

Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray

Tetsuya S. Tanaka*, Saied A. Jaradat*, Meng K. Lim*, George J. Kargul*, Xiaohong Wang*, Marija J. Grahovac*, Serafino Pantano*, Yuri Sano*, Yulan Piao*, Ramaiah Nagaraja*, Hirofumi Doi†, William H. Wood III‡, Kevin G. Becker‡, and Minoru S. H. Ko*[§]

*Laboratory of Genetics and †DNA Array Unit, National Institute on Aging, National Institutes of Health, Baltimore, MD 21224-6820; and ‡Exploratory Research for Advanced Technology (ERATO) Doi Bioasymmetry Project, Japan Science and Technology Corporation (JST), Mihama-ku, Chiba 261-7112, Japan

Communicated by Melvin I. Simon, California Institute of Technology, Pasadena, CA, May 18, 2000 (received for review April 4, 2000)

cDNA microarray technology has been increasingly used to monitor global gene expression patterns in various tissues and cell types. However, applications to mammalian development have been hampered by the lack of appropriate cDNA collections, particularly for early developmental stages. To overcome this problem, a PCR-based cDNA library construction method was used to derive 52,374 expressed sequence tags from pre- and peri-implantation embryos, embryonic day (E) 12.5 female gonad/mesonephros, and newborn ovary. From these cDNA collections, a microarray representing 15,264 unique genes (78% novel and 22% known) was assembled. In initial applications, the divergence of placental and embryonic gene expression profiles was assessed. At stage E12.5 of development, based on triplicate experiments, 720 genes (6.5%) displayed statistically significant differences in expression between placenta and embryo. Among 289 more highly expressed in placenta, 61 placenta-specific genes encoded, for example, a novel prolactin-like protein. The number of genes highly expressed (and frequently specific) for placenta has thereby been increased 5-fold over the total previously reported, illustrating the potential of the microarrays for tissue-specific gene discovery and analysis of mammalian developmental programs.

Systematic studies of gene expression patterns by using cDNA microarrays provide a powerful approach to molecular dissection of cells and tissues by comparing expression levels of tens of thousands of genes at a time (reviewed in refs. 1 and 2). Successful applications have been done with the full gene cohort of yeast (3–6). In mammalian systems, analyses of human fibroblast activation (7) and molecular classification of tumors (8–14) have successfully built on the use of a relatively small microarray containing mainly well-characterized genes. In those instances, the absence of particular genes in the microarray does not affect overall conclusions.

In contrast, inclusion of as many genes as possible in microarrays is desirable for exploratory approaches in which critical genes for biological processes are sought without prior knowledge (15). The problem of maximum coverage is notable for studies of early development, because a recent mouse embryo expressed sequence tag (EST) project (16, 17) has revealed that a significant fraction of the genes expressed in early life are not represented in the collections of genes previously available. We have therefore been assembling a cohort of developmentally active genes to supplement microarray efforts.

Mid-gestation placenta was chosen for a test of the utility of the microarray for several reasons. Methodologically, the placenta can be cleanly separated from the embryo proper, and substantial amounts of RNA can be easily recovered. Biologically, studies of human pregnancy loss and of mouse mutants that render embryos nonviable have motivated widespread interest in the placenta as the indispensable organ for embryo survival (18–20). Furthermore, the extraembryonic lineage

that develops into placenta is the first cell type to differentiate from fertilized eggs, and it has therefore been studied as the first cell differentiation that is unique to mammalian species. Yet the number of genes identified as placenta-specific (or highly expressed in placenta) is very limited, and the mechanisms of placenta development and maintenance are still largely unknown. The assembly and use of the microarray to identify and profile genes expressed in early placenta and embryo are reported here.

Materials and Methods

cDNA Libraries, EST Collections, and Assembly of Unique Gene Set. Procedures for cDNA library construction and EST projects are essentially the same as described previously (16, 17). Sequence information is available at dbEST (<http://www.ncbi.nlm.nih.gov/>) and at the National Institute on Aging web site (<http://lgsun.grc.nia.nih.gov/>). All cDNA clones are available from American Type Culture Collection (ATCC; <http://www.atcc.org/>). Selection of the unique gene set was done by all-against-all similarity searches with BLAST 1.4 algorithm (21) with a cut-off score of 800 to group ESTs into unique genes. A total of 15,264 unique genes were rearranged in 96-well microtiter plates.

Preparation of Microarray. Plasmid DNAs were prepared in 96-well format (R.E.A.L. Prep 96 plasmid kit; Qiagen, Chatsworth, CA). To amplify cDNA inserts, PCR primers (5'-CCAGTC-ACGACGTTGTAAAACGAC-3'; 5'-GTGTGGAATTGT-GAGCGGATAACAA-3') were designed from sequences flanking the multiple cloning site of pSPORT1 plasmid vector (Life Technologies, Rockville, MD), which was used for cDNA cloning. For each cDNA clone, about 10 ng of plasmid DNA was used in 100 μ l PCR reaction containing 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM MgCl₂, 0.5 μ M each primer, 0.8 mM dNTPs, and 3.25 units of *Taq* polymerase (PE Biosystems, Foster City, CA). Amplification was done in a Peltier Thermal Cycler (MJ Research, Cambridge, MA) under the following conditions: initial denaturation at 95°C for 1 min; 38 cycles of denaturation at 94°C for 30 s, annealing at 65°C for 45 s, and extension at 72°C for 3.5 min; and a final extension at 72°C for 3 min. PCR products were precipitated with 50% isopropanol, washed in 70% ethanol, and resuspended in 6.5 μ l TE buffer (10 mM Tris/1 mM EDTA,

Abbreviations: EST, expressed sequence tag; *En*, embryonic day *n*; PRL, prolactin.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. AW537829–AW545916, AW536144–AW537732, AW545922–AW559162, and AF272368).

[§]To whom reprint requests should be addressed at: Laboratory of Genetics, National Institute on Aging, National Institutes of Health, 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820. E-mail: KoM@grc.nia.nih.gov.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

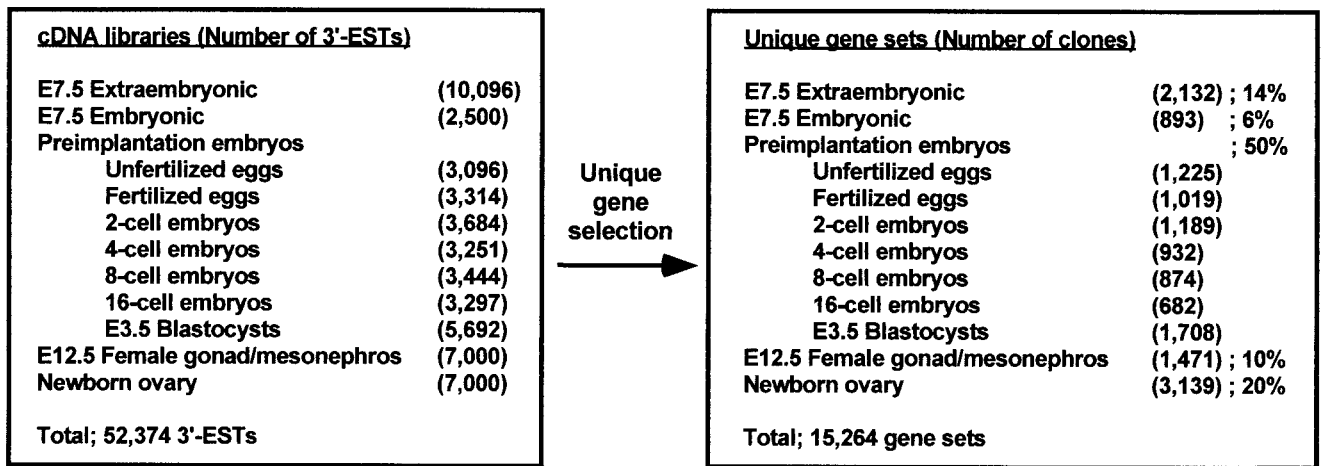


Fig. 1. Construction of 15K mouse cDNA microarray. cDNA libraries used for EST generations and the number of ESTs collected from each library are shown (Left). ESTs were clustered by their sequence similarity, and representative cDNA clones were selected from each gene cluster. Numbers and sources of cDNA clones in the 15,264 (15K) unique gene set are shown (Right).

pH 8.0). To examine the quality and quantity of PCR products, they were run on 1% agarose gels (Embi Tec, San Diego, CA). Greater than 90% of the clones gave single PCR products. PCR products were denatured with 0.1 M NaOH and spotted onto nylon membranes (Schleicher and Schüll, Dassel, Germany) by using a GMS417 Microarrayer (Genetic Microsystems, Woburn, MA). Genes (2,304) were spotted on each membrane with 2.5 × 7.5 cm dimension at 665 μm distance between spots. Diameter of the spotting pin was 300 μm.

Probe Preparation. Hybridization of the microarray with a ³³P-labeled 260-bp vector probe showed relatively uniform signal intensities across individual membranes. Placenta (including yolk sac and amnion) and embryo were dissected out from C57BL/6J at 12.5 days postcoitum [embryonic day (E) 12.5]. Maternal part of the placenta (i.e., decidua and uterus) was stripped away, and the conceptus-derived part of the placenta was used for experiments (see Fig. 2A). Total RNA was extracted by Trizol reagent (Life Technologies); 100–300 μg of total RNA was used as a template to incorporate [α-³³P]dCTP (Amersham Pharmacia, Piscataway, NJ) during cDNA synthesis using SuperScript II reverse transcriptase (SSTII; Life Technologies) with oligo(dT) primer (Amersham Pharmacia). Each labeling reaction contained 10–13 μg of a heat-denatured template RNA, 1× first strand reaction buffer, 0.01 M DTT, 0.5 mM each of dATP, dGTP, and dTTP (Amersham Pharmacia), 40 units RNasin (Boehringer Mannheim) and 50 μCi/mmol [α-³³P]dCTP. The reaction mixture was heated at 65°C for 10 min and then cooled down gradually to 42°C, where 400 units of SSTII was added for 2 h. After a second hour incubation with 400 units more enzyme, template RNA was degraded with 0.2 M NaOH at 65°C for 15 min, and the mixture was then neutralized with 0.3 M Tris-HCl (pH 7.5).

Microarray Hybridization and Data Analyses. Prehybridization was done at 65°C for 4 h with MicroHyb solution (Research Genetics, Huntsville, AL) supplemented with yeast tRNA (1 mg/ml; Sigma) and poly(A) RNA (1 mg/ml; Amersham Pharmacia). Then hybridization was done with purified probe 10% dextran sulfate (Sigma), and mouse Cot1 DNA (70 μg/ml; Life Technologies) added to the prehybridization solution at 65°C for overnight incubation. Membranes were washed twice in 2× SSC/0.1% SDS for 20 min at room

temperature, followed by two washings in 2× SSC/0.1% SDS at 65°C for 30 min. After a final wash with 0.1× SSC/1% SDS for 1 h at 65°C, membranes were exposed to Phosphorscreen (Amersham Pharmacia) for 10 days at room temperature; then the screen was scanned by Storm 860 (Amersham Pharmacia) at a maximum resolution of 50 μm/pixel. Results from three independent hybridizations were obtained for each probe. Images were analyzed by IMAGEQUANT 5.0 (Amersham Pharmacia). Each image was overlaid with grids so that signal intensities of individual spots could be assessed. Local background for each membrane was calculated on the basis of 18 positions with no DNA spotted area. Expression levels of individual genes are represented in arbitrary units after subtracting background. All of the expression data are available at the National Institute on Aging web site (<http://lgsun.grc.nia.nih.gov/array/data1.html>).

Gene Expression Analyses by Northern Blot and *in Situ* Hybridization. For Northern blot analysis, 20 μg of total RNA from E12.5 embryo and placenta were size fractionated on 1% formaldehyde agarose gel and transferred to nylon membranes (Oncor, Gaithersburg, MD). cDNA inserts from enzyme-digested plasmid DNAs were purified and used as probes for hybridization. *In situ* hybridization for one clone, H3018B06, was performed as described (22). Digoxigenin-labeled antisense probe was hybridized on the section of E13.5 embryos with placenta. As a marker for placenta-specific hybridization, 4311 (23) was also used as a probe.

Results and Discussion

Assembly of a 15K Mouse Developmental cDNA Microarray. To collect ESTs from early mouse embryos, a PCR-based cDNA library construction method was used to derive ESTs on a large scale from E7.5 extraembryonic tissue (16) and preimplantation embryos (17). Additional ESTs were generated for this study from E7.5 extraembryonic cDNA libraries (8,088 cDNAs, GenBank nos. AW537829–AW545916), E7.5 embryonic cDNA libraries (1,589 cDNAs, GenBank nos. AW536144–AW537732), and other libraries (13,241 cDNAs, GenBank nos. AW545922–AW559162) (Fig. 1). All of the ESTs were derived from the 3'-ends of cDNAs. All-against-all similarity searches by BLAST program (21) classified the total of 52,374 3'-ESTs into 15,264 unique genes, which were rearranged in a set of 96-well microtiter

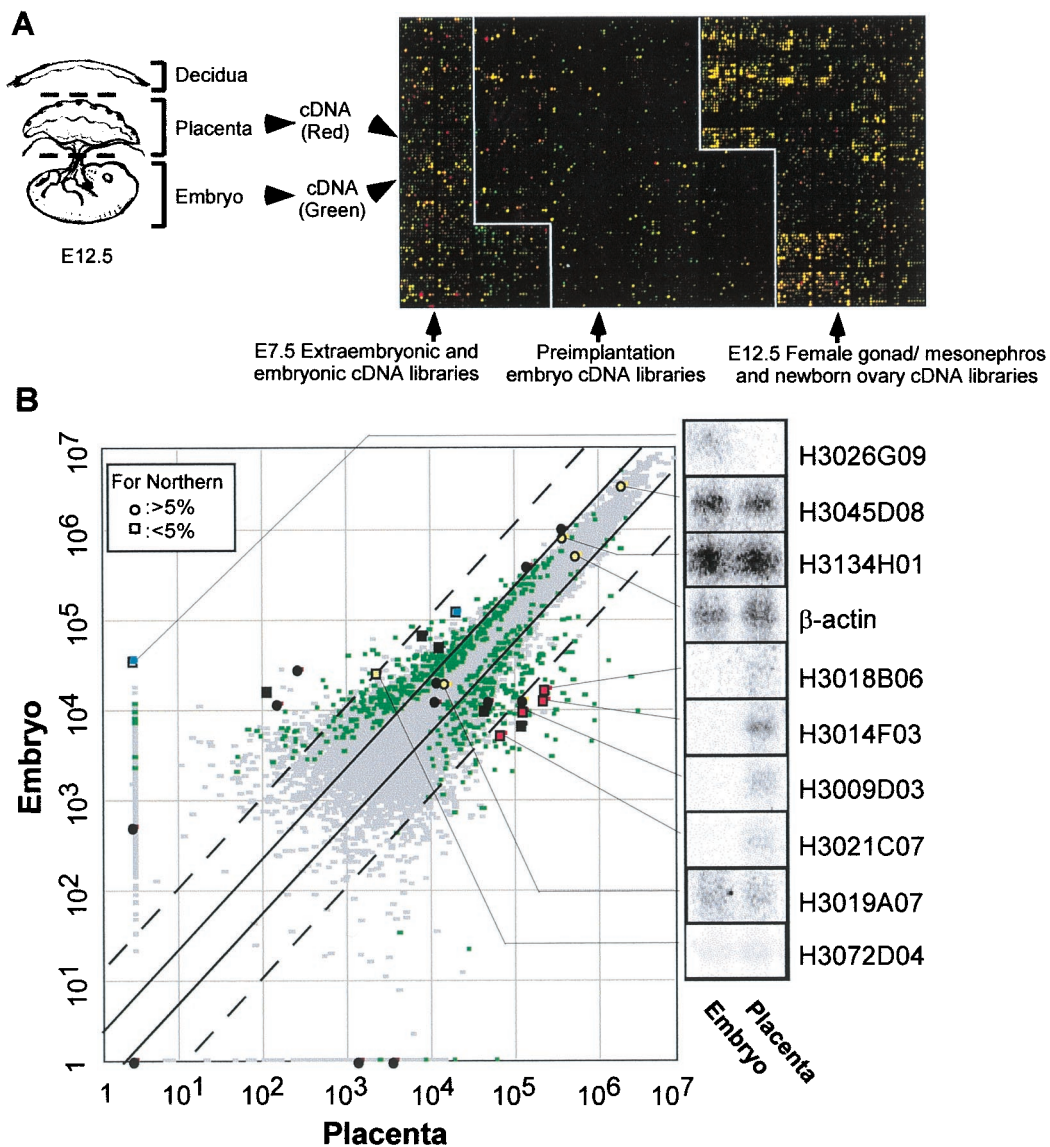


Fig. 2. Expression profiling of 15K genes between embryo and placenta at E12.5. (A) Schematic representation of experimental procedures and pseudocolored image for a typical hybridization. After removal of maternal tissues, embryo and placenta including yolk sac and amnion were dissected out and used for RNA extraction. Radiolabeled cDNAs from each sample were hybridized separately on the 15K microarray. Hybridization was repeated three times independently for each sample. Gene expression levels in placenta were represented as a gradation of red, and those in embryo were represented as a gradation of green. Genes expressed commonly in two samples appear yellow. (B) For each gene, average expression levels (in arbitrary units) were calculated from three independent hybridizations for placenta and embryo, respectively, and displayed on a scatter plot. Genes that show significantly different expression levels between embryo and placenta at the 5% significance level ($P < 0.05$) are displayed as green spots, and the other genes are displayed as gray spots. Solid lines indicate 2-fold expression differences between embryo and placenta; dotted lines indicate 10-fold expression differences. Results of Northern blot analyses of selected genes are shown with coordinates in the scatter plot. A total 28 clones were examined, and Northern blot for 10 clones are shown. Spots were color coded according to their expression patterns on Northern blot: placenta-specific expression (red), embryo-specific expression (blue), commonly expressed genes (yellow), and no detectable expression in both placenta and embryo (black). Genes that show statistically significant differences are represented as squares, and ones that do not are represented as circles.

plates (Fig. 1). This 15K clone set contains cDNA inserts of average 1.5 kb and has been sequenced from both 5'- and 3'-ends to verify clone identity and further characterize individual genes (G.J.K., D. B. Dudekula, Y. Qian, and M.S.H.K. unpublished work). About 3,400 represent known genes, suggesting that one-third to one-half of well-characterized mammalian genes are included. Because the majority of the cDNAs are derived from early mouse embryos and 78% are either completely novel (55%) or are similar but not identical to known genes (23%), the collection provides a unique source for studies of developmental biology and is complementary to other mouse/human EST cohorts.

Expression Profiling of Placenta and Embryo. The global gene expression pattern of E12.5 placenta was compared with that of E12.5 embryo (Fig. 2A). Signal intensities of individual genes were converted to color intensities and superimposed (Fig. 2A). A number of "placenta-specific" (red in Fig. 2A) and "embryo-specific" genes (green in Fig. 2A) are apparent. Rather than applying a current standard procedure, which directly uses ratios of values for each gene (or an arbitrary 2- or 10-fold expression difference) to characterize "differences" in gene expression levels in a single determination, three independent hybridization experiments were first performed for each probe. The multiple probings permit an analysis that satisfies conservative statistical

Table 1. List of genes expressed highly in placenta

Putative ID	Clone name	Expression ratio (PL/EM)	Accession no.	Putative ID	Clone name	Expression ratio (PL/EM)	Accession no.
Endocrine-related protein				Homeobox and transcription-related protein			
Secretin	H3001A02	200	X73580	eHAND	H3020C07	5	S79216
Placental lactogen I (PL-I)	H3008D11	5	M35662	Placenta and embryonic expression gene (pem)	H3027G05	33	NM_008818.1
Placental lactogen II (PL-II)	H3046B04	330	NM_008865.1	Placenta specific homeobox1 (PSX1)	H3120B02	12	NM_008955.1
Calcyclin	H3087H09	35	NM_011313.1	Cellular repressor of E1A-stimulated genes CREG*	H3037D10	50	AF084524
Prolactin-like protein-A precursor (PLP-A)	H3007C10	9	AF011383	Melanocyte-specific gene 1 (msg1)*	H3076B01	16	U65091
Prolactin-like protein G precursor	H3008F06	11	AF020525	Transcriptional repressor SIN3B (Sin3B)*	H3129D02	18	L38622
Prolactin-like protein homologous to the one of Opossum	H3018B06	16		Similar to rat FGF receptor activating protein FRAG1	H3036E12	8	U57715
Proliferin-related protein (PRP)	H3021D07	15	X02594	TBX1*	H3051H03	2	AF171100
Cathepsin L	H3028F04	3	NM_009984	Similar to rat HMGB-box containing protein 1 (Hbp1)*	H3054G12	>1000	NM_013221
Similar to human mRNA for TRAM-protein*	H3100D02	30	X63679				
Similar to human nuclear VCP-like protein NVLp.2 (NVL.2)	H3076C01	13	NM_002533	Others expressed in placenta			
Mammary-derived growth inhibitor (MDGI)*	H3104D07	47	U02883	LPS-binding protein	H3086G08	15	NM_008489.1
Similar to human HEM45 (ISG20)*	H3108A01	27	U88964	Pyruvate carboxylase	H3137A09	26	NM_008797.1
Epithelial protein				Newly defined genes			
Extra-embryonic endodermal cytokeratin type II (EndoA)	H3031C01	35	M21836	Serine protease inhibitor 6 (Spi6)	H3021C07	14	NM_009256.1
Endo B cytokeratin	H3104F03	20	M11686	Protein kinase inhibitor p58*	H3094A04	10	NM_008929.1
Band 3-related protein*	H3137D09	26	J04036	Similar to human unknown protein IT1*	H3137F09	34	AF040964
Similar to human tazarotene-induced gene 2 (TIG2)*	H3137E09	83	U77594	Uroporphyrinogen III synthase (UROS)*	H3137F11	20	U18867.1
Basic transcription element binding protein 2 (Bteb2)	H3102C04	3	NM_009769	Similar to human FKBP-associated protein (FAP48)*	H3101B02	>1000	NM_007070
Oxidative-, stress- and aging-related protein				Unknown			
25-kDa mammalian stress protein*	H3082D09	18	X14686		H3001A06	56	
Carbonic anhydrase IV*	H3095H12	>1000	NM_007607.1		H3003A12	20	
Putative CCR4 protein*	H3101A12	16	U70139		H3005A03	14	
Plasma glutathione peroxidase 3 (Gpx3)	H3125A09	41	NM_008161.1		H3005D09	11	
Cell adhesion- or invasion-related protein					H3009D03	15	
Ctla-2-alpha, proregion of cysteine protease*	H3003A10	114	X15591		H3016D10	13	
Ctla-2-beta, proregion of cysteine protease*	H3022D05	25	X15592		H3036G12	51	
Fat facets homolog (Fam)*	H3003A11	124	U67874		H3046B06	13	
Annexin XI (Anx11)*	H3014B09	12	U65986		H3049B11	13	
Transglutaminase (TGase)	H3137C06	13	M55154		H3050F11	>1000	
Ligatin (Lgtn)*	H3138C09	>1000	U58337		H3087H10	17	
					H3101F01	52	
					H3101G01	10	
					H3101H01	16	
					H3137A10	12	
					H3137C08	29	
					H3137C09	16	
					H3137E08	95	
					H3137F10	10	

*Presence of these genes in placenta has not been previously reported. PL, Average expression level of genes with placenta probes; EM, average expression level of genes with embryo probes.

criteria (24) and allows one to identify 720 genes that are expressed differentially between two samples at the 5% significance level by *t* test (P value < 0.05). These genes are shown as spots in a scatter plot of the average signal intensity of each gene (Fig. 2B).

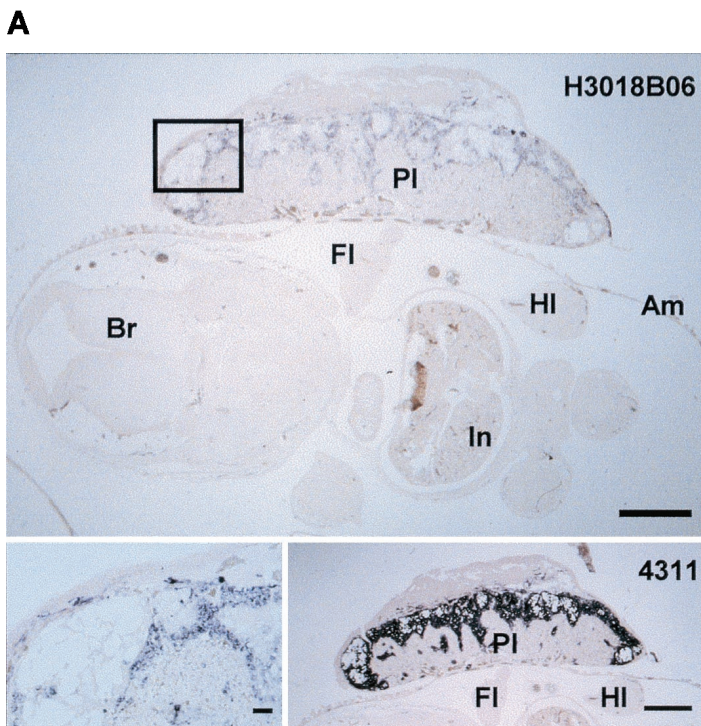
Next, the conventional criteria of 2-fold and 10-fold differences were used to subclassify the significantly different genes into six categories. Sixty-one genes that show 10-fold higher expression in placenta than in embryo are classified as placenta-specific genes; 166 genes with expression levels between 2-fold and 10-fold higher in placenta are classified as predominantly expressed in placenta. Similarly, 57 genes with 10-fold greater expression in embryo are classified as embryo-specific, and 253 whose expression levels are between 2-fold and 10-fold higher in embryo are classified as predominantly expressed in embryo. Expression differences of less than 2-fold have usually been considered at the limit of detection in previous analyses; but, by the statistical criteria applied here, significant differences are reproducibly seen for many differences in the range of 1.2- to 2-fold. The range showing statistically significant differences in less than 2-fold includes 62 genes whose expression is higher in placenta and 121 genes whose expression is higher in embryo. The statistically supported indications are particularly promising for the discovery and analysis of genes that are dosage controlled. Frequently important in early development, they include genes that are imprinted or X inactivated.

On the other hand, even genes showing more than 2-fold differences in their average expression levels sometimes fall into the statistically nonsignificant category (shown in gray in Fig. 2B), because of unknown variables that affect the relative signal from one or another probe. The statistical test of triplicate hybridizations discriminates the likely dependable values from false positives or negatives.

Verification of cDNA Microarray Data. The rankings of gene expression levels are supported by Northern blot analyses of selected genes. Results for 27 genes examined correlated well with cDNA microarray intensities and subclassification (Fig. 2B). Beta-actin, for example, displayed equivalent expression levels on the scatter plot, consistent with Northern blot data. Notably, a 28th gene (H3010F03) again showed the utility of the statistical test. Although it seemed placenta specific, based on average expression levels, it was judged as not significant by the statistical test and indeed showed only very faint signals in both placenta and embryo in Northern blot hybridization tests. The results also illustrated the greater sensitivity of microarray analyses compared with Northern blots. Expression levels 100-fold below the level detected on Northern blots were reproducibly observed in the microarray analyses (Fig. 2B).

Identification of Individual Genes. Details about some of the placenta-specific genes, including putative identifications and functional groupings, are shown in Table 1. Based on current gene classifications, they fall into three categories. One comprises 30 genes that are already known and well-described as placenta-specific genes (19, 20), which include placental lactogen II (25), extraembryonic endodermal cytokeratin type II [endoB (26)], Psx (27), proliferin-related protein, prolactin (PRL)-like protein G precursor, and PRL-like protein A precursor (reviewed in refs. 25 and 28). These results further support the validity of the cDNA microarray-based analyses for identifying placenta-specific genes.

A second category includes 78 genes that are already known, but for which expression in placenta had not been determined. This category includes cytotoxic T lymphocyte-associated protein 2 (*Ctla-2 α* , *Ctla-2 β*) (29), and fat facets (*Faf*). Although *Ctla-2* genes may originate from circulating blood cells



B

```

CAA GCT CTA ATA CGA CTC ACT ATA GGG AAA GCT GGT ACG CCT GCA GGT ACC GGT
CCG GAA TTC CCG GGT CGA CTA TCC CCT GAC GCT GTA GTC AGC CTG CTG TGG AAC
TCC TCA GAG ATG CAG CTA TCT GTA ACT CAT CCT TGC TGC AGG ACC CTG ATT CTG 162
1
M Q L S V T H P C C R T L I L
CTC CTG GTG TCT AAC CTG CTT CTG TGG GAA AGT GAG GCC GTG CTG CCC ATA TGC 216
16 L L V S N L L L W E S E A V L P I C
TCA GTA AGG AAT GGA CGC TGC TTT GGT TCC TTT GAA GAA CTG CTT GAA AGA GCT 270
34 S V R H G R C F G S F E E L L E R A
GTC AGT TTG TCT GAA GAG ATA AGT AAA AAA GCC TTT GAA CTC TCT ACT GCA TTT 324
52 V S L S E E I S K K A F E L F T A F
GAT AGT CAG TAT GCC CAG AGC CAG CAG CTC ATT GTC AAG AGC CTC AAA AAA TGT 378
70 D S Q Y A Q S H Q L I V K S L K K C
CAT ACA TCT TCT CTC GAC CTT CCA AAA CCA GGG AGT CAA GCC ATG CAG ACA CAT 432
88 H T S S L D L P K P G S Q A M Q T H
CCT GTA ACC CTA CTG AAA TTA GCA AGC AAA TTA TTG AGA GCC TGG GAA GTC CCT 486
106 P V T L L K L A S K L L R A W Q V P
CTG AAC CAT CTA GTG AAC AAC CTG CCA TCC TTG AAA ACC GTC TCT CCT TCT ATC 540
124 L N H L V N N L P S L K T V S P S I
CTC TCT AAA GCC AAA GAG ATT GAG GAA AAG ACC AAT GGA CTC CTG GAG GGA GTT 594
142 L S K A K E I E E K S N G L L E G V
AAG AGC ATT CTC ATC CAA ATG CAA AAT GGA GAT ACA GAA GAT GAG AAC TAC CCT 648
160 K S I L I Q M Q N G D T E D E N Y F
GGC TGG TCT GGA CTG GCA TCC TTG AAG TCA GAG ACT GAA GAT ATT CGC CTC TTT 702
178 G W S G L A S L K S E T E D I R L F
GCA TAT TAT AAC ATG ATC CGC TGT GAG GCC AGA GAC ACT CAG AAG GTT GAA ACT 756
196 A Y Y N M I R C E G R D T Q K V E T
GCT CTC AAG ATG GTG AAA TGC AAA ATT TCA AAT GAA AAC AAC TGC TAA GCC CTT 810
214 A L K M V K C K I S N E N N C *
TTC ATC ATA TCT TCG TCT GAG CCA CTG CTT GAT GAT ATA TTG CTG TGA AAC TTC 864
CTT GAA TTT TTT TCT CTG TTA AAT GCA TGC CCA GFG GTG TTT ATC TTC TTT CAC 918
AAT AAA AAT TGA CCC CAT GTA TGT CAC CAT CTG AAA AAA AAA AAA AAA AAA 972
AAA AAA AAA AAA AAA

```

C

```

H3018B06: 24 WESEAVLPICSVNRGRFCFSFEELLERAVLSSEEISKKAFFELTAFDSQYAQSHQLIVKS 83
++S LPIC C S +L +RAV LS I + E+F FD +YAQ I K+
Brush-tailed: 24 FKSVESLPICPSGAVNCQVLSLDLFDRAVMSLHYIHPSSSEMFNEFDERYAQGRGFTKA 83
opossum
H3018B06: 84 LKKCHTSSLDLKPQGSQAMQTHPVTLKLLKSLRAWQVPLNHLVNNLPSLKTVPSPILS 143
+ CHTSSL P+ QA Q H LL L ++LR+W PL HL V + S++ +ILS
Brush-tailed: 84 INSCHTSSLSTPEDKQQAQTHHEDLLNLIILRVLSWNPDLVHLVTEVRSMQEAPDPTILS 143
opossum
H3018B06: 144 KAKEIEEKSNGLLEGVKSILIQMONGDTEENYPGWSGLASLKSETEDIRLFAYYNNMIRC 203
KA EIEE++ LLEG++ I+ Q+ GD E+E Y WSGL SL+ ED RLFA+YN++ C
Brush-tailed: 144 KAMEIEEQNKRLLEGMEKIVGVHPGDRENEVYVSVSGLPSLQMADEDTRLFAFYNNLHC 203
opossum
H3018B06: 204 EGRDTQKVETALKMVCKISNENNC 228
RD+ K++ LK++KC++ +++NC
Brush-tailed: 204 LRRDSHKIDNYLKLKRLIHSNSNC 228
opossum

```

Fig. 3. New placenta-specific gene. (A, Top) *In situ* hybridization of H3018B06 gene. Transverse section of E13.5 embryo with placenta was probed with digoxigenin-labeled RNA transcribed from H3018B06 (bar = 1 mm). (Bottom Left) Magnified view of the region marked with a rectangle in the top panel (bar = 100 μ m). (Bottom Right) Control *in situ* hybridization with a trophoblast-specific gene, 4311 (bar = 1 mm). Am, amion; Br, brain; FI, forelimb; HI, hindlimb; In, intestine; PI, placenta. (B) Nucleotide and putative amino acid sequence of H3018B06. (C) Amino acid sequence alignment between mouse H3018B06 gene and opossum PRL gene.

in placenta, they might also function there. *Faf* is a deubiquitination enzyme specifically required for normal eye development in *Drosophila*; but its mouse homologue, *Fam*, was cloned (30) and implicated in cell–cell adhesion (31), suggesting a possible involvement of this gene in the implantation process.

The third category comprises 181 genes that are new and, together with genes in the second category, will provide a source of discovery of genes with a previously unrecognized important involvement in implantation and placenta development. In fact, the number of genes highly expressed or specific for the placenta are thus increased about 5-fold over the total reported previously.

Global Characteristics of Placenta. The list of putative placenta-specific genes elucidates some characteristic features of placenta (Table 1). Many genes involved in secretion stand out in the list and can be correlated with the status of the placenta as a primary

endocrine organ. As a derivative of trophoctoderm with epithelial character, mid-gestation placenta also actively synthesizes proteins characteristic of epithelia. The presence of a number of cell adhesion and invasion-related proteins is consistent with trophoblast function as the first-line invasion front for the embryo into the maternal uterine epithelia. Interestingly, the additional presence of a number of stress- and aging-related proteins may support the concept of the placenta as a prematurely aging organ (32).

Example of Newly Identified Genes. In an example of further gene characterization based on the microarray data, one new species that shows placenta-specific expression (H3018B06) was further characterized. Northern blot showed strong placenta-specific expression (Fig. 2B), and *in situ* hybridization (Fig. 3A) revealed expression exclusively in spongiotrophoblast and trophoblast giant cells, with no detectable signal in the embryo. This

expression pattern is the same as that of a well-known trophoblast-specific marker gene, clone 4311 (23) (Fig. 3A). The cDNA clone has a 987-bp insert, which suggests that it is essentially full-length (the 1.1-kb transcript observed in Northern blot includes a poly(A) tract of several hundred residues). The clone was completely sequenced (Fig. 3B, GenBank no, AF272368). Computer-assisted searches detected significant homology only to a PRL-like protein from opossum (Fig. 3C). Thus, this gene seems to be a new member of the placental growth-related hormone family (25, 28).

Similarity of Global Gene Expression Patterns Between Placenta and Embryo. Although we have emphasized the 5% of the genes that showed statistically significant differences between the placenta and developing embryo, it is striking that global gene expression patterns are overall rather similar for most genes in these very distinctive portions of the embryo. These results may rationalize the unexpected findings of placental failure in a significant fraction of mice in which nonplacental “tissue-specific” genes had been disrupted (reviewed in ref. 18). It may be relevant that, at E12.5, more than 70% of placental cells are trophoblast-derived trophoblast cells (33), and it is thus very likely that the expression patterns in placenta in this experiment represent trophoblasts. Trophoblast cells first appear in mammals, likely between the Jurassic [≈ 150 million years (Myr)] and Cretaceous (≈ 130 Myr) periods, after which the explosive diversification of mammalian species (“mammalian radiation”) took place (34). To create a functional placenta in such a short evolutionary span,

it may be that genetic pathways and networks were taken directly from ones already extant for basic body plan, with minimum modifications.

Possible Applications of the 15K Mouse Developmental cDNA Microarray. Determining the degree of divergence of placenta from embryo at a particular stage of development is an example of the use of the microarray. Because it includes a large number of genes from pre- and periimplantation mouse embryogenesis, it is adapted to comparable analyses of other stages and tissues during mouse development. For example, it should be possible to characterize genes that change expression as trophoblasts, which have a finite lifespan, form from embryonic cells with indefinite growth potential. Identification of downstream target genes in transgenic/knockout mice and global gene expression analysis of natural/induced mouse mutants are some of the obvious applications of this microarray. Recent findings indicate that some embryonic genes are reactivated in aging heart and muscle, suggesting the potential utility of the microarray for analyses of features of aging as well as growth control and tumorigenesis (35).

Note Added in Proof. The 15K mouse cDNA clone set described in this paper has been made available (see details at <http://lgsun.grc.nia.nih.gov>).

We thank Dr. David Schlessinger for critical reading of the manuscript, Dr. Chris Morrell for the help in statistical analysis, and Dr. Janet Rossant for providing 4311 clone probe.

- Ferea, T. L. & Brown, P. O. (1999) *Curr. Opin. Genet. Dev.* **9**, 715–722.
- Duggan, D. J., Bittner, M., Chen, Y., Meltzer, P. & Trent, J. M. (1999) *Nat. Genet.* **21**, 10–14.
- DeRisi, J. L., Iyer, V. R. & Brown, P. O. (1997) *Science* **278**, 680–686.
- Chu, S., DeRisi, J., Eisen, M., Mulholland, J., Botstein, D., Brown, P. O. & Herskowitz, I. (1998) *Science* **282**, 699–705.
- Spellman, P. T., Sherlock, G., Zhang, M. Q., Iyer, V. R., Anders, K., Eisen, M. B., Brown, P. O., Botstein, D. & Futcher, B. (1998) *Mol. Biol. Cell* **9**, 3273–3297.
- Roberts, C. J., Nelson, B., Marton, M. J., Stoughton, R., Meyer, M. R., Bennett, H. A., He, Y. D., Dai, H., Walker, W. L., Hughes, T. R., Tyers, M., Boone, C. & Friend, S. H. (2000) *Science* **287**, 873–880.
- Iyer, V. R., Eisen, M. B., Ross, D. T., Schuler, G., Moore, T., Lee, J. C. F., Trent, J. M., Staudt, L. M., Hudson, J., Jr., Boguski, M. S., Lashkari, D., Shalon, D., Botstein, D. & Brown, P. O. (1999) *Science* **283**, 83–87.
- DeRisi, J., Penland, L., Brown, P. O., Bittner, M. L., Meltzer, P. S., Ray, M., Chen, Y., Su, Y. A. & Trent, J. M. (1996) *Nat. Genet.* **14**, 457–460.
- Pollack, J. R., Perou, C. M., Alizadeh, A. A., Eisen, M. B., Pergamenschikov, A., Williams, C. F., Jeffrey, S. S., Botstein, D. & Brown, P. O. (1999) *Nat. Genet.* **23**, 41–46.
- Khan, J., Simon, R., Bittner, M., Chen, Y., Leighton, S. B., Pohida, T., Smith, P. D., Jiang, Y., Gooden, G. C., Trent, J. M. & Meltzer, P. S. (1998) *Cancer Res.* **58**, 5009–5013.
- Perou, C. M., Jeffrey, S. S., van de Rijn, M., Rees, C. A., Eisen, M. B., Ross, D. T., Pergamenschikov, A., Williams, C. F., Zhu, S. X., Lee, J. C., Lashkari, D., Shalon, D., Brown, P. O. & Botstein, D. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 9212–9217.
- Alizadeh, A. A., Eisen, M. B., Davis, R. E., Ma, C., Lossos, I. S., Rosenwald, A., Boldrick, J. C., Sabet, H., Tran, T., Yu, X., et al. (2000) *Nature (London)* **403**, 503–511.
- Ross, D. T., Scherf, U., Eisen, M. B., Perou, C. M., Rees, C., Spellman, P., Iyer, V., Jeffrey, S. S., Van De Rijn, M., Waltham, M., et al. (2000) *Nat. Genet.* **24**, 227–235.
- Scherf, U., Ross, D. T., Waltham, M., Smith, L. H., Lee, J. K., Tanabe, L., Kohn, K. W., Reinhold, W. C., Myers, T. G., Andrews, D. T., et al. (2000) *Nat. Genet.* **24**, 236–244.
- Khan, J., Bittner, M. L., Saal, L. H., Teichmann, U., Azorsa, D. O., Gooden, G. C., Pavan, W. J., Trent, J. M. & Meltzer, P. S. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 13264–13269.
- Ko, M. S. H., Threat, T. A., Wang, X., Horton, J. H., Cui, Y., Wang, X., Pryor, E., Paris, J., Wells-Smith, J., Kitchen, J. R., et al. (1998) *Hum. Mol. Genet.* **7**, 1967–1978.
- Ko, M. S. H., Kitchen, J. R., Wang, X., Threat, T. A., Wang, X., Hasegawa, A., Sun, T., Grahovac, M. J., Kargul, G. J., Lim, M. K., et al. (2000) *Development* **127**, 1737–1749.
- Copp, A. J. (1995) *Trends Genet.* **11**, 87–93.
- Cross, J. C., Werb, Z. & Fisher, S. J. (1994) *Science* **266**, 1508–1518.
- Rinkenberger, J. L., Cross, J. C. & Werb, Z. (1997) *Dev. Genet.* **21**, 6–20.
- Altschul, S. F., Gish, W., Miller, W., Myers, E. & Lipman, D. J. (1990) *J. Mol. Biol.* **215**, 403–410.
- Wilkinson, D. G. & Nieto, M. A. (1993) *Methods Enzymol.* **225**, 361–373.
- Lescisin, K. R., Varmuza, S. & Rossant, J. (1988) *Genes Dev.* **2**, 1639–1646.
- Claverie, J. M. (1999) *Hum. Mol. Genet.* **8**, 1821–1832.
- Linzer, D. I. & Fisher, S. J. (1999) *Mol. Endocrinol.* **13**, 837–840.
- Hashido, K., Morita, T., Matsushiro, A. & Nozaki, M. (1991) *Exp. Cell Res.* **192**, 203–212.
- Chun, J. Y., Han, Y. J. & Ahn, K. Y. (1999) *Dev. Dyn.* **216**, 257–266.
- Hamilton, R. T. & Millis, A. J. (1990) *Curr. Top. Dev. Biol.* **24**, 193–218.
- Denizot, F., Brunet, J. F., Roustan, P., Harper, K., Suzan, M., Luciani, M. F., Mattei, M. G. & Golstein, P. (1989) *Eur. J. Immunol.* **19**, 631–635.
- Wood, S. A., Pascoe, W. S., Ru, K., Yamada, T., Hirchenhain, J., Kemler, R. & Mattick, J. S. (1997) *Mech. Dev.* **63**, 29–38.
- Taya, S., Yamamoto, T., Kano, K., Kawano, Y., Iwamatsu, A., Tsuchiya, T., Tanaka, K., Kanai-Azuma, M., Wood, S. A., Mattick, J. S. et al. (1998) *J. Cell Biol.* **142**, 1053–1062.
- Rosso, P. (1976) *Curr. Concepts Nutr.* **4**, 23–41.
- Rossant, J. & Croy, B. A. (1985) *J. Embryol. Exp. Morphol.* **86**, 177–189.
- Easteal, S. (1999) *Bioessays* **21**, 1052–1058; discussion 1059.
- Schlessinger, D. & Ko, M. S. H. (1998) *Genomics* **52**, 113–118.