Screening and analyzing genes associated with Amur tiger placental development

Q. Li[1], T.F. Lu[2], D. Liu[3], P.F. Hu[2], B. Sun[2], J.Z. Ma[1], W.J. Wang[2], K.F. Wang[2], W.X. Zhang[2], J. Chen[3], W.J. Guan[2], Y.H. Ma[2] and M.H. Zhang[1]

[1]College of Wildlife Resources, Northeast Forestry University,
Xiangfang District, Harbin, China
[2]Department of Animal Genetic Resources, Institute of Animal Science,
Chinese Academy of Agricultural Sciences, Haidian District, Beijing, China
[3]Siberian Tiger Park, Songbei District, Harbin, China

Corresponding author: M.H. Zhang
E-mail: zhangminghai2004@126.com

ABSTRACT. The Amur tiger is a unique endangered species in the world, and thus, protection of its genetic resources is extremely important. In this study, an Amur tiger placenta cDNA library was constructed using the SMART cDNA Library Construction kit. A total of 508 colonies were sequenced, in which 205 (76%) genes were annotated and mapped to 74 KEGG pathways, including 29 metabolism, 29 genetic information processing, 4 environmental information processing, 7 cell motility, and 5 organismal system pathways. Additionally, PLAC8, PEG10 and IGF-II were identified after screening genes from the expressed sequence tags, and they were associated with placental development. These findings could lay the foundation for future functional genomic studies of the Amur tiger.

INTRODUCTION

The Amur tiger, also known as the Siberian tiger, is the largest living felid of recent tiger subspecies. There are approximately 400 wild Amur tigers, which are considered to be one of the world's top ten endangered species by the World Wildlife Fund. By constructing a cDNA expression library, treasured endangered biological resources cannot only be protected but can also provide a probe-constructed linkage map. More importantly, a cDNA expression library can be used to separate full-length genes and perform gene function research (Fang et al., 2014). Presently, expressed sequence tag (EST) sequencing analysis of cDNA expression libraries is widely used in gene identification, gene expression mapping, and new gene identification, with great success. Therefore, placenta EST analysis of certain genes can allow researchers to obtain gene structure and conduct further research on these genes (Zhao et al., 2013).

The placenta is an essential organ for maintaining fetal life; thus, understanding the regulation of placental development is very important for the diagnosis and treatment of pregnancy complications. However, it is difficult to obtain Amur tiger placenta, which makes it difficult to study Amur tiger parturition and fetal death (Matsubara, 2014). With this study, constructing an Amur tiger placenta cDNA library can effectively preserve genetic resources, because the genes can be screened for various research purposes, and research on the factors that affect the reproduction of this tiger in northeastern China can be advanced.

MATERIAL AND METHODS

Samples and reagents

Amur tiger placenta was taken from the Siberian Tiger Park. The following reagents were used: SMART™ cDNA Library Construction kit (Clontech, USA), DEPC (Sigma, Germany), TRIzol (Invitrogen, USA), and DNA marker (TaKaRa, Japan).

RNA extraction

Total RNA was extracted using the TRIzol one-step method, and total RNA quality was detected by formaldehyde denaturing gel electrophoresis.

cDNA library construction

The SMART™ cDNA Library Construction kit was used for cDNA library construction. Approximately 2 μg total RNA was prepared for single strand synthesis, and after long-distance PCR (LD-PCR) and *Sfi*I digestion, double-strand DNA was used on a Chroma Spin 400 column to separate large DNA fragments. The resulting DNA was then connected with λTriplEx2, and the junctional complex was packaged using the Gigapack III Gold packaging extract to obtain an unamplified library (Tanaka et al., 2000).

The unamplified library was diluted 1:10 and 1:20 to determine an unamplified library titer. According to the unamplified cDNA library titer, amplification phages were seeded at 4-5 x 10^4/plate in twenty 90-mm plates and cultured overnight with SM buffer, collected and stored. The amplified library titer was determined using the unamplified cDNA library

titer determination method; 700 clones were randomly selected, and fragment insertion size was determined by fragment identification primers (cDNA F: CCATTGTGTTGGTACCCGG; cDNA R: ATACGACTCACTATAGGGCGAATT). The recombination rate was estimated by the positive PCR rate.

EST sequencing and sequence analysis

Large-scale EST sequencing includes template preparation, sequencing reaction and sequence recognition. Template preparation adopts the semi-automatic assembly line method. The main steps of which include: construction of library, randomly picking a large number of colonies (cDNA cloning), bacterial proliferation, and electrophoresis analysis (Zhao et al., 2004). After the λTriplEx2 was translated into pTriplEx2, 508 bacterial monoclonal colonies were randomly sequenced. Contaminated sequences and vector sequences of low quality were removed from the sequencing results. Bioinformatic methods of Gene Ontology (GO) annotation and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were used for analysis (Bi et al., 2010). Through the analysis, we could determine the gene expression status in the placental tissues, expression distribution of various genes, gene distribution channels and screening of special genes from the Amur tiger.

RESULTS

Total RNA extraction and cDNA library construction

Total RNA was extracted and tested; the OD260/OD230 ratio was 2.10, and the OD260/OD280 ratio was 1.98, indicating that the RNA extracted was not contaminated by DNA, proteins or other impurities. Total RNA amounted to 105 μg, and the concentration was 3500 ng/μL. After 1% denaturing gel electrophoresis, three bands were clearly observed, which were 28S, 18S and 5S ribosomal RNA (Figure 1A). The ratio between 28S and 18S peaks was 2:1 (Figure 1B), illustrating that RNA degradation was limited and that the RNA was well preserved.
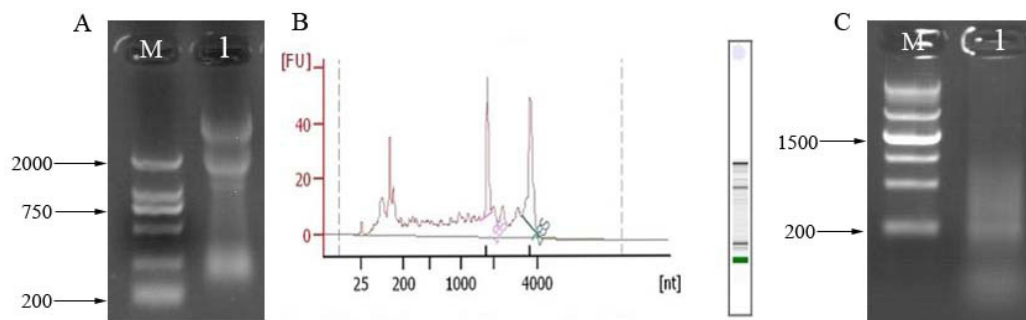

Figure 1. Total RNA from Amur tiger and LD-PCR. **A.** Total RNA from tissues of Amur tiger. **B.** Quality of total RNA was estimated by examining a sample on the Agilent 2100 Bioanalyzer using the Agilent RNA 6000 nano Reagents Port 1. **C.** Products of LD-PCR. *Lane M* = marker; *lane 1* = products of LD-PCR with 22 cycles.

Full-length cDNA was synthesized by reverse transcription and enriched by subsequent LD-PCR. Only non-truncated first strand cDNA was tagged by the SMART IV oligonucleotide sequence 5'-AAGCAGTGGTATCAACGCAGAGTGGCCATTACGGCCGGG-3' and CDS III/3' PCR Primer 5'-ATTCTAGAGGCCGAGGCGGCCGACATG-d (T) $_{30}$N-$_1$N-3' during the initial reverse transcription. PCR amplification products were digested with the restriction enzyme *Sfi*I. Agarose gel analysis demonstrated double-stranded cDNA sizes ranging from 300 bp to 12 kb (Figure 1C). CHROM-400 columns were used to isolate large cDNA fragments, as shown in Figure 2A. cDNA was collected in tubes 8, 9 and 10. Large pieces of cDNA were connected to the carrier and used in the lambda phage transformation. VSC257 host bacteria were selected and screened, and the inserted fragment sizes and titers were determined (Figure 2B). The unamplified cDNA library titer was 7.2 x 10^5/mL, and the amplified library titer was 5.3 x 10^9/mL. λTriplEx2 was converted to pTriplEx2 to recover the gene (Figure 2C). The size of the inserted fragment ranged from 133 bp to 2 kb, and the average inserted fragment was 1 kb in size (Figure 3).
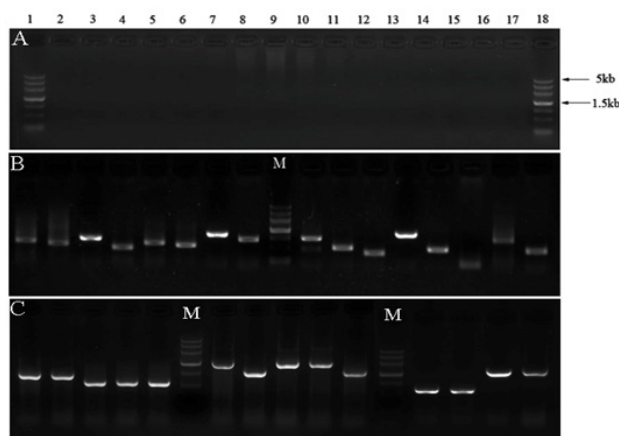

Figure 2. Construction of the Amur tiger placenta cDNA library. **A.** cDNA size fractionation by CHROMA SPIN-400. *Lanes 1* and *18* = marker III; *lanes 2-17* = tube serial number. **B.** Recombinant clones screening within the library. *Lane M* = marker III; the other lanes: PCR products for clones selected randomly. **C.** Converting λTriplEx2 to pTriplEx2 for saving the gene. *Lane M* = marker III.
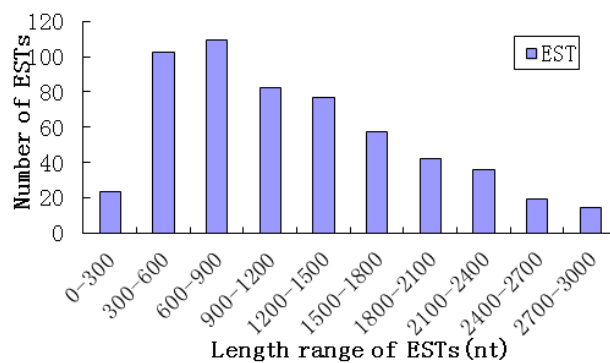

Figure 3. Distribution of the sequence length of the 181 unigenes in the Amur tiger placenta cDNA library.

Sequencing

In total, 508 positive clones were randomly selected for sequencing. Of these, 485 sequences were effective, and 258 sequences were calibrated and assembled. Sequence lengths ranged from 133 to 2992 bp, and the average length was 1186 bp (Table 1, Figure 3).

Table 1. Summary of the Amur tiger placenta cDNA library analysis.

Description	Number
Total number of sequences	508
Number of valid sequences	485
Average expressed sequence tag length (nt)	1186
Number of unigenes	258
Number of annotated genes	205

Sequence analysis

To obtain more information, BLASTX was used to analyze 268 sequences. As shown in Table 2, 205 sequences (76%) were successfully annotated, including 153 highly expressed genes (see Table 2). These 153 sequences were functionally classified by GO analysis. As shown in Figure 4, the 205 annotated genes included 38 functions and were in three main categories: biological processes, cellular components and molecular functions. As shown in Figure 4, genes were first ranked on the basis of the cell compartment followed by the whole cell. The molecular function reaction was the function of an individual gene product. As revealed in Figure 4, the binding activity of DNA, RNA, and protein ranked first in genetic molecule function followed by catalytic activity. The largest proportions of activities were cell processes, and the second largest were metabolic processes, as shown in Figure 4. GO categories determined that the same gene could participate in different biological processes because of different molecular features, which reflects the complexity of life.

Table 2. Most abundant transcripts in the Amur tiger placenta cDNA library.

| Transcript	Length (bp)	Description
Seq1	372	Hemoglobin subunit alpha-like isoform 1
Seq2	315	CDC42 effector protein 3
Seq3	472	BET1-like protein
Seq4	357	Exocyst complex component 3-like protein
Seq5	831	Hypothetical protein
Seq6	735	Cardiac troponin T
Seq7	547	PLP2 proteolipid protein 2
Seq8	515	Schlafen family member 14
Seq9	681	Pituitary tumor-transforming 1 interacting protein
Seq10	431	Prorelaxin
Seq11	381	40S ribosomal protein S11
Seq12	449	40S ribosomal protein S20
Seq13	604	60S ribosomal protein L11
Seq14	714	Kunitz-type protease inhibitor 2
Seq15	1047	Vimentin
Seq16	1017	Mitochondrial import inner membrane translocase subunit Tim17-A
Seq17	583	Heterogeneous nuclear ribonucleoprotein A3 isoform 2
Seq18	977	Conserved protein
Seq19	793	Elongation factor 1-alpha 1
Seq20	731	Outer membrane lipoprotein
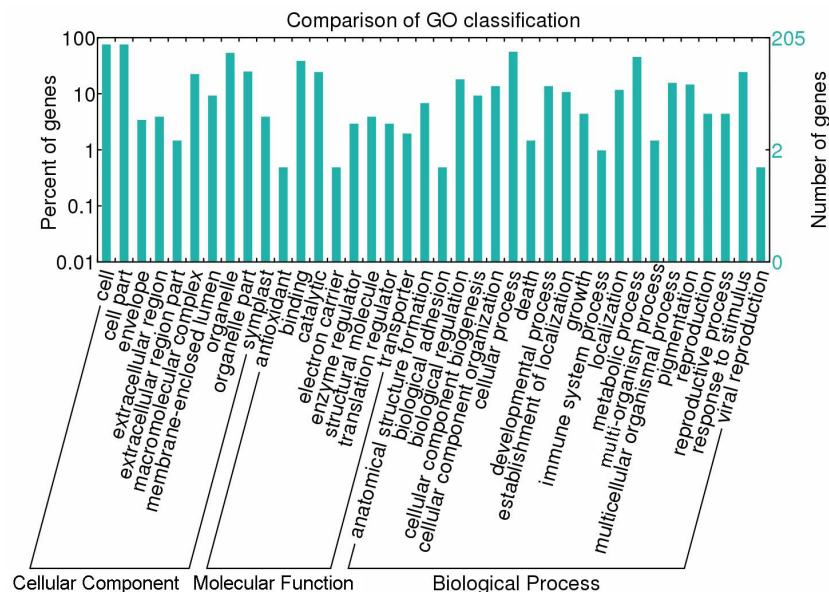
Figure 4. Functional annotation of the Amur tiger placenta cDNA library unigenes, based on the Swissprot database. Gene Ontology (GO) terms at the 2nd level are plotted. ''Biological process'', ''Cellular component'' and ''Molecular function'' are categorized independently in this ontology.

To further explore the annotated gene functions, a pathway analysis was performed using KEGG. A total of 205 unigenes were mapped to 74 KEGG pathways, which included 29 metabolic pathways, 29 genetic information processing pathways, 4 environmental information processing pathways, 7 cell motility pathways, and 5 organismal systems pathways. Table 3 lists the top 20 annotated pathways.

Table 3. Top 20 pathways in the Amur tiger placenta cDNA library, mapped using Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis.

KEGG	No. of clusters	Pathway
P04379	5	Beta3 adrenergic receptor signaling pathway
P00039	5	Metabotropic glutamate receptor group III pathway
P04378	5	Beta2 adrenergic receptor signaling pathway
P00007	4	Axon guidance mediated by semaphorins
P00038	3	JAK/STAT signaling pathway
P04377	3	Beta1 adrenergic receptor signaling pathway
P02739	3	*De novo* pyrimidine deoxyribonucleotide biosynthesis
P00006	3	Apoptosis signaling pathway
P00037	3	Ionotropic glutamate receptor pathway
P04376	3	5HT4 type receptor mediated signaling pathway
P02738	3	De novo purine biosynthesis
P00005	3	Angiogenesis
P04375	3	5HT3 type receptor mediated signaling pathway
P00004	2	Alzheimer disease-presenilin pathway
P00035	2	Interferon-gamma signaling pathway
P04374	2	5HT2 type receptor mediated signaling pathway
P00003	2	Alzheimer disease-amyloid secretase pathway
P00034	2	Integrin signalling pathway
P04373	2	5HT1 type receptor mediated signaling pathway
P00002	2	Alpha adrenergic receptor signaling pathway

Screening genes associated with placental development from ESTs

Early placental growth is regulated by transcriptional activation and trophoblast-specific gene silencing (Rossant and Cross, 2001). In this study, PLAC8, PEG10 and insulin-like growth factor II (IGF-II) were identified after screening genes from ESTs that were associated with placental development.

PLAC8, a pregnancy specific expression gene, is 550 bp in length, including the complete CDS region and 3' untranslated regions. The gene-coding region is from 1 to 297 bp. This gene encodes a protein with 98 nucleotides (Figure 5A), but the biological functions and regulatory pathways of this protein are not clear. However, considering human and mouse PLAC8 functions and other research results, we speculate that PLAC8 plays an important role in the change between embryonic and fetal development during pregnancy, but this hypothesis requires further investigation.

PEG10 (paternally expressed gene 10) is an imprinted gene that is maternally imprinted and paternally expressed. The gene regulates placental formation and trophoblast differentiation. It is also involved in early adipocyte differentiation and the development and metastasis of many types of tumors. The PEG10 gene is 321 bp in length; the entire gene-coding region is full-length (CDS) and encodes a 106-amino acid protein (Figure 5B).

IGF-II is an embryonic growth factor and is also a mitogenic peptide that plays an important role in normal growth and embryonic development. The IGF-II gene is 324 bp in length, and the entire gene is a coding region (CDS) that encodes a 107-amino acid protein (Figure 5C). IGF-II plays an important role in gene imprinting; lost genetic imprinting will lead to fetal overgrowth, congenital malformations, and susceptibility to embryonic tumors.
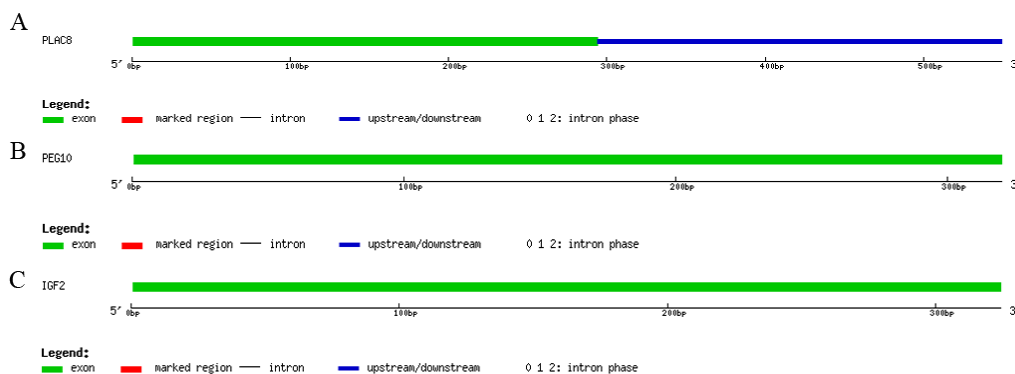

Figure 5. Gene structure of the Amur tiger and develop-related gene. **A.** PLAC8, **B.** PEG10, and **C.** IGF-2 identified from the placenta cDNA library.

DISCUSSION

The Amur tiger is one of top ten endangered species worldwide. The Amur tiger eats the placenta after giving birth, which is a considerable obstacle in the study of Amur tiger breeding and fetal disease. Therefore, establishment of an Amur tiger placenta cDNA library not only saves this endangered species' depleted genetic resources but it is also convenient for gene expression and function studies. For cDNA library construction, template RNA quality

directly affects cDNA synthesis efficiency. Thus, the key to RNA separation and purification is to eliminate RNA degradation by RNase and improve RNA yield and purity.

In this study, we constructed an Amur tiger placenta cDNA library. The titer (2.56 x 109 PFU/mL), recombination efficiency and inserted fragment length (287 to 2992 bp) demonstrated that a high quality cDNA library was successfully constructed (Adams et al., 1993; Gubler and Hoffman, 1983). Furthermore, 205 unigenes were obtained through sequencing and bioinformatic analysis. Most of the identified genes are involved in the regulation of protein activity. Therefore, we speculate that these highly expressed unigenes play vital roles in the Amur tiger placenta at the nymph stage. Pathway analysis revealed that the 205 mapped unigenes are involved in 74 KEGG pathways. This includes 29 metabolic pathways, 29 genetic information processing pathways, 4 environmental information processing pathways, 7 cell motility pathways, and 5 organismal system pathways.

From the ESTs, we selected three genes that were associated with placental development. PLAC8, a gene that is specific for pregnancy, plays an important role in the change between embryonic and fetal development in the Amur tiger during pregnancy. PLAC8 is substantially expressed in trophectoderm, in preimplantation embryos, and in giant trophoblast cells. For controlling cell division and organogenesis, the spongiotrophoblast layer proteins with the PLAC8 domain are critical components of the cellular regulatory scheme at later stages of development. A cross-species comparison demonstrates that the PLAC8 protein in the human genome shares many conserved cysteines and an amino acid sequence with a number of other predicted proteins in a wide range of species, including animals as well as plants such as tomato and rice, which was in accordance with our finding in the current study. From these similarities, a conserved domain of ancient origin has been inferred, but there are no functional data yet.

PEG10, first described by Ono et al. (2001), is a maternally imprinted and paternally expressed gene (Jinquan et al., 2000; Qiuping et al., 2003). PEG10 mRNA is expressed in the brain, kidney, lung, placenta, testis, ovary, spleen, lymphoblasts, endothelial cells, and thymus (Jinquan et al., 2003; Qiuping et al., 2003). PEG10 is downregulated early in hypoxia and highly activated at 11-12 weeks of gestation. To date, little is known about the its regulation and function. PEG10 is differentially regulated during placental development (Kruse et al., 1997). Its overexpression has been reported in hepatocellular carcinoma and in regenerating mouse liver, as well as in embryonic biliary atresia (Nguyen and Taub, 2002). PEG10 decreases cell death in hepatocellular carcinoma and is a novel paternally expressed imprinted gene. PEG10 is from a newly defined imprinted region at human chromosome 7q21 (Ono et al., 2001). PEG10 is located near the SGCE (sarcoglycan) gene, whose mouse homologue was also recently shown to be imprinted (Ono et al., 2001). PEG10 shows parent-of-origin-specific expression in mono-chromosomal hybrids (Ono et al., 2001). Elevated expression levels have been observed in the majority of G2/M phase cells in regenerating mouse liver (Tsou et al., 2003) and human hepatocellular carcinoma cells. Exogenous PEG10 expression confers oncogenic activity and transfection to hepatoma cells with PEG10 antisense, which suppresses its expression and results in cancer cell growth inhibition (Okabe et al., 2003). PEG10 protein associates with SIAH1, a mediator of apoptosis, and its overexpression decreases SIAH1-mediated cell death (Okabe et al., 2003). PEG10 is speculated to be a retrotransposon-derived gene. It is not completely repressed in adult human tissues. Distinct PEG10 expression has not been observed in the liver or a number of other tissues, but it has been observed in the brain, kidney, lung, testis, and placenta. In contrast, PEG10 expression has only been detected in the

placenta of 14 adult mouse tissues (Ono et al., 2003). PEG10 protein is primarily a cytosolic protein, but a tight association with the nuclear membrane of interphase cells has also been observed (Tsou et al., 2003). The exact biochemical and biological properties of PEG10 are not yet clear. However, some experimental data suggest a role for the preferential expression of imprinted genes in regulating hepatocyte growth control (Tsou et al., 2003).

IGF-II is an embryonic growth factor that plays an important role in normal embryonic growth. Imprinted genes play a key role in embryogenesis, fetal growth, differentiation, and the placenta; after birth, these genes influence growth and behavior. Changes in gene imprinting status have been related to genetic diseases (e.g., Beckwith-Wiedemann syndrome and Angelman syndrome) and tumors (such as Wilm's tumor, hepatoblastoma and hepatocellular carcinoma). IGF-II is positively correlated with placental weight at birth. IGF-II plays an important role in cellular trophoblast invasion and placental differentiation and formation in an autocrine manner (Diaz et al., 2007). In the first trimester of pregnancy, IGF-II affects early syncytiotrophoblast proliferation and (or) differentiation by autocrine regulation. At approximately 6 weeks of pregnancy, IGF-II expression continues with villous syncytiotrophoblast penetration and invasion into the maternal decidua, suggesting that IGF-II may play a role in the regulation of trophoblast invasion (Pollheimer et al., 2011).

Glycogen synthesis and glycogen content in placental cavernous trophoblast cells from mice that lack the IGF-II gene were found to be significantly decreased compared with normal mice, suggesting that placental IGF-II may regulate glycogen synthesis and glycogen content, which are important regulators of original sugar synthesis. IGF-II is an embryonic growth factor and is also a mitogenic peptide that plays an important role in normal growth and embryonic development (Hiden et al.. 2012). Studies have demonstrated that IGF-II adjusts the placental nutrient supply to the fetus by influencing placental development concurrent with fetal nutritional needs to influence placental transport. Gene knockouts will lead to fetal growth restriction in the womb.

ACKNOWLEDGMENTS

REFERENCES

Adams MD, Soares MB, Kerlavage AR, Fields C, et al. (1993). Rapid cDNA sequencing (expressed sequence tags) from a directionally cloned human infant brain cDNA library. *Nat. Genet.* 4: 373-380.

Bi YP, Liu W, Xia H, Su L, et al. (2010). EST sequencing and gene expression profiling of cultivated peanut (*Arachis hypogaea* L.). *Genome* 53: 832-839.

Diaz LE, Chuan YC, Lewitt M, Fernandez-Perez L, et al. (2007). IGF-II regulates metastatic properties of choriocarcinoma cells through the activation of the insulin receptor. *Mol. Hum. Reprod.* 13: 567-576.

Fang C, Xu T, Ye C, Huang L, et al. (2014). Method for RNA extraction and cDNA library construction from microbes in crop rhizosphere soil. *World J. Microbiol. Biotechnol.* 30: 783-789.

Gubler U and Hoffman BJ (1983). A simple and very efficient method for generating cDNA libraries. *Gene* 25: 263-269.

Hiden U, Lassance L, Tabrizi NG, Miedl H, et al. (2012). Fetal insulin and IGF-II contribute to gestational diabetes mellitus (GDM)-associated up-regulation of membrane-type matrix metalloproteinase 1 (MT1-MMP) in the human feto-placental endothelium. *J. Clin. Endocrinol. Metab.* 97: 3613-3621.

Jinquan T, Quan S, Jacobi HH, Jing C, et al. (2000). CXC chemokine receptor 3 expression on CD34(+) hematopoietic

progenitors from human cord blood induced by granulocyte-macrophage colony-stimulating factor: chemotaxis and adhesion induced by its ligands, interferon gamma-inducible protein 10 and monokine induced by interferon gamma. *Blood* 96: 1230-1238.

Jinquan T, Jacobi HH, Jing C, Millner A, et al. (2003). CCR3 expression induced by IL-2 and IL-4 functioning as a death receptor for B cells. *J. Immunol.* 171: 1722-1731.

Kruse N, Pette M, Toyka K and Rieckmann P (1997). Quantification of cytokine mRNA expression by RT PCR in samples of previously frozen blood. *J. Immunol. Methods* 210: 195-203.

Matsubara S (2014). Practical consideration of inserting intrauterine balloon during cesarean section for placenta previa. *Acta Obstet. Gynecol. Scand.* 93: 120-121.

Nguyen DH and Taub D (2002). CXCR4 function requires membrane cholesterol: implications for HIV infection. *J. Immunol.* 168: 4121-4126.

Okabe H, Satoh S, Furukawa Y, Kato T, et al. (2003). Involvement of PEG10 in human hepatocellular carcinogenesis through interaction with SIAH1. *Cancer Res.* 63: 3043-3048.

Ono R, Kobayashi S, Wagatsuma H, Aisaka K, et al. (2001). A retrotransposon-derived gene, PEG10, is a novel imprinted gene located on human chromosome 7q21. *Genomics* 73: 232-237.

Ono R, Shiura H, Aburatani H, Kohda T, et al. (2003). Identification of a large novel imprinted gene cluster on mouse proximal chromosome 6. *Genome Res.* 13: 1696-1705.

Pollheimer J, Haslinger P, Fock V, Prast J, et al. (2011). Endostatin suppresses IGF-II-mediated signaling and invasion of human extravillous trophoblasts. *Endocrinology* 152: 4431-4442.

Qiuping Z, Qun L, Chunsong H, Xiaolian Z, et al. (2003). Selectively increased expression and functions of chemokine receptor CCR9 on CD4[+] T cells from patients with T-cell lineage acute lymphocytic leukemia. *Cancer Res.* 63: 6469-6477.

Rossant J and Cross JC (2001). Placental development: lessons from mouse mutants. *Nat. Rev. Genet.* 2: 538-548.

Tanaka TS, Jaradat SA, Lim MK, Kargul GJ, et al. (2000). Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray. *Proc. Natl. Acad. Sci. U. S. A.* 97: 9127-9132.

Tsou AP, Chuang YC, Su JY, Yang CW, et al. (2003). Overexpression of a novel imprinted gene, PEG10, in human hepatocellular carcinoma and in regenerating mouse livers. *J. Biomed. Sci.* 10: 625-635.

Zhao SH, Simmons DG, Cross JC, Scheetz TE, et al. (2004). PLET1 (C11orf34), a highly expressed and processed novel gene in pig and mouse placenta, is transcribed but poorly spliced in human. *Genomics* 84: 114-125.

Zhao Y, Thammannagowda S, Staton M, Tang S, et al. (2013). An EST dataset for *Metasequoia glyptostroboides* buds: the first EST resource for molecular genomics studies in *Metasequoia*. *Planta* 237: 755-770.