Mutations resulting in loss of chitinolytic activity have apparently occurred independently in a variety of lineages.

It is theoretically conceivable that the various types of chi-lectins in mammalian species have interdependently evolved by concerted evolution, a process driven by unequal crossover and gene conversion (reviewed in NEI and ROONEY 2005). The lack of conservation of gene orientation within the AMCase locus (Figure 3) indeed suggests that recombination has occurred. However, phylogenetic analyses show that orthologs are far more closely related than paralogues, which does not substantiate concerted evolution as an important contributing mechanism underlying the diversification of chi-lectins. Instead, on the basis of observed relationships and selective pressures, the evolution of this gene family is in accordance with a form of multigene family evolution now referred to as "birth-and-death evolution under strong purifying selection" (reviewed in NEI and ROONEY 2005).

Our study revealed the existence in mice of a previously unidentified chi-lectin, referred to as a hypothetical protein (BYm). Despite the fact that BYm belongs to the group of AMCase- (or acidic-) lectins, it displays a basic isoelectric point, suggesting substantial specialization. The absence of the gene in animals other than rodents points to a relatively recent gene duplication, nicely illustrating the remarkable ongoing evolution of chitinase(-like) proteins in mammals. Likewise, the occurrence of GP40 seems restricted to artiodactyls. GP40 has been found to be present in dry mammary secretions at times when extensive tissue remodeling occurs (SRIVASTAVA *et al.* 2006); hence its presence may reflect differences in mammary function between artiodactyls and other mammals.

Structural features among orthologs of the chitinase protein family are extremely well conserved among mammals, yet there are differences in expression among species. A striking example in this respect is chitotriosidase in humans and rodents (BOOT et al. 2005a). The marked conservation of structural features of the catalytic domain of chitinases may be imposed by severe restrictions in changes compatible with preservation of catalytic function. Mammalian chitinases appear strongly subjected to negative (purifying) selection, substantially more than the functionally similar lysozymes. This is nicely illustrated by the fact that comparison of substitution rate ratios between lysozyme from Rhesus macaque and from hominoids point strongly toward positive selection whereas such comparisons for chitotriosidase and AMCase result in  $\omega$ -values far <1. A large number of constrained sites in both chitinases should mask any signal of positive selection when the  $\omega$ -values are averaged over all sites. Site-specific models for  $\omega$ calculation did not give values >1, although positive selection in the diversification of the two chitinases and their respective chi-lectins clearly occurred. To detect a very strong indicator of positive selection ( $\omega > 1$ ), it is necessary to look not only at a specific site but also at a specific time interval (SHARP 1997; ZHANG et al. 2005).

The high homology among members of the chitinase(-like) protein family may cause confusion when comparing genes of different species (BOOT *et al.* 2005b; RAES *et al.* 2005; REESE *et al.* 2007). There are considerable differences between species regarding presence or absence of particular chi-lectins. One example in this respect is Ym1. This protein has been extensively studied since it is secreted by alternatively activated macrophages in mice under a variety of inflammatory conditions (CHANG *et al.* 2001; WELCH *et al.* 2002; ZHAO *et al.* 2005; IWASHITA *et al.* 2006; REESE *et al.* 2007). The implications for human pathology, however, may be considered limited since no true human ortholog of Ym1 exists.

In conclusion, our investigation of family 18 glycosyl hydrolases has revealed that active chitinases and chilectins are widespread and conserved in the mammalian kingdom. An ancient gene duplication first allowed the specialization of two active chitinases, chitotriosidase and AMCase, and subsequent gene duplications followed by loss-of-enzymatic-function mutations, have led to the evolution of a broad spectrum of chi-lectins in mammals.

We gratefully acknowledge Roy Erkens of the University of Utrecht, The Netherlands, for his skillful technical assistance and useful discussions relating to the phylogenetic analyses.

#### LITERATURE CITED

- BADARIOTTI, F., M. KYPRIOTOU, C. LELONG, M. P. DUBOS, E. RENARD et al., 2006 The phylogenetically conserved molluscan chitinase-like protein 1 (Cg-Clp1), homologue of human HC-gp39, stimulates proliferation and regulates synthesis of extracellular matrix components of mammalian chondrocytes. J. Biol. Chem. 281: 29583–29596.
- BEUTLER, E., and G. A. GRABOWSKI, 1995 Gaucher disease, pp. 2641–2670 in *The Metabolic and Molecular Bases of Inherited Disease*, Ed. 7, edited by C. R. SCRIVER, A. L. BEAUDET, W. S. SLY and D. VALLE. McGraw-Hill, New York.
- BOOT, R. G., G. H. RENKEMA, A. STRIJLAND, A. J. VAN ZONNEVELD and J. M. AERTS, 1995 Cloning of a cDNA encoding chitotriosidase, a human chitinase produced by macrophages. J. Biol. Chem. 270: 26252–26256.
- BOOT, R. G., G. H. RENKEMA, M. VERHOEK, A. STRIJLAND, J. BLIEK *et al.*, 1998 The human chitotriosidase gene. Nature of inherited enzyme deficiency. J. Biol. Chem. **273**: 25680–25685.
- BOOT, R. G., E. F. BLOMMAART, E. SWART, K. GHAUHARALI-VAN DER VLUGT, N. BIJL *et al.*, 2001 Identification of a novel acidic mammalian chitinase distinct from chitotriosidase. J. Biol. Chem. **276**: 6770–6778.
- BOOT, R. G., A. P. BUSSINK, M. VERHOEK, P. A. DE BOER, A. F. MOORMAN et al., 2005a Marked differences in tissue-specific expression of chitinases in mouse and man. J. Histochem. Cytochem. 53: 1283– 1292.
- BOOT, R. G., A. P. BUSSINK and J. M. AERTS, 2005b Human acidic mammalian chitinase erroneously known as eosinophil chemotactic cytokine is not the ortholog of mouse YM1. J. Immunol. 175: 2041–2042.
- BUHI, W. C., 2002 Characterization and biological roles of oviductspecific, oestrogen-dependent glycoprotein. Reproduction 123: 355–362.
- BUSSINK, A. P., M. VAN EIJK, G. H. RENKEMA, J. M. AERTS and R. G. BOOT, 2006 The biology of the Gaucher cell: the cradle of human chitinases. Int. Rev. Cytol. 252: 71–128.
- CHANG, N. C., S. I. HUNG, K. Y. HWA, I. KATO, J. E. CHEN et al., 2001 A macrophage protein, Ym1, transiently expressed during inflammation is a novel mammalian lectin. J. Biol. Chem. 276: 17497–17506.
- FELSENSTEIN, J., 2006 PHYLIP (Phylogeny Inference Package), Version 3.65. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.
- FUJIMOTO, W., K. KIMURA and T. IWANAGA, 2002 Cellular expression of the gut chitinase in the stomach of frogs Xenopus laevis and Rana catesbeiana. Biomed. Res. 23: 91–99.
- FUSETTI, F., H. VON MOELLER, D. HOUSTON, H. J. ROZEBOOM, B. W. DIJKSTRA et al., 2002 Structure of human chitotriosidase. Implications for specific inhibitor design and function of mammalian chitinase-like lectins. J. Biol. Chem. 277: 25537–25544.

- FUSETTI, F., T. PIJNING, K. H. KALK, E. Bos and B. W. DIJKSTRA, 2003 Crystal structure and carbohydrate-binding properties of the human cartilage glycoprotein-39. J. Biol. Chem. 278: 37753–37760.
- GASTEIGER, E., A. GATTIKER, C. HOOGLAND, I. IVANYI, R. D. APPEL et al., 2003 ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res. 31: 3784–3788.
- GOUET, P., E. COURCELLE, D. I. STUART and F. METOZ, 1999 ESPript: multiple sequence alignments in PostScript. Bioinformatics 15: 305–308.
- GOUET, P., X. ROBERT and E. COURCELLE, 2003 ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. Nucleic Acids Res. **31**: 3320–3323.
- HAKALA, B. E., C. WHITE and A. D. RECKLIES, 1993 Human cartilage gp-39, a major secretory product of articular chondrocytes and synovial cells, is a mammalian member of a chitinase protein family. J. Biol. Chem. **268**: 25803–25810.
- HENRISSAT, B., 1991 A classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem. J. 280: 309–316.
- HOLLAK, C. E., S. VAN WEELY, M. H. VAN OERS and J. M. AERTS, 1994 Marked elevation of plasma chitotriosidase activity: a novel hallmark of Gaucher disease. J. Clin. Invest. 93: 1288–1292.
- HOMER, R. J., Z. ZHU, L. COHN, C. G. LEE, W. I. WHITE *et al.*, 2006 Differential expression of chitinases identify subsets of murine airway epithelial cells in allergic inflammation. Am. J. Physiol. Lung Cell. Mol. Physiol. **291:** 502–511.
- HOUSTON, D. R., A. D. RECKLIES, J. C. KRUPA and D. M. VAN AALTEN, 2003 Structure and ligand-induced conformational change of the 39-kDa glycoprotein from human articular chondrocytes. J. Biol. Chem. **278**: 30206–30212.
- HU, B., K. TRINH, W. F. FIGUEIRA and P. A. PRICE, 1996 Isolation and sequence of a novel human chondrocyte protein related to mammalian members of the chitinase protein family. J. Biol. Chem. 271: 19415–19420.
- HUGENTOBLER, S., D. G. MORRIS, M. T. KANE and J. M. SERENA, 2004 In situ oviduct and uterine pH in cattle. Theriogenology 61: 1419–1427.
- IWASHITA, H., S. MORITA, Y. SAGIYA and A. NAKANISHI, 2006 Role of eosinophil chemotactic factor by T lymphocytes on airway hyperresponsiveness in a murine model of allergic asthma. Am. J. Respir. Cell Mol. Biol. 35: 103–109.
- JEUNIAUX, C., 1961 Chitinase: an addition to the list of hydrolases in the digestive tract of vertebrates. Nature **192**: 135–136.
- JIN, H. M., N. G. COPELAND, D. J. GILBERT, N. A. JENKINS, R. B. KIRKPATRICK et al., 1998 Genetic characterization of the murine Ym1 gene and identification of a cluster of highly homologous genes. Genomics 54: 316–322.
- JOHANSEN, J. S., 2006 Studies on serum YKL-40 as a biomarker in diseases with inflammation, tissue remodelling, fibroses and cancer. Dan. Med. Bull. 53: 172–209.
- JOHANSEN, J. S., H.S. JENSEN and P. A. PRICE, 1993 A new biochemical marker for joint injury: analysis of YKL-40 in serum and synovial fluid. Br. J. Rheumatol. 32: 949–955.
- MESSIER, W., and C. B. STEWART, 1997 Episodic adaptive evolution of primate lysozymes. Nature **385**: 151–154.
- MONZINGO, A. F., E. M. MARCOTTE, P. J. HART and J. D. ROBERTUS, 1996 Chitinases, chitosanases, and lysozymes can be divided into procaryotic and eucaryotic families sharing a conserved core. Nat. Struct. Biol. 3: 133–140.
- NEI, M., and A. P. ROONEY, 2005 Concerted and birth-and-death evolution of multigene families. Annu. Rev. Genet. 39: 121–152.
- RAES, G., R. VAN DEN BERGH, P. DE BAETSELIER, G. H. GHASSABEH, C. SCOTTON *et al.*, 2005 Arginase-1 and Ym1 are markers for murine, but not human, alternatively activated myeloid cells. J. Immunol. 174: 834–845.
- RECKLIES, A. D., C. WHITE and H. LING, 2002 The chitinase 3-like protein human cartilage glycoprotein 39 (HC-gp39) stimulates proliferation of human connective-tissue cells and activates both extracellular signal-regulated kinase- and protein kinase B-mediated signalling pathways. Biochem. J. **365**: 119–126.
- REESE, T. A., H. E. LIANG, A. M. TAGER, A. D. LUSTER, N. VAN ROOIJEN et al., 2007 Chitin induces accumulation in tissue of innate immune cells associated with allergy. Nature 447: 92–96.
- RENKEMA, G. H., R. G. BOOT, A. O. MUIJSERS, W. E. DONKER-KOOPMAN and J. M. AERTS, 1995 Purification and characterization of

human chitotriosidase, a novel member of the chitinase family of proteins. J. Biol. Chem. **270:** 2198–2202.

- RENKEMA, G. H., R. G. BOOT, A. STRIJLAND, W. E. DONKER-KOOPMAN, M. VAN DEN BERG *et al.*, 1997 Synthesis, sorting, and processing into distinct isoforms of human macrophage chitotriosidase. Eur. J. Biochem. **244**: 279–285.
- RENKEMA, G. H., R. G. BOOT, F. L. AU, W. E. DONKER-KOOPMAN, A. STRIJLAND *et al.*, 1998 Chitotriosidase, a chitinase, and the 39-kDa human cartilage glycoprotein, a chitin-binding lectin, are homologues of family 18 glycosyl hydrolases secreted by human macrophages. Eur. J. Biochem. **251**: 504–509.
- SHARP, P. M., 1997 In search of molecular Darwinism. Nature **385**: 111–112.
- SRIVASTAVA, D. B., A. S. ETHAYATHULLA, J. KUMAR, N. SINGH, S. SHARMA et al., 2006 Crystal structure of a secretory signalling glycoprotein from sheep at 2.0A resolution. J. Struct. Biol. 156: 505–516.
- SUN, Y. J., N. C. CHANG, S. I. HUNG, A. C. CHANG, C. C. CHOU *et al.*, 2001 The crystal structure of a novel mammalian lectin, Ym1, suggests a saccharide binding site. J. Biol. Chem. **276**: 17507–17514.
- SUZUKI, M., W. FUJIMOTO, M. GOTO, M. MORIMATSU, B. SYUTO *et al.*, 2002 Cellular expression of gut chitinase mRNA in the gastrointestinal tract of mice and chickens. J. Histochem. Cytochem. 50: 1081–1089.
- SWANSON, W. J., Z. YANG, M. F. WOLFNER and C. F. AQUADRO, 2001 Positive Darwinian selection drives the evoluion of several female reproductive proteins in mammals. Proc. Natl. Sci. USA 98: 2509–2514.
- TERWISSCHA VAN SCHELTINGA, A. C., S. ARMAND, K. H. KALK, A. ISOGAI, B. HENRISSAT *et al.*, 1995 Stereochemistry of chitin hydrolysis by a plant chitinase/lysozyme and X-ray structure of a complex with allosamidin: evidence for substrate assisted catalysis. Biochemistry **34**: 15619–15623.
- THARANATHAN, R. N., and F. S. KITTUR, 2003 Chitin: the undisputed biomolecule of great potential. Crit. Rev. Food Sci. Nutr. 43: 61–87.
- THOMPSON, J. D., T. J. GIBSON, F. PLEWNIAK, F. JEANMOUGIN and D. G. HIGGINS, 1997 The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 24: 4876–4882.
- TSURUHA, J., K. MASUKO-HONGO, T. KATO, M. SAKATA, H. NAKAMURA et al., 2002 Autoimmunity against YKL-39, a human cartilage derived protein, in patients with osteoarthritis. J. Rheumatol. 29: 1459–1466.

- VAN AALTEN, D. M., D. KOMANDER, B. SYNSTAD, S. GASEIDNES, M. G. PETER et al., 2001 Structural insights into the catalytic mechanism of a family 18 exo-chitinase. Proc. Natl. Acad. Sci. USA 98: 8979–8984.
- VAN DAMME, E. J., R. CULERRIER, A. BARRE, R. ALVAREZ, P. ROUGE et al., 2006 A novel family of lectins evolutionarily related to class V chitinases: an example of neofunctionalization in legumes. Plant Physiol. 144: 662–672.
- VAN EIJK, M., C. P. VAN ROOMEN, G. H. RENKEMA, A. P. BUSSINK, L. ANDREWS *et al.*, 2005 Characterization of human phagocytederived chitotriosidase, a component of innate immunity. Int. Immunol. **17**: 1505–1512.
- VERHEIJDEN, G. F., A. W. RIJNDERS, E. BOS, C. J. COENEN-DE ROO, C. J. VAN STAVEREN *et al.*, 1997 Human cartilage glycoprotein-39 as a candidate autoantigen in rheumatoid arthritis. Arthritis Rheum. 40: 1115–1125.
- WARD, J. M., M. YOON, M. R. ANVER, D. C. HAINES, G. KUDO *et al.*, 2001 Hyalinosis and Ym1/Ym2 gene expression in the stomach and respiratory tract of 12984/SvJae and wild-type and CYP1A2null B6, 129 mice. Am. J. Pathol. **158**: 323–332.
- WELCH, J. S., L. ESCOUBET-LOZACH, D. B. SYKES, K. LIDDIARD, D. R. GREAVES *et al.*, 2002 TH2 cytokines and allergic challenge induce Ym1 expression in macrophages by a STAT6-dependent mechanism. J. Biol. Chem. **277**: 42821–42829.
- WIERENGA, R. K., 2001 The TIM-barrel fold: a versatile framework for efficient enzymes. FEBS Lett. 492: 193–198.
- YANG, Z., 1997 PAML: a program package for phylogenetic analysis by maximum likelihood. Comput. Appl. Biosci. 13: 555–556.
- YANG, Z., 1998 Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. Mol. Biol. Evol. 15: 568–573.
- ZHANG, J., R. NIELSEN and Z. YANG, 2005 Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. Mol. Biol. Evol. 22: 2472–2479.
- ZHAO, J., H. ZHU, C. H. WONG, K. Y. LEUNG and W. S. WONG, 2005 Increased lungkine and chitinase levels in allergic airway inflammation: a proteomics approach. Proteomics 5: 2799–2807.
- ZHU, Z., T. ZHENG, R. J. HOMER, Y. K. KIM, N. Y. CHEN *et al.*, 2004 Acidic mammalian chitinase in asthmatic Th2 inflammation and IL-13 pathway activation. Science **304**: 1678–1682.

Communicating editor: N. TAKAHATA