Published in final edited form as:
Genetica. 2010 July ; 138(7): 695–708. doi:10.1007/s10709-010-9438-z.

Mammalian carboxylesterase 3: comparative genomics and proteomics

Roger S. Holmes,
Department of Genetics, Southwest Foundation for Biomedical Research, San Antonio, TX 78227, USA rholmes@sfbrgenetics.org

Southwest National Primate Research Center, Southwest Foundation for Biomedical Research, San Antonio, TX, USA

School of Biomolecular and Physical Sciences, Griffith University, Nathan, QLD, Australia

Laura A. Cox, and
Department of Genetics, Southwest Foundation for Biomedical Research, San Antonio, TX 78227, USA

Southwest National Primate Research Center, Southwest Foundation for Biomedical Research, San Antonio, TX, USA

John L. VandeBerg
Department of Genetics, Southwest Foundation for Biomedical Research, San Antonio, TX 78227, USA

Southwest National Primate Research Center, Southwest Foundation for Biomedical Research, San Antonio, TX, USA

Abstract

At least five families of mammalian carboxylesterases (CES) catalyse the hydrolysis or transesterification of a wide range of drugs and xenobiotics and may also participate in fatty acyl and cholesterol ester metabolism. In this study, *in silico* methods were used to predict the amino acid sequences, secondary and tertiary structures, and gene locations for *CES3* genes and encoded proteins using data from several mammalian genome projects. Mammalian *CES3* genes were located within a *CES* gene cluster with *CES2* and *CES6* genes, usually containing 13 exons transcribed on the positive DNA strand. Evidence is reported for duplicated *CES3* genes for the chimp and mouse genomes. Mammalian CES3 protein subunits shared 58–97% sequence identity and exhibited sequence alignments and identities for key CES amino acid residues as well as extensive conservation of predicted secondary and tertiary structures with those previously reported for human CES1. The human genome project has previously reported *CES3* mRNA isoform expression in several tissues, particularly in colon, trachea and in brain. Predicted human CES3 isoproteins were apparently derived from exon shuffling and are likely to be secreted extracellularly or retained within the cytoplasm. Mouse *CES3*-like transcripts were localized in specific regions of the mouse brain, including the cerebellum, and may play a role in the detoxification of drugs and xenobiotics in neural tissues and other tissues of the body. Phylogenetic analyses demonstrated the relationships and potential evolutionary origins of the mammalian CES3 family of genes which were related to but distinct from other mammalian *CES* gene families.

Correspondence to: Roger S. Holmes.

Keywords

Mammals; Amino acid sequence; Carboxylesterase; Evolution; Gene duplication

Introduction

At least five families of mammalian carboxylesterases (CES; E.C.3.1.1.1) have been reported (Holmes et al. 2008a, b, 2009a, b, c, d) including CES1, the major liver enzyme (Munger et al. 1991; Shibita et al. 1993; Ghosh 2000; Holmes et al. 2009a); CES2, the major intestinal enzyme (Langmann et al. 1997; Schewer et al. 1997; Holmes et al. 2009a); CES3, expressed in liver, colon and brain (Sanghani et al. 2004; Holmes et al. 2009a, b, c); CES5 (also called CES7 or 'cauxin'), a major urinary protein of the domestic cat (Miyazaki et al. 2003, 2006; Holmes et al. 2008b); and CES6, a predicted CES-like enzyme in brain (Clark et al. 2003; Holmes et al. 2009d). These enzymes catalyze hydrolytic and transesterification reactions with xenobiotics, anticancer pro-drugs and narcotics (Satoh and Hosokawa 1998, 2006; Satoh et al. 2002; Ohtsuka et al. 2003; Redinbo and Potter 2005), and detoxify organophosphates and insecticides (Ahmad and Forgash 1976; Leinweber 1987). CES may also catalyze several cholesterol and fatty acid metabolic reactions (Ghosh 2000; Tsujita and Okuda 1993; Becker et al. 1994; Hosokawa et al. 2007; Diczfalusy et al. 2001), the conversion of lung alveolar surfactant (Ruppert et al. 2006) and assist with the assembly of low density lipoprotein particles in liver (Wang et al. 2007).

Structures for several human and animal *CES* genes have been determined, including human (Becker et al. 1994; Langmann et al. 1997; Ghosh 2000; Marsh et al. 2004) and rodent *CES1* and *CES2* 'like' genes (Ghosh et al. 1995; Dolinsky et al. 2001; Hosokawa et al. 2007). Predicted gene structures have also been described for other mammalian *CES* gene families, including *CES1*, *CES2* and *CES5* genes, which are localized in two contiguous *CES* gene clusters on human (chromosome 16), horse (chromosome 3), cow (chromosome 18) and opossum (chromosome 1) genomes (Holmes et al. 2008a, 2009a, b, c). Mammalian CES genes usually contain 12–14 exons of DNA encoding CES enzyme sequences which may undergo exon shuffling generating several CES isoproteins in each case (Thierry-Mieg and Thierry-Mieg 2006). There are similarities for the exon boundaries for each of the mammalian *CES* genes (Holmes et al. 2008a, b, 2009a, b, c, d) as well as significant sequence identities, especially for key regions previously identified for the major human liver CES1 (Bencharit et al. 2003, 2006; Fleming et al. 2005). Three-dimensional structural analyses of human CES1 have identified three major ligand binding sites, including the promiscuous active site, 'side door' and 'Z-site', where substrates, fatty acids and cholesterol analogues, respectively, are bound; and the 'gate', reportedly functioning in product release (Bencharit et al. 2003, 2006; Fleming et al. 2005).

This paper reports predicted gene structures and amino acid sequences for several mammalian *CES3* genes and proteins, the predicted secondary and tertiary structures for mammalian CES3 protein subunits, and the structural and evolutionary relationships for these genes and enzymes with those for other mammalian *CES* gene families. This study also tests an hypothesis that mammalian *CES3* genes are members of a distinct *CES* gene family to those previously reported, including *CES1* (encoding the major liver isozyme) (Munger et al. 1991), *CES2* (encoding the major intestinal CES) (Schewer et al. 1997), *CES5* (also called 'cauxin' or *CES7)* (Holmes et al. 2008b) and *CES6* (transcribed in human brain and also called *CES8)* (Holmes et al. 2009d).

Methods

In silico mammalian *CES* gene and protein identification

BLAST (Basic Local Alignment Search Tool) studies were undertaken using web tools from the National Center for Biotechnology Information (NCBI) (http://blast.ncbi.nlm.nih.gov/Blast.cgi) (Altschul et al. 1990). Protein BLAST analyses used mammalian CES amino acid sequences previously described (Table 1). Non-redundant protein sequence databases for several mammalian genomes were examined using the blastp algorithm, including human (*Homo sapiens*) (International Human Genome Sequencing Consortium 2001); chimp (*Pan troglodytes*) (Chimpanzee Sequencing and Analysis Consortium 2005); orangutan (*Pongo abelii*) (http://genome.wustl.edu); the cow (Bos *Taurus*) (Bovine Genome Project 2008); horse (*Equus caballus*) (Horse Genome Project 2008); mouse (*Mus musculus*) (Mouse Genome Sequencing Consortium 2002); and opossum (*Monodelphis domestica*) (Mikkelsen et al. 2007). This procedure produced multiple BLAST 'hits' for each of the protein databases which were individually examined and retained in FASTA format, and a record kept of the sequences for predicted mRNAs and encoded CES-like proteins. These records were derived from annotated genomic sequences using the gene prediction method: GNOMON and predicted sequences with high similarity scores for human CES6. Predicted CES3-like protein sequences were obtained in each case and subjected to *in silico* analyses of predicted protein and gene structures.

BLAT analyses were subsequently undertaken for each of the predicted CES3 amino acid sequences using the UC Santa Cruz web browser [http://genome.ucsc.edu/cgi-bin/hgBlat] (Kent et al. 2003) with the default settings to obtain the predicted locations for each of the mammalian *CES3* genes, including predicted exon boundary locations and gene sizes. BLAT analyses were similarly undertaken of *CES1, CES2, CES5* and *CES6*-like human genes using previously reported sequences for encoded CES subunits in each case (see Table 1). Structures for human CES3 isoforms (splicing variants) were obtained using the AceView website (http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/index.html?human) to examine predicted gene and protein structures using the human *CES3* cDNA sequence (Gen-Bank NM024922) to interrogate this database of human mRNA sequences (Thierry-Mieg and Thierry-Mieg 2006).

BLAT (BLAST-Like Alignment Tool) *in silico* studies were undertaken using the UC Santa Cruz web browser [http://genome.ucsc.edu/cgi-bin/hgBlat] (Kent et al. 2003) with the default settings. UniProtKB/Swiss-Prot Database [http://au.expasy.org] and GenBank [http://www.ncbi.nlm.nih.gov/Genbank/] sequences for human, orangutan and mouse CES sequences (Table 1) were used to interrogate human, chimp, orangutan, horse, cow, mouse and opossum genome sequences. Gene locations, predicted gene structures and CES protein subunit sequences were observed for each CES examined for those regions showing identity with the respective mammalian *CES* gene products.

Predicted structures and properties of mammalian CES subunits

Predicted secondary and tertiary structures for human and other mammalian CES-like subunits were obtained using the PSIPRED v2.5 web site tools [http://bioinf.cs.ucl.ac.uk/psipred/psiform.html] (McGuffin et al. 2000) and the SWISS MODEL web tools [http://swissmodel.expasy.org/], respectively (Guex and Peitsch 1997; Kopp and Schwede 2004). The reported tertiary structure for the rabbit CES1 4-piperidino–piperidine complex (Bencharit et al. 2003) served as the reference for the predicted human CES3 tertiary structure, with a modeling range of residues 34–553. Theoretical isoelectric points and molecular weights for mammalian CES3 subunits were obtained using Expasy web tools (http://au.expasy.org/tools/pi_tool.html). SignalP 3.0 web tools were used to predict the

presence and location of signal peptide cleavage sites (http://www.cbs.dtu.dk/services/SignalP/) for each of the predicted mammalian CES sequences (Emmanuelsson et al. 2007). The NetNGlyc 1.0 Server was used to predict potential *N*-glycosylation sites for human CES and mammalian CES3 subunits (http://www.cbs.dtu.dk/services/NetNGlyc/). Predictions of subcellular locations for mammalian CES3 'like' enzymes and for human CES1, CES2, CES5 and CES6 were conducted using PSORT 11 (http://psort.ims.u-tokyo.ac.jp/form2.html) (Horton and Nakai 1997).

Mouse brain CES3 expression

The Allen Mouse Brain Atlas was interrogated for cellular transcription of the mouse *CES3* gene using the GenBank BC061004 sequence (see Table 1) (Lein et al. 2007). Data for sagittal sections of mouse brain were examined and recorded for mouse CES3 transcripts using web tools available on the Allen Brain Atlas web site (http://www.brain-map.org).

Phylogenetic analyses and sequence divergence

ClustalW-derived alignments of CES protein sequences were assembled using BioEdit v.5.0.1 and the default settings (Hall 1999). Alignment ambiguous regions, including the amino and carboxyl termini, were excluded prior to phylogenetic analysis yielding alignments of 448 residues for comparisons of mammalian CES1, CES2, CES3, CES5 and CES6 and *Drosophila melanogaster* CES6 (or Est6) sequences (Table 1). Evolutionary distances were calculated using the Kimura option (Kimura 1983) in TREECON (Van de Peer and De Wachter 1994). Phylogenetic trees were constructed from evolutionary distances using the neighbor-joining method (Saitou and Nei 1987) and were rooted using the *Drosophila melanogaster* CES6 (Est6) sequence. Tree topology was re-examined by the boot-strap method (1,000) of resampling (Felsenstein 1985).

Results and discussion

Alignments of human CES subunits and mammalian CES3 amino acid sequences

The deduced amino acid sequences for horse, cow and mouse CES3-like subunits are shown in Fig. 1 together with the previously reported sequences for human CES1 (Munger et al. 1991;Shibita et al. 1993); human CES2 (Langmann et al. 1997;Schewer et al. 1997); human CES3 (Sanghani et al. 2004); human CES5 (Holmes et al. 2008b); and human CES6 (Holmes et al. 2009d) (Table 1). Alignments of the human and other mammalian CES3 subunits examined in this figure showed between 68 and 77% sequence identities, suggesting that these CES protein subunits are products of the same family of genes, whereas sequence alignments of mammalian CES3 subunits with human CES1, CES2, CES5 and CES6 subunits exhibited lower levels of sequence identities with human CES3 (47, 50, 48 and 45%, respectively), indicating that these are members of distinct CES families (Table 2).

The amino acid sequence for human CES3 contained 571 residues whereas other predicted mammalian CES3 subunits contained fewer amino acids: 570 residues for cow and horse CES3; 554 residues for mouse CES3; and 568 residues for mouse CES3L (Fig. 1). Previous studies on human CES1 have identified key residues which contribute to the catalytic, subcellular localization, oligomeric and regulatory functions for this enzyme (sequence numbers refer to human CES1). These include the catalytic triad for the active site (Ser228; Glu345; His458) (Cygler et al. 1993); microsomal targeting sequences, including the hydrophobic N-terminus signal peptide (von Heijne 1983;Zhen et al. 1995;Potter et al. 1998) and the C-terminal endoplasmic reticulum (ER) retention sequence (His-Ile-Glu-Leu) (Robbi and Beaufay 1991); disulfide bond forming residues (Cys95/Cys123 and Cys280/Cys291) (Lockridge et al. 1987); and ligand binding sites, including the 'Z-site' (Gly358), the 'side door' (Val424-

Met425-Phe426) and the 'gate' (Phe551) residues. Identical residues were observed for each of the mammalian CES3 subunits for the active site triad and disulfide bond forming residues for human CES1 although several changes were observed for the corresponding residues for CES3 subunits of other key human CES1 residues: the 'side-door' identified for human CES1 (Val424-Met425-Phe426) (Bencharit et al. 2003,2006) has a Val416-Phe417-Ile418 sequence for human CES3; the 'Z-site' (Gly358 for human CES1) has been changed to Ser355 for human CES3, although other mammalian CES3 sequences examined have retained a glycine residue at this position; the hydrophobic N-terminal sequence for mammalian CES3 sequences has undergone major changes, however, this region retains in each case a predicted signal peptide property; the mammalian CES3 C-terminal sequences have undergone significant changes in sequence (QEDL, QEEL or PEEL) in comparison with human CES1 (HIEL). This may be responsible for a change in the predicted microlocalization for the CES3 subunits (secreted or extracellular enzymes), as compared with human CES1, which is retained by the endoplasmic reticulum (Table 1). The 'gate' residue has also undergone changes for the mammalian CES3 sequences examined from Phe551 (CES1) to Tyr (human and horse CES3) or Leu (cow and mouse CES3).

Other key human CES1 sequences included two charge clamps which are responsible for subunit–subunit interaction, namely residues Lys78/Glu183 and Glu72/Arg193 (Bencharit et al. 2003, 2006; Fleming et al. 2005). Predicted mammalian CES3 subunit sequences for these sites indicated that only one of the charge clamps would be retained for human CES3 subunits: Arg88/Glu191 (Fig. 1), whereas other mammalian CES3 subunits examined showed no potential for forming charge clamps at the corresponding sequence positions. Previous reports have shown that human and baboon CES2 subunits, which lack these charge clamp residues, behave as monomers (Pindel et al. 1997; Holmes et al. 2009a). With the exception of human CES3, the mammalian CES3 sequences have undergone amino acid substitutions for those residues which contribute to the oligomeric subunit structure for human CES1. The identified human CES1 charge clamp residues Lys78/Glu183 and Glu72/Arg193 have other residues at the corresponding sites (Arg88/191Lys and Gln82/Ala194 for the horse CES3 sequence) (Fig. 1), and it is likely that these mammalian CES3 subunits are also monomeric in subunit structure. This may have a major influence on the kinetics and biochemical roles for mammalian CES3 since three dimensional studies have suggested that ligand binding to the human CES1 'Z-site' shifts the trimer–hexamer equilibrium towards the trimer facilitating substrate binding and enzyme catalysis (Redinbo and Potter 2005). This property is unlikely to be shared by mammalian CES3 which is expected to display distinct kinetic properties to the oligomeric CES1 enzyme.

The *N*-glycosylation site reported for human CES1 at Asn79-Ala80-Thr81 (Kroetz et al. 1993; Bencharit et al. 2003, 2006; Fleming et al. 2005) has not been retained for any of the mammalian CES3 sequences, although additional potential *N*-glycosylation sites were observed at one or several positions (Table 3). Given the reported role of the *N*-glycosylated carbohydrate group in contributing to mammalian CES stability and maintaining catalytic efficiency (Kroetz et al. 1993), this property may be shared by the mammalian CES3 subunits as well, especially for those containing multiple potential sites for *N*-glycosylation, such as the mouse CES3L sequence, which contains four such sites.

Predicted secondary and tertiary structures for mammalian CES3 subunits

Analyses of predicted secondary structures for mammalian CES3 sequences were compared with the previously reported secondary structure for human CES1 (Bencharit et al. 2003, 2006) and with those predicted for human CES2, CES5 and CES6 (Holmes et al. 2008b, 2009d) (Fig. 1). Similar α-helix β-sheet structures were observed for all of the mammalian CES3 and human CES gene products examined. Consistent structures were apparent near key

residues or functional domains including the α-helix within the N-terminal signal peptide; the β-sheet and α-helix-helix structures near the active site Ser228 (human CES1) and 'Z-site' (Glu354/Gly356, respectively); the α-helices bordering the 'side door' site; and the α-helix containing the 'gate' residue (Phe551 for human CES1). In addition, two random coil regions (residues 51–115 and 169–188 for human CES1) were predominantly retained for all forms of mammalian CES3 examined which have been shown for human CES1 to contain two charge clamps sites: (Lys79/Glu183 and Glu73/Arg186); an *N*-glycosylation site at Asn79-Ala80-Thr81; a second potential *N*-glycosylation site for human CES2, and one of the disulfide bridges (87Cys/117Cys) reported for human CES1. The human CES5 secondary structure, however, predicted an additional helix at the hydrophobic C-termini, in comparison with mammalian CES3 C-terminal sequences. In addition, the predicted 3-D structure for human CES3 is similar to the reported for human CES1 (Bencharit et al. 2003, 2006) (Fig. 2). These were based on structures reported for residues 21–553 of human CES1 and a predicted structure for residues 34–553 of human CES3. The single helix and following two β-sheet structures at the C-termini (shown in red at the top) for the human CES1 and CES3 subunits are readily apparent, as are four β-sheet structures at the N-termini of the CES subunits (shown in blue at the base), in each case. It is apparent that human CES3 is likely to be highly similar in its secondary and tertiary structure to that reported for human CES1.

Predicted gene locations and exonic structures for mammalian CES3 genes

Table 1 summarizes the predicted locations for mammalian CES3 genes based upon BLAT interrogations of several mammalian genomes using the reported sequences for human CES3, mouse CES3 and CES3L, and the predicted sequences for horse and cow CES3 and the UC Santa Cruz Web Browser (Kent et al. 2003). The predicted mammalian CES3 genes were located in a CES2/CES3/CES6 gene cluster in each case, although the gene order underwent changes for different species. In addition, mouse showed evidence of duplicate CES3 genes, with predicted CES3 and CES3L genes being located within a four CES gene cluster on chromosome 8: CES2-CES3-CES3L-CES6 (Table 1). Chimp CES3 also provided evidence for duplicate forms of this gene on chromosome 16 but with a changed order of genes: CES6-CES3L-CES2-CES3 (Table 1). BLAT interrogations of mammalian genomes with the corresponding CES sequences suggested that the CES2-CES3-CES6 gene cluster was syntenic for chromosomes 16 (human, chimp and orangutan), 3 (horse), 18 (cow), 8 (mouse) and 1 (opossum), respectively.

With the exception of the opossum CES3 gene, mammalian CES3 'like' genes were apparently transcribed on the positive strand, as for the human CES2 and CES6 genes. In contrast, human CES1 and CES5 genes were transcribed on the negative strand (Table 1). Figure 1 summarizes the predicted exonic start sites for cow, horse and mouse CES3 genes with each having 13 exons, in identical or similar positions to those predicted for the human CES3 gene (Holmes et al. 2009a). The predicted chimp CES3L gene, however, exhibited only 11 predicted exons. Human CES1 (Munger et al. 1991;Shibita et al. 1993), CES2 (Tang et al. 2008), CES5 (Ota et al. 2004) and CES6 (Holmes et al. 2009d) genes contained 14, 12, 13 and 14 exons, respectively (Fig. 1), in similar positions to those predicted for the mammalian CES3 genes (Fig. 1).

Figure 3 shows the predicted nucleotide sequences for exons 2–4 and introns 2 and 3 of mouse CES3 and CES3L genes. The sequences for the exons were 98% identical with nearly identical amino acid sequences involving five amino acid substitutions for these genes, whereas mouse CES3 and CES3L introns 2 and 3 showed distinct nucleotide sequences which were 85 and 94% identical, respectively. In addition, BLAT analyses of the mouse genome revealed distinct but closely localized predicted positions for these genes, being separated by ~26 kilobases of DNA on chromosome 8 (Table 1). These data support an hypothesis for two closely related

and localized CES3-like genes in the mouse, for which the subunits have been previously designated as esterase-31 (also called CES3) (Aida et al. 1993) and esterase-31L (also called CES3L) (Mouse Genome Sequencing Consortium 2002).

Mouse brain CES3 and CES3L expression

Figure 4 shows a sagittal section of mouse brain examining the distribution of CES3 and CES3L transcripts (GenBank mRNA BC061004 and BC019047) within different regions of the brain (provided by the Allen Institute for Brain Science) (Lein et al. 2007). The cerebellum folds exhibited highest staining levels for both transcripts although staining was observed throughout the brain, including the hippocampus, amygdalar nuclei, the olfactory bulb, the cortex, the pons and the medulla regions.

The differential tissue and subcellular distributions of CES family members may assist in determining the differential roles of CES family members within mammalian organisms. Mammalian liver is predominantly responsible for drug and xenobiotic clearance from the body with CES1 and CES2 (with CES1 > CES2) playing major roles, following absorption of drugs and xenobiotics into the circulation (Pindel et al. 1997; Imai 2006). Mammalian intestine (with CES2 > CES1) is predominantly responsible for first pass clearance of several drugs and xenobiotics, with the activity occurring mostly in the ileum and jejunum and processed via CES2 (Imai et al. 2003). CES1 and CES2 also serve different roles in prodrug activation, as shown for the anti-cancer drug irinotecan (CPT-11) which is converted to its active form SN-38 predominantly by CES2 (Humerickhouse et al. 2000; Xu et al. 2002). In contrast with *CES1* and *CES2*, mammalian *CES5* is predominantly expressed in peripheral tissues, including brain, kidney, lung and testis (Thierry-Mieg and Thierry-Mieg 2006), and in cat, is a secreted form of CES enzyme appearing in cat urine (Miyazaki et al. 2003, 2006). Mammalian CES5 may serve in two major roles within mammalian fluids and peripheral tissues, including regulating the production of a pheromone precursor in urine (Miyazaki et al. 2003, 2006) and contributing to lipid and cholesterol transfer processes within male reproductive fluids (Ecroyd et al. 2006).

The metabolic role(s) for mammalian CES3 has not been extensively investigated although the enzyme is capable of activating prodrugs such as irinotecan (Sanghani et al. 2004) and the gene is expressed as distinct isoforms in tissues of the body, including the colon, placenta and neural tissues, such as the cerebellum and hippocampus (Fig. 5; Table 4) (Thierry-Mieg and Thierry-Mieg 2006). Human CES3 mRNAs encode isoforms which are predicted to be either secreted or extracellular (CES3a–c), cytosolic (CES3d) or extracellular (CES3e–u), and may play distinct roles in carboxyl-ester metabolism in the body. A comparison of the predicted cellular locations for mammalian CES3 reveals that it is likely to be extracellular (including the cell wall) in primate (human, chimp and orangutan) and mouse tissues, whereas horse and bovine CES3 are predicted to be localized in the endoplasmic reticulum, which is comparable with human CES1 and CES2. The extracellular locations predicted for human and mouse CES3 may reflect distinct roles for this enzyme in drug metabolism within peripheral tissues of the body.

Phylogeny and divergence of *CES3* and other mammalian CES sequences

Phylogenetic trees (Fig. 6) were constructed from ClustalW-derived alignments of human, chimp, orangutan, horse, cow, mouse and opossum CES3-like amino acid sequences with human CES1, CES2, CES5 and CES6 and *Drosophila melanogaster* CES6 sequences. The dendrogram was rooted using the latter *Drosophila* CES6 sequence and showed clustering of all of the CES3-like sequences which were distinct from the other human *CES* gene families. In addition, the CES3 and CES3L mouse sequences showed clustering with the mouse (and rat) CES3-like genes, which is consistent with these genes being products of recent duplication

events during mammalian evolution. Holmes et al. (2008a) have previously demonstrated that the gene duplication events which generated the ancestral mammalian CES1, CES2, CES3, CES5 and CES6 genes predated the common ancestor for marsupial and eutherian mammals [estimated at ~ 173–193 million years ago by Woodburne et al. (2003)], and appeared to coincide with the early diversification of tetrapods approximately 350–360 million years ago (Donoghue and Benton 2007). The mammalian *CES3* family of genes may therefore be considered to be of an ancient origin, prior to the appearance of mammals in evolutionary time.

The mammalian *CES* gene complex: differential roles and comparison with other gene family complexes

The mammalian *CES3* gene family is localized within a *CES2*-*CES3*-*CES6* gene complex on eutherian and marsupial mammalian genomes, with a consistent gene order for most genomes examined, with the exception of the chimpanzee (*Pan troglodytes*) and orangutan (*Pongo abelii*) genomes (Table 1). The other two *CES* gene families (*CES1* and *CES5*) are also usually located on the same mammalian chromosome, but within a separate distantly located gene complex (see Holmes et al. 2008b,2009d). The existence of these two mammalian *CES* gene complexes may reflect the evolutionary origins of these genes by gene duplication processes, as well as the timing and sequence of gene duplication events generating the *CES1*-*CES5* and *CES2*-*CES5*-*CES6* gene complexes prior to mammalian radiation and evolution.

Mammalian *CES* gene families have, however, diverged significantly in terms of their differential gene regulation and in the properties of encoded CES isozymes, which may reflect their differential and specialized roles in the body. In addition to being enzymes of broad substrate specificities in carboxyl-ester metabolism, CES family members perform several specialized roles in the body, including: CES1: liver drug clearance (see Satoh and Hosokawa 2006; Hosokawa et al. 2007), lung surfactant conversion (Ruppert et al. 2006); and macrophage cholesterol ester metabolism (Ghosh 2000); CES2: intestine drug clearance (see Marsh et al. 2004; Imai 2006); CES3: colon and neural drug metabolism (Sanghani et al. 2004); CES5: pheromone and lipid transfer metabolism (Miyazaki et al. 2003, 2006; Ecroyd et al. 2006); and CES6: predicted roles in drug and xenobiotic metabolism in the brain (Holmes et al. 2009d). The mammalian *CES* gene complexes stand in contrast to the *CES* gene families reported in a model plant, *Arabidopsis thaliana*, for which 20 genes were identified (Marshall et al. 2003). With the exception of three clusters of tandemly duplicated genes, these were present in four of five plant chromosomes.

The mammalian *CES* gene complexes are comparable to those previously reported for other mammalian genes and enzymes, including alcohol dehydrogenase (*ADH*), which occurs as a gene complex on the same chromosome in human (*Homo sapiens*) (chromosome 4) (see Jornvall et al. 2000) and mice (*Mus musculus*) (chromosome 3) (Holmes et al. 1983) genomes, and comprises at least five families of *ADH* genes, for which encoded isozymes perform differential roles on a wide range of alcohol substrates (see Hoog et al. 2003). The number and diversity of genes found within the *CES* gene complex is in contrast, however, with the large numbers of genes reported for the major histocompatibility complex (*MHC*), for which a 3.6 Mb *MHC* region on human chromosome 6 contains 140 genes, which are divided into three classes, with differential roles in the immune system of the body (see MHC Sequencing Consortium 1999).

Conclusion

The results of the present study indicate that mammalian *CES3* genes and encoded CES3 enzymes represent a distinct *CES* gene and enzyme family which share key conserved sequences and structures that have been reported for human CES1 and have family specific sequences consistent with the oligomeric and monomeric subunit structures for CES1 and

CES3, respectively. This study also reports that chimp and mouse genomes contain at least two *CES3* 'like' genes, which are located in tandem with the *CES2* and *CES6* genes, and with the more distantly located *CES1* and *CES5* genes on chromosomes 16 and 8, respectively. Predicted secondary structures and tertiary structures for mammalian CES3 subunits showed a high degree of conservation with human CES1. Phylogeny studies using several mammalian CES3 subunits (human, chimp, orangutan, mouse, horse, cow and opossum) indicated that *CES3*' like' genes have apparently appeared during mammalian evolution well prior to the eutherian and marsupial common ancestor more than 170 million years ago, which is consistent with previous studies (Holmes et al. 2008a).

Acknowledgments

This project was supported by NIH Grants P01 HL028972 and P51 RR013986. In addition, this investigation was conducted in facilities constructed with support from Research Facilities Improvement Program Grant Numbers 1 C06 RR13556, 1 C06 RR15456, 1 C06 RR017515. We gratefully acknowledge the assistance of Dr B. Patel in studying the phylogeny of CES3 and related CES gene families.

References

Ahmad S, Forgash AJ. Nonoxidative enzymes in the metabolism of insecticides. Drug Metab Rev 1976;5:141–164.

Aida K, Moore R, Negishi M. Cloning and sequencing of a novel, male-predominant carboxylesterase in mouse liver. Biochim Biophys Acta 1993;1174:72–74. [PubMed: 7916639]

Altschul F, Vyas V, Cornfield A, Goodin S, Ravikumar TS, Rubin EH, Gupta E. Basic local alignment search tool. J Mol Biol 1990;215:403–410. [PubMed: 2231712]

Becker A, Bottcher A, Lackner KJ, Fehringer P, Notka F, Aslandis C, Schmithz C. Purification, cloning and expression of a human enzyme with acyl coenzyme A: cholesterol acyltransferase activity, which is identical to liver carboxylesterase. Arterioscler Thromb 1994;14:1346–1355. [PubMed: 8049197]

Bencharit S, Morton CL, Xue Y, Potter PM, Redinbo MR. Structural basis of heroin and cocaine metabolism by a promiscuous human drug-processing enzyme. Nat Struct Biol 2003;10:349–356. [PubMed: 12679808]

Bencharit S, Edwards CC, Morton CL, Howard-Williams EL, Kuhn P, Potter PM, Redinbo MR. Multisite promiscuity in the processing of endogenous substrates by human carboxylesterase 1. J Mol Biol 2006;363:201–214. [PubMed: 16962139]

Bovine Genome Project. 2008. http://hgsc.bcm.tmc.edu/projects/bovine

Chimpanzee Sequencing and Analysis Consortium. Initial sequence of the chimpanzee genome and comparison with the human genome. Nature 2005;437:69–87. [PubMed: 16136131]

Clark HF, Gurney AL, Abaya E, Baker K, Baldwin D, Brush J, Chen J, Chow B, Chui C, Crowley C, Currell B, Deuel B, Dowd P, Eaton D, Foster J, Grimaldi C, Gu Q, Hass PE, Heldens S, Huang A, Kim HS, Klimowski L, Jin Y, Johnson S, Lee J, Lewis L, Liao D, Mark M, Robbie E, Sanchez C, Schoenfeld J, Seshagiri S, Simmons L, Singh J, Smith V, Stinson J, Vagts A, Vandlen R, Watanabe C, Wieand D, Woods K, Xie MH, Yansura D, Yi S, Yu G, Yuan J, Zhang M, Zhang Z, Goddard A, Wood WI, Godowski P, Gray A. The secreted protein discovery initiative (SPDI), a large-scale effort to identify novel human secreted and transmembrane proteins: a bioinformatics assessment. Genome Res 2003;13:226–2270.

Cygler M, Schrag JD, Sussman JL, Harel M, Silman I, Gentry MK, Dostor BP. Relationship between sequence conservation and three-dimensional structure in a large family of esterases, lipases and related proteins. Protein Sci 1993;2:366–382. [PubMed: 8453375]

Diczfalusy MA, Bjorkkem I, Einarsson C, Hillebrant CG, Alexson SE. Characterization of enzymes involved in formation of ethyl esters of long-chain fatty acids. J Lipid Res 2001;42:1025–1032. [PubMed: 11441128]

Dolinsky VW, Sipione S, Lehner R, Vance DE. The cloning and expression of murine triacylglycerol hydrolase cDNA and the structure of the corresponding gene. Biochim Biophys Acta 2001;1532:162–172. [PubMed: 11470237]

Donoghue PCJ, Benton MJ. Rocks and clocks: calibrating the tree of life using fossils and molecules. Trends Genet 2007;22:424–431.

Ecroyd H, Belghazi M, Dacheux JL, Miyazaki M, Yamashita T, Gatti JL. An epididymal form of cauxin, a carboxylesterase-like enzyme, is present and active in male reproductive fluids. Biol Reprod 2006;74:439–447. [PubMed: 16251497]

Emmanuelsson O, Brunak S, von Heijne G, Nielson H. Locating proteins in the cell using TargetP, SignalP and related tools. Nat Protoc 2007;2:953–971. [PubMed: 17446895]

Felsenstein J. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 1985;39:783–791.

Fleming CD, Bencharit S, Edwards CC, Hyatt JL, Tsurkan L, Bai F, Fraga C, Morton CL, Howard-Williams EL, Potter PM, Redinbo MR. Structural insights into drug processing by human carboxylesterase 1: tamoxifen, Mevaststin, and inhibition by Benzil. J Mol Biol 2005;352:165–177. [PubMed: 16081098]

Ghosh S. Cholesteryl ester hydrolase in human monocyte/macrophage: cloning, sequencing and expression of full-length cDNA. Physiol Genomics 2000;2:1–8. [PubMed: 11015575]

Ghosh S, Mallonee DH, Grogan WM. Molecular cloning and expression of rat hepatic neutral cholesteryl ester hydrolase. Biochim Biophys Acta 1995;1259:305–312. [PubMed: 8541339]

Guex N, Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modelling. Electrophoresis 1997;18:2714–2723. [PubMed: 9504803]

Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 1999;41:95–98.

Holmes, RS.; Duley, JA.; Burnell, JN. The alcohol dehydrogenase gene complex on chromosome 3 of the mouse. In: Rattazzi, MC.; Scandalios, JG.; Whitt, GS., editors. Isozymes: current topics in biological and medical research, vol 8. New York: Alan R Liss; 1983. p. 155-174.

Holmes RS, Chan J, Cox LA, Murphy WJ, VandeBerg JL. Opossum carboxylesterases: sequences, phylogeny and evidence for CES duplication events predating the marsupial-eutherian common ancestor. BMC Evol Biol 2008a;8:54. [PubMed: 18289373]

Holmes RS, Cox LA, VandeBerg JL. Mammalian carboxylesterase 5: comparative biochemistry and genomics. Comp Biochem Physiol Part D 2008b;3:195–204.

Holmes RS, Glenn JP, VandeBerg JL, Cox LA. Baboon carboxylesterases 1 and 2: sequences, structures and phylogenetic relationships with human and other primate carboxylesterases. J Med Primatol 2009a;38:27–38. [PubMed: 19187434]

Holmes RS, VandeBerg JL, Cox LA. Bovine carboxylesterases: evidence for two *CES1* and five families of *CES* genes on chromosome 18. Comp Biochem Physiol Part D 2009b;4:11–20.

Holmes RS, VandeBerg JL, Cox LA. Horse carboxylesterases: evidence for six *CES1* and four families of *CES* genes on chromosome 3. Comp Biochem Physiol Part D 2009c;4:54–65.

Holmes RS, VandeBerg JL, Cox LA. A new class of mammalian carboxylesterase *CES6*. Comp Biochem Physiol Part D 2009d;4:209–217.

Hoog J-O, Stromberg P, Hedberg JJ, Griffiths WJ. The mammalian alcohol dehydrogenases interact in several pathways. Chem Biol Interact 2003;143–144:175–181.

Horse Genome Project. 2008. http://www.uky.edu/Ag/Horsemap/

Horton P, Nakai K. Better prediction of cellular localization sites with the k nearest neighbors classifier. Proc Int Conf Intell Syst Mol Biol 1997;5:147–152. [PubMed: 9322029]

Hosokawa M, Furihata T, Yaginuma Y, Yamamoto N, Kayano N, Fujii A, Nagahara Y, Satoh T, Chiba K. Genomic structure and transcriptional regulation of the rat, mouse and human carboxylesterase genes. Drug Metab Rev 2007;39:1–15. [PubMed: 17364878]

Humerickhouse R, Lohrbach K, Li L, Bosron WF, Dolan ME. Characterization of CPT-11 hydrolysis by human liver carboxylesterase isoforms h-CE1 and hCE-2. Cancer Res 2000;60:1189–1192. [PubMed: 10728672]

Imai T. Human carboxylesterase isozymes: catalytic properties and rational drug design. Drug Metab Pharmacokinet 2006;21:173–185. [PubMed: 16858120]

Imai T, Yoshigae Y, Hosokawa M, Chiba K, Otagiri M. Evidence for the involvement of a pulmonary first-pass effect via carboxylesterase in the disposition of a propanolol ester derivative after intravenous administration. J Pharmacol Exp Ther 2003;307:1234–1242. [PubMed: 14534358]

International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. Nature 2001;409:860–921. [PubMed: 11237011]

Jornvall H, Hoog J-O, Persson B, Pares X. Pharmacogenetics of the alcohol dehydrogenase system. Pharmacology 2000;61:184–191. [PubMed: 10971204]

Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. The human genome browser at UCSC. Genome Res 2003;12:994–1006.

Kimura, M. The neutral theory of molecular evolution. Cambridge: Cambridge University Press; 1983.

Kopp J, Schwede T. The SWISS-MODEL repository of annotated three-dimensional protein structure homology models. Nucleic Acids Res 2004;32:D230–D234. [PubMed: 14681401]

Kroetz DL, McBride OW, Gonzalez FJ. Glycosylation-dependent activity of Baculovirus-expressed human liver carboxylesterases: cDNA cloning and characterization of two highly similar enzyme forms. Biochemistry 1993;32:11606–11617. [PubMed: 8218228]

Langmann T, Becker A, Aslanidis C, Notka F, Ulrich H, Schwer H, Schcmitz G. Structural organization and characterization of the promoter region of a human carboxylesterase gene. Biochim Biophys Acta 1997;1350:65–74. [PubMed: 9003459]

Lein ES, Hawrylycz MJ, Ao N, Ayres M, Bensinger A, Bernard A, Boe AF, Boguski MS, Brockway KS, Byrnes EJ, Chen L, Chen L, Chen T-M, Chin MC, Chong J, Crook BE, Czaplinska A, Dang CN, Datta S, Dee NR, Desaki AL, Desta T, Diep E, Dolbeare TA, Donelan MJ, Dong H-W, Dougherty JG, Duncan BJ, Ebbert AJ, Eichele G, Estin LK, Faber C, Facer BA, Fields R, Fischer SR, Fliss TP, Frensley C, Gates SN, Glattfelder KG, Halverson KR, Hart MR, Hohmann JG, Howell MP, Jeung DP, Johnson RA, Karr PT, Kawal R, Kidney JM, Knapik RH, Kuan CL, Lake JH, Laramee AR, Larsen KD, Lau C, Lemon TA, Liang AJ, Liu Y, Luong LT, Michaels J, Morgan JJ, Morgan RJ, Mortrud MT, Mosqueda NF, Ng LL, Ng R, Orta GJ, Overly CC, Pak TH, Parry SE, Pathak SD, Pearson OC, Puchalski RB, Riley ZL, Rockett HR, Rowland SA, Royall JJ, Ruiz MJ, Sarno NR, Schaffnit K, Shapovalova NV, Sivisay T, Slaughterbeck CR, Smith SC, Smith KA, Smith BI, Sodt AJ, Stewart NN, Stumpf K-R, Sunkin SM, Sutram M, Tam A, Teemer CD, Thaller C, Thompson CL, Varnam LR, Visel A, Whitlock RM, Wohnoutka PE, Wolkey CK, Wong VY, Wood M, Yaylaoglu MB, Young RC, Youngstrom BL, Yuan XF, Zhang B, Zwingman TA, Jones AR. Genome wide atlas of gene expression in the mouse brain. Nature 2007;145:168–176. [PubMed: 17151600]

Leinweber FJ. Possible physiological roles of carboxyl ester hydrolases. Drug Metab Rev 1987;18:379–439. [PubMed: 3286170]

Lockridge O, Adkins S, La Due BN. Location of disulfide bonds within the sequence of human serum cholinesterase. J Biol Chem 1987;262:12945–12952. [PubMed: 3115973]

Marsh S, Xiao M, Yu J, Ahluwalia R, Minton M, Freimuth RR, Kwok P-Y, McLeod HL. Pharmacogenomic assessment of carboxylesterases 1 and 2. Genomics 2004;84:661–668. [PubMed: 15475243]

Marshall SD, Putterill JJ, Plummer KM, Newcomb RD. The carboxylesterase gene family from *Arabidopsis thaliana*. J Mol Evol 2003;57:487–500. [PubMed: 14738307]

McGuffin LJ, Bryson K, Jones DT. The PSIPRED protein structure prediction server. Bioinformatics 2000;16:404–405. [PubMed: 10869041]

MHC Sequencing Consortium. Complete sequence and map of a human major histocompatibility complex. Nature 1999;401:921–923. [PubMed: 10553908]

Mikkelsen TS, Wakefield MJ, Aken B, Amemiya CT, Chang JL, Duke S, Garber M, Gentles AJ, Goodstadt L, Heger A, Jurka J, Kamal M, Mauceli E, Searle SMJ, Sharpe T, Baker ML, Batzer MA, Benos PV, Belov K, Clamp M, Cook A, Cuff J, Das R, Davidow L, Deakin JE, Fazzari MJ, Glass JL, Grabherr M, Greally JM, Gu W, Hore TA, Huttley GA, Kleber M, Jirtle RL, Koina E, Lee JT, Mahony S, Marra MA, Miller RD, Nicholls RD, Oda M, Papenfuss AT, Parra ZE, Pollock DD, Ray DA, Schein JE, Speed TP, Thompson K, VandeBerg JL, Wade CM, Walker JA, Waters PD, Webber C, Weidman JR, Xie X, Zody MC, Marshall Graves JA, Ponting CP, Breen M, Samollow PB, Lander ES, Lindblad-Toh K. Broad Institute Genome Sequencing Platform, Broad Institute Whole Genome Assembly Team. Genome of the marsupial *Monodelphis domestica* reveals innovation in noncoding sequences. Nature 2007;447:167–177. [PubMed: 17495919]

Miyazaki M, Kamiie K, Soeta S, Taira H, Yamashita T. Molecular cloning and characterization of a novel carboxylesterase-like protein that is physiologically present at high concentrations in the urine of domestic cats (*Felis Catus*). Biochem J 2003;370:101–110. [PubMed: 12401131]

Miyazaki K, Yamashita T, Suzuki Y, Saito Y, Soeta S, Taira H, Suzuki A. A major urinary protein of the domestic cat regulates the production the production of felinine, a putative pheromone precursor. Chem Biol 2006;13:10171–10179.

Mouse Genome Sequencing Consortium. Initial sequencing and comparative analysis of the mouse genome. Nature 2002;420:520–562. [PubMed: 12466850]

Munger JS, Shi GP, Mark EA, Chin DT, Gerard C, Chapman HA. A serine esterase released by human alveolar macrophages is closely related to liver microsomal carboxylesterases. J Biol Chem 1991;266:18832–18838. [PubMed: 1918003]

Ohtsuka K, Inoue S, Kameyama M. Intracellular conversion of irinotecan to its active form, SN-38, by native carboxylesterase in human non-small cell lung cancer. Lung Cancer 2003;41:87–198.

Ota T, Suzuki Y, Nishikawa T, Otsuki T, Sugiyama T, Irie R, Wakamatsu A, Hayashi K, Sato H, Nagai K, Kimura K, Makita H, Sekine M, Obayashi M, Nishi T, Shibahara T, Tanaka T, Ishii S, Yamamoto J, Saito K, Kawai Y, Isono Y, Nakamura Y, Nagahari K, Murakami K, Yasuda T, Iwayanagi T, Wagatsuma M, Shiratori A, Sudo H, Hosoiri T, Kaku Y, Kodaira H, Kondo H, Sugawara M, Takahashi M, Kanda K, Yokoi T, Furuya T, Kikkawa E, Omura Y, Abe K, Kamihara K, Katsuta N, Sato K, Tanikawa M, Yamazaki M, Ninomiya K, Ishibashi T, Yamashita H, Murakawa K, Fujimori K, Tanai H, Kimata M, Watanabe M, Hiraoka S, Chiba Y, Ishida S, Ono Y, Takiguchi S, Watanabe S, Yosida M, Hotuta T, Kusano J, Kanehori K, Takahashi-Fujii A, Hara H, Tanase TO, Nomura Y, Togiya S, Komai F, Hara R, Takeuchi K, Arita M, Imose N, Musashino K, Yuuki H, Oshima A, Sasaki N, Aotsuka S, Yoshikawa Y, Matsunawa H, Ichihara T, Shiohata N, Sano S, Moriya S, Momiyama H, Satoh N, Takami S, Terashima Y, Suzuki O, Nakagawa S, Senoh A, Mizoguchi H, Goto Y, Shimizu F, Wakebe H, Hishigaki H, Watanabe T, Sugiyama A, Takemoto M, Kawakami B, Yamazaki M, Watanabe K, Kumagai A, Itakura S, Fukuzumi Y, Fujimori Y, Komiyama M, Tashiro H, Tanigami A, Fujiwara T, Ono T, Yamada K, Fujii Y, Ozaki K, Hirao M, Ohmori Y, Kawabata A, Hikiji T, Kobatake N, Inagaki H, Ikema Y, Okamoto S, Okitani R, Kawakami T, Noguchi S, Itoh T, Shigeta K, Senba T, Matsumura K, Nakajima Y, Mizuno T, Morinaga M, Sasaki M, Togashi T, Oyama M, Hata H, Watanabe M, Komatsu T, Mizushima-Sugano J, Satoh T, Shirai Y, Takahashi Y, Nakagawa K, Okumura K, Nagase T, Nomura N, Kikuchi H, Masuho Y, Yamashita R, Nakai K, Yada T, Nakamura Y, Ohara O, Isogai T, Sugano S. Complete sequencing and characterization of 21, 243 full-length human cDNAs. Nat Genet 2004;36:40–45. [PubMed: 14702039]

Pindel EV, Kedishvili NY, Abraham TL, Brezinski MR, Zhang A, Dean RA, Bosron WF. Purification and cloning of a broad substrate specificity human liver carboxylesterase that catalyzes the hydrolysis of cocaine and heroin. J Biol Chem 1997;272:14769–14775. [PubMed: 9169443]

Potter PM, Wolverton JS, Morton CL, Wierdl M, Danks MK. Cellular localization domains of a rabbit and human carboxylesterase: influence on irinotecan (CPT-11) metabolism by the rabbit enzyme. Cancer Res 1998;58:3627–3632. [PubMed: 9721871]

Redinbo MR, Potter PN. Mammalian carboxylesterases: from drug targets to protein therapeutics. Drug Discov Today 2005;10:313–320. [PubMed: 15749280]

Robbi M, Beaufay H. The COOH terminus of several liver carboxylesterases targets these enzymes to the lumen of the endoplasmic reticulum. J Biol Chem 1991;266:20498–20503. [PubMed: 1939102]

Ruppert C, Bagheri A, Markart P, Schmidt R, Seegar W, Gunther A. Liver carboxylesterase cleaves surfactant protein (SP-B) and promotes surfactant subtype conversion. Biochem Biophys Res Commun 2006;348:1449–1454. [PubMed: 16919595]

Saitou N, Nei M. The neighbour-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 1987;4:406–426. [PubMed: 3447015]

Sanghani SP, Quinney SK, Fredenberg TB, Davis WI, Murray DJ, Bosron WF. Hydrolysis of irinotecan and its oxidative metabolites, 7-ethyl-10-[4-N(5-aminopentanoic acid)-1-piperidino] carbonyloxycampothecin and 7-ethyl-10-[4-(1-piperidino)-1 amino]-carbonyloxycamptothecin, by human carboxylesterases CES1A1, CES2, and a newly expressed carboxylesterase isoenzyme, CES3. Drug Metab Dispos 2004;32:505–511. [PubMed: 15100172]

Satoh T, Hosokawa M. The mammalian carboxylesterases: from molecules to functions. Ann Rev Pharmacol Toxicol 1998;38:257–288. [PubMed: 9597156]

Satoh T, Hosokawa M. Structure, function and regulation of carboxylesterases. Chem-Biol Interact 2006;162:195–211. [PubMed: 16919614]

Satoh H, Taylor P, Bosron WF, Sanghani P, Hosokawa M, Du PB. Current progress on esterases: from molecular structure to function. Drug Metab Dispos 2002;30:488–493. [PubMed: 11950776]

Schewer H, Langmann T, Daig R, Becker A, Aslandis C, Schmidt G. Molecular cloning and characterization of a novel putative carboxylesterase, present in human intestine and liver. Biochem Biophys Res Commun 1997;233:117–120. [PubMed: 9144407]

Shibita F, Takagi Y, Kitajima M, Kuroda T, Omura T. Molecular cloning and characterization of a human carboxylesterase gene. Genomics 1993;17:76–82. [PubMed: 8406473]

Tang X, Wu H, Wu Z, Wang G, Zhu D. Carboxylesterase 2 is downregulated in colorectal cancer following progression of the disease. Cancer Invest 2008;26:178–181. [PubMed: 18259949]

Thierry-Mieg D, Thierry-Mieg J. AceView: a comprehensive cDNA-supported gene and transcripts annotation. Genome Biol 2006;7:S12. http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/index.html?human. [PubMed: 16925834]

Tsujita T, Okuda H. Palmitoyl-coenzyme A hydrolyzing activity in rat kidney and its relationship with carboxylesterase. J Lipid Res 1993;34:1773–1781. [PubMed: 7902406]

Van De Peer Y, de Wachter R. TreeCon for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. Comput Appl Sci 1994;10:569–570.

von Heijne G. Patterns of amino acids near signal-sequence cleavage sites. Eur J Biochem 1983;133:17–21. [PubMed: 6852022]

Wang H, Gilham D, Lehner R. Proteomic and lipid characterization of apo-lipoprotein B-free luminal lipid droplets from mouse liver microsomes: implications for very low density lipoprotein assembly. J Biol Chem 2007;282:33218–33226. [PubMed: 17848546]

Woodburne MO, Rich TH, Springer MS. The evolution of tribospheny and the antiquity of mammalian clades. Mol Phylogenet Evol 2003;28:360–385. [PubMed: 12878472]

Xu G, Zhang W, Ma MK, MacLeod HL. Human carboxylesterase 2 is commonly expressed in tumor tissue and is correlated with the activation of irinotecan. Clin Cancer Res 2002;8:2605–2611. [PubMed: 12171891]

Zhen L, Rusiniak ME, Swank RT. The beta-glucuronidase propeptide contains a serpin-related octamer necessary for complex formation with egasyn esterase and for retention within the endoplasmic reticulum. J Biol Chem 1995;270:11912–11920. [PubMed: 7744842]

Fig. 1.

Amino acid sequence alignments for mammalian CES3 and human CES1, CES2, CES5 and CES6 subunits. See Table 1 for sources of CES sequences; * shows identical residues for CES subunits;: similar alternate residues;. dissimilar alternate residues; residues involved in endoplasmic reticulum processing at N-(Signal peptide) and C-termini (MTSmicrosomal (endoplasmic reticulum) targeting sequence) (in *red bold*); *N*-glycosylation residues at 79NAT (Human CES1) and potential *N*-glycosylation sites (in *blue bold*); active site (AS) triad residues Ser; Glu; and His (in *shaded pink*). 'Side door', 'Gate' residues and Cholesterol binding Gly residue (Z site) for human CES1 (in *shaded dark green*); Disulfide bond Cys residues for human CES1 (in *shaded blue*): S–S. Charge clamp residues identified for human CES1 (in *shaded purple*); Helix (Human CES1) or predicted helix (in *shaded yellow*); Sheet (Human CES1) or predicted sheet (*shaded light grey*);. Bold font shows known or predicted exon junctions. Exon numbers refer to human CES1 gene. Color designatations were not available for the printed version.

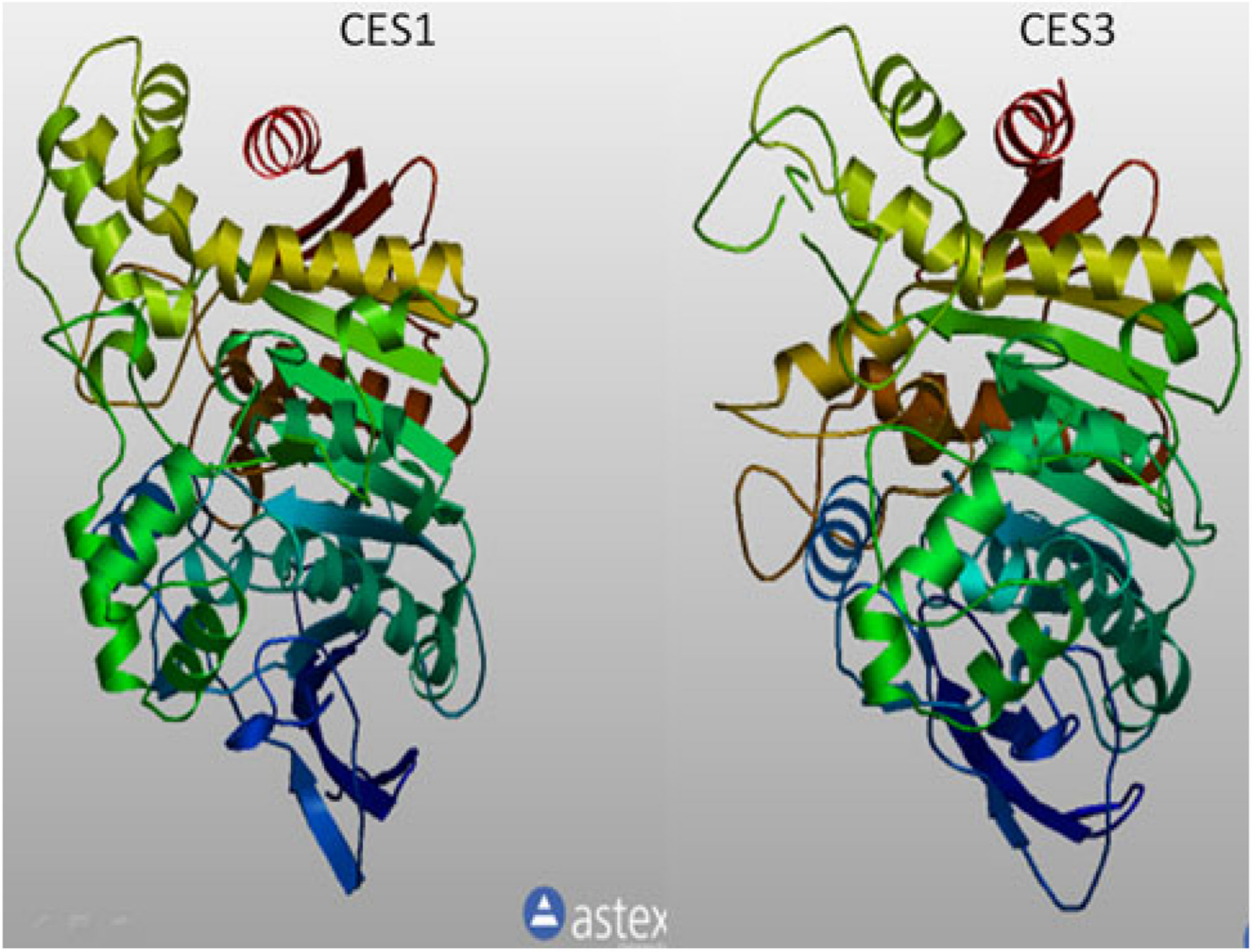

Fig. 2.
Tertiary structure for human CES1 subunit and predicted tertiary structure for human CES 3 subunit. The structure for human CES1 is taken from Bencharit et al. (2003). Predicted human CES3 3-D structure was obtained using the SWISS MODEL web site http://swissmodel.expasy.org/workspace/index.php? And the reported amino acid sequence (Sanghani et al. 2004). The rainbow color code describes the 3-D structures from the N- (*blue*) to C-termini (*red color*) (Color figure online)

Fig. 3.

Nucleotide sequence alignments for mouse CES3 'Like' genes: *CES3* (*Es31*) and *CES3L* (*Es31L*): predicted exons 2–4-; and predicted introns 2 and 3. Exon sequences are in *CAPITALS* and intron sequence in *lower case*. Note the different percentages of sequence identities for exon 2–4 (98%); introns 2–3 (92%). Nucleotide substitutions are *shaded*. The deduced amino acid sequence for exons 2–4 of mouse CES3 is shown as well as any deduced amino acid substitutions observed for CES3L

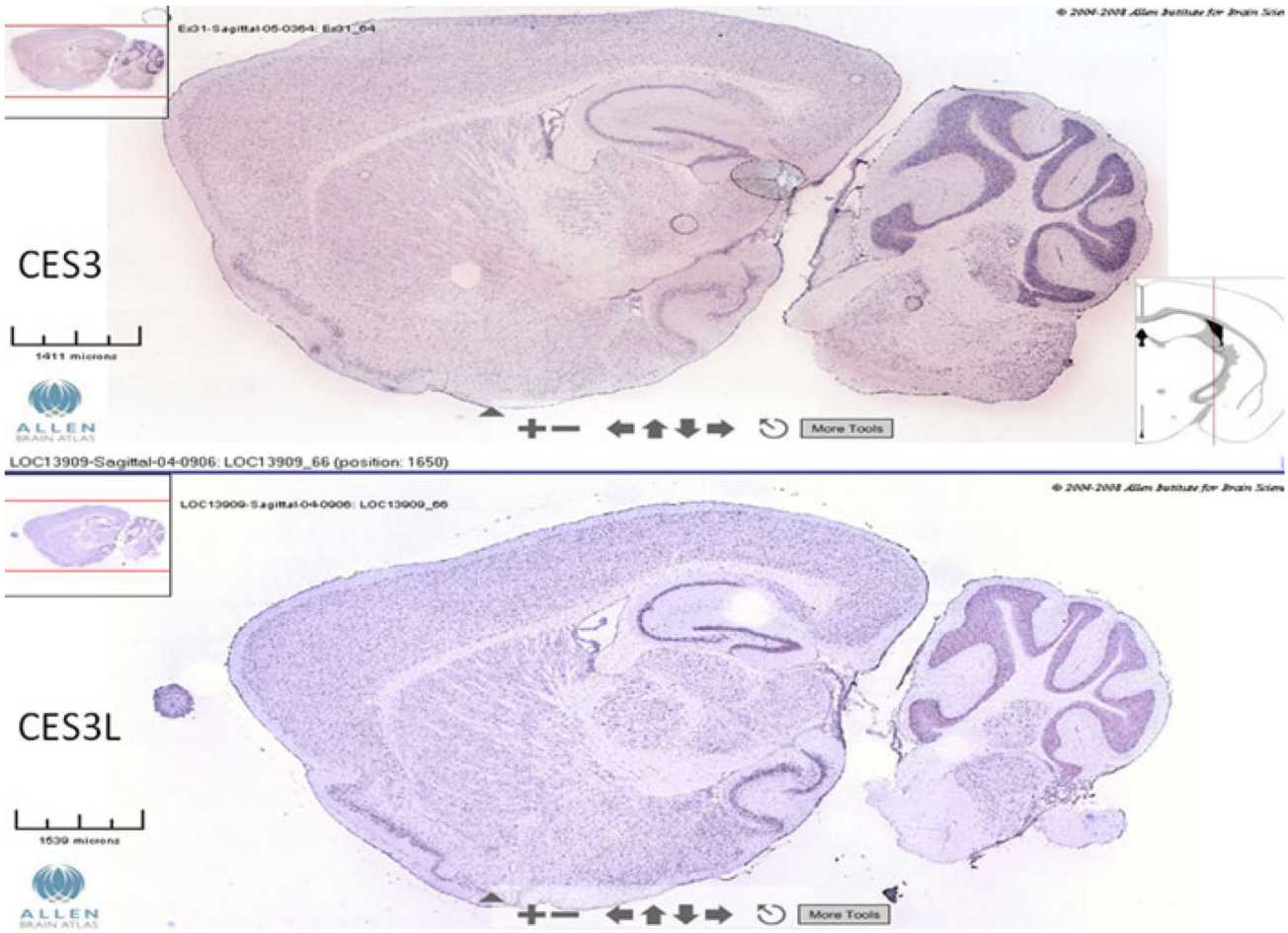

Fig. 4.

Sagittal section of mouse brain showing the distribution of mouse CES3 and CES3L transcripts. The two distinct regions of the mouse brain (C57BL6/6J strain) included the cerebellum (*right*) and the cerebrum (*left*). Note major staining of CES3 and CES3L transcripts in the cerebellum folds and the hippocampus and amygdalar nuclei regions of the cerebrum. The sections were obtained from the Allen Brain Atlas web site (http://www.brain-map.org) (Lein et al. 2007) using the GenBank IDs BC061005 and BC019147, respectively

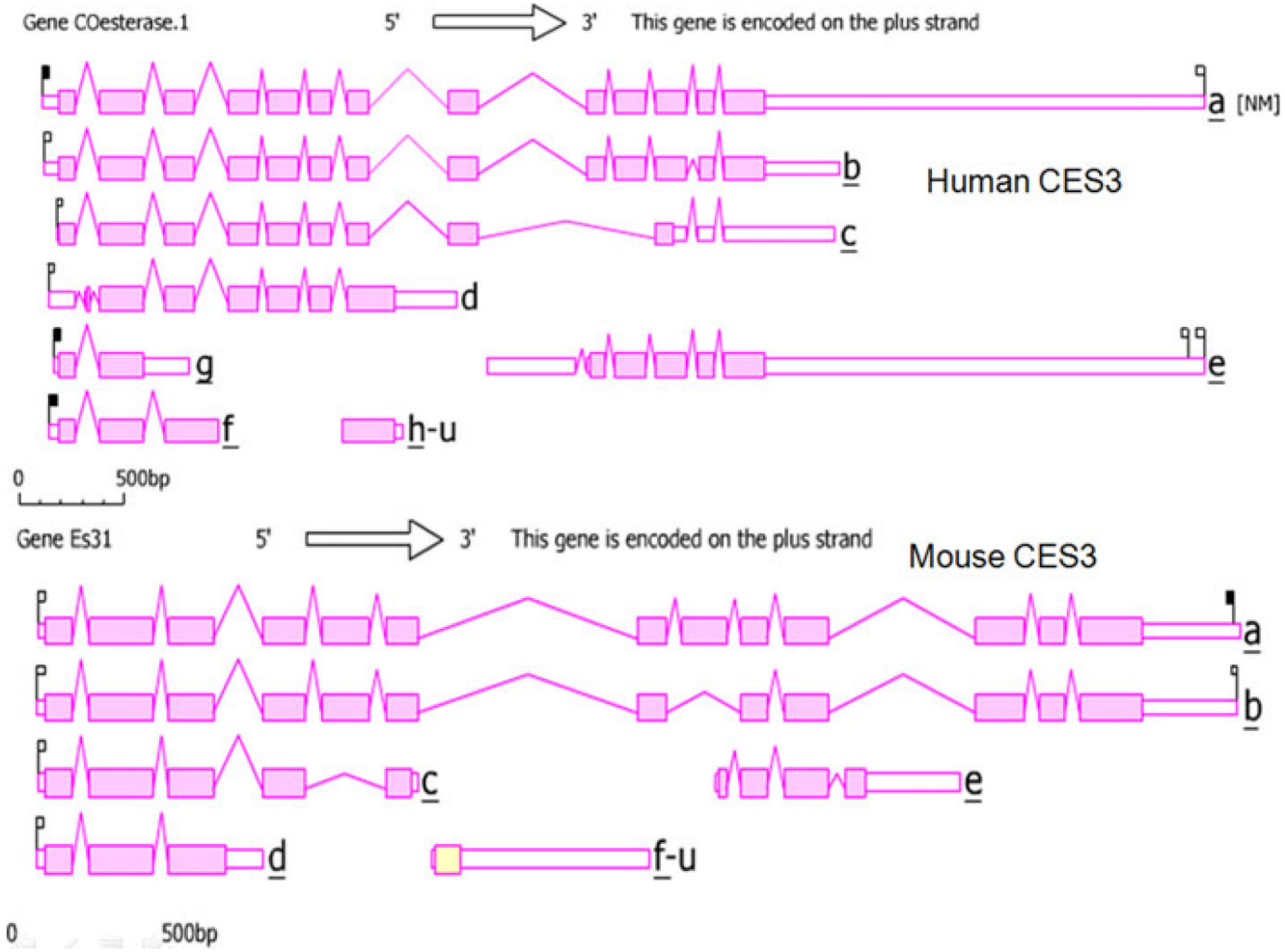

Fig. 5.

Gene structures and splicing variants for the human CES3 (NM024922) and the mouse CES3 (BC061004) genes. Derived from AceView website (Thierry-Mieg and Thierry-Mieg 2006) http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/. Mature isoform variants (*a*, *b*, *c*, etc.) are shown with capped 5′- and 3′-ends for the predicted mRNA sequences. Gene COesterase 1 refers to human CES3 gene; Gene Es31 refers to mouse CES3 gene

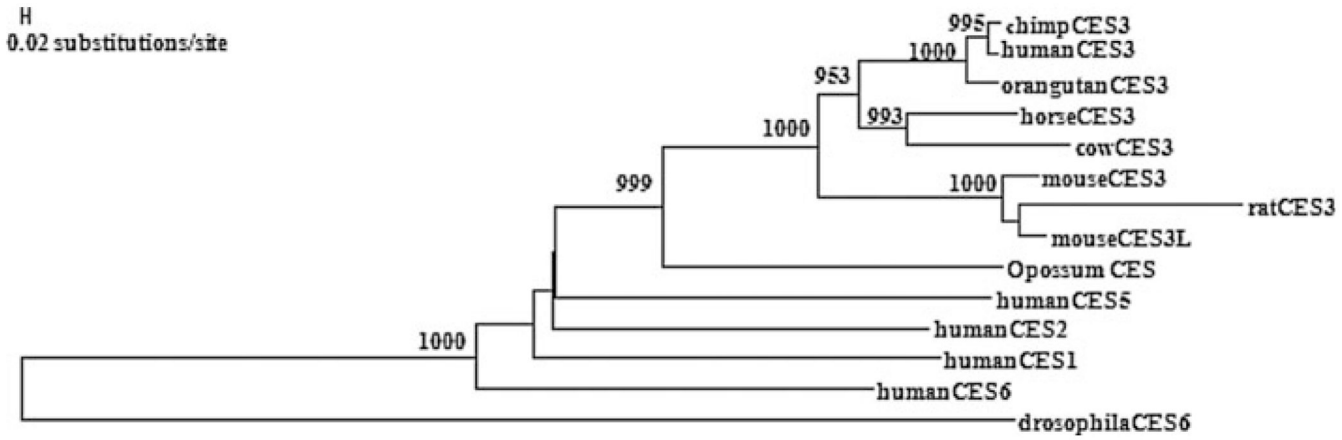

Fig. 6.

Phylogenetic tree of mammalian CES3, human CES1, CES2, CES5 and CES6 and fruit fly (*Drosophila melanogaster*) CES6 amino acid sequences. The tree is labeled with the gene name and the name of the animal. Note the major cluster for mammalian CES3 'like' sequences and the separation of these sequences from human CES1, CES2, CES5 and CES6 sequences, and from the *Drosophila melanogaster* CES6 (Est6) sequence, which served as the 'root' for the tree. Bootstrap values >995 (of 1,000 bootstrap analyses conducted) are shown. *Bar*, 2 changes per 100 amino acid residues. The gene duplication events generating the five distinct gene families (*CES1, CES2, CES3, CES5* and *CES6*) have been previously estimated to have occurred ~328–378 million years ago (Holmes et al. 2008a)

Table 1

Mammalian *CES3* and human *CES1*, *CES2*, *CES5* and *CES6* genes and subunits

Animal	*CES* gene	Other gene name	GenBank ID *RefSeq NCBI Accession No.	UNIPROT ID	No of amino acids	Chromosome location	Strand	CES gene order[a]	Gene size (bps)	No. of exons	PI[b]	Subunit MW	Predicted subcellular location
Human	*CES3*	*CES31*	BC053670	Q6UWW8	571	16: 65,552,639–65,566,552	Positive	CES2-**CES3**-CES6	13,913	13	5.4	62,282	Extracellular[c]
Chimp	*CES3*	*CES31*	*XP_511224		571	16: 66,682,109–66,695,987	Positive	CES6-CES3L-CES2-**CES3**	13,879	13	5.7	62,471	Extracellular[c]
Chimp	*CES3L*	*CES31L*	*XP_001173132		479	16: 5,892,348–5,905,878	Positive	CES6-**CES3L**-CES2-CES3	13,531	11	6.6	52,573	Cytoplasmic
Orangutan	*CES3*	*CES31*	CR857194	Q5RCL7	569	16: 54,242,602–54,266,316	Positive	**CES3**-CES2-CES6	23,714	13	6.4	62,409	Extracellular[c]
Horse	*CES3*	*CES31*	*XP_001496251		570	3: 17,417,834–17,428,268	Positive	CES2-**CES3**-CES6	10,434	13	6.3	62,720	ER[d]
Cow	*CES3*	*CES31*	*XP_590749		570	18: 33,671,230–33,683,956	Positive	CES2-**CES3**-CES6	12,727	13	5.6	62,616	ER[d]
Mouse	*CES3*	*CES31*	BC061004	Q63880	554	8: 107,572,572–107,582,000	Positive	^CES2-**CES3**-CES3L-CES6	9,429	13	5.9	61,510	Extracellular[c]
Mouse	*CES3L*	*CES31L*	BC019147		568	8: 107,607,684–107,617,159	Positive	^CES2-CES3-**CES3L**-CES6	9,476	13	5.7	63,007	Extracellular[c]
Opossum	*CES3*	*CES31*	*XP_001372405		571	1: 677,687,773–677,741,087	Negative	^CES2-**CES3**-CES6	53,315	13	5.6	63,154	ER[d]
Human	*CES1*		L07765	P23141	567	16: 54,394,265–54,424,576	Negative	CES4-**CES1**-CES5	30,311	14	6.2	62,521	ER[d]
Human	*CES2*	*CE-2*	BX538086	000748	559	16: 65,525,828–65,536,493	Positive	**CES2**-CES3-CES6	10,665	12	5.7	61,807	ER[d]
Human	*CES5*	*cauxin '*	BC069501	Q6NT32	575	16: 54,437,867–54,466,634	Negative	CES4-CES1-**CES5**	28,767	13	6	63,936	ER[d]
Human	*CES6*	*CES8*	FLJ37464	Q6UX55	545	16: 65,580,134–65,602,403	Positive	CES2-CES3-**CES6**	22,269	14	9.4	60,366	Mitochondrial
Drosophila	*CES6*	*Este*	AY095091	P08171	544	3L:12,181,959–12,183,641	Positive	**Est6**-EstP	1,747	2	5.8	61,125	Extracellular[c]

*RefSeq NCBI sequence;

^the mouse and opossum genomes contain multiple contiguous CES2 genes

[a]CES gene in bold refers to the specific gene among the contiguous CES genes

[b]pI refers to predicted isoelectric point

[c]Extracellular and/or cell wall location

[d]ER: endoplasmic reticulum location

Table 2

Percentage identities for mammalian CES3 and human CES1, CES2, CES5 and CES6 subunit amino acid sequences

CES gene	Human CES3	Chimp CES3	Orang CES3	Chimp CES3L	Horse CES3	Cow CES3	Mouse CES3	Mouse CES3L	Opossum CES3	Human CES1	Human CES2	Human CES5	Human CES6
Human CES3	100	97	94	92	77	73	68	68	58	47	50	48	45
Chimp CES3	97	100	95	92	78	75	68	68	58	47	50	48	44
Orang CES3	94	95	100	93	78	74	68	68	58	47	50	48	44
Chimp CES3L	92	92	93	100	76	73	67	67	58	46	49	48	43
Horse CES3	77	78	78	76	100	79	68	68	61	49	50	49	43
Cow CES3	73	75	74	73	79	100	66	65	60	47	47	47	41
Mouse CES3	68	68	68	67	68	66	100	91	57	45	48	48	43
Mouse CES3L	68	68	68	67	68	65	91	100	56	45	47	47	43
Opossum CES3	58	58	58	58	61	60	57	56	100	49	49	48	45
Human CES1	47	47	47	46	49	47	45	45	49	100	51	49	46
Human CES2	50	50	50	49	50	47	48	47	49	51	100	48	44
Human CES5	48	48	48	48	49	47	48	47	48	49	48	100	47
Human CES6	45	44	44	43	43	41	43	43	45	46	44	47	100

Table 3

Predicted *N*-glycosylation sites for mammalian CES3 and human CES1, CES2, CES5 and CES6 subunits

Mammal	CES gene	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6	Site 7	Site 8	Site 9	Site 10	Site 11	Site 12	Site 13	Site 14	Site 15	No of sites
Human	CES1	79NAT															1
	CES2			111NMT			276NLS										2
	CES5							28 INAS				358NLS				513NLT	3
	CES6					214NVT			276NST				370NIT				3
Human	CES 3		105NSS														1
Cow	CES 3		105NSS		212NIT												2
Horse	CES 3				212NIT				284NSS						423NFS		3
Mouse	CES 3									295NIS	301NDS			381NGS			3
Mouse	CES3L									309NIS	323NDS			395NGS	421NFS		4

The identified *N*-glycosylation site is for human CES1 (Kroetz et al. 1993). Amino acid residues are shown for potential *N*-glycosylation sites: *N*, Asn; *A*, Ala; *T*, Thr; *S*, Ser; *M*, Met; *L*, Leu; *D*, Asp; *G*, Gly; *F*, Phe; *I*, Ile

Table 4

Tissue distribution of human mRNA transcripts and predicted encoded CES3 isoforms

Human CES3 isoform	No. of amino acids	No. of exons	Gene size (kb)	N-signal peptide (location)	Predicted subcellular location	Tissue distribution (mRNA frequency)	Predicted CES type B domain	GenBank ID
CES3a	571	13	13.9	Yes (26–27)	Secreted or extracellular	Colon (8), breast, trachea (4) brain, placenta, rectum (2)	Yes	NM024922
CES3b	568	13	12.2	Yes (26–27)	Secreted or extracellular	Colon (10) trachea (6) brain, breast (5) placenta (4) testis (3)	Yes	BC053670
CES3c	379	11	12.1	Yes (26–27)	Secreted or extracellular	Colon (10) trachea (6) brain breast (5) placenta (4) testis (3)	Yes	CR618303
CES3d	326	8	5.5	No	Cytosol	Ovary, teratocarcinoma, stomach, trachea (1)	Yes	*DB232966
CES3e	210	6	7.7	No	Extracellular	Colon (5), nerve tumor, uterus, brain (2)	Yes	AK025389
CES3f	181	3	2.5	Yes (26–27)	Extracellular	Bladder, uterus (1)	Yes	DA040317
CES3g	96	2	2.3	Yes (26–27)	Extracellular	Trachea (1)	Yes	DB215224
CES3h–u	85	1	0.3	No	Extracellular	Stomach (1)	No	AW813514

Predicted CES3 gene sizes (kilobases); CES type B domain refers to the presence of essential amino acids at the CES active site; N-signal peptide location refers to the predicted site of hydrolysis for the N-terminal peptide involved in intracellular CES3 processing. Note the higher frequencies for CES3a, CES3b and CES3c expressed in colon, trachea and breast tissue

*Refers to a partial mRNA sequence