Differences in tissue-specific and embryonic expression of mouse *Ceacam1* and *Ceacam2* genes

Eric HAN*, Dillon PHAN*, Piao LO*, Matthew N. POY†, Richard BEHRINGER‡, Sonia M. NAJJAR† and Sue-Hwa LIN*[1]

*Department of Molecular Pathology, The University of Texas M. D. Anderson Cancer Center, Houston, TX 77030, U.S.A., †Department of Pharmacology and Therapeutics, Medical College of Ohio, Toledo, OH 43614, U.S.A., and ‡Department of Molecular Genetics, The University of Texas M. D. Anderson Cancer Center, Houston, TX 77030, U.S.A.

The intercellular adhesion molecule CEACAM1, also known as C-CAM1 (where CAM is cell-adhesion molecule), can function as a tumour suppressor in several carcinomas, including those of the prostate, breast, bladder and colon. This suggests that CEACAM1 may play an important role in the regulation of normal cell growth and differentiation. However, there is no direct evidence to support this putative function of CEACAM1. To elucidate its physiological function by targeted gene deletion, we isolated the *Ceacam* genes from a mouse 129 Sv/Ev library. Although there is only one *Ceacam1* gene in humans and one in rats, two homologous genes (*Ceacam1* and *Ceacam2*) have been identified in the mouse. Our sequence analysis revealed that the genes encoded nine exons and spanned approx. 16–17 kb (*Ceacam1*) and 25 kb (*Ceacam2*). The genes were highly similar (79.6%). The major differences in the protein-coding regions were located in exons 2, 5 and 6 (76.9%, 87.0% and 78.5% similarity respectively). In addition, introns 2, 5 and 7 were also significantly different, being 29.7%, 59.8% and 64.5% similar respectively. While most of these differences were due to nucleotide substitutions, two insertions of 418 and 5849 bp occurred in intron 2 of *Ceacam2*, and another two insertions of 1384 and 197 bp occurred in introns 5 and 7 respectively. To determine whether functional redundancy exists between *Ceacam1* and *Ceacam2*, we examined their expression in 16 mouse tissues by using semi-quantitative reverse transcription–PCR. As in human and rat, in the mouse *Ceacam1* mRNA was highly abundant in the liver, small intestine, prostate and spleen. In contrast, *Ceacam2* mRNA was only detected in kidney, testis and, to a lesser extent, spleen. Reverse transcription–PCR using testis RNA indicated that *Ceacam2* in the testis is an alternatively spliced form containing only exons 1, 2, 5, 6, 8 and 9. In the mouse embryo, *Ceacam1* mRNA was detected at day 8.5, disappeared between days 9.5 and 12.5, and re-appeared at day 19. On the other hand, no *Ceacam2* mRNA was detected throughout embryonic development. The different tissue expression patterns and regulation during embryonic development suggest that the CEACAM1 and CEACAM2 proteins, although highly similar, may have different functions both during mouse development and in adulthood.

Key words: C-CAM, cell-adhesion molecule, gene expression, tumour suppressor.

INTRODUCTION

The intercellular adhesion molecule CEACAM1, also known as C-CAM1 [1] (where CAM is cell-adhesion molecule), is a member of the Ig gene family [2]. In addition to its cell-adhesion function, CEACAM1 is also a tumour suppressor in prostate [3], breast [4], bladder [5] and colon [6] carcinomas. Consistent with its role as a tumour suppressor, loss of CEACAM1 expression was observed in hepatomas [7], colon carcinomas [8–10] and endometrial [11] and prostate [12,13] cancers. This suggests that CEACAM1 may regulate cell growth and differentiation.

To elucidate the physiological functions of CEACAM1, we plan to generate mice with targeted *Ceacam* deletions. Genetic manipulation of mouse genes *in vivo* is a powerful approach for understanding the function of a gene, both during embryonic development and in adult tissues. This method requires full knowledge of the genomic structure of the gene of interest. Unlike humans and rats, which each have one *Ceacam1* gene [14–16], two *Ceacam1*-like genes, *Ceacam1* and *Ceacam2*, were identified in BALB/c mice [17]. This poses potential problems in any attempt to manipulate these genes. First, if the two genes have the same function, deletion of one gene may not produce any phenotypic alteration. Secondly, if there is no significant difference in the sequences of the genes, any genetic manipulation may not be specific for a single gene, or it might be difficult to ascertain which gene was altered. It is, therefore, essential to determine the complete sequences and expression profiles of these related *Ceacam* genes in the 129 Sv mouse strain before genetic manipulation is performed.

We report herein the isolation and sequencing of the full-length mouse *Ceacam1* and *Ceacam2* genes. They both encode nine exons, with significant sequence differences in certain regions. These sequence differences will allow selective targeting of one *Ceacam* gene compared with another. In addition, the tissue-specific distributions of CEACAM1 and CEACAM2 were different, and only *Ceacam1* was expressed during embryonic development. The differences in their tissue and developmental expression patterns suggest that these two genes may have different functions *in vivo*.

MATERIALS AND METHODS

Isolation and characterization of *Ceacam1* and *Ceacam2*

A 129 Sv/Ev mouse genomic library in λLEX (provided by Dr Li-Yuan Yu-Lee, Baylor College of Medicine, Houston, TX,

U.S.A.) was screened with the 1.6 kb full-length mouse *Ceacam1* cDNA [18], which had been labelled by using the Klenow fragment of DNA polymerase I, random hexanucleotide primers and [α-^{32}P]dCTP [19]. A total of 24 positive clones were identified in the initial screening of this genomic library. PCR with Oligo 127 (5′-GTGTCACTCTAGGCTACAGGAAAT-3′) and Oligo 122 (5′-GAGGCCAGCTCCATGTCTCTGCTG-3′), which are specific to the 5′ region of mouse *Ceacam1* [20], showed that seven of the clones had the N-terminal *Ceacam1* sequence. Similarly, PCR with Oligo 120 (5′-GAAGTCTGGCGGATCT-GGCTCCTT-3′) and Oligo 131 (5′-TTGAAGTTCAGGACA-GTGTATGCG-3′) showed that seven clones had the 3′ *Ceacam1* sequence. These clones were isolated by secondary and tertiary screening and characterized by restriction mapping. The nucleo-tide sequences of the exons and introns were determined by primer walking using specific oligonucleotide primers. Sequencing was performed by the DNA Sequencing Core Fa-cilities at M. D. Anderson Cancer Center with an automated fluorescent DNA sequencer (Applied Biosystems Inc., Ramsey, NJ, U.S.A.).

Analysis of *Ceacam1* and *Ceacam2* expression using a mouse cDNA panel

A mouse cDNA panel containing first-strand cDNA prepared from mouse tissues and normalized for β-actin expression was purchased from OriGene Technologies Inc. (Rockville, MD, U.S.A.) and used to analyse the expression of *Ceacam1* and *Ceacam2* in various mouse tissues. Oligonucleotides specific to exon 2 of *Ceacam1* (Oligo 116, 5′-AATCTGCCCCTGGCG-CTTGGAGCC-3′; Oligo 179, 5′-AAATCGCACAGTCGCC-TGAGTACG-3′) and to exon 2 of *Ceacam2* (Oligo 117, 5′-AATATGATGAAGGGAGTCTTGGCC-3′; Oligo 180, 5′-AAATTGTCCAGTCAGGACCCTACG-3′) were used as primers to detect specific mRNAs for *Ceacam1* and *Ceacam2* by PCR. PCR cycling conditions were as follows: (1) pre-denaturation (94 °C, 3 min) for one cycle; (2) denaturation (94 °C, 30 s), annealing (60 °C, 30 s) and extension (72 °C, 2 min) for 35 cycles; and (3) final extension (72 °C, 5 min). The predicted size of the PCR products was 246 bp. These PCR products were analysed by agarose-gel electrophoresis, transferred on to a nylon membrane and hybridized with ^{32}P-labelled oligo-nucleotide probes specific to *Ceacam1* (Oligo 181; 5′-AACAC-TACGGCTATAGACAAA-3′) and *Ceacam2* (Oligo 182; 5′-TCTACTACGTCTACAAATGCT-3′).

Ceacam2 cDNA sequence from mouse testis

RNA was prepared from mouse testis by using RNAzol B (TEL-TEST Inc. Friendswood, TX, U.S.A.). The cDNAs coding for *Ceacam2* were obtained from the testis RNA by reverse transcription–PCR (RT-PCR) with Oligo 630 (5′-GAATTCA-AGCTTAAGAAGCTAGCAGGCAGCAGAGAC-3′), which contains nucleotides −36 to +1 of exon 1, and Oligo 631 (5′-GCGGCCGCCTAATGATGATGATGATGATGATGCT-TCTTTTTTACTTCTGAATAAAC-3′), which is complemen-tary to the end of the coding sequence in exon 9 plus seven histidine codons. RT-PCR was performed with Oligo 630 and Oligo 631 according to the procedures provided by the manu-facturer (Amersham/Pharmacia Corp., Arlington Heights, IL, U.S.A.). The PCR product was subcloned into pCRII-topo (Invitrogen, San Diego, CA, U.S.A.) and its sequence was determined.

RESULTS

Isolation and characterization of mouse *Ceacam1* and *Ceacam2* genes

The seven clones with the 5′ sequence of *Ceacam1* had different restriction maps, as did another seven clones with the 3′ end. The different restriction maps seemed to reflect different genes. Grouping these clones according to their restriction maps and DNA sequences revealed two distinct sequences, representing *Ceacam1* and *Ceacam2*. The overlapping genomic clones that spanned *Ceacam1* and *Ceacam2* are shown in Figure 1. DNA sequence analysis of both strands of these two genes revealed that *Ceacam1* and *Ceacam2* each contain nine exons each, with sizes of approx. 18.3 and 24.65 kb respectively (Figure 1).

The sizes of the exons and introns and the intron/exon boundary sequences for *Ceacam1* and *Ceacam2* are shown in Tables 1 and 2 respectively. The first exon, which codes for the first two-thirds of the signal sequence, is 304 bp in both *Ceacam1* and *Ceacam2*. In contrast, exons 2–5 are each approx. 300 bp, and each codes for one Ig-like domain in both *Ceacam1* and *Ceacam2*. The transmembrane domain is encoded by exon 6 and the cytoplasmic domain by exons 7–9. Consistent with the RNA splicing rule [21], each intron starts with GT at the 5′ end and ends with AG at the 3′ end (Tables 1 and 2).

Comparison of *Ceacam1* and *Ceacam2* genes

The overall similarity between *Ceacam1* and *Ceacam2*, including all exons and introns, is approx. 79.6 %. The similarities between
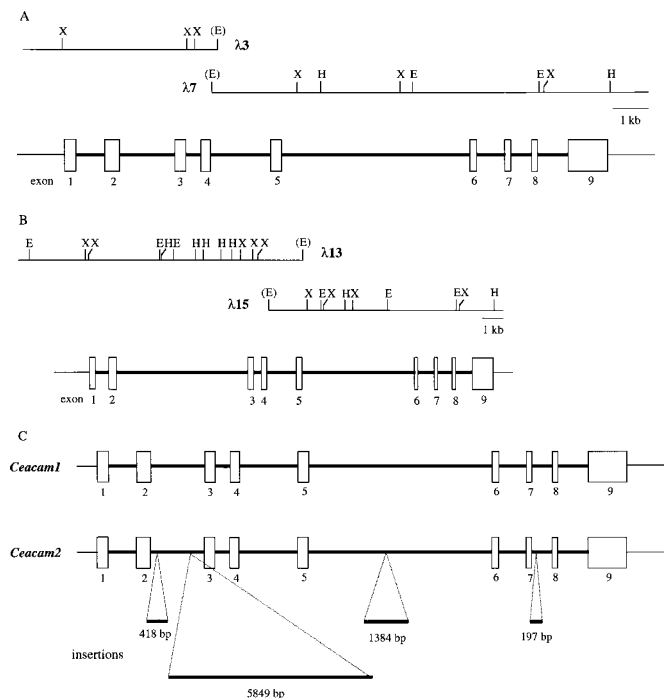

Figure 1   Structures of the mouse *Ceacam1* and *Ceacam2* genes

Numbered open boxes indicate exons. The thick lines connecting the exons represent introns. (**A**) The maps of two overlapping λ clones (λ3 and λ7) that contain *Ceacam1*. (**B**) The maps of two overlapping λ clones (λ13 and λ15) that contain *Ceacam2*. The restriction maps are aligned with the intron/exon map. E, *Eco*RI; H, *Hin*dIII; X, *Xba*I. The enzyme sites in parentheses are from the multiple cloning region of the λ phage vector. (**C**) Comparison of the mouse *Ceacam1* and *Ceacam2* genes.

Table 1   Intron/exon boundaries of *Ceacam1*

Exon	Size (bp)	Exon 3′			Intron	Size (bp)		Exon 5′	
1	304	ACA Thr	G A	gtaaggagatattcc …	1	767	… tctcttccctcttag	CC la	TCA Ser
2	360	CAC His	C P	gtaagtaattatctg …	2	1540	… ctattatctgcacag	CC ro	ATA Ile
3	285	ATC Ile	T T	gtgagtaacttcttc …	3	467	… ttctgtttgctccag	AT yr	GGT Gly
4	255	CTT Leu	G G	gtaagtggatctctg …	4	1727	… tctgtttgtccacag	AG ln	CCA Pro
5	276	ATA Ile	T P	gtgagtgacctcgcc …	5	5358	… ttctttccctgacag	TT he	GAC Asp
6	121	GGC Gly	GG Gl	gtaggacagtctttc …	6	796	… ctcatatttatttag	G y	GGA Gly
7	53	CAC His	A A	gtaagtaaagccaat …	7	587	… ttcttcctcccctag	AT sn	CTG Leu
8	32	AAC Asn	AAG Lys	gtgagcactgccact …	8	948	… ctctcatcctttcag	GTG Val	GAT Asp
9	1180								

Table 2   Intron/exon boundaries of *Ceacam2*

Exon	Size (bp)	Exon 3′			Intron	Size (bp)		Exon 5′	
1	304	ACA Thr	G A	gtaaggagatattcc …	1	766	… tctcttccctcttag	CC la	TCA Ser
2	360	CAC His	A T	gtaagtaattctctg …	2	7807	… acacagtgcagacag	CC hr	CTA Leu
3	285	ATC Ile	T T	gtgagtaacttcttt …	3	465	… tcttgtttgctccag	AT yr	GGT Gly
4	255	CTT Leu	G G	gtaagtggatctctg …	4	1717	… tctgtttgtccacag	AG lu	CCA Pro
5	276	ATA Ile	T P	gtgagtgaccttgcc …	5	6737	… ttctttccctgacag	TT he	GAC Asp
6	118	CGC Arg	TG Tr	gtaggacagtctttc …	6	798	… ctcatatttatttag	G p	GGA Gly
7	53	CAC His	A A	gtaagtaaagccaat …	7	771	… ctcttcctcccctag	AT sn	CTG Leu
8	32	AAC Asn	AAG Lys	gtgagcactgccact …	8	973	… ctctcatattttcag	GTG Val	GAT Asp
9	1180								

Table 3   Intron/exon lengths and similarities of *Ceacam1* and *Ceacam2*

	Length (bp)				Length (bp)		
Exon	*Ceacam1*	*Ceacam2*	Similarity (%)	Intron	*Ceacam1*	*Ceacam2*	Similarity (%)
1	304	304*	88.9	1	767	766	97.4
2	360	360	76.9	2	1540	7807	29.7
3	285	285	97.9	3	467	465	94.7
4	255	255	99.6	4	1726	1717	97.9
5	276	276	87.0	5	5358	6737	59.8
6	121	118	78.5	6	796	798	98.5
7	53	53	100	7	587	771	64.5
8	32	32	100	8	948	973	94.3
9	1180	1180†	86.8				

* Includes 240 bp of promoter and 5′ untranslated sequence.
† Includes 889 bp of 3′ untranslated sequence.

Exon 2 (nucleotide)

```
              843       853       863       873       883       893
Ceacam1   CCTCACTTTTAGCCTCCTGGAGCCCTGCCACCACTGCTGAAGTCACCATTGAGGCTGTGC
          ::::::::::::::::::::::::::::: :::: :: ::  :::: : :
Ceacam2   CCTCACTTTTAGCCTCCTGGAGCCCTCCCACCACTGCACAAGTGACTGTTATGGCTTTTC
              840       850       860       870       880       890

              903       913       923       933       943       953
Ceacam1   CGCCCCCAGGTTGCTGAAGCAACAATGTTCTTCTACTTGTTCACAATCTGCCCCTGGCGC
          : : :::: : :::::::: ::::::::::: ::::: ::::: ::::: ::     ::
Ceacam2   CACTCCACGCTGCTGAAGGCAACAATGTTATTCTAGTTGTTTACAATATGATGAAGGGAG
              900       910       920       930       940       950

              963       973       983       993      1003      1013
Ceacam1   TTGGAGCCTTTGCCTGGTACAAGGGAAACACTACGGCTATAGACAAAGAAATTGCACGAT
          :  :::::: :::: ::::::::: ::::::: ::: : : ::::::: :::::
Ceacam2   TCTTGGCCTTTAGCTGGCACAAGGGATCTACTACGTCTACAAATGCTGAAATTGTACGAT
              960       970       980       990      1000      1010

             1023      1033      1043      1053      1063      1073
Ceacam1   TTGTACCAAATAGTAATATGAATTTCACGGGGCAAGCATACAGCGGCAGAGAGATAATAT
          ::::: ::   : :::::: ::  : :  :::: : : :::: :::::::::: : :::
Ceacam2   TTGTAACAGGCACTAATAAGACTATAAAAGGGCCTGTACACAGTGGCAGAGAGGCACTAT
             1020      1030      1040      1050      1060      1070

             1083      1093      1103      1113      1123      1133
Ceacam1   ACAGCAATGGATCCCTGCTCTTCCAAATGATCACCATGAAGGATATGGGAGTCTACACAC
          ::::::::::::::::::::::: :::::: : ::::::::::::::: :::::::::::::
Ceacam2   ACAGCAATGGATCCCTGCTCATCCAAAGGGTCACCATGAAGGATACGGGAGTCTACACAA
             1080      1090      1100      1110      1120      1130

             1143      1153      1163      1173      1183      1193
Ceacam1   TAGATATGACAGATGAAAACTATCGTCGTACTCAGGCGACTGTGCGGATTTCATGTACACC
          :::: ::::::::: :: :::::::::::: ::  :: ::     ::  :: :::::::::
Ceacam2   TAGAAATGACAGATCAAGACTATCGTCGTAGGGTCCTGACTGGACAATTTCATGTACACA
             1140      1150      1160      1170      1180      1190
```

Exon 2 (protein)

```
              32        42        52        62        72        82
Ceacam1   SLLASWSPATTAEVTIEAVPPQVAEDNNVLLLVHNLPLALGAFAWYKGNTTAIDKEIARF
          ::::::::: ::: :: : :   :: ::: : :        :: : :: ::    :: ::
Ceacam2   SLLASWSPPTTAQVTVMAFPLHAAEGNNVILVVYNMMKGVLAFSWHKGSTTSTNAEIVRF
              32        42        52        62        72        82

              92       102       112       122       132
Ceacam1   VPNSNMNFTGQAYSGREIIYSNGSLLFQMITMKDMGVYTLDMTDENYRRTQATVRFHVH
          :   :    :   :::: ::::::: : :::: :::: ::: ::: ::  : ::::
Ceacam2   VTGTNKTIKGPVHSGREALYSNGSLLIQRVTMKDTGVYTIXMTDQDYRRRVLTGQFHVH
              92       102       112       122       132
```
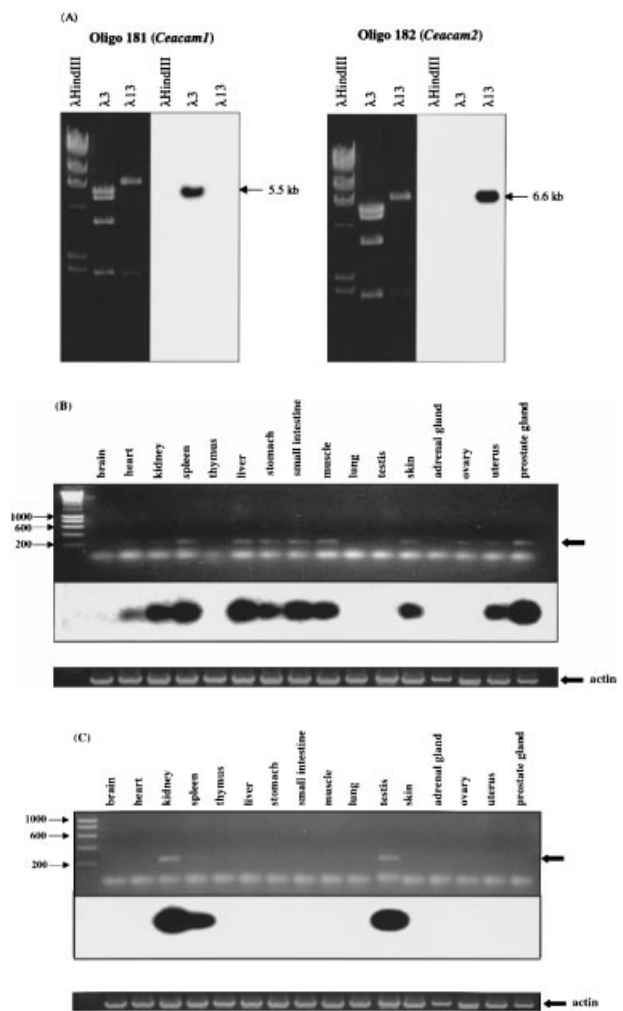
Figure 2   Comparison of the nucleotide and protein sequences of exon 2 of *Ceacam1* and *Ceacam2*

Oligonucleotides for PCR analyses are underlined. Oligonucleotides for hybridization (Oligo 181 and Oligo 182) are boxed. The nucleotides are numbered with the A of the start ATG as the first nucleotide according to the genomic sequence. The amino acids are numbered with the start methionine (ATG) as the first amino acid.

Figure 3   Expression of *Ceacam1* and *Ceacam2* mRNAs in various mouse tissues

(**A**) Specificity of Oligos 181 and 182 in distinguishing *Ceacam1* from *Ceacam2*. (**B**, **C**) Tissue-specific expression of *Ceacam1* (**B**) and *Ceacam2* (**C**) analysed by RT-PCR and hybridization with Oligo 181 and Oligo 182 respectively. The positions of the PCR products are indicated by the arrows on the right.

the exons and introns of *Ceacam1* and *Ceacam2* are shown in Table 3. The major differences between these two genes are in exon 2, intron 2, intron 5, exon 6 and intron 7, which are 76.9 %, 29.7 %, 59.8 %, 78.5 % and 64.5 % similar respectively. The nucleotide and the deduced amino acid sequences of exon 2 for *Ceacam1* and *Ceacam2* are shown in Figure 2. It is apparent that most of the sequence substitutions in exons 2, which code for the first Ig domain, do result in amino acid substitutions in *Ceacam1* and *Ceacam2*. Thus the amino acid identity between the first Ig domains of CEACAM1 and CEACAM2 is 57.1 %, which is significantly lower than the similarity of other domains. In addition, insertions of 418 bp and 5849 bp in intron 2, and of 1384 bp and 197 bp in introns 5 and 7 respectively, are found in *Ceacam2* compared with *Ceacam1* (Table 3 and Figure 1C).

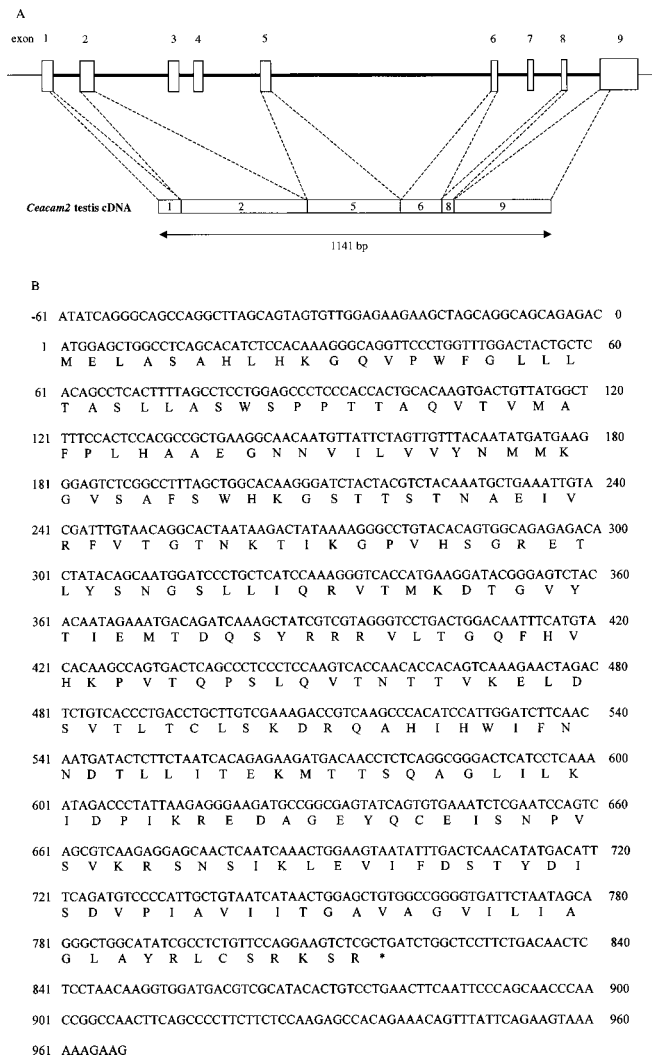## Expression of *Ceacam1* and *Ceacam2* mRNAs in mouse tissues

To determine the tissue-specific distribution of *Ceacam1* and *Ceacam2*, we performed semi-quantitative RT-PCR on a panel

of mouse cDNAs from various tissues by using *Ceacam1*- and *Ceacam2*-specific oligonucleotide pairs. The PCR products were hybridized further with *Ceacam1*- and *Ceacam2*-specific oligo-nucleotides derived from exon 2, the least similar exon (Figure 2). The specificities of these two oligonucleotides were confirmed by Southern blot analysis. As shown in Figure 3(A), Oligo 181 hybridized to DNA from λ3 (*Ceacam1*) but not from λ13 (*Ceacam2*), whereas Oligo 182 only hybridized to λ13 DNA. Both *Ceacam1* and *Ceacam2* mRNAs were detected, although not in the same tissues. This observation supports the notion that both *Ceacam1* and *Ceacam2* are expressed in mouse (Figures 3B and 3C). High levels of *Ceacam1* mRNA were detected in liver, small intestine, prostate and spleen (Figure 3B), similar to the tissue expression pattern in humans [22] and rats [23,24]. *Ceacam1* mRNA was also detected in heart, kidney, stomach, muscle, skin and uterus. In contrast, *Ceacam2* mRNAs were detected only in kidney, testis and, to a lesser extent, spleen (Figure 3C). These
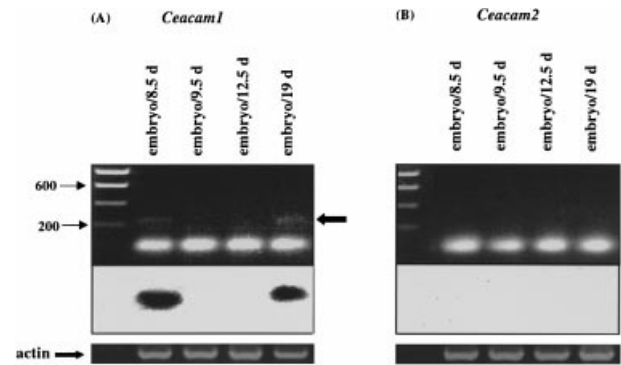
Figure 4   *Ceacam2*

(**A**) cDNA structure. The numbered boxes indicate exons. The thick lines connecting the exons represent introns. The broken lines show the portions of the exon sequences included in testis *Ceacam2* cDNA. (**B**) cDNA and protein sequences.

Figure 5   Expression of *Ceacam1* (A) and *Ceacam2* (B) mRNAs during embryonic development

(**A**) Analysis of RT-PCR products of *Ceacam1* on an agarose gel and gene-specific hybridization using Oligo 181. (**B**) Analysis of RT-PCR products of *Ceacam2* on an agarose gel and gene-specific hybridization using Oligo 182. Marker sizes are indicated by thin arrows, and the size of the PCR product is indicated by thick arrow.

and its sequence was determined (Figure 4). The mouse testis cDNA contained exons 1, 2, 5, 6, 8 and 9, and had an open reading frame of 273 amino acids (Figure 4B).

Expression of *Ceacam1* and *Ceacam2* during mouse embryonic development

The expression of *Ceacam1* and *Ceacam2* was examined in mouse embryos at days 8.5, 9.5, 12.5 and 19 of embryonic development by RT-PCR hybridization as described above. Using *Ceacam1*-specific Oligo 181, we observed age-related differences in mRNA levels: we detected the mRNA at day 8.5, it had disappeared at days 9.5–12.5, and it re-appeared at day 19 (Figure 5A). In contrast, no hybridization signal was detected when *Ceacam2*-specific Oligo 182 was used (Figure 5B). This result suggests that the expression of *Ceacam1*, but not that of *Ceacam2*, is developmentally regulated.

DISCUSSION

Because targeted gene deletion is performed in the 129 Sv mouse, we determined in the present studies the complete genomic structure and DNA sequences of the mouse *Ceacam1* and *Ceacam2* genes in a 129Sv/Ev mouse genomic library. We also examined the expression patterns of these two genes in adult mouse tissues and during embryonic development. Several conclusions can be made from this study. First, the *Ceacam1* and *Ceacam2* genes contain sufficient sequence differences that targeted gene deletion specific to either gene should be feasible. Secondly, functional redundancy may not be a problem when only one of these two genes is deleted, because, although *Ceacam1* and *Ceacam2* are highly homologous, they have different tissue expression patterns. Thirdly, *Ceacam1* is probably more important than *Ceacam2* in the mouse, because the tissue expression pattern of *Ceacam1* is similar to those of the single *Ceacam* genes in humans and rats. Conservation of the expression profile among these different species suggests that the function of

observations demonstrate that the *Ceacam1* and *Ceacam2* genes are differentially expressed in mouse tissues.

Ceacam2 cDNA

Although it was detected in testis, kidney and spleen, *Ceacam2* mRNA may not contain open reading frames for protein translation. As testis expressed only *Ceacam2*, the cDNAs coding for *Ceacam2* were obtained from testis RNA by RT-PCR with Oligo 630, which contained the 5′ untranslated region of exon 1, and Oligo 631, complementary to the exon 9 coding region at the termination codon TGA. The 1 kb PCR product obtained only hybridized to *Ceacam2*-specific Oligo 182, and not to *Ceacam1*-specific Oligo 181, suggesting that the fragment codes for *Ceacam2*. This result is consistent with the tissue distribution of *Ceacam2* (Figure 3C). The PCR product was then subcloned

Ceacam1 may be essential. Fourthly, *Ceacam2* plays no role in embryonic development; it is not expressed in the mouse embryo.

The results of the present study provide important information for the design of gene targeting strategies for functional studies of *Ceacam* genes, by either deleting or introducing mutations into the *Ceacam* genes in the mouse germline. The differences in the two DNA sequences can be used to target a specific gene. As *Ceacam1* lacks the 418 and 5849 bp insertions in *Ceacam2* intron 2 and the 1384 bp insertion present in *Ceacam2* intron 5 (Figure 1C), gene-targeting vectors containing introns 2 or 5 of *Ceacam1* or *Ceacam2* could be used to achieve selective homologous recombination in the desired gene. If mutant mice in which only one *Ceacam* gene is deleted survive, then they can be crossed further to generate mice deficient in both *Ceacam* genes. However, as *Ceacam1* and *Ceacam2* are both on chromosome 7, it is possible that they are too close to allow construction of double-gene-knockout mice by crossing the single-gene-knockout mice.

The existence of the highly similar second *Ceacam* gene, *Ceacam2*, in mice raises the possibility that *Ceacam2* is a pseudogene. However, our results indicate that *Ceacam2* is probably not a pseudogene, since (1) *Ceacam2* contains a complete set of exons and introns typical of a *Ceacam* gene; (2) *Ceacam2* is transcribed in mouse, as shown by the presence of *Ceacam2* mRNA in several tissues; and (3) *Ceacam2* mRNA contains an open reading frame of 273 amino acids in testis, as has been shown in CMT-93 mouse rectal carcinoma cells [17]. These findings raise the interesting question of the function of *Ceacam2*. One of the functions of CEACAM1 is inhibition of tumour growth [3,6]. This suggests that this protein may play an important role in regulating cell growth and differentiation. Structural and functional analyses of rat CEACAM1 have revealed that the tumour-suppressive function requires a long cytoplasmic domain, generated by alternative splicing [4,6,24]. Because the cytoplasmic domains of CEACAM1 and CEACAM2 are identical, CEACAM2 may also have growth-suppressive activity. However, the role of CEACAM2 in testis, which is composed of rapidly dividing cells in spermatogenesis, is not clear. Thus the *in vivo* function of CEACAM2 may not be revealed until its gene is deleted in the mouse.

Knocking out a gene may have no phenotypic effect if related genes have similar functions. We showed that *Ceacam1* and *Ceacam2* are expressed in different tissues in the mouse. *Ceacam1* mRNA was detected in tissues rich in epithelial cells, which is consistent with expression of the CEACAM1 homologue in rats and humans [25–27]. In contrast, *Ceacam2* mRNA was abundantly expressed in testis, which does not express *Ceacam1* mRNA. In addition, *Ceacam2* mRNA was undetectable in the mouse embryo, whereas *Ceacam1* mRNA was developmentally regulated. Thus it appears that the two mouse *Ceacam* genes are not functionally redundant. However, we cannot rule out the possibility that the loss of *Ceacam1* expression may up-regulate expression of *Ceacam2*. These studies will have to await targeted gene deletion of these two related genes.

REFERENCES

1   Beauchemin, N., Draber, P., Dveksler, G., Gold, P., Gray-Owen, S., Grunert, F., Hammarstrom, S., Holmes, K. V., Karlsson, A., Kuroki, M. et al. (1999) Redefined nomenclature for members of the carcinoembryonic antigen family. Exp. Cell Res. **252**, 243–249

2   Lin, S. H., Culic, O., Flanagan, D. and Hixson, D. C. (1991) Immunochemical characterization of two isoforms of rat liver ecto-ATPase that show an immunological and structural identity with a glycoprotein cell-adhesion molecule with M_r 105000. Biochem. J. **278**, 155–161

3   Hsieh, J. T., Luo, W., Song, W., Wang, Y., Kleinerman, D. I., Van, N. T. and Lin, S. H. (1995) Tumor suppressive role of an androgen-regulated epithelial cell adhesion molecule (C-CAM) in prostate carcinoma cell revealed by sense and antisense approaches. Cancer Res. **55**, 190–197

4   Luo, W., Wood, C., Earley, K., Hung, M.-C. and Lin, S.-H. (1997) Suppression of tumorigenicity of breast cancer cells by an epithelial cell adhesion molecule (C-CAM1): The adhesion and growth suppression are mediated by different domains. Oncogene **14**, 1697–1704

5   Kleinerman, D. I., Dinney, C. P. N., Zhang, W.-W., Lin, S.-H., Van, N. T. and Hsieh, J.-T. (1996) Suppression of human bladder cancer growth by increased expression of C-CAM1 gene in an orthotopic model. Cancer Res. **56**, 3431–3435

6   Kunath, T., Ordonez-Garcia, C., Turbide, C. and Beauchemin, N. (1995) Inhibition of colonic tumor cell growth by biliary glycoprotein. Oncogene **11**, 2375–2382

7   Thompson, N. L., Lin, S. H., Panzica, M. A. and Hixson, D. C. (1994) Cell CAM 105 isoform RNA expression is differentially regulated during rat liver regeneration and carcinogenesis. Pathobiology **62**, 209–220

8   Neumaier, M., Paululat, S., Chan, A., Matthaes, P. and Wagener, C. (1993) Biliary glycoprotein, a potential human cell adhesion molecule, is down-regulated in colorectal carcinomas. Proc. Natl. Acad. Sci. U.S.A. **90**, 10744–10748

9   Nollau, P., Scheller, H., Kona-Horstmann, M., Rohde, S., Hagenmuller, F., Wagener, C. and Neumaier, M. (1997) Expression of CD66a (human C-CAM) and other members of the carcinoembryonic antigen gene family of adhesion molecules in human colorectal adenomas. Cancer Res. **57**, 2354–2357

10 Rosenberg, M., Nedellec, P., Jothy, S., Fleiszer, D., Turbide, C. and Beauchemin, N. (1993) The expression of mouse biliary glycoprotein, a carcinoembryonic antigen-related gene, is down-regulated in malignant mouse tissues. Cancer Res. **53**, 4938–4945

11 Bamberger, A. M., Riethdorf, L., Nollau, P., Naumann, M., Erdmann, I., Gotze, J., Brummer, J., Schulte, H. M., Wagener, C. and Loning, T. (1998) Dysregulated expression of CD66 (BGP, C-CAM), an adhesion molecule of the CEA family, in endometrial cancer. Am. J. Pathol. **152**, 1401–1406

12 Kleinerman, D. I., Troncoso, P., Lin, S. H., Pisters, L. L., Sherwood, E. R., Brooks, T., von Eschenbach, A. C. and Hsieh, J. T. (1995) Consistent expression of an epithelial cell adhesion molecule (C-CAM) during human prostate development and loss of expression in prostate cancer: implication as a tumor suppressor. Cancer Res. **55**, 1215–1220

13 Pu, Y. S., Luo, W., Lu, H.-H., Gingrich, J., Greenberg, N. and Lin, S.-H. (1999) Differential expression of C-CAM protein in prostate cancer progression in a transgenic adenocarcinoma mouse model. J. Urol. **162**, 892–896

14 Brandriff, B. F., Gordon, L. A., Tynan, K. T., Olsen, A. S., Mohrenweiser, H. W., Fertitta, A., Carrano, A. V. and Trask, B. J. (1992) Order and genomic distances among members of the carcinoembryonic antigen (CEA) gene family determined by fluorescence in situ hybridization. Genomics **12**, 773–779

15 Najjar, S. M., Boisclair, Y. R., Nabih, Z. T., Philippe, N., Imai, Y., Suzuki, Y., Suh, D. S. and Ooi, G. T. (1996) Cloning and characterization of a functional promoter of the rat pp120 gene, encoding a substrate of the insulin receptor tyrosine kinase. J. Biol. Chem. **271**, 8809–8817

16 Cheung, P. H., Culic, O., Qiu, Y., Earley, K., Thompson, N., Hixson, D. C. and Lin, S. H. (1993) The cytoplasmic domain of C-CAM is required for C-CAM-mediated adhesion function: studies of a C-CAM transcript containing an unspliced intron. Biochem. J. **295**, 427–435

17 Nedellec, P., Dveksler, G. S., Daniels, E., Turbide, C., Chow, B., Basile, A. A., Holmes, K. V. and Beauchemin, N. (1994) Bgp2, a new member of the carcinoembryonic antigen-related gene family, encodes an alternative receptor for mouse hepatitis viruses. J. Virol. **68**, 4525–4537

18 Williams, R. K., Jiang, G.-S. and Holmes, K. V. (1991) Receptor for mouse hepatitis virus is a member of the carcinoembryonic antigen family of glycoprotein. Proc. Natl. Acad. Sci. U.S.A. **88**, 5533–5536

19 Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd edn, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

20 Nedellec, P., Turbide, C. and Beauchemin, N. (1995) Characterization and transcriptional activity of the mouse biliary glycoprotein 1 gene, a carcinoembryonic antigen-related gene. Eur. J. Biochem. **231**, 104–114

21 Senapathy, P., Shapiro, M. B. and Harris, N. L. (1990) Splice junctions, branch point sites, and exons: sequence statistics, identification, and applications to genome project. Methods Enzymol. **183**, 252–278

22   Prall, F., Nollau, P., Neumaier, M., Haubeck, H. D., Drzeniek, Z., Helmchen, U., Loning, T. and Wagener, C. (1996) CD66a (BGP), an adhesion molecule of the carcinoembryonic antigen family, is expressed in epithelium, endothelium, and myeloid cells in a wide range of normal human tissues. J. Histochem. Cytochem. **44**, 35–41

23   Odin, P., Asplund, M., Busch, C. and Obrink, B. (1988) Immunohistochemical localization of cell CAM 105 in rat tissues: appearance in epithelia, platelets, and granulocytes. J. Histochem. Cytochem. **36**, 729–739

24   Estrera, V. T., Luo, W., Phan, D., Hixson, D. and Lin, S.-H. (1999) The cytoplasmic domain of C-CAM1 is necessary and sufficient in suppression of the growth of prostate cancer cells. Biochem. Biophys. Res. Commun. **263**, 797–803

25   Cheung, P. H., Thompson, N. L., Earley, K., Culic, O., Hixson, D. and Lin, S. H. (1993) Cell-CAM105 isoforms with different adhesion functions are coexpressed in adult rat tissues and during liver development. J. Biol. Chem. **268**, 6139–6146

26   Thompson, J. A., Grunert, F. and Zimmermann, W. (1991) Carcinoembryonic antigen gene family: molecular biology and clinical perspectives. J. Clin. Lab. Anal. **5**, 344–366

27   Hinoda, Y., Neumaier, M., Hefta, S. A., Drzeniek, Z., Wagener, C., Shively, L., Hefta, L. J. F., Shively, J. E. and Paxton, R. (1988) Molecular cloning of a cDNA coding biliary glycoprotein I: Primary structure of a glycoprotein immunologically crossreactive with carcinoembryonic antigen. Proc. Natl. Acad. Sci. U.S.A. **85**, 6959–6963