The carcinoembryonic antigen (CEA) gene family encodes a large family of glycoproteins. Some are probably involved in the homeostasis/development of epithelial cells and granulocyte activation, while others e.g. the pregnancy-specific glycoproteins, are expressed in the placenta and are essential for a positive outcome of pregnancy. In this paper, we have characterized cea5, a member of the murine CEA gene family. RNase protection and in situ hybridization analyses revealed that Ceas mRNA is exclusively synthesized in primary and secondary trophoblast giant cells of the placenta only during early stages of development. Full-length Cea5 cDNA was obtained by a reverse transcription-polymerase chain reaction using day 10.5 post-coitum placental RNA. The 1.6-kilobase pair (kb) Cea5 mRNA encodes a secreted glycoprotein with a predicted size of 30 kDa. It is composed of a leader peptide (L), one immunoglobulin (Ig) variable or N, and one Ig constant-like or A domain. This domain organization is unique within the human and murine CEA families. Two overlapping cosm id clones covering 54 kb of the cea5 gene locus were mapped. cea5 consists of three exons (L, N, A/3 untranslated region exon) located within a 4-kb region. rncGM2, the rat cea5 counterpart, exhibits the same restricted expression pattern. This together with their exceptional conservation within the rat and murine CEA families and their absence from the human CEA family suggests that cea5 and rncGM2 are of functional importance for rodent placental development.

The carcinoembryonic antigen (CEA) family was originally discovered in humans and was named after CEA, the first family member to be discovered and later characterized as a human tumor marker (1). Protein sequencing and gene cloning have since revealed the existence of a CEA gene family that belongs to the immunoglobulin superfamily (2). CEA-related proteins are highly glycosylated and are composed of one immunoglobulin variable (IgV)-like and between zero and six immunoglobulin constant (IgC)-like domains. They are either membrane-bound or actively secreted molecules. The human CEA gene family has been intensively studied and consists of 29 genes, 18 of which are expressible (3). Based on sequence similarities, three subgroups can be discerned. The CEA-subgroup consists of 12 genes, of which 8 are expressed, mainly in epithelial cells and/or in granulocytes (3, 4). The pregnancy-specific glycoprotein (PSG) subgroup comprises 11 expressible genes, all of which are active in the syncytiotrophoblast cells of the placenta. A third subgroup contains 6 genes, none of which have been shown to be active. All 29 genes are clustered on the long arm of chromosome 19 (3, 5).

Although the in vivo functions of CEA family members are still unknown, several CEA subgroup members have been shown to convey homo- and heterophilic cell adhesion, which could play a role in the development of epithelia during organogenesis (6, 7). Some CEA-related molecules appear to be involved in granulocyte adhesion to activated endothelial cells in presenting sialyl-Lei carbohydrate structures to E-selectin and in mediating signals for granulocyte activation (8–10). Ectopic expression in rat myeloblasts revealed that CEA subgroup members might also play a role in the regulation of differentiation (11). Down-regulation of two CEA subgroup members, the biliary glycoprotein (BGP) and CEA gene family member 2 (CGM2), during the development of colorectal carcinomas indicates a causative role of their products in tumorigenesis (12–15). Indeed, it has been demonstrated that rat and murine BGP homologs can function as tumor suppressors in human prostate cancer cells, in breast tumor cell lines, and in murine colorectal tumor cell lines (16–18). For one PSG subgroup member (PSG11), it was shown that an arginine-glycine-aspartic acid (RGD) tripeptide-containing motif mediates binding to receptors on promyocytic cell lines (19). This indicates a possible role of PSG in regulating the maternal immune system during pregnancy. In addition, murine BGPs (Bgp1 and Bgp2) function as mouse hepatitis virus receptors (20, 21).

Rigorous proof is still lacking that the in vitro properties of CEA family members are relevant for normal development and tumorigenesis. As a basis for in vivo functional investigations, e.g., through gene inactivation experiments, the rodent CEA gene family has also been analyzed. Although CEA gene families exist in both the rat and mouse, in comparison with their human counterparts, they vary at both the sequence level and in their domain structures (22–32). These differences make it impossible to determine homologs for most family members between orders and, in some instances, even between rat and mouse species (for a review see Ref. 33). Sequence analyses indicate that prior to mammalian radiation only one or two
CEA-like genes existed, which have rapidly expanded through gene duplication events, but independently in the different mammalian orders, leading to these differences (33). Therefore, the CEA gene family represents a paradigm for molecular evolution in progress, both between and within mammalian orders.

Based mainly upon expression studies, the murine gene family can also be divided into a CEA subgroup, consisting to date of 3 genes and a PSG subgroup, where 15 genes have so far been characterized (31).2 Similar subgroups exist in the rat (33). The only gene homologs so far identified within the human, rat, and mouse CEA gene families are BGP-related genes based on an identical exon arrangement and similar expression pattern (21, 34–38). In general, orthologous gene pairs between the CEA families cannot be assigned through sequence comparison, even between rat and mouse (33). However, to date one exception is known: the putative counterparts rnCGM2 (rat) and cea5 (mouse), are highly conserved at the amino acid sequence level of their N domains (93% identity), whereas no counterpart exists in humans (26, 39). Therefore, these two genes may be rodent-specific and could serve important rodent-specific function(s). Until now, the structure, expression pattern, and function of Cea5 and rnCGM2 remained unknown.

For this reason, we have isolated and characterized a full-length cdNA for the murine CEA gene family member cea5, for which only partial genomic sequences were so far available (26). This gene encodes a CEA-related protein with a novel domain organization and cannot be unequivocally classified either to the murine CEA or PSG subgroups based on sequence comparisons. As found for PSG genes, cea5 was found to be expressed exclusively in the placenta. However, in contrast to the PSG genes, cea5 is expressed in primary and secondary trophoblast giant cells, whereas PSG transcripts are found in secondary trophoblast giant cells and in the spongiosotrophoblast (40, 41). Consequently, following the disappearance of the giant cells, cea5 shows an inverse expression kinetics to PSG genes in being predominantly active only during early stages of placental development. Northern blot analyses revealed a similar expression pattern for rnCGM2 in the rat, supporting the idea that cea5 and rnCGM2 represent homologs and are indeed functionally important.

MATERIALS AND METHODS

RNA Isolation, RNase Protection Assay, and Northern Blot Analysis—BALB/c mice (Charles River) or Wistar rats (animal facility of the Institute of Biochemistry and Molecular Biology, University of Freiburg) were anesthetized by inhalation of isoflurane and killed by cervical dislocation. Tissues were immediately frozen in liquid nitrogen (see below) with XbaI and KpnI. As an rnCGM2 probe, a 289-bp HindIII/BamHI genomic DNA fragment comprising about 50 bp of intron 1 and two-thirds of the N domain exon (exon 2) was used (39). Both fragments were labeled with [α-32P]dATP (MegaPrime™ DNA labeling system, Amersham Corp.). The final wash was performed at high stringency in 0.1 × SSPE (1 × SSPE: 180 mM NaCl, 10 mM sodium phosphate, pH 7.4, 1 mM EDTA), 0.1% sodium dodecyl sulfate (SDS) twice for 20 min at 65 °C. The filters were rehybridized with a 32P-labeled 1.2-kb mouse β-actin probe (43) to check for equal loading and the integrity of the RNAs.

Isolation and Mapping of Cea5 cdNA and Genomic Clones—Total RNA (3 μg) from placenta (day 10.5 of pregnancy) was reverse-transcribed with an oligo(dT)-containing oligonucleotide as the 3'-anchor primer (26) using standard procedures. Cea5 cDNA was amplified by PCR (annealing: 52 °C, 45 s; extension: 72 °C, 3 min; denaturation: 94 °C, 30 s; 30 cycles) using oligonucleotides BK4 or DF4 (position 1–20) as 5'-primers and as 3'-primers the 3'-anchor primer and M8-ko (position 926–905), respectively (for primer positions see Fig. 3A). The nucleotide sequences of primers BK4 and DF4 were deduced from partial genomic cea5 sequences (26). The products were subcloned into pUC18 after blunt-ending (SureClone ligation kit, Pharmacia Biotech Inc.).

The cea5 gene-containing cosmids clone 14.23 was isolated from a C57BL/6 mouse liver cosmid library in PWE15 (Stratagene). 3 × 106 cosmids clones were plated in duplicate onto Hybond-N+ (Amersham Corp.) and hybridized with the 32P-labeled 5.5-kb N exon probe. The final wash was in 4 × SSPE, 0.1% SDS twice for 20 min at 65 °C. The extent of overlap with cosC3 (26) was determined by comparing the sizes of the DNA fragments of 14.23 and cosC3 after digestion with various restriction endonucleases.

Determination of Transcriptional Start Sites—The major transcriptional start sites of cea5 were identified using the 5'-RACE system for the rapid amplification of cdNA ends (Life Technologies, Inc.) according to the manufacturer's protocol. Reverse transcription of 1 μg of total placental RNA (day 10.5 of pregnancy) was performed with GSP1 (position 555–535 in Fig. 3A) as primer. 5'-RACE products were amplified (annealing: 48 °C, 30 s; extension: 72 °C, 2 min; denaturation: 94 °C, 30 s; 35 cycles) using oligonucleotides BK4 or DF4 (position 1–20) as 5'-primers and as 3'-primers the 3'-anchor primer and M8-ko (position 926–905), respectively (for primer positions see Fig. 3A). The nucleotide sequences of primers BK4 and DF4 were deduced from partial genomic cea5 sequences (26). The products were subcloned into pUC18 after blunt-ending (SureClone ligation kit, Pharmacia Biotech Inc.).

DNA fragments were subcloned into pBluescript. Genomic and cdNA fragments were sequenced as double-stranded templates on both strands with flanking primers to the multiple cloning site region of the vectors. The specific activity of the β-actin probe was reduced to 1/10 by addition of unlabeled UTP to compensate for the differences in abundance of β-actin and Cea5 mRNA. Hybridization and RNA isolation procedures were performed following the manufacturer's protocol. The protected probe fragments were separated on 8% urea, 6% polyacrylamide gels, which were washed with 10% ethanol, 10% acetic acid and then in 10% ethanol, dried, and exposed to Kodak XAR films.

Northern blot analyses were performed using standard procedures. For detection of Cea5 mRNA, the partial 1.3-kb Cea5 cdNA fragment starting at position 103 and covering the N and A domain and the 3'-untranslated region (Fig. 3A) was excised from the pUC18 plasmid (see below) with XbaI and KpnI. As an rnCGM2 probe, a 289-bp HindIII/BamHI genomic DNA fragment comprising about 50 bp of intron 1 and two-thirds of the N domain exon (exon 2) was used (39). Both fragments were labeled with [α-32P]dATP (MegaPrime™ DNA labeling system, Amersham Corp.). The final wash was performed at high stringency in 0.1 × SSPE (1 × SSPE: 180 mM NaCl, 10 mM sodium phosphate, pH 7.4, 1 mM EDTA), 0.1% sodium dodecyl sulfate (SDS) twice for 20 min at 65 °C. The filters were rehybridized with a 32P-labeled 1.2-kb mouse β-actin probe (43) to check for equal loading and the integrity of the RNAs.

Sequence Determination and Analysis—Suitable DNA fragments were subcloned into pBluescript. Genomic and cdNA fragments were sequenced as double-stranded templates on both strands with flanking universal and internal oligonucleotides as primers using a T7 polymerase sequencing kit (Pharmacia). The program PC-GENE (IntelliGenetics) was used to calculate the degree of similarity of the amino acid sequences after optimal alignment.

In situ Hybridization—In situ hybridization analyses were performed essentially as described before (44), except that the RNA probes were not fragmented by hydrolysis. Embryos and placenta of day 9.5 of development were embedded in toto in Tissue Freezing Medium (Jung) diluted 1:1 with water by freezing in ethanol/dry ice. The tissues were cryosectioned (7-μm sections) and placed on SuperFrost/Plus microscopical slides (Robo). RNA probes were synthesized in the presence of digoxigenin-labeled UTP (DIG Labeling Kit; Boehringer Mannheim) after subcloning of a 0.6-kb EcoRI Cea5 cdNA fragment (EcoRI cuts in the 3'-untranslated region as well as in the oligo(dT) primer used for cdNA synthesis) into the pBluescript vector and linearization of the plasmid with either XhoI (antisense probe) or BamHI (sense probe) using T7 and T3 RNA polymerase, respectively.

2 W. Zimmermann, B. Fischer, A. Olsen, P. Nédellec, and J. Thompson, submitted for publication.
RESULTS

Expression Pattern of Cea5—To be able to clone Cea5 cDNA, we identified Cea5 mRNA-containing tissues. To this end, we used an RNase protection assay, which allows the specific detection of Cea5 mRNA in the presence of mRNAs of other closely related CEA family members. A complementary $^{32}$P-labeled RNA of 418 nucleotides was hybridized with total RNA from various organs of adult BALB/c mice, including those known to express other members of the murine CEA family, like colon, small intestine, and bone marrow. The probe was derived from the N domain exon (360 nucleotides) of the previously isolated cea5 gene (26) and is flanked by sequences from the multiple cloning site of the vector which was used as template for probe synthesis. The expected 360-nucleotide protected fragment could not be detected with RNA from any of the adult tissues tested (Fig. 1A, lanes 1–8; Fig. 1B, lanes 3–7; Fig. 1C, lane 15). The intactness of the RNAs was verified using a $^{32}$P-labeled $\beta$-actin antisense RNA probe which yielded a protected fragment with an expected size of 250 nucleotides in all samples (Fig. 1A, lanes 11–18; Fig. 1B, lanes 10–14; Fig. 1C, lane 3–15). Since the CEA-related PSGs are expressed in placentae, we also assayed total RNA from placentae of different developmental stages for the presence of Cea5 mRNA. Indeed, the protected Cea5 probe fragment of the expected size was obtained with RNA from placentae of early development. The largest amounts were found in placentae of day 10.5 of embryonic development (E10.5). The steady state levels of Cea5 mRNA declined continuously during pregnancy and were barely detectable at E15.5 (Fig. 1C, lanes 3, 5, 7, 9, 11, and 13). On the other hand, RNAs from maternal decidual and uterine tissues as well as from day 13.5 fetuses were negative (Fig. 1C, lanes 4, 6, 8, 10, 12, and 14).

Expression of rnCGM2, the Rat Cea5 Counterpart—We also investigated the expression pattern of rnCGM2 (39), the presumed cea5 homolog of the rat, to assess the functional significance of cea5 expression in placenta as well as the orthologous relationship of the two genes. Northern blot analysis using part of exon 2 of the rnCGM2 gene as a probe revealed that rnCGM2 exhibits an expression pattern very similar to that found for cea5. An rnCGM2 mRNA species of about 1.1 kb was detected predominantly in E12.5 and smaller amounts in E10.5 and E17.5 rat placentae (Fig. 2, lanes 2, 5, and 7). Fainter signals (relative to the $\beta$-actin content) were seen with total RNA from...
E12.5 and E17.5 deciduae (Fig. 2, lanes 3 and 6). No rncGM2 mRNAs were found in E12.5 embryos nor in adult tissues (colon, small intestine, stomach, lung, spleen, testis, and skeletal muscle) tested (Fig. 2, lanes 4, and 8–14).

Cloning of Cea5 cDNA—Initial experiments failed to amplify Cea5 mRNA by polymerase chain reaction (PCR) from E10.5 placental RNA after reverse transcription (RT) using the Cea5-specific primer BK3 (Fig. 4A) from the putative 5'-untranslated region of cea5. We, therefore, employed a Cea5-specific primer (BK4) from the 5'-end of the N exon (Fig. 3A) and a (dT)14-containing 3'-anchor primer to amplify Cea5 cDNA. A DNA fragment of about 1.3 kb was obtained and sequenced after subcloning.

To amplify the complete coding region of the Cea5 mRNA, an oligonucleotide (M8-ko) was synthesized based on the sequence of the 3'-untranslated region of the partial Cea5 cDNA clone and used together with an oligonucleotide (DF4; Fig. 3A) located closer to the putative translational start of cea5 than BK3. The composite Cea5 cDNA sequence is shown in Fig. 3A.

Cea5 mRNA codes for a putative mature protein (after removal of the leader peptide of 33 amino acids) of 201 amino acids (Mr 52,225) and is composed of one IgV-like N domain and one IgC-like A domain (Fig. 3, A and B). No hydrophobic domain capable of membrane anchorage is present. The amino acid sequence of the A domain is most closely related to the A2 domains of the murine Bgps (e.g. Bgp1 A2: 57% identity) and less to the A domains of the murine PSGs (e.g. Cea2 A: 48% identity). The least similarity (30% identity) is observed with the A1 domain of Bgp1 (Fig. 3D). The Cea5 protein is predicted to contain two potential N-glycosylation sites which would re-
Murine CEA Gene Family Member cea5

31373

FIG. 4. Comparison of the transcriptional start site sequence (A) and the amino acid sequence of the N domain exon (B) of cea5 and other rodent CEA gene family members. A, the nucleotide sequences (numbered at the right margin) upstream of the boxed start codon of cea5 and cea10 (31) were aligned. Dashes were inserted to allow for optimal alignment. Identical nucleotides are indicated by asterisks. The transcriptional start sites of cea5 are marked by an arrow. The region homologous to the footprint 1 of the human CEA gene promoter (52) is underlined in the cea5 nucleotide sequence. The consensus binding sequence for the transcription factor, upstream regulatory factor (USF) is underlined in the cea10 sequence. The positions of two oligonucleotides used for the amplification of Cea5 cDNA are marked. Since Cea5 cDNA could be amplified by RT-PCR using the DF4 but not with the BK3 oligonucleotide as 5’-primer, additional transcriptional start site(s) must be located between the positions of the two oligonucleotides. B, the amino acid sequences in one-letter code (numbered at the right margin) deduced from the exon 2 nucleotide sequences of cea5 and its rat homolog rncGM2 (39) were aligned and identical amino acids marked by asterisks. The presumed start of the mature protein is indicated by an arrow. The amino acids (R and D) possibly involved in formation of the salt bridge are marked by a bracket. The non-conservative amino acid changes in the Ce5 sequence in the close neighborhood of the salt bridge are indicated by + and − above the sequence.

result in an M, of 28,000–30,000 assuming a mean increase in the molecular mass of 3,000–4,000 per carbohydrate addition site (45). The major transcriptional start site of cea5 was determined by PCR (5’-RACE) and found to be located at position 30 of the Cea5 cDNA (as estimated by the frequency of the 5’-RACE clones obtained starting at this site), only 12 nucleotides upstream of the putative translational start site of Cea5 (Fig. 3A). This unusually short 5’-untranslated region is caused by a deletion as evidenced by comparison with the corresponding region of cea10, another member of the murine CEA gene family (Fig. 4A; 31). A minor start site must be located in the vicinity of the region between the BK3 and DF4 primers used for PCR amplification of Cea5 cDNA (Fig. 4A), since only DF4 yielded an amplification product. Indeed, transcriptional start sites of members of the human CEA gene family (e.g. CEA and NCA) have been mapped to the corresponding regions (26). A prominent Cea5 mRNA species with a size of 1.6 kb migrating slightly below the 18 S rRNA could be detected in placenta using the 1.3-kb cDNA fragment as a probe (Fig. 3C). An additional faint hybridization signal corresponding to a 3.5-kb RNA species migrating between the 28 S and 18 S rRNAs was observed. As expected from the Rnase protection experiments (Fig. 1C), the steady state levels decline during placental development (Fig. 3C). No cross-reaction was observed with PSG mRNAs, which are increasingly abundant during later stages of development.

Cea5 Is Expressed in Trophoblast Giant Cells of the Placenta—To identify the cell type that expresses Cea5 mRNA, we performed in situ hybridization analyses on uteri of BALB/c mice at day 9.5 of gestation. Only trophoblast cells were labeled when an antisense probe derived from the 3’-untranslated region of the Cea5 mRNA was used (Fig. 5A). Both primary and secondary trophoblast giant cells, which are readily identifiable by their large nuclei as well as their location, were stained (Fig. 5, A and B). Primary trophoblast giant cells are located distally to the region of invasive trophoblast growth, and secondary giant cells are found more proximal to this region (46). Besides these cells, a trophoblastic cell population with smaller nuclei directly underneath the layer of secondary trophoblast giant cells was also labeled (Fig. 5B). In contrast, fetal spongiotrophoblast cells were negative (Fig. 5B). The embryo, the maternal part of the placenta (decidua), and other maternal tissues did not exhibit any hybridization signal. No signal was obtained with a corresponding sense probe (Fig. 5C).

FIG. 5. Expression of cea5 in trophoblast cells. Unfixed uteri of pregnant BALB/c mice (day 9.5 of development) were cryosectioned and hybridized in situ with a digoxigenin-labeled antisense (A and B) or with a corresponding sense Cea5 RNA probe as a negative control (C). The hybridized RNA probe was visualized by reaction with an antidigoxigenin alkaline phosphatase-conjugated antibody and subsequent incubation with 5-bromo-4-chloro-3-indolyl phosphate and nitro blue tetrazolium salt which results in a dark blue staining of Cea5 mRNA-containing cells. A, only fetal primary (pgi) and secondary trophoblast giant cells (agi; arrowheads) and the spongiotrophoblast (sp) is visible. They are characterized by their smaller nuclei (arrows). C, no staining was seen on a near-adjacent section with a sense probe. Scale bars, 1 mm (in A); 0.4 mm (in B and C).
therefore, searched for the presence of potential mRNA splice variants of \textit{cea5} in placental RNA by RT-PCR using various oligonucleotides from the N and A domain region in combination with the oligo(dT) anchor primer. No additional PCR products to the one described above were found (data not shown).

To analyze whether \textit{cea5} is closely linked to other members of the CEA gene family we screened a BALB/c cosmid library for overlapping clones using the N domain exon sequence as a probe. One clone (14.23) with an insert of 41 kb was identified and shown to extend 20 kb further upstream (Fig. 6A). Digestion of 14.23 DNA with \textit{Bam}HI, \textit{Eco}RI, \textit{Hin}dIII, and \textit{Sst}I, respectively, and subsequent hybridization with the exon 2 \textit{Cea5} probe (see above) under non-stringent conditions did not reveal additional N domain-related exons (not shown). Therefore, no \textit{Cea}-related gene seems to be present in the 5'-vicinity of \textit{cea5}.

\textbf{DISCUSSION}

\textit{Unique Structure and Expression Pattern of Cea5—}
The murine \textit{Cea5}, described here, represents a soluble and possibly secreted member of the CEA family, exhibiting a novel domain organization (L-N-A), thus far not found for other members of the \textit{Cea5}, \textit{Bgp1}, \textit{Bgp2}, \textit{Cea2}, and \textit{Cea6} and human CEA families (\textit{BgpA} and \textit{g}) were aligned. \textit{BgpA} and \textit{BgpG} represent different splice variants of the \textit{Bgp} gene. Amino acids identical to the ones found in \textit{Cea5} are underlined. Dashes were introduced to allow optimal alignment. Exon and intron sequences are shown in \textit{capital} and \textit{lowercase} letters, respectively. Potential and actually used splice donor sites are boxed. The following sequence sources were used: \textit{Bgp1a}, \textit{Bgp2} (21, 34), \textit{BgpG} (W211; Ref. 47), \textit{BgpA} (36), \textit{Cea2}, and \textit{Cea6} (26).

\begin{table}[h]
\centering
\begin{tabular}{llllll}
\hline
Name & Size & Splice donor & Size & Splice acceptor & Size Name \\
\hline
5'-UTL 76 & bp & & & & bp \\
\textit{CtgCtCACAG} & 1650 & \textit{Gatatttattagtttttctctcactag} & \textit{CCTTCTCTTTA} & \textit{360 N} & \textit{labheLeuLeu} \\
\textit{Graagtgtgtcgtgacctcctggg} & 950 & \textit{atttctcgtaccttccctctcag} & \textit{AGCCCGGTGCGA} & \textit{920 A3'-UT} & \textit{luProValAla} \\
\hline
\end{tabular}
\caption{Exon/intron organization of the BALB/c \textit{cea5} gene}
\end{table}

The critical bases of the splice donor and splice acceptor sequences are in italic. 5'-3',5'-3'-untranslated region. For other abbreviations see legend to Fig. 3. The size of the unknown intron was determined by PCR using oligonucleotides from the N and A domain exons as primers and \textit{cosC3} as template.
the murine, rat, or human CEA families. Northern blot and RT-PCR analyses indicate that only one major Ceae5 mRNA species exists. However, splicing might occur at the 3'-end of the A domain where a potential splice donor site is present, similar to the ones that are functional in corresponding exons of the human and murine BGP/bgp genes. The 3.5-kb RNA species found to hybridize with the Ceae5 probe in day 10.5 placenta might represent such an alternative splice variant. However, we were unable to demonstrate the existence of such a splice variant by RT-PCR. Sequence comparison of the IgV- and IgC-related domain (A type) of Ceae5 with corresponding domains of members of the CEA and PSG subgroups does not allow unequivocal assignment to either of these groups. Determination of the mutational rates of silent sites (third positions of most codons) in the coding sequences (48) of members of the murine CEA family revealed that ceae5 and rnCGM2 are old genes, the exons of which diverged from bgp and from the PSG subgroup ancestors more than 65 million years ago, around the time of primate and rodent separation. The expansion of the CEA and PSG subgroups started some 20 million years later (33).

The slightly higher similarity of the Ceae5 A domain to the A2 domains of the murine bgp genes, bgp1 and bgp2, would suggest assignment to the CEA subgroup. However, no murine or human CEA subgroup member is exclusively expressed in placenta as is ceae5. This is characteristic for members of the PSG subgroup of both species (33, 41). In contrast to PSGs, Ceae5 mRNA levels decrease during placental development. This is also found for cea10 which codes for a soluble member of the murine CEA subgroup and is closely related to bgp1 and bgp2 (31). This decline of expression follows the disappearance of the trophoblast giant cells located at the maternal/fetal interphase (49) which indicates expression of ceae5 (and cea10) in this cell population. Indeed, in situ hybridization studies confirm expression of ceae5 in primary and secondary trophoblast giant cells and in a trophoblast cell population with smaller nuclei located in the vicinity of the secondary trophoblast giant cells. Thus, Ceae5 mRNA complements the set of known markers of trophoblast development (50). In contrast, rodent PSGs are expressed in the secondary but not in the primary trophoblast cells and additionally in spongiospermatoblast cells (40). In contrast to ceae5, cea10 is also transcribed in the gastrointestinal tract and bone marrow, which might be due to the presence of an upstream stimulatory factor (USF) binding motif (CACGCTG). This motif is present in all members of the murine CEA family that are expressed in intestinal epithelia and has been shown to bind USF and to be important for the expression of the human CEA and BGP genes (51, 52). There is no conserved USF motif in the homologous promoter region of ceae5 (Fig. 4A), which supports the hypothesis that USF is involved in directing expression to intestinal epithelial cells (31) but is not necessary for trophoblast expression.

The chromosomal location of ceae5 also stresses its uniqueness within the murine and human CEA gene families. It is displaced several centimorgans, probably as a single gene (no CEA-related gene has been identified in the immediate neighborhood (Fig. 6A)) from the region where other CEA subgroup members (cea10/bgp1) reside on the centromere-proximal region of mouse chromosome 7 (53). The latter region is syntenic to the human CEA/PSG locus at 19q13.2, where all gene family members are clustered within a region of about 1 mega-base pair (54). It is interesting to note, that in mice the PSG gene cluster also became detached from the CEA subgroup (53). Taken together, ceae5 seems to represent a separate entity, belonging to neither the CEA nor the PSG subgroup.

Function of Ceae5—The function of Ceae5 is still unclear. It probably differs, however, from that of the placentaly expressed PSGs for two reasons. First, there is no RGD-like motif in the N domain of Ceae5 (33), and second, Ceae5 exhibits inverse expression kinetics and reveals a different expression pattern with regard to the cell population during placental development (Fig. 1C and Fig. 5 and Refs. 40 and 41). Ceae5-like molecules seem to be important for both mouse and rat based on the conservation of the N domain amino acid sequence in the rat Ceae5 homolog, rnCGM2 (39), which has 93% of its amino acid sequence in common with Ceae5 (26). This extremely high conservation is unparalleled in other rat and murine CEA family members. In comparison, Bgp1/Bgp1β and their rat homologs C-CAM1/C-CAM1β share only 49–54% of their N domain amino acid sequences. Interestingly, the few amino acid substitutions observed in Ceae5 and rnCGM2 represent non-conservative replacements of uncharged amino acids by charged ones. However, they probably do not lead to major structural changes because they seem to be neutralized by an opposite charge in their presumed direct spatial neighborhood (Fig. 4B). The extremely high conservation of Ceae5 and rnCGM2 together with their apparent rodent specificity and their similar expression patterns indicates a function that is exclusively needed during early placental development in rodents. Indeed, a number of structural and functional differences exist between rodent and primate placentae, necessitating differences at the molecular level. The study of placentaly expressed members of evolutionary young gene families, like the CEA gene family, is expected to further our understanding of the molecular mechanisms involved in the establishment of structural and functional differences in mammal placentation.

Acknowledgments—We thank Sabine von Kleist for continuous support and Haymo Kurz (Institute of Anatomy, University of Freiburg) for helpful discussions. The expert technical assistance by Beate Fischer and Martina Weiss and the photographic work by Hans-Joachim Grammel are gratefully acknowledged.

REFERENCES
1. Shively, J. E., and Beatty, J. D. (1985) Crit. Rev. Oncol. Hematol. 2, 355–399
2. Thompson, J. A., Grunert, F., and Zimmermann, W. (1991) J. Clin. Lab. Anal. 5, 344–366
3. Tegund, S., Olsen, A., Khan, W. N., Frangsmyr, L., and Hammarstrom, S. (1994) Genomics 23, 669–684
4. Thompson, J. A. (1995) Tumor Biol. 16, 10–16
5. Olsen, A., Tegund, S., Nelson, D., Gordon, L., Copeland, A., Georgescu, A., Carrano, A., and Hammarstrom, S. (1994) Genomics 23, 659–668
6. Benchimol, S., Fuks, A., Joby, B., Beaufachmin, N., Shirota, K., and Stanners, C. P. (1998) Cell 97, 527–334
7. Okawa, S., Inuzuka, C., Kuroki, M., Arakawa, F., Matsuoka, Y., Kosaki, G., and Nakazato, H. (1998) J. Clin. Invest. 102, 7995–8001
8. Kuipers, T. W., Hoogerwerf, M., van der Laan, L. W., Nagel, G., van der Schout, C., Grunert, F., and Roos, D. (1992) J. Cell Biol. 118, 457–466
9. Kuipers, T. W., van der Schoot, C. E., Hoogerwerf, M., and Roos, D. (1993) J. Immunol. 151, 4934–4940
10. Stoeck, S. C., and Kerr, M. A. (1992) Biochem. J. 288, 23–27
11. Screaton, R. A., Penn, L. Z., and Stanners, C. P. (1997) J. Cell Biol. 137, 939–952
12. Neumaier, M., Paululat, S., Chan, A., Mattheas, P., and Wagener, C. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 10744–10748
13. Rosenberg, M., Nedelec, P., Joby, B., Platenier, D., Tubide, C., and Beaufachmin, N. (1993) Cancer Res. 53, 4938–4945
14. Kleinerman, D. I., Troncoso, P., Lin, S.-H., Pisters, L. L., Sherwood, E. R., Brooks, W. E., and Eschenbach, A. C. (1993) Cancer Res. 53, 1215–1220
15. Thompson, J., Seitz, M., Chastre, E., Ditter, M., Aldrian, C., Gespach, C., and Zimmermann, W. (1997) Cancer Res. 57, 1756–1764
16. Haeh, J.-T., Luo, W., Song, W., Wang, Y., Kleinerman, D. I., Van N. T. N., and Lin, S.-H. (1995) Cancer Res. 55, 190–197
17. Turbide, C., Kunath, T., Danies, E., and Beaufachmin, N. (1997) Cancer Res. 57, 2781–2788
18. Luo, W. P., Wood, C. G., Earley, K., Hung, M. C., and Lin, S. H. (1997) Oncogene 14, 1697–1704
19. Rutherford, K. J., Chou, Y. J., and Mansfield, W. B. (1995) Mol. Endocrinol. 9, 1297–1305
20. Dveksler, G. S., Pensiero, M. N., Cardellichio, C. B., Williams, R. K., Jiang, G.-S., Holmes, K. V., and Deifenebach, C. W. (1991) J. Viral. 65, 6681–6691
21. Nedelec, P., Dveksler, G. S., Danies, E., Turbide, C., Chow, B., Basile, A. A., Holmes, K. V., and Beaufachmin, N. (1994) J. Viral. 68, 4525–4537
22. Lin, S.-H., and Guidotti, G. (1989) J. Biol. Chem. 264, 14408–14411
23. Rehstock, S., Lucas, K., Thompson, J. A., and Zimmermann, W. (1990) J. Biol.
24. Turbide, C., Rojas, M., Stanners, C. P., and Beauchemin, N. (1991) *J. Biol. Chem.* 266, 309–315
25. McCuig, K., Turbide, C., and Beauchemin, N. (1992) *Cell Growth Differ.* 3, 165–174
26. Rudert, F., Saunders, A. M., Rebstock, S., Thompson, J. A., and Zimmermann, W. (1992) *Mamm. Genome* 3, 262–273
27. Chen, H., Plouzek, C. A., Liu, J.-L., Chen, C.-L., and Chou, J. Y. (1992) *DNA Cell Biol.* 11, 139–148
28. Edlund, M., Gahrnsvoll, H., Bock, E., and Öbrink, B. (1993) *Eur. J. Biochem.* 213, 1109–1116
29. Chen, H., Chen, C.-L., and Chou, J. Y. (1994) *Biochemistry* 33, 9615–9626
30. Chen, D. S., Asanake, M., Yokomori, K., Wang, F., Hwang, S. B., Li, H. P., and Lai, M. M. C. (1994) *J. Biol. Chem.* 269, 3202–3207
31. Kromer, B., Finkenzeller, D., Wessels, J., Dveksler, G., Thompson, J., and Zimmermann, W. (1995) *Eur. J. Biochem.* 229, 455–464
32. Kuroki, M., Arakawa, F., Matsuo, Y., Oikawa, S., Nakazato, H., and Matsuzaka, Y. (1996) *Biochem. J.* 315, 799–806
33. Zimmermann, W. (1997) *Cell Adhesion and Communication Mediated by the CEA Family: Basic and Clinical Perspectives* (Stanners, C. P., ed.) Harwood Academic Publishers, Amsterdam, in press
34. Nedellec, P., Turbide, C., and Beauchemin, N. (1995) *Eur. J. Biochem.* 231, 104–114
35. Barnett, T. R., Kretschmer, A., Austen, D. A., Goebel, S. J., Hart, J. T., Elting, J. J., and Kamarck, M. E. (1989) *J. Cell Biol.* 108, 267–276
36. Barnett, T. R., Drake, L., and Pickle, W., II (1993) *Mol. Cell. Biol.* 13, 1273–1282
37. Najjar, S. M., Accili, D., Philippe, N., Margolis, R., and Taylor, S. I. (1993) *J. Biol. Chem.* 268, 1201–1206
38. Chomczynski, P., and Sacchi, N. (1987) *Anal. Biochem.* 162, 156–159
39. Kodelja, V., Lucas, K., Barnert, S., von Kleist, S., Thompson, J. A., and Zimmermann, W. (1989) *J. Biol. Chem.* 264, 6906–6912
40. Rebstock, S., Lucas, K., Weiss, M., Thompson, J., and Zimmermann, W. (1993) *Dev. Dyn.* 198, 171–181
41. Kuroki, M., Arakawa, F., Wang, W., Wessels, J., Dveksler, G., Thompson, J., and Zimmermann, W. (1996) *Eur. J. Biochem.* 242, 280–287
42. Chomczynski, P., and Sacchi, N. (1983) *Anal. Biochem.* 162, 156–159
43. Alonso, S. A., Minty, A., Bourlet, Y., and Buckingham, M. (1986) *J. Mol. Biol.* 23, 11–22
44. Schaeren-Wiemers, N., and Gerfin-Moser, A. (1993) *Histochemistry* 100, 431–440
45. Neumaier, M., Zimmermann, W., Shively, L., Hinoda, Y., Rigs, A. D., and Shively, J. E. (1988) *J. Biol. Chem.* 263, 3202–3207
46. Cross, J. C., Werb, Z., and Fisher, S. J. (1994) *Science* 266, 1508–1518
47. Kuroki, M., Arakawa, F., Matsuo, Y., Oikawa, S., Nakazato, H., and Matsuzaka, Y. (1991) *Biochem. Biophys. Res. Commun.* 176, 578–585
48. Li, W.-H., Wu, C.-I., and Luo, C.-C. (1985) *J. Biol. Chem.* 260, 12095–12099
49. Kaufman, M. H. (1992) *The Atlas of Mouse Development*, pp. 469–477, Academic Press, London
50. Carney, E. W., Prideaux, V., Lye, S. J., and Rossant, J. (1993) *Mol. Reprod. Dev.* 34, 357–368
51. Hauck, W., Nedellec, P., Turbide, C., Stanners, C. P., Barnett, T. R., and Beauchemin, N. (1994) *Eur. J. Biochem.* 223, 529–541
52. Hauck, W., and Stanners, C. P. (1995) *J. Biol. Chem.* 270, 3602–3610
53. Stubbs, L., Carver, E. A., Shinnick, M. E., Kim, J., Geisler, J., Generoso, E. E., Stanford, B. G., Dunn, W. C., Mohrenweiser, H., Zimmermann, W., Watt, S. M., and Ashworth, L. K. (1996) *Genomics* 35, 499–508
54. 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) *Genomics* 12, 773–779