Table 2 continued

Mouse CES gene (proposed)	Chr 8 coordinates	Gene size (bp)	Exons	Strand[a]	Subunit MW	Amino acids	GenBank ID	MGI ID_YZ	Current MGI symbol_YZ	Current gene symbols	NCBI transcript	Vega ID	Ensembl ID	UNIPROT ID	Tissue expression (relative)[b]
Ces5a	96,038,095–96,059,607	21,512	13	+ve	64,167	575	AB186393	MGI:1915185	*Ces7*	*Ces7*	NM_001003951	None	ENSMUSG00000058019	Q8R0W5	Prostate [0.03]

RefSeq, GenBank, UNIPROT, MGI, Vega, and Ensembl IDs provide the sources for the gene and protein sequences; gene sizes are given as base pairs of nucleotides

http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/

ps pseudogene (*Ces2d-ps*)

[a] +ve and −ve = transcription strand

[b] The relative gene expression level for mouse *Ces* genes in comparison with the expression of an average mouse gene is given in brackets

CES1-like and CES2-like subunits exhibited higher levels of sequence identities with the CES family homolog in each case [66–78% identities for human and mouse CES1-like subunits and 64–72% for human and mouse CES2-like subunits, respectively (data not shown)], suggesting that these are members of the same mammalian CES families, in each case. Similar results were observed for comparisons of human CES3, CES4A (previously CES6 or CES8), and CES5A (previously CES7) with the corresponding mouse CES homolog sequences, with 65, 72, and 69% identities being observed, respectively. This supports the designation of these *CES* genes as members of the same family, in each case.

The amino acid sequences for the human CES subunits examined contained 567 (CES1), 559 (CES2), 571 (CES3), 561 (CES4A), and 575 (CES5A) residues (Fig. 1). Previous studies on human CES1 have identified key residues that contribute to the catalytic, oligomeric, subcellular localization and regulatory functions for this enzyme (sequence numbers refer to human CES1). These included the catalytic triad for the active site (Ser221; Glu354; His468) (Cygler et al. 1993); disulfide bond-forming residues (Cys87/Cys116 and Cys274/Cys285) (Lockridge et al. 1987); microsomal targeting sequences, including the hydrophobic N-terminus signal peptide (Potter et al. 1998; von Heijne 1983; Zhen et al. 1995) and the C-terminal endoplasmic reticulum (ER) retention sequence (His-Ile-Glu-Leu) (Robbi and Beaufay 1983); and ligand-binding sites, including the "Z-site" (Gly356), the "side door" (Val424-Met425-Phe426), and the "gate" (Phe550) residues (Bencharit et al. 2003, 2006; Fleming et al. 2005). Identical residues were observed for each of the human CES subunit families for the active site triad and disulfide bond-forming residues, although changes were observed for some key residues for CES1 subunits, including the "side-door" and "gate" of the active site, with family-specific sequences or residues in each case. The "Z-site" (Gly356 for human CES1) has been retained for human CES2 and CES5A sequences, but substituted for CES3 (Ser) and CES4A (Asn). The hydrophobic N-terminal sequence for human CES sequences has undergone major changes, although this region retains a predicted signal peptide property. The human CES C-terminal tetrapeptide sequences have also changed, although CES2 (HTEL) and CES3 (QEDL) are similar in sequence with human CES1 (HIEL), which plays a role in the localization of human CES1 within endoplasmic reticulum membranes (Robbi and Beaufay 1983).

Other key human CES1 sequences included two charge clamps that are responsible for subunit-subunit interaction, namely, residues Lys78/Glu183 and Glu72/Arg186, which contribute to the trimeric and hexameric structures for this enzyme (Bencharit et al. 2003, 2006; Fleming et al. 2005).

Table 3 Rat *Ces* genes and subunits

Rat CES gene (proposed)	Chromosomes 19 (and 1) coordinates	Gene size (bp)	Exons strand[a]	Subunit MW	Amino acids	GenBank ID	RGD ID	Ortholog	Current gene symbols	NCBI RefSeq ID	Ensembl transcript ID	UNIPROT ID	Tissue expression [relative]
Ces1a	19:15,025,350–15,051,534	26,185	14 +ve	62,362	563		RGD:1583671	Mouse Gm4976	*LOC679817*	XM_001054575	ENSRNOT00000060929	D4AA05	[0.01]
Ces1c	19:14,981,539–15,021,040	39,502	14 +ve	60,501	550	BC088251	RGD:2571	Mouse Es1	*Es1*	NM_017004	ENSRNOT00000024622	P10959	Liver [0.2]
Ces1d	19:14,928,590–14,966,890	38,301	14 +ve	62,150	565	BC061789	RGD:70896	Mouse Ces3	*Ces3*	NM_133295	ENSRNOT00000021812	P16303	Liver, lung [0.4]
Ces1e	19:14,887,969–14,924,191	36,223	14 +ve	61,715	561	X81395	RGD:621508	Mouse Es22	*Ces1, Es22*	NM_031565	ENSRNOT00000020775	Q924V9	Liver [0.1]
Ces1f	19:14,849,955–14,876,723	26,769	14 +ve	62,495	561	BC128711	RGD:1642419	None specified	*LOC100125372*	NM_001103359	ENSRNOT00000024187	Q64573	Kidney, liver [0.1]
Ces2a	19:37,855–44,723	6,869	13 −ve	61,802	558	AY834877	RGD:708353	Mouse Ces6	*Ces6*	NM_144743	ENSRNOT00000015451	Q8K3RO	Liver [0.05]
Ces2c	1:267,887,436–267,894,795	7,360	12 +ve	62,170	561	AB010632	RGD:621510	Mouse Ces2	*Ces2l*	NM_133586	ENSRNOT00000045656	O70631, O70177	Brain, liver [0.1]
Ces2e	19:65,698–80,142	14,445	12 +ve	62,410	557	D50580	RGD:621563	Mouse Ces5	*Ces5*	NM_001100477	ENSRNOT00000015724	O35535	Liver [0.01]
Ces2g	19:34,883,500–34,890,289	6,790	12 +ve	62,909	560	CH473972	RGD:1308358	Mouse 2210023G05Rik	*2210023G05Rik*		ENSRNOT00000048385	D3ZXQ0	Kidney, liver [0.06]
Ces2h	19:34,910,987–34,925,261	14,275	12 +ve	62,280	557	BC107806	RGD:1560889	Gm5744	*Ces2*	NM_001044258	ENSRNOT00000019072	Q32Q55	Intestine [0.08]
Ces2i	1:267,807,848–267,815,235	7,388	11 +ve	62,072	559	XM212849	RGD:1565045	Mouse Ces2	*RGD1565045*	XM_001074128	ENSRNOT00000015997	D3ZE31	Not available
Ces2j	19:215,376–222,512	7,137	12 +ve	61,795	556		RGD:1591368	Mouse Ces2	*LOC685645*	XM_001074128	ENSRNOT00000061734	D3ZP14	[0.01]
Ces3a	19:34,929,247–34,937,264	8,018	14 +ve	62,393	563		RGD:1588734	Human CES3			ENSRNOT00000040499		Not available
Ces4a	19:34,948,579–34,965,647	17,069	14 +ve	63,446	563		RGD:1307418	Mouse Ces8	*Ces8*	NM_001106176	ENSRNOT00000019169	D4AE76	[0.01]
Ces5a	19:11,910,831–11,938,412	27,582	11 +ve	64,401	575	AF479659	RGD:1549717	Mouse Ces7	*Ces7*	NM_001012056	ENSRNOT00000049452	Q5GRG2	[0.01]

RefSeq, GenBank, UNIPROT, RGD, Vega, and Ensembl IDs provide the sources for the gene and protein sequences; gene sizes are given as base pairs of nucleotides; the relative gene expression level for rat *Ces* genes in comparison with the expression of an average rat gene is given in brackets

http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/

[a] +ve and −ve = transcription strand direction

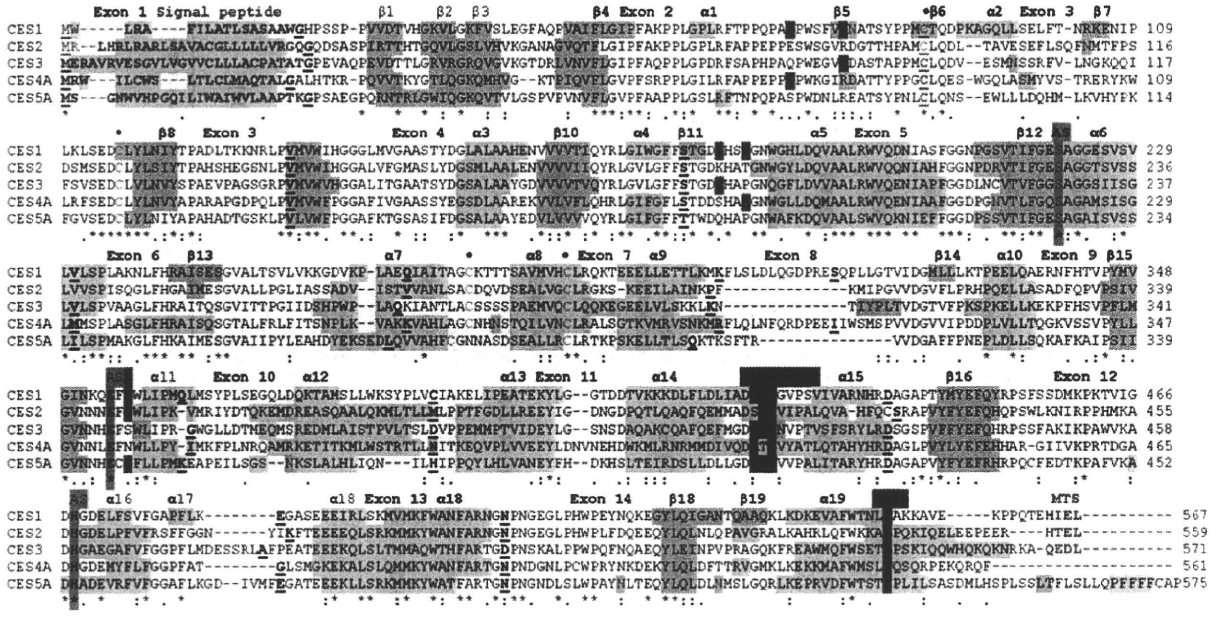

Fig. 1 Amino acid sequence alignments for human *CES1*, *CES2*, *CES3*, *CES4A*, and *CES5A* subunits. See Table 1 for CES isoform sequences aligned. *Asterisk* identical residues for CES subunits; *colon* similar alternate residues; *dot* dissimilar alternate residues. Signal peptide sequences for CES1 (1–17), CES2 (1–25), CES3 (1–27), CES4A (1–19), and CES5A (1–24) and C- termini (MTS) microsomal targeting sequences for CES1 (564–567), CES2 (556–569), and CES3 (568–571) are shown in *red*. Active site (AS) triad residues (human CES1) Ser221, Glu354, and His468 are highlighted in *green*. "Side door" (Val424-Met425-Phe426), "Gate" (Phe550), and cholesterol binding residue ("Z site") (Gly356) for human CES1 (Fleming et al. 2005) are highlighted in *khaki*. Disulfide bond Cys residues for human CES1 (*filled circle*) are shown in *blue*. Charge clamp residues identified for human CES1 (Glu72···Arg186; Lys78···Glu183) (Fleming et al. 2005) are highlighted in *purple*. Confirmed (CES1) (Asn79-Ala80-Thr81) [site 1] or predicted N-glycosylation sites for human CES2 (Asn111-Met112-Thr113) [site 3]; CES3 (Asn105-Ser106-Ser107) [site 2]; CES4A (Asn213-Val214-Thr215) [site 4], Asn276-Ser-277-Thr278) [site 5], and Asn388-Ile389-Thr390) [site 7]; and CES5A (Asn363-Lys364-Ser365) [site 6], (Asn513-Leu514-Thr515) [site 8], and (Asn524-Met525-Ser526 [site 9] are highlighted in blue. α-Helix (human CES1 or predicted) and β-sheet (human CES1 or predicted) regions are highlighted in *yellow* and *gray*, respectively. α-Helices and β-sheets are numbered according to the reported human CES1 3D structure (Fleming et al. 2005). *Bold underlined font* shows known or predicted exon start sites; exon numbers refer to the human *CES1* gene (see Langmann et al. 1997). (Color figure online)

Other human CES subunit sequences for these charge clamp sites included substitutions with neutral amino acids for the human CES2 and CES5A sequences, while the CES3 and CES4A sequences retained one potential clamp site (Fig. 1). Pindel et al. (1997) and Holmes et al. (2009b) have reported monomeric subunit structures for human and baboon CES2, which is consistent with the absence of charge clamps for this enzyme. This could have a major influence on the kinetics and biochemical roles for human CES isozymes since three-dimensional studies have indicated that ligand binding to the human CES1 "Z-site" shifts the trimer–hexamer equilibrium toward the trimer that facilitates substrate binding and enzyme catalysis (Redinbo and Potter 2005). The *N*-glycosylation site for human CES1 at Asn79-Ala80-Thr81 (Bencharit et al. 2003, 2006; Fleming et al. 2005; Kroetz et al. 1993) was not retained for any of the other human CES sequences, although potential *N*-glycosylation sites were observed at other positions, including CES2 (site 3), CES3 (site 2), CES4A (sites 4, 5, and 7), and CES5A (sites 6, 8, and 9) (Table 4). Given the reported role of the *N*-glycosylated carbohydrate group contributing to CES1 stability and maintaining catalytic efficiency (Kroetz et al. 1993), the *N*-glycosylation sites predicted for other human CES subunits may perform similar functions or indeed may serve new functions specific to a particular CES family.

Predicted secondary structures for human CES2 (Holmes et al. 2009b), CES3 (Holmes et al. 2010), CES4A (Holmes et al. 2009a), and CES5A (Holmes et al. 2008a) sequences were compared with those reported for human CES1, and similar α-helix β-sheet structures were observed for all of the CES subunits examined (Bencharit et al. 2003, 2006) (Fig. 1). This was especially apparent near key residues or functional domains such as the α-helix within the N-terminal signal peptide, the β-sheet and α-helix structures near the active site Ser221 (human CES1) and "Z-site" (Glu354/Gly356, respectively), the α-helices bordering the "side door" site, and the α-helix containing the "gate" residue (Phe550 for human CES1). The human CES5A sequence, however, contained a predicted helix at

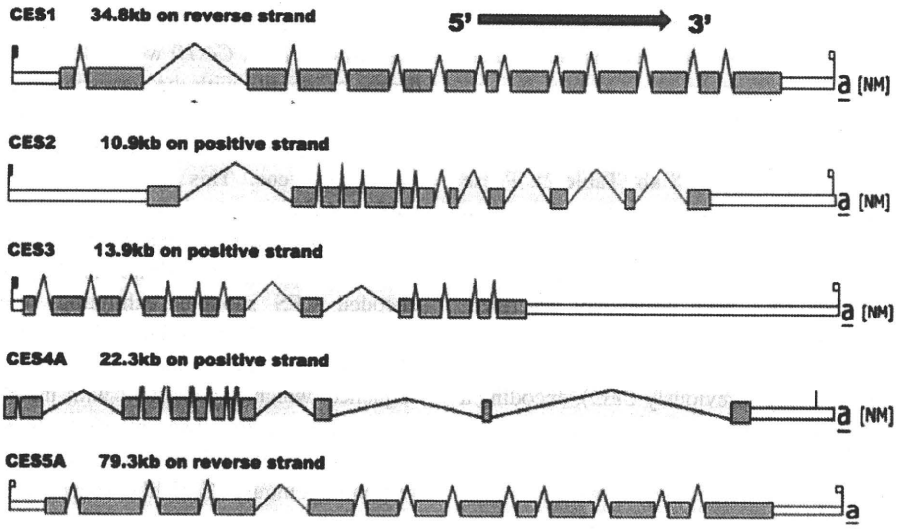

Fig. 2 Gene structures and major isoforms for human *CES1*, *CES2*, *CES3*, *CES4A*, and *CES5A* genes. Derived from AceView website http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/ (Thierry-Mieg and Thierry-Mieg 2006). Mature isoform variants (a) are shown with capped 5′ and 3′ ends for the predicted mRNA sequences. Exons are in *solid color*. 5′ and 3′ untranslated regions of the genes are shown as *open boxes*. Introns are shown as a line. The 5′ → 3′ transcription directions are shown. a refers to the major transcript isoform for each human *CES* gene. Note that each *CES* gene structure is drawn to a different scale and that the respective gene sizes are shown: *CES1*, 34.8 kb; *CES2*, 10.9 kb; *CES3*, 13.9 kb; *CES4A*, 22.3 kb; and *CES5A*, 79.3 kb. (Color figure online)

the hydrophobic C-terminus not observed for other CES subunits which may perform a family-specific function. Predicted 3D structures have been previously described for each of the human CES subunits (Holmes et al. 2008a, 2009a, b, 2010); they were similar to the human CES1 structure (Bencharit et al. 2003, 2006).

Mouse *Ces* genes and enzymes

Table 2 summarizes the proposed names, locations, and overall structures for the *Ces* genes observed for the mouse genome (July 2007 mouse [*Mus musculus*] genome data obtained from the Build 37 assembly by NCBI and the Mouse Genome Sequencing Consortium) (http://www.ncbi.nlm.nih.gov was used in this study). The italicized gene name *Ces* is consistent with other mouse gene nomenclature and is preferred to the *CES* stem used for human genes. At least 20 mouse *Ces* genes are recognized on the Mouse Genome Database http://www.informatics.jax.org/) (MGI) and further described in terms of their locations on mouse chromosome 8, the number of predicted exons for each gene, predicted strand for transcription, number of amino acid residues and subunit molecular weights (MWs) for the encoded CES subunits, and identification symbols from MGI (e.g., MGI3648919 for *Ces1a*), NCBI (Reference Sequences were identified from the National Center for Biotechnology Information database) (http://www.ncbi.nlm.nih.gov/), Vega (the VErtebrate Genome Annotation database) (http://vega.sanger.ac.uk/index.html), UNIPROT (Universal Protein Resource) (http://www.ebi.ac.uk/uniprot/), and Ensembl (Genome Database) (http://www.ensembl.org/) database sources.

Eight *Ces1*-like genes are located in tandem within a 360-kb segment of mouse chromosome 8, with an average gene size of 28 kb. The names for these genes (*Ces1a*, *Ces1b*,..., *Ces1h*) are allocated in the same order as their locations on the mouse genome (Table 3). The *Ces1*-like gene cluster is also located near the mouse *Ces5a* gene, which is comparable to the *CES1P1–CES1–CES5A* cluster observed for human chromosome 16. Each of these genes contained 13 or 14 exons predicted for transcription on the negative strand and with encoded CES subunits exhibiting distinct but similar amino acid sequences (554–567 residues). The subunits were 63–85% identical with each other and with the human CES1 sequence, which is consistent with these being members of the mouse *Ces1* gene family. Mouse *Ces1*-like genes included several that have been previously investigated, including *Ces1c* (previously called *Es1*), encoding a major mouse plasma esterase with 554 amino acid residues and also exhibiting lung surfactant convertase activity (Genetta et al. 1988; Krishnasamy et al. 1998); *Ces1d* (previously *Ces3*), encoding a mouse liver enzyme with 565 residues and exhibiting triacylglycerol hydrolase activity (Dolinsky et al. 2001); *Ces1e* (previously called *Es22* or *egasyn*), encoding a liver CES with 562 residues and exhibiting β-glucuronidase-binding properties (Ovnic et al. 1991); and *Ces1g* (previously

Ces1), encoding a liver CES with 565 amino acid residues and exhibiting lipid metabolizing activity (Table 4) (Ellingham et al. 1998).

Eight *Ces2*-like genes were also observed in a second 286-kb gene cluster on mouse chromosome 8, with an average gene size of approximately 8 kb (Table 2). These genes were named according to their sequence of position on the mouse genome (*Ces2a*, *Ces2b*,..., *Ces2h*) and included a pseudogene designated *Ces2d-ps*. Three of these mouse *Ces2*-like genes have been previously described, including *Ces2c* (previously *Ces2*), encoding an inducible liver acyl-carnitine hydrolase enzyme with 561 residues (Furihata et al. 2003); *Ces2e* (previously *Ces5*), encoding a liver and intestinal enzyme with 560 amino acid residues (The MGC Project Team 2004); and *Ces2a* (previously *Ces6*), encoding a liver and colon enzyme with 558 residues (The MGC Project Team 2004). The *Ces2*-like cluster was located alongside two *Ces3*-like mouse genes (*Ces3a* and *Ces3b*) and a *Ces4a* gene (Table 3); this is comparable to the *CES2-CES3-CES4A* gene cluster on human chromosome 16 (Table 1). The *Ces3a* gene (previously mouse *esterase 31* or *Est31*) is expressed strongly in male mouse livers and encodes a 554-residue CES3-like subunit (Aida et al. 1993), whereas the *Ces3b* gene (previously *Es31L* or *EG13909*) is also expressed in liver and encodes a 568-residue subunit (The MGC Project Team 2004). The *Ces4a* gene (previously called *EST8* or *Ces8*) encodes an enzyme predicted for secretion in epidermal cells with 563 amino acid residues and showing 72% identity with human CES4A (The MGC Project Team 2004).

Rat *Ces* genes and enzymes

Table 3 summarizes the proposed names, locations, and structures for *Ces* genes observed for the rat genome [the November 2004 rat (*Rattus norvegicus*) genome assembly based on version 3.4 produced by the Baylor Human Genome Sequencing Center (Gibbs et al. 2004) was used in this study]. Fifteen rat *Ces* genes were identified on the Rat Genome Database (RGD) (http://rgd.mcw.edu/) and further characterized by their locations on rat chromosomes 1 and 19, the number of predicted exons for each gene, the predicted strand for transcription, current gene symbols, the number of amino acid residues and subunit MWs for the encoded CES subunits, and the identification symbols from RGD (e.g., RGD1583671 for *Ces1a*), NCBI Reference Sequences (http://www.ncbi.nlm.nih.gov/), Vega (http://vega.sanger.ac.uk/index.html), UNIPROT (http://www.ebi.ac.uk/uniprot/), and Ensembl (http://www.ensembl.org/) database sources.

Five *Ces1*-like genes were located in tandem within a 201-kb segment of rat chromosome 19, with an average gene size of 33 kb (Table 3). The names for these genes (*Ces1a*, *Ces1c*,..., *Ces1f*) were allocated according to their degree of identity with the corresponding mouse *Ces1*-like genes (Table 3). The genes were located in tandem in the same order as the mouse *Ces1*-like genes and were near the rat *Ces5a* gene. This is comparable to the *CES1P1–CES1A–CES5A* gene cluster observed for human chromosome 16. The rat *Ces1*-like genes contained 14 exons and were predicted for transcription on the positive strand, with encoded CES subunits exhibiting similar amino acid sequences (550–565 residues). The subunits were 65–73% identical with each other and with the human CES1 sequence, which is consistent with membership of the rat *Ces1* gene family. The encoded rat *Ces1*-like subunit sequences showed higher levels of identity with the corresponding mouse *Ces1*-like sequences (81–92% for rat and mouse CES1a, CES1c, CES1d, CES1e, and CES1f amino acid sequences). At least three rat *Ces1*-like genes have been previously described, including *Ces1c* (previously called *Es1*), encoding a rat plasma esterase (Sanghani et al. 2002; Vanlith et al. 1993); *Ces1d* (previously *Ces3*), encoding a rat liver enzyme with 565 residues and exhibiting cholesteryl ester hydrolase activity (Ghosh et al. 1995; Robbi et al. 1990); and *Ces1e* (previously called *ES-3* or *egasyn*), encoding a rat liver Ces with 561 residues and having β-glucuronidase-binding properties (Robbi and Beaufay 1994).

Seven rat *Ces2*-like genes were observed on the rat genome and were localized on two chromosomes: chromosome 1 (*Ces2c* and *Ces2i*) and chromosome 19 in three locations: *Ces2a* and *Ces2e*; *Ces2j*; and *Ces2g* and *Ces2h* (Table 3). The genes were named according to the degree of sequence identity with the corresponding mouse *Ces2*-like genes. Rat *Ces2*-like genes have been previously investigated, including *Ces2c* (previously *Ces2*), encoding an inducible liver acyl-carnitine hydrolase enzyme with 561 residues (Furihata et al. 2003); *Ces2e* (previously *Ces5*), encoding a liver and intestinal enzyme with 560 amino acid residues (The MGC Project Team 2004); and *Ces2a* (previously *Ces6*), encoding a liver and colon enzyme with 558 residues.(The MGC Project Team 2004). The rat *Ces2*-like cluster was located alongside a *Ces3*-like gene (*Ces3a* and *Ces3b*) and a *Ces4a* gene (Table 3), which is comparable to the *CES2A-CES3A-CES4A* gene cluster on human chromosome 16 (Table 1).

Functions of mammalian *CES* families

Mammalian CES families exhibit broad substrate specificities, and specific roles for these enzymes have been difficult to establish because of the promiscuity of the CES active site toward a wide range of substrates and the

Table 4 Functions and substrates for human *CES* and mouse and rat *Ces* genes and enzymes

Mammal	*CES* (*Ces*) gene	Current gene symbol(s)	Substrates and function (hydrolysis or detoxification)
Human	*CES1*	*CES1, hCE-1, CES1A1, HU1*	Heroin, cocaine[1–3], methyl phenidate[4], temocapril[5], CPT-11[6], flurbiprofen[7]
		CES1	Fatty acid ethyl ester synthase[8], sarin[9], ciclesonide[10], cholesteryl ester hydrolase[11], triacylglycerol hydrolase[11]
	CES2	*CES2, hCE-2, HU2*	Procaine[5], heroin, cocaine[1–3], temacapril[5], CPT-11,6 flurbiprofen[7], doxazolidine[12]
	CES3	*CES3*	CPT-11[6]
Mouse	*Ces1c*	*Es1, Ces-N*	Lung surfactant convertase[13], CPT-11[14]
	Ces1d	*Ces3*	Triacylglycerol hydrolase[15]
	Ces1e	*Es22, egasyn*	β-glucuronidase binding in the liver endoplasmic reticulum[16], retinyl ester hydrolase[26]
	Ces1f	*CesML1, TGH-2*	Triacylglycerol hydrolase[27], monoacylglycerol hydrolase[27], cholesteryl ester hydrolase[27], phospholipase[27]
	Ces1g	*Ces1*	Lipid metabolism[17]
	Ces2c	*Ces2*	Inducible liver acylcarnitine hydrolase[18]
Rat	*Ces1c*	*Es1*	Retinyl palmitate[19]
	Ces1d	*Ces3*	Cholesterol ester hydrolase[20], triacylglycerol hydrolase[27], retinyl ester hydrolase[28]
	Ces1e	*ES-3*	β-glucuronidase binding in the liver endoplasmic reticulum[21]
	Ces2a	*Ces6*	Intestinal first pass metabolism[22]
	Ces2c	*Ces2*	Inducible liver acylcarnitine hydrolase[18], intestinal first pass metabolism[22]
	Ces2e	*Ces5*	Intestinal first pass metabolism[22]
Cat	*CES5A*	*CES7, cauxin*	3-Methylbutanol-cysteinylglycine hydrolysis in urine releasing pheromone[23]
Rat, sheep	*CES5A*	*CES7, cauxin*	Lipid transfer reactions in epididymis[24]

[1] Pindel et al. 1997, [2] Bencharit et al. 2003, [3] Satoh and Hosokawa 2006, [4] Sun et al. 2004, [5] Takai et al. 1997, [6] Humerickhouse et al. 2000, Xu et al. 2002, Ohtsuka et al. 2003, Morton et al. 2005, [7] flurbiprofen derivatives serve as substrates, Imai 2006, Taketani et al. 2007, Hosokawa 2008, [8] Diczfalusy et al. 2001, [9] Hemmert et al. 2010, [10] Mutch et al. 2007, [11] Becker et al. 1994, [12] Barthel et al. 2008, [13] Krishnasamy et al. 1998, Ruppert et al. 2006, [14] Morton et al. 2005, [15] Dolinsky et al. 2005, [16] Ovnic et al. 1991, [17] Ellingham et al. 1998, Ko et al. 2009, [18] Furihata et al. 2003, [19] Sanghani et al. 2002, [20] Ghosh et al. 1995, Okazaki et al. 2008, [21] Robbi and Beaufay 1994, [22] Masaki et al. 2007, [23] Miyazaki et al. 2006, [24] Ecroyd et al. 2006, Zhang et al. 2009, [25] Gilham et al. 2005, [26] Schreiber et al. 2009, [27] Lehner and Vance 1999, [28] Okazaki et al. 2006, [29] Linke et al. 2005

existence of multiple forms with overlapping specificities (Fleming et al. 2005; Imai 2006; Leinweber 1987; Redinbo and Potter 2005; Satoh and Hosokawa 1998, 2006). Table 4 summarizes current knowledge concerning substrates and functions reported for human, mouse, and rat *CES* gene family members.

Studies on human CES1 have examined its role in the metabolism of various drugs, including narcotics such as heroin and cocaine (Bencharit et al. 2003; Pindel et al. 1997), warfare nerve agents (Hemmert et al. 2010), psychostimulants (Sun et al. 2004), analgesics (Takai et al. 1997), and chemotherapy drugs (Sanghani et al. 2004). Mammalian liver is predominantly responsible for drug clearance from the body, with CES1 and CES2 (with CES1 > CES2) playing major roles, following absorption of drugs into the circulation (Imai 2006; Pindel et al. 1997). Mammalian intestine (with CES2 > CES1) plays a major role in first-pass clearance of several drugs, predominantly via CES2 in the ileum and jejunum (Imai et al. 2003). CES1 and CES2 also have different roles in prodrug activation, as shown for the anticancer drug irinotecan (CPT-11), which is converted to its active form SN-38 predominantly by CES2 (Humerickhouse et al. 2000). Recent modeling studies have shown that the human CES2 active site cavity is lined with negatively charged residues; this may explain the preference of this enzyme for neutral substrates (Vistoli et al. 2010). The role for human CES3 has not been studied extensively, although the enzyme is capable of activating prodrugs such as irinotecan (Sanghani et al. 2004). There are no reports concerning the metabolic role(s) for human CES4A, and functional studies on mammalian CES5 function are limited to feline species, where the enzyme is secreted into cat urine and apparently regulates the production of a cat-specific amino acid "felinine," a putative pheromone precursor (Miyazaki et al. 2006).

Evolution of mammalian *CES* gene families

Recent comparative and evolutionary studies (Holmes et al. 2008b; Williams et al. 2010) have concluded that there are at least five major mammalian *CES* gene families. In addition, the gene duplication events that generated the ancestral mammalian *CES1*, *CES2*, *CES3*, *CES4*, and *CES5* genes have apparently predated the common ancestor for marsupial and eutherian mammals (Holmes et al. 2008b) which has been estimated at approximately 173–193 million years ago (Woodburne et al. 2003) and may coincide with the early diversification of tetrapods approximately 350–360 million years ago (Donoghue and Benton 2007). The mammalian *CES* gene families are ancient in their genetic origins and were established prior to the appearance of mammals during evolution. Further *CES/Ces* gene duplication events have subsequently occurred during mammalian evolution, however, especially for rodent species, for which the mouse and rat *Ces1*-like and *Ces2*-like genes have apparently undergone successive duplication events. At least three of these are likely to have occurred in the common ancestor for rat and mouse during rodent evolution since several homolog genes and proteins were recognized, including *Ces1c* (previously *Es1*), *Ces1d* (*Ces3*), *Ces1e* (*Es22*), *Ces2a* (*Ces6*), *Ces2c* (*Ces2*), and *Ces2e* (*Ces5*) (Tables 3, 4). With the exception of the rat *Ces2*-like genes, which were located in multiple clusters on chromosomes 1 and 19, human, mouse, and rat *CES* genes were localized within two clusters of genes on the same chromosome, namely, *Ces1–Ces5A* (with multiple *Ces1*-like genes) and *Ces2–Ces3–Ces4A* (with multiple *Ces2*-like genes in mouse and rat). The presence of two *Ces3*-like genes in the mouse suggests that a further duplication event also took place in this species.

Conclusions

This article has examined human, mouse, and rat carboxylesterase genes and encoded subunits and has proposed a new nomenclature system, identifying each of five gene families (designated as *CES1*, *CES2*,..., *CES5* for human genes and *Ces1*, *Ces2*,..., *Ces5* for mouse and rat genes) and allocating a unique gene name for each of the genes. The italicized root symbol "*CES*" for human and "*Ces*" for mouse and rat genes followed by a number for the family were used, which is consistent with current practice. When multiple genes were identified for a gene family or where a gene required a name that clashed with an existing name, a capital letter (for human genes) (e.g., *CES4A*) or a lower-case letter (for mouse and rat genes) (e.g., *Ces1a*, *Ces1b*) was added after the number. A human *CES* pseudogene was named, using a capital "P" and a number (e.g., *CES1P1*), whereas mouse and rat *Ces* pseudogenes were named with a unique lower-case letter followed by "-ps" (e.g., *Ces2d-ps*). This new nomenclature will also assist in naming multiple CES genes and proteins from other mammalian species. As an example, Holmes et al. (2009c) and Williams et al. (2010) have reported multiple *CES1*-like genes on the horse genome that may be designated in accordance with the recommended nomenclature as *CES1A*, *CES1B*, *CES1C*, and so on, in order of the tandem locations of these genes on chromosome 3. Transcript isoforms of *CES* gene transcripts were named by following the gene name with the GenBank ID for the specific transcript. This nomenclature will assist our understanding of the genetic relatedness and the CES family origins for individual human, mouse, and rat *CES* genes and proteins and facilitate future research into the structure, function, and evolution of these genes. It will also serve as a model for naming *CES* genes from other mammalian species.

Acknowledgments This research was supported by NIH Grants P01 HL028972 and P51 RR013986 (to LAC); R01 ES07965 (to BY); and CA108775, and a Cancer Center Core Grant CA21765, the American Lebanese and Syrian Associated Charities (ALSAC) and St. Jude Children's Research Hospital (SJCRH) (to PMP); and a program project grant HG000330 entitled 'Mouse Genome Informatics' from the National Human Genome Research Institute of the NIH (to LJM). Acknowledgement is also given to members of the Redinbo laboratory and NIH grants CA98468 and NS58089 (to MRR).

References

Aida K, Moore R, Negishi M (1993) Cloning and nucleotide sequence of a novel, male-predominant carboxylesterase in mouse liver. Biochim Biophys Acta 1174:72–74

Barthel BL, Torres RC, Hyatt JL, Edwards CC, Hatfield MJ et al (2008) Identification of human intestinal carboxylesterase as the primary enzyme for activation of a doxazoline carbamate prodrug. J Med Chem 51:298–304

Becker A, Bottcher A, Lackner KJ, Fehringer P, Notka F et al (1994) Purification, cloning and expression of a human enzyme with acyl coenzyme A: cholesterol acyltransferase activity, which is identical to liver carboxylesterase. Arterioscler Thromb 14:1346–1355

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

Bencharit S, Edwards CC, Morton CL, Howard-Williams EL, Kuhn P et al (2006) Multisite promiscuity in the processing of endogenous substrates by human carboxylesterase 1. J Mol Biol 363:201–214

Berning W, De Looze SM, von Deimling O (1985) Identification and development of a genetically closely linked carboxylesterase gene family of the mouse liver. Comp Biochem Physiol 80:859–865

Cygler M, Schrag JD, Sussman JL, Harel M, Silman I et al (1993) Relationship between sequence conservation and three-dimensional structure in a large family of esterases, lipases and related proteins. Protein Sci 2:366–382

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

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

Donoghue PCJ, Benton MJ (2007) Rocks and clocks: calibrating the tree of life using fossils and molecules. Trends Genet 22:424–630

Ecroyd H, Belghazi M, Dacheux JL, Miyazaki M, Yamashita T et al (2006) An epididymal form of cauxin, a carboxylesterase-like enzyme, is present and active in mammalian male reproductive fluids. Biol Reprod 74:439–447

Ellingham P, Seedorf U, Assmann G (1998) Cloning and sequencing of a novel murine liver carboxylesterase cDNA. Biochim Biophys Acta 1397:175–179

Fleming CD, Bencharit S, Edwards CC, Hyatt JL, Tsurkan L et al (2005) Structural insights into drug processing by human carboxylesterase 1: tamoxifen, Mevastatin, and inhibition by Benzil. J Mol Biol 352:165–177

Fukami T, Nakajima M, Maruichi T, Takahashi S, Takamiya M et al (2008) Structure and characterization of human carboxylesterase 1A1, 1A2 and 1A3 genes. Pharm Genomics 18:911–920

Furihata T, Hosokawa M, Nakata F, Satoh T, Chiba K (2003) Purification, molecular cloning, and functional expression of inducible liver acylcarnitine hydrolase in C57BL/6 mouse, belonging to the carboxylesterase multigene family. Arch Biochem Biophys 416:101–109

Genetta TL, D'Eustachio P, Kadner SS, Finlay TH (1988) cDNA cloning of esterase 1, the major esterase activity in mouse plasma. Biochem Biophys Res Commun 151:1364–1370

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

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

Gibbs RA, Weinstock GM, Metzker ML, Muzny DM, Sodergren EJ et al (2004) Genome sequence of the Brown Norway rat yields insights into mammalian evolution. Nature 428:493–521

Gilham D, Alam M, Gao W, Vance DE, Lehner R (2005) Triacylglycerol hydrolase is localized to the endoplasmic reticulum by an unusual retrieval sequence where it participates in VLDL assembly without utilizing VLDL lipids as substrates. Mol Biol Cell 16:984–996

Hemmert AC, Otto TC, Wierdl M, Edwards CC, Fleming CD et al (2010) Human carboxylesterase 1 stereoselectively binds the nerve agent cyclosarin and spontaneously hydrolyzes the nerve agent sarin. Mol Pharmacol 77:508–516

Holmes RS, Cox LA, VandeBerg JL (2008a) Mammalian carboxylesterase 5: comparative biochemistry and genomics. Comp Biochem Physiol D Genomics Proteomics 3:195–204

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

Holmes RS, VandeBerg JL, Cox LA (2009a) A new class of mammalian carboxylesterase *CES6*. Comp Biochem Physiol Part D Genomics Proteomics 4:209–217

Holmes RS, Glenn JP, VandeBerg JL, Cox LA (2009b) Baboon carboxylesterases 1 and 2: sequences, structures and phylogenetic relationships with human and other primate carboxylesterases. J Med Primatol 38:27–38

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

Holmes RS, Cox LA, VandeBerg JL (2010) Mammalian carboxylesterase 3: comparative genomics and proteomics. Genetica 138(7):695–708

Hosokawa M (2008) Structure and catalytic properties of carboxylesterase isozymes involved in metabolic activation of prodrugs. Molecules 13:412–431

Hosokawa M, Furihata T, Yaginuma Y, Yamamoto N, Kayano N et al (2007) Genomic structure and transcriptional regulation of the rat, mouse and human carboxylesterase genes. Drug Metab Rev 39:1–15

Hosokawa M, Furihata T, Yaginuma Y, Yamamoto N, Watanabe N et al (2008) Structural organization and characterization of the regulatory element of the human carboxylesterase (*CES1A1* and *CES1A2*) genes. Drug Metab Pharmacokinet 23:73–84

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

Imai T (2006) Human carboxylesterase isozymes: catalytic properties and rational drug design. Drug Metab Pharmacokinet 21: 173–185

Imai T, Yoshigae Y, Hosokawa M, Chiba K, Otagiri M (2003) 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 307:1234–1242

Ko KW, Erickson B, Lehner R (2009) Es-x/Ces1 prevents triacylglycerol accumulation in McArdle-RH7777 hepatocytes. Biochim Biophys Acta 1791:1133–1143

Krishnasamy R, Teng AL, Dhand R, Schultz RM, Gross NJ (1998) Molecular cloning, characterization and differential expression pattern of mouse lung surfactant convertase. Am J Physiol Lung Mol Cell Biol 275:L969–L975

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

Langmann T, Becker A, Aslanidis C, Notka F, Ulrich H et al (1997) Structural organization and characterization of the promoter region of a human carboxylesterase gene. Biochim Biophys Acta 1350:65–74

Lehner R, Vance DE (1999) Cloning and expression of a cDNA encoding a hepatic microsomal lipase that mobilizes stored triacylglycerol. Biochem J 343:1–10

Leinweber FJ (1987) Possible physiological roles of carboxyl ester hydrolases. Drug Metab Rev 18:379–439

Linke T, Dawson H, Harrison EH (2005) Isolation and characterization of a microsomal retinyl ester hydrolase. J Biol Chem 280:23287–23294

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

Marsh S, Xiao M, Yu J, Ahluwalia R, Minton M et al (2004) Pharmacogenomic assessment of carboxylesterases 1 and 2. Genomics 84:661–668

Masaki K, Hashimoto M, Imai T (2007) Intestinal first-pass metabolism via carboxylesterase in rat jejunum and intestine. Drug Metab Dispos 35:1089–1095

Miyazaki M, Kamiie K, Soeta S, Taira H, Yamashita T (2003) 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 370:101–110

Miyazaki M, Yamashita T, Suzuki Y, Saito Y, Soeta S et al (2006) A major urinary protein of the domestic cat regulates the production of felinine, a putative pheromone precursor. Chem Biol 13:1070–1079

Morton CL, Iacono L, Hyatt JL, Taylor KR, Cheshire PJ et al (2005) Activation and antitumor activity of CPT-11 in plasma esterase-deficient mice. Cancer Chemother Pharmacol 56:629–636

Munger JS, Shi GP, Mark EA, Chin DT, Gerard C et al (1991) A serine esterase released by human alveolar macrophages is closely related to liver microsomal carboxylesterases. J Biol Chem 266:18832–18838

Mutch E, Nave R, McCracken N, Zech K, Williams FM (2007) The role of esterases in the metabolism of ciclesinide to deisobutyrl-ciclesonide in human tissue. Biochem Pharmacol 73:1657–1664

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

Okazaki H, Igarashi M, Nishi M, Tajima M, Sekiya M et al (2006) Identification of a novel member of the carboxylesterase family that hydrolyzes triacylglycerol. A potential role in adipocyte lipolysis. Diabetes 55:2091–2097

Okazaki H, Igarashi M, Nishi M, Sekiya M, Tajima M et al (2008) Identification of neutral cholesterol hydrolase, a key enzyme removing cholesterol from macrophages. J Biol Chem 283:33357–33364

Ovnic M, Swank RT, Fletcher C, Zhen L, Novak EK et al (1991) Characterization and functional expression of a cDNA encoding egasyn (esterase-22): the endoplasmic reticulum-targeting protein of beta-glucuronidase. Genomics 11:956–967

Pindel EV, Kedishvili NY, Abraham TL, Brezinski MR, Zhang A et al (1997) Purification and cloning of a broad substrate specificity human liver carboxylesterase that catalyzes the hydrolysis of cocaine and heroin. J Biol Chem 272:14769–14775

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

Redinbo MR, Potter PM (2005) Mammalian carboxylesterases: from drug targets to protein therapeutics. Drug Discov Today 10:313–320

Rhead B, Karolchik D, Kuhn RM, Hinrichs AS, Zweig AS et al (2010) The UCSC Genome Browser database: update 2010. Nucl Acids Res 38:D613–D619

Robbi M, Beaufay H (1983) Purification and characterization of various esterases from rat liver. Eur J Biochem 137:293–301

Robbi M, Beaufay H (1994) Cloning and sequencing of rat liver carboxylesterase ES-3 (egasyn). Biochem Biophys Res Commun 203:1404–1411

Robbi M, Beaufay H, Octave JN (1990) Nucleotide sequence of cDNA coding for rat liver pI 6.1 esterase (ES-10), a carboxylesterase located in the lumen of the endoplasmic reticulum. Biochem J 269:451–458

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

Sanghani SP, Davis WI, Dumaual NG, Mahrenholz A, Bosron WF (2002) Identification of microsomal rat liver carboxylesterases and their activity with retinyl palmitate. Eur J Biochem 269:4387–4398

Sanghani SP, Quinney SK, Fredenberg TB, Davis WI, Murray DJ et al (2004) 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 32:505–511

Satoh T, Hosokawa M (1995) Molecular aspects of carboxylesterase isoforms in comparison with other esterases. Toxicol Letters 82–83:439–445

Satoh T, Hosokawa M (1998) The mammalian carboxylesterases: from molecules to functions. Ann Rev Pharmacol Toxicol 38:257–288

Satoh T, Hosokawa M (2006) Structure, function and regulation of carboxylesterases. Chem Biol Interact 162:195–211

Satoh T, Taylor P, Bosron WF, Sanghani P, Hosokawa M et al (2002) Current progress on esterases: from molecular structure to function. Drug Metab Dispos 30:488–493

Schewer H, Langmann T, Daig R, Becker A, Aslandis C et al (1997) Molecular cloning and characterization of a novel putative carboxylesterase, present in human intestine and liver. Biochem Biophys Res Commun 233:117–120

Schreiber R, Taschler U, Wolinski H, Seper A, Tamegger SN et al (2009) Esterase 22 and beta-glucuronidase hydrolyze retinoids in mouse liver. J Lipid Res 50:2514–2523

Shibita F, Takagi Y, Kitajima M, Kuroda T, Omura T (1993) Molecular cloning and characterization of a human carboxylesterase gene. Genomics 17:76–82

Sun Z, Murry DJ, Sanghani SP, Davis WI, Kedishvilli NY et al (2004) Methylphenadate is stereoselectively hydrolyzed by human carboxylesterase CES1A1. J Pharmcol Exp Ther 310:469–476

Takai S, Matsuda A, Usami Y, Adachi T, Sugiyama T et al (1997) Hydrolytic profile for ester- or amide-linkage by carboxylesterases pI 5.3 and 4.5 from human liver. Biol Pharm Bull 20:869–873

Taketani M, Shii M, Ohura K, Ninomiya S, Imai T (2007) Carboxylesterase in the liver and small intestine of experimental animals and human. Life Sci 81:924–932

Tanimoto K, Kaneyasu M, Shimokuni T, Hiyama K, Nishiyama M (2007) Human carboxylesterase 1A2 expressed from carboxylesterase 1A1 and 1A2 genes is a potent predictor of CPT-11 cytotoxicity in vitro. Pharm Genomics 17:1–10

The MGC Project Team (2004) The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC). Genome Res 14:2121–2127

Thierry-Mieg D, Thierry-Mieg J (2006) AceView: a comprehensive cDNA-supported gene and transcripts annotation. Genome Biol 7 (Suppl 1):S12-S14

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

Vanlith HA, Haller M, Vanhoof IJM, Vanderwouw MJA, Vanzutphen BFM et al (1993) Characterization of rat plasma esterase ES-1A concerning its molecular and catalytic properties. Arch Biochem Biophys 301:265–274

Vistoli G, Pedretti A, Mazzolari A, Testa B (2010) Homology modelling and metabolism prediction of human carboxylesterase-2 using docking analyses by GriDock: a parallelized tool based on AutoDock 4.0. J Comput Aided Mol Des 24(9):771-787

von Heijne G (1983) Patterns of amino acids near signal-sequence cleavage sites. Eur J Biochem 133:17–21

Wang H, Gilham D, Lehner R (2007) 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 282:33218–33226

Williams ET, Wang H, Wrighton SA, Qian YW, Perkins EJ (2010) Genomic analysis of the carboxylesterases: identification and classification of novel forms. Mol Phylogenet Evol 57(1):23–34

Woodburne MO, Rich TH, Springer MS (2003) The evolution of tribospheny and the antiquity of mammalian clades. Mol Phylogenet Evol 28:360–385

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

Yan B, Matoney L, Yang D (1999) Human carboxylesterases in term placenta: enzymatic characterization, molecular cloning and evidence for the existence of multiple forms. Placenta 20:517–525

Yoshimura M, Kimura T, Ishii M, Ishii K, Matsuura T et al (2008) Functional polymorphisms in carboxylesterase1A2 (*CES1A2*) gene involves specific protein 1 (Sp1) binding sites. Biochem Biophys Res Commun 369:939–942

Zhang L, Hu Z, Zhu C, Liu Q, Zhou Y et al (2009) Identification and characterization of an epididymis-specific gene, *Ces7*. Acta Biochim Biophys Sin 41:809–815

Zhen L, Rusiniak ME, Swank RT (1995) 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 270:11912–11920

BJCP British Journal of Clinical Pharmacology

DOI:10.1111/j.1365-2125.2010.03695.x

Association of *carboxylesterase 1A* genotypes with irinotecan pharmacokinetics in Japanese cancer patients

Kimie Sai,[1] Yoshiro Saito,[2] Naoko Tatewaki,[3] Masakiyo Hosokawa,[4] Nahoko Kaniwa,[2] Tomoko Nishimaki-Mogami,[1] Mikihiko Naito,[1] Jun-ichi Sawada,[1,14] Kuniaki Shirao,[6,15] Tetsuya Hamaguchi,[6] Noboru Yamamoto,[6] Hideo Kunitoh,[6,16] Tomohide Tamura,[6] Yasuhide Yamada,[6] Yuichiro Ohe,[6,10] Teruhiko Yoshida,[7] Hironobu Minami,[8,17] Atsushi Ohtsu,[9,12] Yasuhiro Matsumura,[11] Nagahiro Saijo[13,18] & Haruhiro Okuda[4]

[1]Division of Functional Biochemistry and Genomics, [2]Division of Medicinal Safety Science, [3]Project Team for Pharmacogenetics, [4]Division of Organic Chemistry, National Institute of Health Sciences, 1-18-1 Kamiyoga, Setagaya-ku, Tokyo 158-8501, [5]Laboratory of Drug Metabolism and Biopharmaceutics, Faculty of Pharmaceutical Sciences, Chiba Institute of Science, Shiomi-Cho, Choshi City, Chiba 288-0025, [6]Division of Internal Medicine, National Cancer Center Hospital, [7]Genomics Division, National Cancer Center Research Institute, 5-1-1 Tsukiji, Chuo-Ku, Tokyo 104-0045, [8]Division of Oncology/Hematology, [9]Division of GI Oncology, Digestive Endoscopy, [10]Division of Internal Medicine, Investigative Treatment Division, [11]Research Center for Innovative Oncology, [12]Deputy Director, National Cancer Center Hospital East, 6-5-1 Kashiwanoha, Kashiwa, Chiba 277-8577, [13]Pharmaceuticals and Medical Devices Agency, 3-3-2 Kasumigaseki, Chiyoda-Ku, Tokyo 100-0013, [14]Department of Medical Oncology, OITA University Faculty of Medicine, 1-1 Idaigaoka, Hasama-machi, Yufu 879-5593, [15]Department of Respiratory Medicine, Mitsui Memorial Hospital, 1 Kandasurugadai, Chiyoda-Ku, Tokyo 101-8643, [16]Medical Oncology, Department of Medicine, Kobe University Hospital and Graduate School of Medicine, 7-5-2 Kusunoki-cho, Chuo-Ku, Kobe 650-0017 and [17]Kinki University School of Medicine, Osaka-Sayama, Osaka 589-8511, Japan

Correspondence
Dr Kimie Sai PhD, Division of Functional Biochemistry and Genomics, National Institute of Health Sciences, 1-18-1 Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan.
Tel.: + 81 3 3700 9478
Fax: + 81 3 3707 6950
E-mail: sai@nihs.go.jp

Keywords
CES1, genetic polymorphism, haplotype, irinotecan

Received
30 November 2009

Accepted
14 March 2010

WHAT IS ALREADY KNOWN ABOUT THIS SUBJECT

WHAT THIS STUDY ADDS

AIMS

Human carboxylesterase 1 (CES1) hydrolyzes irinotecan to produce an active metabolite SN-38 in the liver. The human *CES1* gene family consists of two functional genes, *CES1A1* (*1A1*) and *CES1A2* (*1A2*), which are located tail-to-tail on chromosome 16q13-q22.1 (*CES1A2-1A1*). The pseudogene *CES1A3* (*1A3*) and a chimeric *CES1A1* variant (*var1A1*) are also found as polymorphic isoforms of *1A2* and *1A1*, respectively. In this study, roles of *CES1* genotypes and major SNPs in irinotecan pharmacokinetics were investigated in Japanese cancer patients.

METHODS

CES1A diplotypes [combinations of haplotypes A (*1A3-1A1*), B (*1A2-1A1*), C (*1A3-var1A1*) and D (*1A2-var1A1*)] and the major SNPs (−75T>G and −30G>A in *1A1*, and −816A>C in *1A2* and *1A3*) were determined in 177 Japanese cancer patients. Associations of *CES1* genotypes, number of functional *CES1* genes (*1A1*, *1A2* and *var1A1*) and major SNPs, with the AUC ratio of (SN-38 + SN-38G)/irinotecan, a parameter of *in vivo* CES activity, were analyzed for 58 patients treated with irinotecan monotherapy.

RESULTS

The median AUC ratio of patients having three or four functional *CES1* genes (diplotypes A/B, A/D or B/C, C/D, B/B and B/D; $n = 35$) was 1.24-fold of that in patients with two functional *CES1* genes (diplotypes A/A, A/C and C/C; $n = 23$) [median (25th–75th percentiles): 0.31 (0.25–0.38) *vs.* 0.25 (0.20–0.32), $P = 0.0134$]. No significant effects of *var1A1* and the major SNPs examined were observed.

CONCLUSION

This study suggests a gene-dose effect of functional *CES1A* genes on SN-38 formation in irinotecan-treated Japanese cancer patients.

Introduction

Human carboxylesterases (CESs) are members of the α/β-hydrolase-fold family and are localized in the endoplasmic reticulum of many different cell types. These enzymes efficiently catalyze the hydrolysis of a variety of ester- and amide-containing chemicals as well as drugs (including prodrugs) to the respective free acids. They are involved in detoxification or metabolic activation of various drugs, environmental toxicants and carcinogens. CESs also catalyze the hydrolysis of endogenous compounds such as short- and long-chain acyl-glycerols, long-chain acyl-carnitine, and long-chain acyl-CoA esters. The two major CES families CES1 and CES2 have been identified in human tissues. CES1 is abundant in the liver and lung but not in the intestine, while CES2 is highly expressed in the intestine and kidney but has low expression in the liver and lung [1].

Human CES1 and CES2 are involved in producing a topoisomerase I inhibitor SN-38, an active metabolite of irinotecan which is clinically used for colorectal, lung and other cancers [2]. SN-38 is further inactivated by UDP-glucuronosyltransferase 1As (UGT1As) to produce SN-38 glucuronide (SN-38G). Irinotecan is also converted by cytochrome P450 3A4 (CYP3A4) to an inactive compound 7-ethyl-10-[4-N-(5-aminopentanoic acid)-1-piperidino]carbonyloxycamptothecin (APC) (Figure 1).

Recent pharmacogenetic studies on irinotecan have revealed significant associations of *UGT1A1* polymorphisms **28* [–54_39A(TA)$_6$TAA>A(TA)$_7$TAA or –40_39insTA] and **6* [211G>A (G71R)], the latter being specifically detected in East Asians, with reduced clearance of SN-38 resulting in severe neutropenia [3–8]. These findings have led to the clinical application of genetic testing for *UGT1A1*28* in the United States (since August 2005) and for *UGT1A1*6* and **28* in Japan (since March 2009). In addition, possible additive effects of genotypes of the transporters for irinotecan and its metabolites, such as *ABCB1*, *ABCC2*, *ABCG2* and *SLCO1B1*, have been suggested [9–12]. We previously analyzed *CES2* polymorphisms in a Japanese

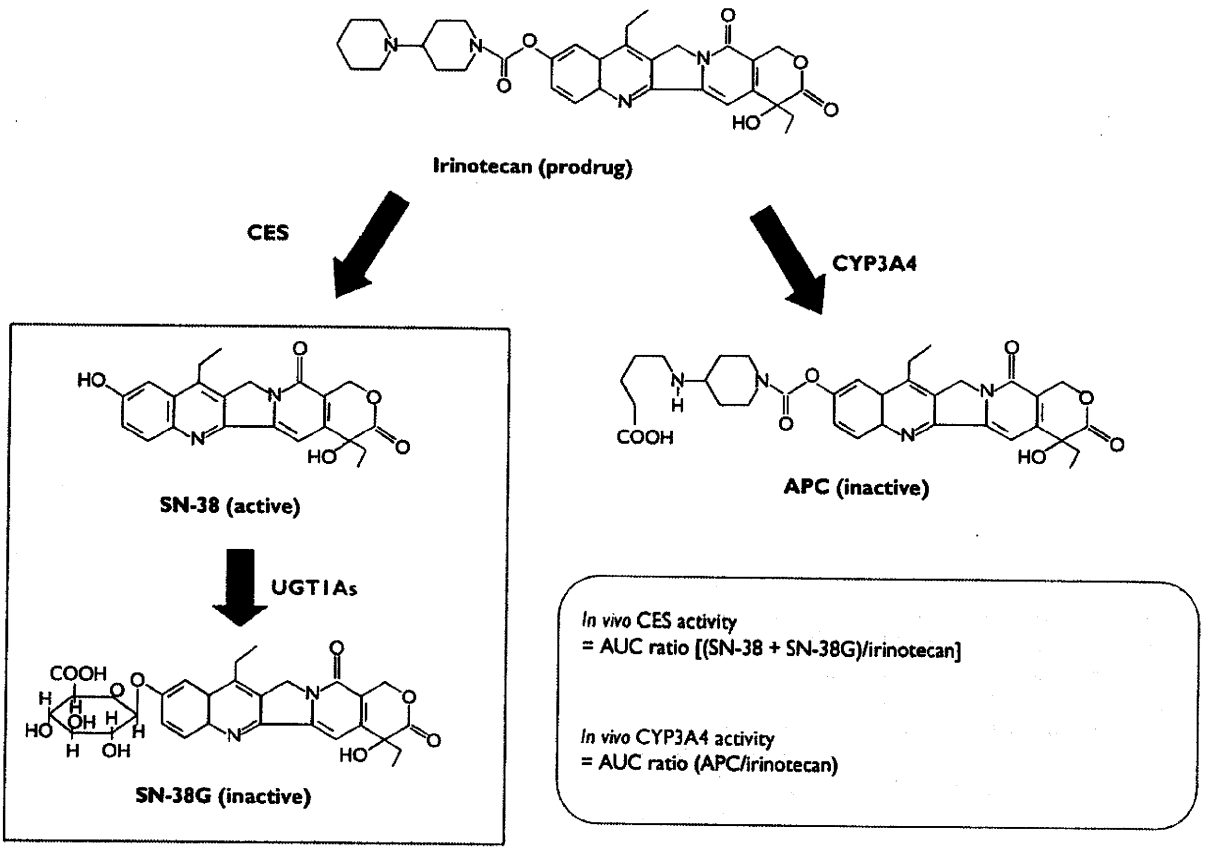

Figure 1

Metabolic pathway of irinotecan. The prodrug irinotecan is hydrolyzed by carboxylesterase (CES) to produce an active metabolite SN-38, and subsequently detoxified by UDP-glucuronosyltransferase 1As (UGT1As) to produce an inactive metabolite SN-38 glucuronide (SN-38G). Irinotecan is also metabolized by cytochrome P450 3A4 (CYP3A4) to produce another inactive metabolite APC

population and identified minor genetic variations which were associated with lower expression/function *in vitro* and *in vivo* [13, 14]. However, major *CES2* haplotypes (**1b* and **1c*) did not affect irinotecan pharmacokinetics (PK) [14]. Since CES1 is expressed at higher levels in the liver, a major organ for activating irinotecan, it is possible that *CES1* genotypes affect the plasma concentrations of irinotecan metabolites. However, their clinical relevance to irinotecan pharmacokinetics/pharmacodynamics has not yet been fully investigated.

Functional human *CES1* genes include *CES1A1* (*1A1*) and *CES1A2* (*1A2*), which are inversely located (tail-to-tail) on chromosome 16q13-q22.1 (*1A2-1A1*). Both *1A1* and *1A2* consist of 14 exons encoding 567 amino acids, and they have 98% homology with 5 nucleotide (4 amino acid) differences in exon 1, which encodes a signal peptide [1]. Recent studies also identified *CES1A1* variants (*var1A1*), in which exon 1 was replaced with exon 1 of *CES1A2*, and a pseudogene *CES1A3* (*1A3*; formerly referred to as *CES4*) replacing *CES1A2* [15, 16]. The *1A3* sequence from the promoter region to exon 1 is the same as that of *CES1A2*, but contains a stop codon in exon 3. The sequence downstream from exon 11 is highly homologous with that of *1A1* (NT_010498) [16]. Ethnic differences in these *CES1* genes (*1A1*, *var1A1*, *1A2* and *1A3*) have been reported [16].

Expression levels of CES1A2 mRNA were lower than those of CES1A1 mRNA in several tissues. This CES1A1 up-regulation could be mediated by additional Sp1 and C/EBP binding sites in the promoter region [17]. Transcript levels of CES1A2 derived from *var1A1* were reported to be higher than those from the original *1A2* [15, 16]. These findings suggest that polymorphisms in the upstream region of *CES1A1* or *var1A1* could affect their expression.

In addition to structural variations of the *CES1* gene family, several single nucleotide polymorphisms (SNPs) and small deletion/insertion variants were found. –816C in the *CES1A2* promoter region was reported to be associated with enhanced CES1A2 expression and imidapril efficacy [18]. Furthermore, –816A>C was found to be linked with several SNPs (–62T>C, –47G>C, –46G>T, –41C>G, –40A>G, –37G>C, –34del/G and –32G>T) in the proximal promoter region, leading to two additional Sp1 binding sites, and these additional sites were suggested to increase transcription of *1A2* [19].

In this context, this study investigated the clinical significance of *CES1* genotypes in irinotecan therapy. For this purpose, we analyzed the *CES1* genotypes (combinations of four *CES1A* isoforms) and major SNPs in the *CES1A1* exon 1 with its adjacent region and in the *CES1A2* and *1A3* promoter regions, which could be important for CES1 expression or function, in Japanese cancer patients treated with irinotecan, and then examined the associations of these *CES1* genotypes or SNPs with irinotecan PK.

Methods

Patients

Genetic analysis of 177 Japanese cancer patients who received irinotecan therapy at the National Cancer Center in Japan was performed. The patients were the same as those described in our previous study [7], where details on eligibility criteria for irinotecan therapy, patient profiles and irinotecan regimens were described. Since the AUC ratio [(SN-38 + SN-38G) : irinotecan], a parameter of *in vivo* CES activity, was influenced by irinotecan regimens [14], 58 patients receiving irinotecan monotherapy (100 mg m^{-2} weekly or 150 mg m^{-2} biweekly) from the 177 patients were primarily used for analysis of the association between *CES1* genotypes and irinotecan PK parameters. The patient set was the same as used in our previous study on *CES2* [14]. This study was approved by the ethics committees of the National Cancer Center and the National Institute of Health Sciences, and written informed consent was obtained from all participants.

Determination of CES1 genotypes and SNPs

For describing the *CES1* gene family, haplotypes A to D designated by Fukami *et al.* [16] were used (Figure 2): haplotype A, *CES1A3-CES1A1* (*1A3-1A1*); haplotype B, *CES1A2-CES1A1* (*1A2-1A1*); haplotype C, *CES1A3-CES1A1* variant (*1A3-var1A1*); and haplotype D, *CES1A2-CES1A1* variant (*1A2-var1A1*). To determine the diplotypes, combinations of haplotypes A to D, we sequenced *1A1/var1A1* exon 1 and its flanking region and the *1A2/1A3* promoter region of 177 patients. These regions are indicated in Figure 2, and a list of primers/probes is shown in Table 1.

For discrimination between *1A1* and *var1A1*, their exon 1s and flanking regions were sequenced (Figure 2a). Briefly, the first PCR was performed using 25 ng of genomic DNA with 0.625 units of Ex-Taq (Takara Bio. Inc., Shiga, Japan) and 0.2 μM of primers, Ces1-FP and Ces1-RP (Table 1a, first PCR). The PCR conditions were 94°C for 5 min, followed by 30 cycles of 94°C for 30 s, 60°C for 1 min, and 72°C for 2 min, and then a final extension at 72°C for 7 min. Then, the second PCR was performed with the primers, Ces1_seqF and Ces1_seqR (Table 1a, second PCR) under the same reaction conditions described above. The PCR products were treated with a PCR Product Pre-Sequencing Kit (USB Co., Cleveland, OH, USA) and directly sequenced on both strands using an ABI BigDye Terminator Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) with the sequencing primers listed in Table 1a (sequencing). Excess dye was removed by a DyeEx96 kit (Qiagen, Hilden, Germany), and the eluates were analyzed on an ABI Prism 3730 DNA Analyzer (Applied Biosystems). The conditions of the PCR and sequencing procedures described in the following section were the same as described above unless otherwise noted.

1A2 and *1A3* were discriminated by the restriction fragment length polymorphism (RFLP) method for exon 5

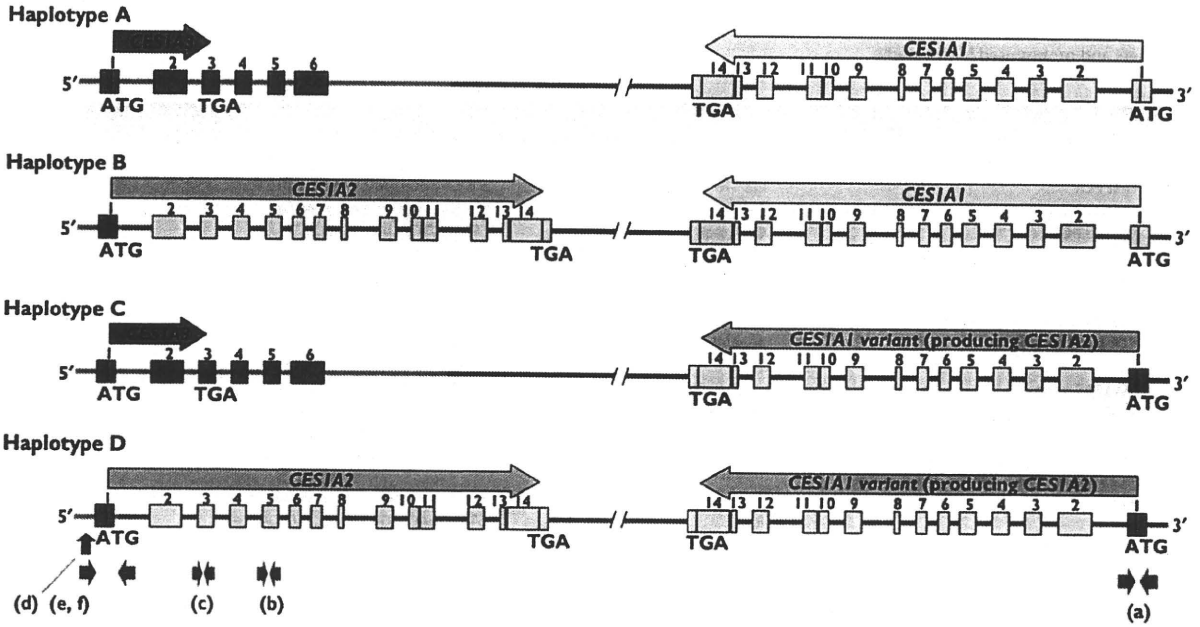

Figure 2

CES1 gene structure and haplotypes. The regions used for haplotype determination in this study are indicated with arrows (a–f)

reported by Fukami *et al.* [16] (Figure 2b). Briefly, the PCR was performed using a primer set (1A-int4F and 1A-int5AS) (Table 1b), and then the PCR products were digested with *Pvu*II to produce *CES1A3*-derived fragments (409 bp and 248 bp). UV intensity of the fragments stained with ethidium bromide was measured after electrophoresis (2% agarose gel). The number of *1A3* (0, 1 or 2) was also confirmed by direct sequencing of exon 5 using the same primer set. To verify that the *1A3* sequence is derived from the pseudogene, we confirmed the existence of a stop codon at codon 105 of *1A3* exon 3 (Figure 2c) in 11 randomly selected patients (heterozygous or homozygous) by amplification and sequencing using primers listed in Table 1c.

Genotyping for –816A>C in the *1A2* and *1A3* promoter region (Figure 2d) was conducted by the TaqMan method of Geshi *et al.* [18] (Table 1d) in all patients. We also examined attribution of –816C to *1A2* or *1A3* by specific amplifications from 5′-regions to intron 1 of the *1A2* and *1A3* (Figure 2e,f) in 23 randomly selected heterozygous patients. For specific amplifications, primers CES1A3-1A2_F1 and CES1A2 R1 for *CES1A2* (Table 1e) and primers CES1A3-1A2_F1 and CES1A3 R1 for *1A3* (Table 1f, first PCR) were used with 0.05 U μl^{-1} LA-Taq with GC buffer I (Takara Bio. Inc.); and for *1A3*, the second PCR using primers CES1A3-1A2_F2 and CES1A3 R2 (Table 1f, second PCR) was also conducted with 0.05 U μl^{-1} Ex-taq. Then, direct sequencing of the *1A2* and *1A3* PCR products was performed. Complete linkage among –816A>C and several SNPs in the proximal promoter region (between –62 to –32) [19] was confirmed for 11 randomly selected subjects.

All variations were confirmed by sequencing PCR products generated from new amplifications from genomic DNA. GenBank NT_010498.15 was used as the reference sequence for *CES1A1*, *CES1A3* and the promoter region of *CES1A2*, and AB119998.1 was used for exon 1 and its downstream region of *CES1A2*. The translational initiation site was designated as +1 to describe the polymorphism positions. Diplotype configuration was estimated with the LDSUPPORT software [20]. The diplotypes A/D and B/C could not be distinguished.

Pharmacokinetic data and association analysis

The area under the concentration–time curve (AUC) values for irinotecan and its metabolites, SN-38, SN-38G and APC, were previously obtained [4, 21]. The AUC ratio of SN-38 plus SN-38G to irinotecan [$AUC_{(SN\text{-}38\ +\ SN\text{-}38G)}/AUC_{irinotecan}$] was used as a parameter reflecting *in vivo* CES activity [14]. The AUC ratio of APC to irinotecan [$AUC_{APC}/AUC_{irinotecan}$] was used as a parameter for *in vivo* CYP3A4 activity [21].

Statistical significance (two-sided, $P < 0.05$) for associations between AUC ratios (or AUC/dose) and *CES1* genotypes or SNPs was determined by the Mann-Whitney test or the Jonckheere-Terpstra (JT) test using Prism version 4.0 (GraphPad Prism Software Inc. San Diego, CA, USA) and StatXact version 6.0 (Cytel Inc., Cambridge, MA). Correla-

Table 1

Primers and probes used in this study

Region (indicated in Figure 2)	Primer	Primer sequence	Reference
(a) *CES1A1* exon 1 and promoter region			This study
First PCR	Ces1-FP	5′-CCAGGCAAAACCTAGGAGTG-3′	
	Ces1-RP	5′-AGTACAGGGCGATCTCAGGA-3′	
Second PCR	Ces1_seqF	5′-GTATTTCCTTAGCCAGCGGTA-3′	
	Ces1_seqR	5′-CAGAGCCGGACCTGTTGT-3′	
Sequencing	Ces1_SF2	5′-AGAGCCTGGAAAGCTATGAAAA-3′	
	Ces1_SR	5′-TTTCTACGCATCTGCGCCCACC-3′	
(b) *CES1A1, 1A2* and *1A3* exon 5			[16]
PCR and sequencing	1A-int4F	5′-GCTCAGTAAATAGTTGCCAGTT-3′	
	1A-int5AS	5′-TCTCATCAGCATCACATCAAG-3′	
(c) *CES1A3* exon 3			This study
PCR and sequencing	CES1A3-15183F	5′-CAGGGAAGATCGTTGTATTGGTTT-3′	
	CES1A3-15974R	5′-TTCCTTCCACCACTAACATTATTG-3′	
Sequencing (additional primer)	CES1A3-15823R	5′-AAGATGTTCATTAAAGATGCACAG-3′	
(d) *CES1A2* and *1A3* -816A>C genotyping			[18]
PCR	F	5′-CCTTAATTTGGTGATTTCACATTGC-3′	
	R	5′-CAAGACATGGTTCAGCTTCTCAAG-3′	
TaqMan probe	FAM	5′-CATCACCCCTACTGC-3′	
	VIC	5′-CATCACACCTACTGCT-3′	
(e) *CES1A2* promoter region			This study
PCR	CES1A3-CES1A2_F1	5′-ATGATTTCCAGCTTCATCTACA-3′	
	CES1A2_R1	5′-GAGAGAACGTTCCCATGTCTTT-3′	
(f) *CES1A3* promoter region			This study
First PCR	CES1A3-CES1A2_F1	5′-ATGATTTCCAGCTTCATCTACA-3′	
	CES1A3_R1	5′-GCTTGAGTTTTCTTTACAGACA-3′	
Second PCR	CES1A3-CES1A2_F2	5′-AACAGTTTATAACCTGTTATTTTT-3′	
	CES1A3_R2	5′-TGCTTTGGATAAAGACAAGATGTT-3′	
Seqeuncing of *CES1A2/1A3* promoter region			
	CES1A3-CES1A2_F2	5′-AACAGTTTATAACCTGTTATTTTT-3′	
	CES1A3-CES1A2_R1	5′-CACACTTCCAATCTCAGGTAAA-3′	
	CES1A3-CES1A2_F3	5′-TTATGCCACAAGCAGTTGGGCG-3′	
	CES1A3-CES1A2_R2	5′-TCCAAGTCCAATTCCAAGTACGGA-3′	

NT_010498.15 was used as the reference sequence for *CES1A1*, *CES1A3* and the promoter region of *CES1A2*, and AB119998.1 was used for exon 1 and its downstream region of *CES1A2*.

tions between the AUC ratios [$AUC_{(SN\text{-}38\ +\ SN\text{-}38G)}/AUC_{irinotecan}$] and [$AUC_{APC}/AUC_{irinotecan}$] were analyzed by Spearman's rank correlation test. Multiplicity adjustment was not applied to bivariate analysis, and contributions of the candidate genetic markers to the AUC ratios [$AUC_{(SN\text{-}38\ +\ SN\text{-}38G)}/AUC_{irinotecan}$] were further determined by multiple regression analysis after logarithmic transformation of the AUC ratio. The variables examined were age, sex, body surface area, history of smoking or drinking, performance status, serum biochemistry (GOT, ALP, creatinine) at baseline, *CES1* genotypes and SNPs, *CES2*2* [100C>T(R34W)] or **5* [1A>T (M1L)] [13, 14], *UGT1A1*6* or **28* [7, 8], and the transporter haplotypes, *ABCB1*2* [2677G>T(A893A)], *ABCC2*1A* (−1774delG), *ABCG2*IIB* [421C>A (Q141K) and IVS12+49G>T] and *SLCO1A1*15-17* [521T>C (V174A)] [10]. The variables in the final models were selected by the forward and backward stepwise procedure at a significance level of 0.10 using JMP version 7.0.0 (SAS Institute, Inc., Cary, NC, USA). *UGT1A1*6 or *28* was grouped as '+' for stratifying patients: for example, homozygous *UGT1A1 *6 or *28* was depicted as *UGT+/+*.

Results

Genotypes and SNPs of CES1 gene family in Japanese

Frequencies of individual *CES1* genes and *CES1* diplotypes stratified according to the number of functional *CES1* genes are summarized in Table 2. The frequencies of the patients with two, three and four functional *CES1* genes were 44%, 47% and 9%, respectively, in all 177 patients.

By sequencing *1A1* and *var1A1* exon 1s and their flanking region, we detected four novel variations; three in the 5′-flanking region and one in the 5′-untranslated region (5′-UTR) (Table 3): −258C>T (allele frequency: 0.014), −233C>A (0.003), −161A>G (0.006) and −30G>A (0.042). Eleven nucleotide substitutions from the 5′-UTR to intron 1 at allele frequencies of 0.294–0.299 were closely linked with *var1A1* (Table 3). The SNP −816A>C found in the *1A2* and *1A3* promoter regions was genotyped by a TaqMan method [18], and the allele frequency of −816A>C in 177 subjects was 0.249 (Table 4). It was noted that −816C was detected only in patients with *1A3* (*1A3/1A2* and *1A3/1A3*),

Table 2

Frequency of *CES1* genes and diplotypes in Japanese cancer patients

CES1 diplotype	Number of *CES1* gene *1A1*	*var1A1*	*1A2*	*1A3*	Total*	Frequency (n = 177)†		Frequency (monotherapy: n = 58)†	
A/A	2	0	0	2	2	0.203	0.441	0.138	0.397
A/C	1	1	0	2		0.220		0.241	
C/C	0	2	0	2		0.017		0.017	
A/B	2	0	1	1	3	0.237	0.469	0.293	0.534
A/D or B/C	1	1	1	1		0.192		0.190	
C/D	0	2	1	1		0.040		0.052	
B/B	2	0	2	0	4	0.040	0.090	0.017	0.069
B/D	1	1	2	0		0.034		0.052	
D/D	0	2	2	0		0.017		0.000	
Frequency (n = 354)‡	0.703	0.297	0.325	0.675					
(monotherapy: n = 116)‡	0.690	0.310	0.336	0.664					

*Number of functional genes. †Number of subjects. ‡Number of chromosomes.

but not in the *1A2* homozygotes (*1A2/1A2*). In the *1A2/1A3* patients, 38 of the 39 patients having –816C were heterozygous for –816C (Table 4). These findings suggested a close association between –816C with *1A3*. Following specific amplifications of the regions from 5′-regions to intron 1 in *1A2* and *1A3* (Figure 2e,f) of 23 patients randomly selected from the 38 patients with –816A/C and *1A2/1A3*, we confirmed that –816C resided in the *1A3* gene (data not shown). Thus, –816A>C is the major SNP of *1A3* but very rare in *1A2*. In addition, the SNPs, –62T>C, –47G>C, –46G>T, –41C>G, –40A>G, –37G>C, –34del/G and –32G>T, in the proximal promoter region reported to be linked with –816A>C [19] were found to be completely linked with *1A3* (data not shown).

Association of CES1 genotypes with in vivo CES activity

CES1 diplotypes In patients treated with irinotecan monotherapy, we found the AUC ratios of patients with haplotypes A or C (having the *1A3* pseudogene) were lower than those without A or C, indicating functional *CES1* gene number dependency. The median AUC ratio of patients having three or four functional *CES1* genes was 1.24-fold of that in patients with two functional *CES1* genes [median (25th–75th percentiles): 0.31 (0.25–0.38) *vs.* 0.25 (0.20–0.32), P = 0.0134, Mann-Whitney test)] (Figure 3a). No significant differences were observed between *1A1* and *var1A1* (among *1A1/1A1*, *var1A1/1A1* and *var1A1/var1A1*). As we previously reported, the *CES2* variations, *CES2*5* [1A>T(M1L)] and *CES2*2* [100C>T(R34W)] [13, 14] showed low CES activity as indicated in Figure 3a.

Platinum-containing regimens themselves enhance renal excretion of irinotecan and its metabolites, especially SN-38G. No significant effect of *CES1* gene number on the AUC ratio was observed. However, it was noted that the median renal excretion ratio [(SN-38 + SN-38G)/irinotecan] in patients with four functional *CES1* genes was 1.37-fold higher than that in patients with two or three functional genes (P = 0.0217, Mann-Whitney test) (data not shown).

To exclude the possibility that the higher AUC ratio observed above (Figure 3a) was biased by CYP3A4, another metabolic enzyme for irinotecan, we analyzed the association between the (SN-38 + SN-38G)/irinotecan AUC ratio and the APC/irinotecan AUC ratio, an *in vivo* parameter of CYP3A4 activity [21], in patients treated with irinotecan monotherapy. The result showed no correlation between the two parameters (Spearman r = 0.126, P = 0.345).

CES1 SNPs Next, associations of the two *1A1* SNPs, –75G>T and –30G>A (Table 3) and *1A3*-816A>C with the AUC ratio [(SN-38 + SN-38G)/irinotecan] were analyzed. The effects of the SNPs were analyzed in patients stratified by the functional *CES1* gene number and also in all the patients receiving monotherapy. A –75G>T-dependent increase in the AUC ratio was observed in the whole group of patients (P = 0.027, JT test) (Figure 3b), and this trend was remarkable in patients with three or four functional *CES1* genes. No significant effect of –30G>A was observed (Figure 3c). As for –816C in *1A3*, no association between this SNP and the AUC ratio was evident in patients with two or three functional *CES1* genes (Figure 3d). In the platinum-containing regimens, no significant effects of these SNPs on the AUC ratio or the renal recovery ratio were observed (data not shown).

Multivariate analysis The contribution of *CES1* genotypes to the AUC ratio was further analyzed by multivariate analysis, using the patient background factors and polymorphisms including the haplotypes of *CES2*, *UGT1A1* and transporters as variables [7, 8, 10, 13, 14]. The final model revealed a significant association of the functional *CES1* gene number (n = 3 or 4) with the AUC ratio. Contributions of smoking history, irinotecan dose, hepatic and renal function were also detected while that of *ABCB1*2* (+/+) was

Table 3

Summary of genetic variations of *CES1A1* and *var 1A1* exon 1s and their flanking regions detected in this study

SNP identification				Position					
This study	NCBI (dbSNP)	JSNP	Location	NT_010498.15	From the translational initiation site or the nearest exon	Nucleotide change and flanking sequences (5′ to 3′)	Amino acid change	Allele frequency (n = 354)*	*CES1A1* variant (*CES1A2* type)
MPJ6_CS1001†			**5′-flank**	**9481424**	**−258**	**ttgggcaagtttacagctct<u>C/T</u>tgtaatctgacagtagagtc**		**0.014**	
MPJ6_CS1002†			**5′-flank**	**9481399**	**−233**	**atctgacagtagagtccaga<u>C/A</u>tggtttgatgaaagagggta**		**0.003**	
MPJ6_CS1003†			**5′-flank**	**9481327**	**−161**	**tagaagcccagggagatctg<u>A/G</u>ggaaagggagggcttttctg**		**0.006**	
MPJ6_CS1004	rs3815583	IMS-JST175949	Exon1(5′-UTR)	9481241	−75	aactctgggcggggctgggc<u>G/T</u>ccagggctggacagcacagt		0.41	
MPJ6_CS1005	rs28429139		Exon1(5′-UTR)	9481212	−46	ggacagcacagtccctctga<u>A/G</u>ctgcacagagacctcgcagg		0.299	*var1A1*
MPJ6_CS1006	rs28494177		Exon1(5′-UTR)	9481205	−39	acagtccctctgaactgcac<u>A/G</u>gagacctcgcaggccccgag		0.299	*var1A1*
MPJ6_CS1007†			Exon1(5′-UTR)	9481196	−30	ctgaactgcacagagacctc<u>G/A</u>caggccccgagaactgtcgc		0.042	
MPJ6_CS1008	rs28520463		Exon1(5′-UTR)	9481187	−21	acagagacctcgcaggcccc<u>G/C</u>agaactgtcgcccttccacg		0.297	*var1A1*
MPJ6_CS1009	rs28499065		Exon1(5′-UTR)	9481186	−20	cagagacctcgcaggcccc<u>g/G</u>gaactgtcgcccttccacga		0.297	*var1A1*
MPJ6_CS1010	rs28515828		Exon1(5′-UTR)	9481168	−2	cgagaactgtcgcccttcca<u>C/G</u>gatgtggctccgtgccttta		0.299	*var1A1*
MPJ6_CS1011			Exon 1	9481156	11	ccccttccacgatgtggctcc<u>G/C</u>tgcctttatcctggccactc	Arg4Pro	0.297	*var1A1*
MPJ6_CS1012			Exon 1	9481152	15	tccacgatgtggctccgtgc<u>C/T</u>tttatcctggccactctctc	Ala5Ala	0.297	*var1A1*
MPJ6_CS1013			Exon 1	9481151	16	ccacgatgtggctccgtgcc<u>T/C</u>ttatcctggccactctctct	Phe6Leu	0.297	*var1A1*
MPJ6_CS1014			Exon 1	9481148	19	cgatgtggctccgtgcctt<u>A/G</u>tcctggccactctctctgct	Ile7Val	0.297	*var1A1*
MPJ6_CS1015	rs28563878		Exon 1	9481133	34	tgcctttatcctggccactct<u>C/G</u>ctgcttccgcggcttggggt	Ser12Ala	0.297	*var1A1*
MPJ6_CS1016	rs12149359		Intron 1	9481099	IVS1+16	ttggggtgagtccttctgaa<u>A/G</u>tcaaaatgcggggcacttttt		0.294	*var1A1*

*Number of chromosomes. †Novel variation detected in this study.

Table 4

Frequency of *CES1A2(/1A3)* promoter SNP −816A>C in Japanese cancer patients

CES1A2 and *1A3* Genotype	−816A>C Genotype	Number of subjects	Allele frequency
1A2/1A2	A/A	16	0/32 (0%)
	A/C	0	
	C/C	0	
1A2/1A3	A/A	44	40/166 (24.1%)
	A/C	38	
	C/C	1	
1A3/1A3	A/A	41	48/156 (30.8%)
	A/C	26	
	C/C	11	
Total		177	88/354 (24.9%)

not significant (Table 5). The *CES1* genotypes explained 22.6% of variability in the final model among all the variables and 11.3% of total variability in the AUC ratio.

Effects of CES1 genotypes on SN-38 AUC and toxicity

To clarify the clinical importance of *CES1* genotyping for irinotecan therapy, the effects of *CES1* genotypes or SNPs on AUC levels of the active metabolite SN-38 and neutropenia were examined in the non-*UGT*+/+ patients. In this non-*UGT*+/+ population, significantly higher AUC ratios of (SN-38 + SN-38G)/irinotecan were also observed in the patients with three or four functional *CES1* genes (P = 0.0234, Mann-Whitney test) as observed in all the patients treated with irinotecan monotherapy (Figure 3a). With increased number of functional *CES1* genes, an increasing trend of SN-38 AUC/dose was observed in patients receiving irinotecan monotherapy (1.4-fold for four genes *vs.* two genes; P = 0.080, JT test) (Figure 4). However, multiple regression analysis revealed no statistically significant contribution of *CES1* genotypes to SN-38 AUC/dose although *UGT1A1*6 or *28* and *ABCB1*2/*2* showed significant contributions [10]. Regarding neutropenia, a higher incidence (though statistically insignificant) for grade 3/4 neutropenia in patients with four functional *CES1* genes was observed (50% for four genes and 16% for two or three genes, P = 0.09, Fisher's exact test). The effects of the SNPs (−75G>T, −30G>A and −816A>C) on SN-38 AUC or incidence grade 3/4 neutropenia were not significant (data not shown). In platinum-containing regimens, no significant effects of the *CES1* genotypes on SN-38 AUC/dose or incidence of grade 3/4 neutropenia were detected in the non-*UGT*+/+ patients (data not shown).

Discussion

Recent pharmacogenetic studies on irinotecan have shown the clinical significance of *UGT1A1 *6* and **28* in Japanese patients [7, 8] and *UGT1A1*28* in Caucasians [5, 6] for severe neutropenia. Subsequent studies have revealed additional genetic factors including transporters [10–12]. However, the clinical importance of genotypes of the irinotecan-activating enzymes *CES1* and *CES2* is still uncertain.

Since the hydrolytic activity of CES2 for irinotecan was reported to be much higher than that of CES1 [2], most studies have focused on the clinical significance of *CES2* polymorphisms in irinotecan therapy [13, 14, 22]. We previously identified minor *CES2* genetic variations in Japanese, including *CES2*2* [100C>T (R34W)] and *CES2*5* [1A>T (M1L)] which caused low *in vitro* expression/function of CES2 [13, 14] and also exhibited reduced *in vivo* CES activity in irinotecan-treated patients [14] (also see Figure 3a). However, the major *CES2* haplotypes in Japanese, **1b* (IVS10-108G>A and 1749A>G, frequency = 0.233) and **1c* (−363C>G, IVS10-108G>A and IVS10-87G>A, frequency = 0.027), did not show any significant effects on irinotecan PK [14]. No clinical significance of *CES2* polymorphisms has been reported in Caucasians [22]. Neither *CES1* nor *CES2* SNPs affecting their mRNA expression in normal colonic mucosa were found in European and African populations [23]. Since precise structures of the *CES1* genes and their promoter regions had not been elucidated, evaluation of the roles of the *CES1* genotypes in irinotecan therapy has been rather difficult.

In the present study, the frequencies of individual *CES1* genes (*1A1*, *var1A1*, *1A2* and *1A3*) (Table 2) were almost comparable with the previous report in the Japanese population (0.748, 0.252, 0.313 and 0.687, respectively) [16]. To our knowledge, the present study is the first report suggesting a possible effect of *CES1* genotypes on irinotecan PK. This study showed that the AUC ratio [(SN-38 + SN-38G)/irinotecan], and probably *in vivo* CES activity, was elevated depending on the number of functional *CES1* genes (*1A1*, *var1A1* and *1A2*) in patients treated by irinotecan monotherapy (100 or 150 mg m^{-2} irinotecan) (Figure 3a). This gene-dose effect was not clearly shown in the platinum-containing combination therapy (60–70 mg m^{-2} irinotecan), where renal excretion of irinotecan and its metabolites (especially SN-38G) is highly enhanced by a large volume of infusion fluid. However, the median renal excretion ratio [(SN-38 + SN-38G)/irinotecan] in patients with four functional genes was 1.37-fold higher than that in patients with two or three functional genes in the platinum-containing therapy (data not shown), supporting a partial but significant contribution of the CES1s to activate irinotecan. The present study showed no significant differences in the AUC ratios between *1A1* and *var1A1* (Figure 3a), indicating a common upstream region may be involved in regulation of gene expression of *1A1* and *var1A1*. The previous reports showed the expression levels of CES1A2 were lower than those of CES1A1 [17] and suggested that CES1A2 mRNA was derived mainly from transcription of *var1A1* rather than the original *1A2* [15, 16]. The present study, on the other hand, has suggested that the

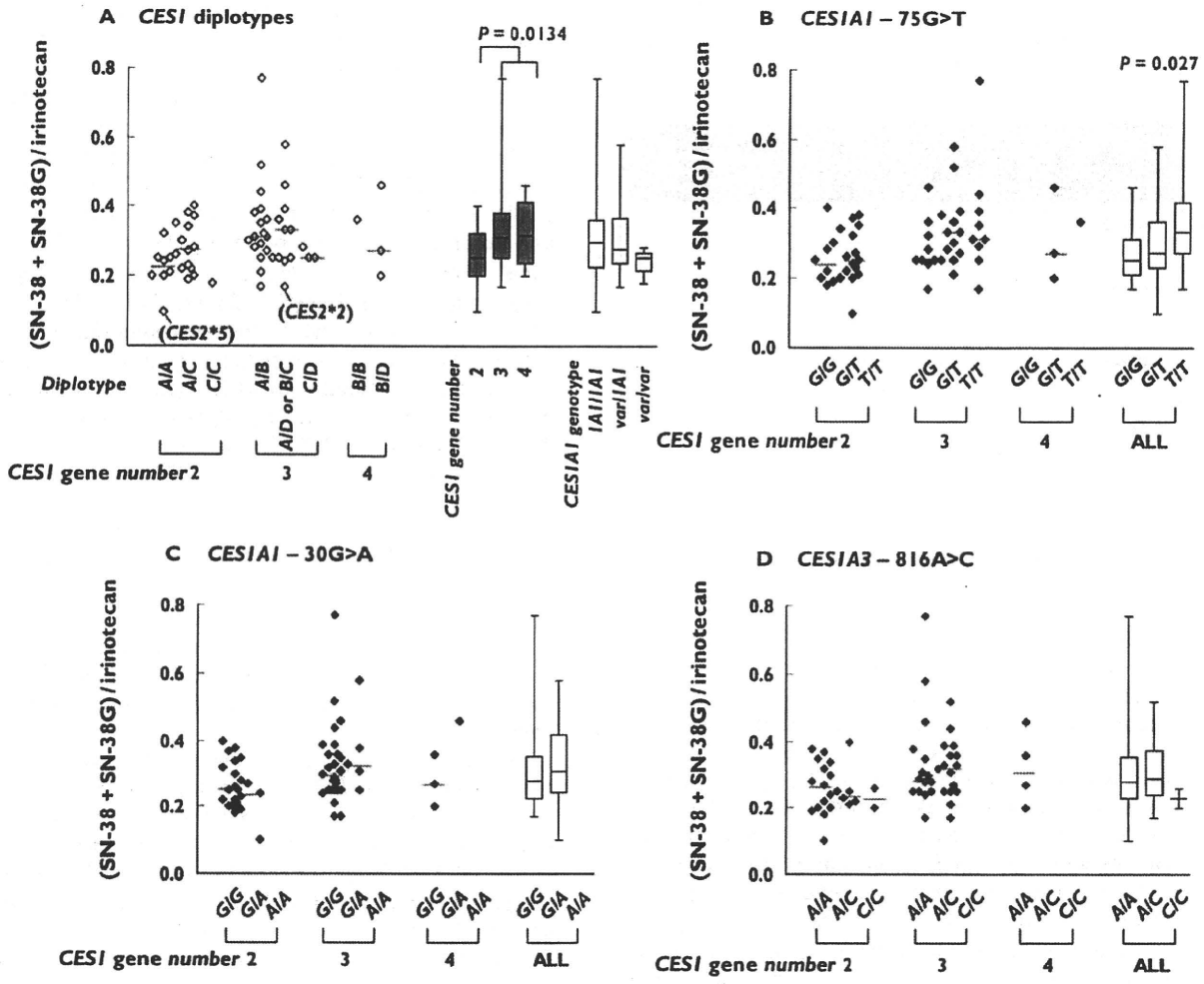

Figure 3

Association of *CES1* diplotypes (A) or SNPs (B–D) with AUC ratio [(SN-38 + SN-38G)/irinotecan], an *in vivo* index of CES activity, in Japanese cancer patients treated with irinotecan monotherapy ($n = 58$). '*CES1* gene number' means the number of functional genes (*1A1*, *var1A1* and *1A2*). Higher AUC ratios were observed in patients with three or four functional *CES1* genes than with two functional genes ($P = 0.0134$, Mann-Whitney test) in (A). Patients with *CES2*5* [*CES2* 1A>T (M1L)] (*CES2*5*) and *CES2*2* [*CES2* 100C>T (R34W)] (*CES2*2*) were found to have reduced CES activity in our previous study [13, 14]

Table 5

Multiple regression analysis of AUC ratio [(SN-38 + SN-38G)/irinotecan]* in Japanese cancer patients treated with irinotecan monotherapy

Variable	Coefficient	SE	*P* value
Smoking	0.073	0.034	0.0375
Initial dose of irinotecan (mg m^{-2})	−0.002	0.001	0.0005
Serum GOT and ALP†	0.082	0.027	0.0038
Serum creatinine (mg dl^{-1})	0.130	0.062	0.0399
ABCB1*2‡ (+/+)	0.042	0.024	0.0831
***CES1* functional gene ($n = 3$ or 4)**	0.038	0.016	0.0215

$r^2 = 0.500$, Intercept = −0.248, $n = 58$. * Values after logarithmic conversion were used. † Grade 1 or greater for both GOT and ALP. ‡ *ABCB1*2* [2677G>T (A8935)].

1A2 transcript could contribute to the total CES activity because the [(SN-38 + SN-38G)/irinotecan] AUC ratios of patients without *1A2* (with two functional *CES1* genes) were lower than those with *1A2* (with three or four functional genes) (Figure 3a). However, it must be noted that the increase in the AUC ratio by three or four functional *CES1* genes was only 20% compared with two functional genes (Figure 3a), and that such alterations might be masked by other non-genetic factors. In fact, hepatic and renal function, irinotecan dosage and smoking history were found to be potent contributors to this parameter (Table 5).

−816A>C SNP in *1A2* was reported to be associated with imidapril efficacy and a higher promoter activity for

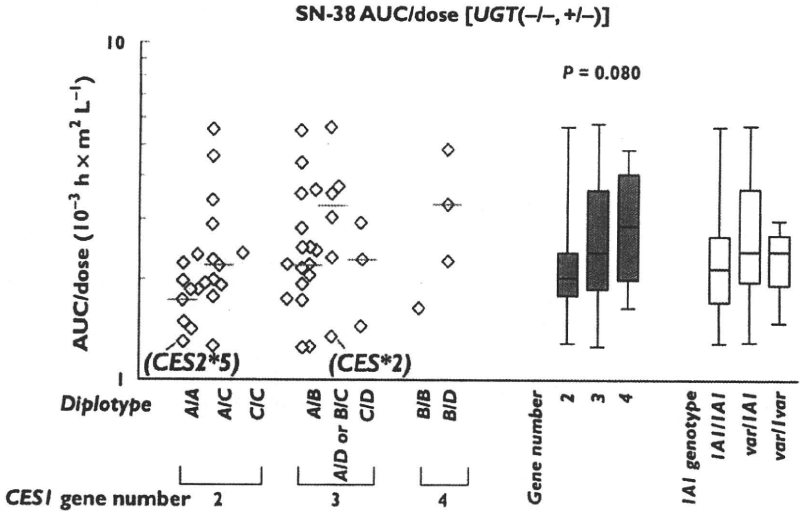

Figure 4

Association of *CES1* genotypes with SN-38 AUC/dose in *UGT*(−/− and +/−) patients treated with irinotecan monotherapy ($n = 51$). '*CES1* gene number' means the number of functional genes (*1A1*, *var1A1* and *1A2*). One patient with an outlying value who had *ABCB1*2* [2677G>T (A893S)] and **14* [2677G>T (A893S) and 1345G>A 230 (E448K)] was excluded from this analysis [10]. A slightly increasing trend in SN-38 AUC(/dose) was observed depending on functional *CES1* gene number. ($P = 0.080$, Jonckheere-Terpstra test). The patients with *CES2*5* [*CES2* 1A>T (M1L)] (*CES2*5*) and *CES2*2* [*CES2* 100C>T (R34W)] (*CES2*2*) [13, 14] are marked

CES1A2 [18] and had strong linkage with SNPs in the proximal promoter region (between −62 to −32) which resulted in additional Sp1 binding sites in the *1A2* promoter region [19]. However, our current study showed no significant effect of −816A>C on the AUC ratio. This can be explained by our finding that −816C and several linked SNPs were mostly located on the *CES1A3* psuedogene but not the functional *1A2* gene.

We newly detected three SNPs (−258C>T, −233C>A and −161A>G) in the 5′-flanking region and one SNP (−30 G>A) in the 5′-UTR of *CES1A1* (Table 3). The effect of −30 G>A on the AUC ratio was not significant (Figure 3c). The frequencies of three other SNPs in the 5′-flanking region were very low (0.003–0.014) which made statistical analysis difficult. These SNPs are not located in the putative transcriptional regulatory regions of *CES1A1*, the binding sites of transcription factors Sp1 and C/EBP [17]. The AUC ratios of the patients with these SNPs were within the 25th–75th percentiles except that slightly higher values were shown in the two -258T patients who received platinum-combination therapy (data not shown). Thus, clinical impact of these SNPs would be small.

With respect to the clinical importance of *CES1* genotyping for irinotecan therapy, the effects of *CES1* genotypes on the AUC level of the active metabolite SN-38 and incidence of grade 3/4 neutropenia should be considered. Since the patients homozygous for *UGT1A1*6 or *28* (*UGT+/+*: **6/*6*, **6/*28* and **28/*28*) showed higher SN-38 AUC/dose levels and severe neutropenia [7], we examined the effects of *CES1* genotypes and SNPs in the non-*UGT+/+* patients. Increasing trends of SN-38 AUC/dose (Figure 4) and incidence of grade 3/4 neutropenia were observed depending on the functional *CES1* gene number in patients with irinotecan monotherapy although statistical significance was not obtained. For the platinum-containing regimens, no significant effects of *CES1* genotypes were shown. Thus, although possible effects of the *CES1* genotypes on neutropenia could not be excluded in irinotecan monotherapy, this study was still insufficient to establish the clinical importance of *CES1* genotyping in irinotecan therapy. Since the sample size will be twice that of the present study to detect a statistically significant decrease of absolute neutrophil counts in the patients with four functional *CES1* genes, future clinical data obtained in a larger number of patients could clarify this point.

In conclusion, this study suggests that the total number of functional *CES1A* genes could influence the formation of the active metabolite of irinotecan in Japanese cancer patients.

Competing interests

HK has received lecture honorarium from Yakult Honsha, the manufacturer of irinotecan. HM has been paid by Yakult Honsha, the manufacturer of irinotecan, for speaking and research.

This study was supported in part by the Program for the Promotion of Fundamental Studies in Health Sciences of the National Institute of Biomedical Innovation, and by the