Chinese Science Bulletin



# Renewal and preliminary study of expressed sequence tags database on human fetal liver aged 22 wk of gestation

CHEN TingGui[1,2], WU SongFeng[1], ZHOU GangQiao[1], ZHU YunPing[1†] & HE FuChu[1†]

[1] State Key Laboratory of Proteomics, Beijing Proteome Research Center, Beijing Institute of Radiation Medicine, Beijing 102206, China;
[2] Key Laboratory of Chemical Biology and Molecular Engineering of Ministry of Education, Shanxi University, Taiyuan 030006, China

**With the developments of international human transcriptome data and our ESTs of human fetal liver aged 22 weeks (wk) of gestation (HFL22w), the former research must be renewed. In this work, the EST data were firstly clustered by blasting against the ESTs of HFL22w, UniGene, DoTS, MGC and Twinscan-predicted human transcriptome. Then, after EST assembly and gene identification, the known genes were classified by GO (gene ontology), and the unknown genes were predicted by Pfam and ScanProsite to clarify their functions. In the end, the relations of 5 tissues including fetal liver, adult liver, bone marrow, thymus and lymph node that possess hemopoiesis or can indicate fetal liver characteristics were analyzed by hierarchical clustering. The results show that: (i) By comparing the 5 newest human transcriptome databases, we can largely reduce the probability that the ESTs belonging to unconnected parts of one gene were probably divided into different clusters, so it is recommended to blast against the newest databases when clustering EST data; (ii) some previous unknown ESTs had been identified as function-known genes, and 1379 genes were identified as fully new sequences possessed in our lab; (iii) through GO classification, we got a rough understanding of HFL22w, and obtained 6 cell migration genes and 6 hemopoiesis genes; (iv) prediction of gene function had enabled us to obtain 277 profiles, among them, there are 5 categories distributed in more than 10 genes; (v) five tissue relations analyzed by hierarchical clustering are related to their functions; (vi) We have built the world's largest EST database on HFL22w. Renewal and preliminary analysis of EST database on HFL22w will help to understand hemopoiesis and cell migration mechanism, and promote future research on human fetal liver.**

human fetal liver, expressed sequence tags, cluster and assembly, gene ontology, hierarchical clustering

The liver, as the largest gland in the human body, carries on metabolic processes of carbohydrates, fats, proteins, vitamins and hormones, and other important physiological functions, such as secretion, detoxification, phagocytosis, defense and homeostasis. HFL22w, intriguingly consisting of hepatic parenchymal cells and hematopoietic stem/progenitor cells, not only has these characteristics, but also has hemopoiesis and is a turning point between immigration and emigration of hematopoietic system. Therefore, the fetal liver at this stage are worthy of investigation.

For this reason, a system for large scale sequencing of cDNA library about HFL22w was established, and 13077 ESTs (expressed sequence tags) were sequenced and analyzed before[1]. With the growth of ESTs from 13077 to 20282, a new analysis needed to be processed.

Received April 3, 2008; accepted June 7, 2008
doi: 10.1007/s11434-008-0429-8
†Corresponding author (email: zhuyp@hupo.org.cn, hefc@nic.bmi.ac.cn)
Supported by the National High Technology Research and Development Program of China (Grant Nos. 2006AA02A312, 2006AA02Z334), National Basic Research Program of China (Grant No. 2006CB910803, 2006CB910706), National Natural Science Foundation of China (Grant No. 30621063) and Beijing Municipal Key Project (Grant No. H030330280590)

In this work, the ESTs were firstly blasted against 5 human transcriptome databases developed recently from Internet and our laboratory, and then after assembly by Phrap (http://www.genome.washington.edu/UWGC/analysistools/Phrap.cfm), the EST contigs were identified by blasting against GenBank database (ftp://ftp.ncbi.nih.gov/blast/db/FASTA/). After that, the known genes were subdivided into functional categories based on GO (gene ontology)[2], and the unknown contigs were predicted by Pfam (http://www.sanger.ac.uk/Software/Pfam/index.shtml) and ScanProsite (http://au.expasy.org/prosite/) to determine their possible functions. Finally, the tissues that have hemopoiesis functions or can indicate specialty of human fetal liver from the microarray data (aged 15 − 24 weeks (wk) of gestation) (http://wombat.gnf.org/index.html) were analyzed by hierarchical cluster[3]. All of the above related databases were downloaded on May 15, 2005.

## 1 Methods

### 1.1 DNA sequencing

Bacteria growth and plasmid extractions of the HFL22w cDNA library (ClonTech) were performed by a QIAprep 96 Turbo Miniprep Kit (QIAGEN). Sequencing reactions were performed on a GeneAmp PCR System 9700 thermal reactor (Perkin-Elmer) by using a BigDye Terminator Cycle Sequencing Kit (Perkin-Elmer) with T7 or SP6 primers. After removing the unincorporated dye terminators from sequencing reactions with DyeEx Spin Kits (QIAGEN), the reaction products were electrophoresed on an ABI 377-XL DNA sequencer (Perkin-Elmer-Applied Biosystems), and raw sequence data were automatically recorded. Standards: EST length $\geqslant$100 bp (base pair), error rate $\leqslant$1%.

### 1.2 Clustering, assembling and identifying ESTs

Each EST was firstly blasted against the ESTs of HFL22w, UniGene (ftp://ftp.ncbi.nih.gov/repository/UniGene/), DoTS (database of transcribed sequences) (http://www.allgenes.org), MGC (mammalian gene collection) (http://mgc.nci.nih.gov)[4] and Twinscan-predicted human transcriptome (http://genes.cs.wustl.edu/)[5]. Standards: EST full-length match, Identities $\geqslant$98%. Secondly, the clusters were assembled by Phrap with default parameters. Thirdly, each assembled sequence was matched with nt database from NCBI by BLAST (basic local alignment search tool). The EST, whose score was no less than 200 and whose matched sequence was a known gene, was considered to be function-known gene (that is, known gene in this paper), otherwise, it was considered to be a function-unknown gene (that is, unknown gene).

### 1.3 Gene function classification

The function classifications of our known genes were performed by DAVID (Database for Annotation, Visualization and Integrated Discovery) (http://apps1.niaid.nih.gov/david/)[6] whose level is ALL. And the others were default parameters.

### 1.4 ORF prediction and protein function prediction

For the unknown genes, the ORFs (open reading frame) were predicted by ATGpr (http://www.hri.co.jp/atgpr/)[7] and ORF finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html). Standard: the highest reliability. The functions (aa $\geqslant$ 30) were predicted by Pfam and ScanProsite[8]. Databases: Swiss-Prot, TrEMBL and PDB (Protein Data Bank); the others: default parameters.

### 1.5 Relation analysis of 5 tissues

Since fetal liver, bone marrow, thymus and lymph node have hemopoiesis functions, and adult liver can indicate specialty of human fetal liver, we selected these 5 tissues with microarray data to analyze their relations by hierarchical cluster (statistics: agglomeration schedule and proximity matrix; method: between-group linkage; measure: Pearson correlation). The microarray data were supplied by Mimmi Brown from Genomics Institute of the Novartis Research Foundation (San Diego), the human fetal liver is 15−24 wk old of gestation). Standard: AD (average difference)$\geqslant$200.

## 2 Results

### 2.1 General situation of data

By large scale sequencing, we totally got 20282 ESTs from a 3′-directed cDNA library of HFL22w. The shortest EST was 100 bp, the longest EST was 889 bp, and the total average length is 398 bp. Excluding 2091 ESTs coming from repeated sequencing, 1517 ESTs coming from other species, and 4306 ESTs coming from genomic sequences, mitochondrial genomic sequences and repetitive sequences, 12368 ESTs were considered good ones.

## 2.2 Results of gene identification

In 12368 ESTs, 8097 ESTs (65.47%) (Group I) were matched to 2483 known genes, and 4271 ESTs (34.53%) (Group II) were identified as 2416 genes that exhibited no significant homology to known genes. The changes of gene numbers are shown in Table 1. Table 1 indicates that although the EST data increased, the total gene numbers decreased. The explanation was as follows: before assembly, we first blasted each EST against 5 human transcriptome databases developed recently from Internet and our laboratory to reduce the probability that the ESTs belonging to unconnected parts of one gene were divided into different clusters, so we could largely improve veracity of ESTs clustering.

**Table 1** Changes of gene numbers about HFL22w

| Identification | Before EST data added | | After EST data added | |
|---|---|---|---|---|
| | Gene numbers | EST numbers | Gene numbers | EST numbers |
| Known | 1729 | 5819 | 2483 | 8097 |
| Unknown | 4768 | 5460 | 2416 | 4271 |
| Total | 6497 | 11279 | 4899 | 12368 |

## 2.3 Frequency information

The results are listed in Table 2. Although there were only 266 highly expressed genes, they included lots of ESTs (5410, 43.74%). Among them, HBGG (Hemoglobin G gamma globin) (frequency: 775; the same below), HSA (human serum albumin) (572), ALB (albumin) (347) and ferritin L chain mRNA (112) frequencies were greater than 100; their total frequencies occupied 33.38% of ESTs. The frequencies of moderately expressed genes were less than that of highly expressed genes, but the gene numbers (589, 12.02%) were more than those of the former. The lowly expressed genes had the most gene numbers (4044, 82.55%).

**Table 2** Total ESTs frequency distribution about HTL22w

| Type | Frequency | Gene numbers (%) | ESTs numbers (%) |
|---|---|---|---|
| High | >5 | 266(5.43%) | 5410(43.74%) |
| Moderate | 5 | 92(1.88%) | 460(3.72%) |
| | 4 | 137(2.79%) | 548(4.43%) |
| | 3 | 360(7.35%) | 1080(8.73%) |
| | subtotal | 589(12.02%) | 2088(16.88%) |
| Low | 2 | 826(16.86%) | 1652(13.36%) |
| | 1 | 3218(65.69%) | 3218(26.02%) |
| | subtotal | 4044(82.55%) | 4870(39.38%) |
| Total | | 4899(100%) | 12368(100%) |

## 2.4 Fully new sequences

Through above-mentioned methods, the ESTs of HFL22w were fully clustered, assembled and identified. Here, we only analyzed the function-unknown genes (Group II). Among them, 1037 genes (42.92% of 2416 genes) were matched to UniGene, DoTS, MGC and Twinscan-predicted human transcriptome, the remaining 1379 sequences (57.08% of 2416 genes) were fully new sequences. Among these new sequences, some probably possess very important functions, so they will be experimented later in our lab. The results are shown in Table 3.

## 2.5 Known gene function classification

In 2483 known genes, 931 genes were divided into 424 categories by DAVID, but the others could not be classified. The classification results are listed in Supplement 1. Some important biological processes of HFL22w are listed in Table 4.

**Table 3** The clustered results of function-unknown genes on HFL22w

| Databases | Gene numbers | Gene numbers (%) |
|---|---|---|
| UniGene | 17 | |
| DoTS | 441 | 459(19.00%) |
| Twinscan | 1 | |
| UniGene, DoTS | 378 | |
| UniGene, Twinscan | 1 | 387(16.02%) |
| DoTS, MGC | 4 | |
| DoTS, Twinscan | 4 | |
| UniGene, DoTS, MGC | 96 | |
| UniGene, DoTS, Twinscan | 31 | 129(5.34%) |
| DoTS, MGC, Twinscan | 2 | |
| UniGene, DoTS, MGC, Twinscan | 62 | 62(2.57%) |
| Fully new sequences | 1379 | 1397(57.08%) |
| Total | 2416 | 2416(100%) |

**Table 4** Some important biological processes of HFL22w

| Category | Gene numbers | Gene numbers (%) |
|---|---|---|
| Metabolism | 581 | 62.41 |
| Cell growth and/or maintenance | 325 | 34.91 |
| Biosynthesis | 125 | 13.43 |
| Development | 116 | 12.46 |
| Cell proliferation | 107 | 11.49 |
| Defense response | 69 | 7.41 |
| Immune response | 63 | 6.77 |
| Cell differentiation | 11 | 1.18 |
| Endocytosis | 7 | 0.75 |
| Cell migration | 6 | 0.64 |
| Hemopoiesis | 6 | 0.64 |

To understand cell migration and hemopoiesis about HFL22w, we list them in Tables 5 and 6. In Table 5, apoE (apolipoprotein E) is a multifunctional protein and is synthesized in several areas of the body[9,10]. Approximately three-fourths of the plasma apoE is synthesized in the liver. Hepatic parenchymal cells are the main place producing and secreting apoE; SEMA6A takes part in apoptosis, axon guidance, development and neurogenesis biological process, and at the same time, participates in cell surface receptor linked signal transduction; platelet factor 4 has immune cell chemotaxis function; it is chemotactic for neutrophils and monocytes[11].

**Table 5** Genes related to cell migration of HFL22w

| Clone | Frequency | Gene name |
|---|---|---|
| C3092 | 10 | Apolipoprotein E |
| D2575 | 3 | Apolipoprotein E precursor |
| HA0052 | 1 | Sema domain, transmembrane domain (TM), and cytoplasmic domain, (semaphorin) 6A |
| E2638 | 1 | Fasciculation and elongation protein zeta 1 (zygin I) |
| D4894 | 1 | Platelet factor 4 (chemokine (C-X-C motif) ligand 4) |
| E4827 | 1 | Reticulon 4 |

RTN (reticulon) is a gene family localized on the membrane of endoplasmic reticulum. Up to now, RTN1, RTN2, RTN3 and RTN4 of this family have been found[12].

In Table 6, ACIN1 (apoptotic chromatin condensation inducer 1) plays a part in erythrocyte differentiation[13] and positive regulation of monocyte differentiation[14]; NOTCH2 (Notch homolog 2) encodes a member of the Notch family. In *Drosophila*, notch interacting with its cell-bound ligands (delta, serrate) establishes an intercellular signaling pathway that plays a key role in development[15]; ERAF (erythroid associated factor) has many functions, and can participate in protein folding, hemoglobin metabolic process and hemopoiesis[16]; PF4 (platelet factor 4) participates in negative regulation of angiogenesis and megakaryocyte differentiation[17,18].

**Table 6** Genes related to hemopoiesis of HFL22w

| Clone | Frequency | Gene name |
|---|---|---|
| F0503 | 2 | Apoptotic chromatin condensation inducer in the nucleus |
| D13517 | 1 | Histone deacetylase 4 |
| HA0172 | 1 | Notch homolog 2 (Drosophila) |
| HA0006 | 6 | WW domain-containing protein 1 |
| F0555 | 6 | Erythroid associated factor |
| D4894 | 1 | Platelet factor 4 (chemokine (C-X-C motif) ligand 4) |

The results on molecular function and cellular component are omitted. They can be obtained by analyzing the original data with DAVID program. The original data are shown in Supplement 2.

### 2.6 Prediction of unknown gene function

Through ATGpr and ORF finder, we built an ORFome database including known genes and unknown genes, which would promote our genomic and proteomic researches on HFL22w, such as gene function study, protein identification and protein-protein interactions. Up to now, we have studied some novel genes, including ARFGAP1[19,20], NPDC1[21], MAGE-D[22], NDRG2, NDRG3 and NDRG4[23], E9730[24], Ceap-11 and Ceap-16[25], etc. The ORF statistic results of 2416 unknown genes and 1037 validated sequences are shown in Table 7. From the results, we can see that the larger novel genes are more difficult to find than the smaller genes.

**Table 7** ORF statistic results of 2416 unknown genes and 1037 validated sequences

| aa length | ORF numbers (%) | ORF numbers (%) |
|---|---|---|
| ≥1000 | 2(0.08) | 2(0.19) |
| 500－1000 | 150(6.21) | 40(3.86) |
| 100－500 | 616(25.50) | 217(20.93) |
| 30－100 | 1153(47.72) | 543(52.36) |
| <30 | 495(20.49) | 235(22.66) |
| Total | 2416(100) | 1037(100) |

Through ScanProsite and Pfam, 1921 proteins were predicted and 277 profiles were obtained. The categories distributed in no less than 10 genes are shown in Table 8.

As we know, the categories that are widely existent in HFL22w are relevant to their functions. Cys2His2 type zinc-finger protein is the most frequently used type of transcription factorit accounts for about 3% of genes in the human genome. Zinc-fingers mostly interact with DNA, and new evidence indicates that zinc-fingers are more widely used to recognize RNA[26]. At the same time, some zinc finger proteins may regulate hematopoietic differentiation toward erythroid (GATA-1 and EKLF), megakaryocytic (GATA-1 and GATA-2), lymphoid (GATA-3 and Ikaros), granulocytic (MZF-1), as well as monocytic (EGR-1) lineages[27]. Moreover, many zinc finger genes are involved in the pathogenesis of hematological malignancies.

G-β (β-transducin) is one of the three subunits (α, β, and γ) of the guanine nucleotide-binding proteins (G proteins), while G proteins acts as an intermediary in the signal transduction[28]. Structurally, G-β consists of eight tandem repeats of about 40 residues, each containing a central Trp-Asp motif (that is, WD-40 repeat). Such a repetitive segment exists in a number of proteins, and possesses diverse functions[29,30].

Ub (Ubiquitin), among eukaryotes, is highly conserved. It has many functions including cell division, embryogenesis, regulation of transcription, programmed cell death, etc. At the same time, proteins that are to be degraded are first tagged by conjugating with Ub and these tagged proteins are then recognized and shuttled to the proteasome for degradation[31]. So it is not surprising that Ub profile is existent in eukaryotes widely.

Ankyrin repeats (ANKs) are tandemly repeated modules of about 33 amino acids. They occur in a large number of diverse proteins which are mainly from eukaryotes, such as p53-binding protein 53BP2[32], cyclin-dependent kinase inhibitor p19Ink4d[33], transcriptional regulator GABP-β[34], and NF-kappaB inhibitory protein IkB-α[35]. As we know, many ANK act as protein-protein interaction domains.

### 2.7 Relations of 5 tissues

The relations analyzed by hierarchical cluster are shown in Figure 1. From the figure, we can see that 5 tissues were divided into two clusters: I (4, 5, 2), II (1, 3). It can be explained as follows: Thymus, lymph node and bone marrow all belong to lymphoid tissue, so they were divided into one cluster, while fetal liver and adult liver carry out metabolism, secretion, detoxifcation, phagocytosis, defense and homeostasis functions, etc., so they were divided into another cluster. At the same time, we have noticed that the relation between fetal liver and adult liver is closer than that of fetal liver and bone marrow. The coefficients are 0.687 and 0.543 (Table 9), respectively in that fetal liver and adult liver have lots of similar functions, although fetal liver and bone marrow have hemopoiesis. From the analysis, we can understand the relationship among these tissues.

**Table 9** Relations among fetal liver, bone marrow, liuer, thymus and lymph node

| Stage | Culster combined | | Coeffi-cients | Stage cluster First appears Next | | |
|---|---|---|---|---|---|---|
| | Cluster 1 | Cluster 2 | | Cluster 1 | Cluster 2 | Stage |
| 1 | 4 | 5 | 0.904 | 0 | 0 | 2 |
| 2 | 2 | 4 | 0.741 | 0 | 1 | 4 |
| 3 | 1 | 3 | 0.687 | 0 | 0 | 4 |
| 4 | 1 | 2 | 0.543 | 3 | 2 | 0 |

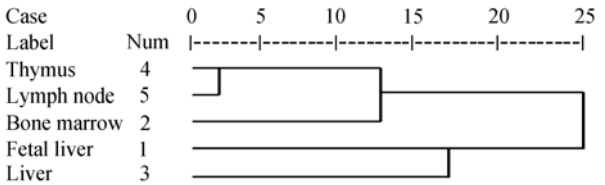


**Figure 1** Hierarchical cluster of 5 tissue (fetal liver, bone marrow, liver, thymus and lymph node).

**Table 8** The categories (gene numbers≥10) in 2416 unknown genes on HFL22w

| Categories | Sub-categories | Numbers | Total |
|---|---|---|---|
| Zinc finger | Zinc finger SP-RING-type profile | 1 | 39 |
| | Zinc finger RING-type signature | 1 | |
| | Zinc finger RING-type profile | 2 | |
| | Zinc finger PHD-type signature | 2 | |
| | Zinc finger PHD-type profile | 2 | |
| | Zinc finger matrin-type profile | 1 | |
| | Zinc finger FYVE/FYVE-related type profile | 1 | |
| | Zinc finger CCHC-type profile | 1 | |
| | Zinc finger C2H2 type domain signature | 14 | |
| | Zinc finger C2H2 type domain profile | 10 | |
| | Zinc finger BED-type profile | 1 | |
| | Zinc finger AN1-type profile | 3 | |
| Trp-Asp (WD) repeats | Trp-Asp (WD) repeats signature | 6 | 21 |
| | Trp-Asp (WD) repeats profile | 7 | |
| | Trp-Asp (WD) repeats circular profile | 8 | |
| Ribosomal protein | Omitted | 15 | 15 |
| Ubiquitin | Ubiquitin-interacting motif (UIM) repeat profile | 3 | 14 |
| | Ubiquitin-conjugating enzymes family profile | 2 | |
| | Ubiquitin-conjugating enzymes active site | 2 | |
| | Ubiquitin-associated domain (UBA) profile | 3 | |
| | Ubiquitin domain profile | 2 | |
| | Ubiquitin carboxyl-terminal hydrolases family 2 signature 1 | 1 | |
| | Ubiquitin carboxyl-terminal hydrolases family 2 profile | 1 | |
| Ankyrin repeat | Ankyrin repeat region circular profile | 6 | 12 |
| | Ankyrin repeat profile | 6 | |

## 3 Summaries and discussions

Comparison between EST and the newest databases

largely improved the veracity of EST clustering. As we know, when we sequence ESTs, we cannot pick out all clones of a tissue and sequence them. If we can obtain the ESTs of this tissue from other laboratories, we will find that some of them probably connect with each other, or may be fully new sequences. So it is recommended to get the newest databases from Internet and compare them when clustering ESTs.

Through assembly and identification, some former unknown genes have become known genes, and 1379 fully new sequences have been identified. These processes largely lessen the tedious work of one-by-one EST analysis, and the identification of fully new sequences lays the foundation for original researches on HFL22w.

GO classification is used to classify gene functions in order to avoid man-made errors, and is convenient for the comparisons among different tissue profiles. This analysis got a rough understanding of HFL22w, and more importantly, revealed some genes about hemopoiesis and cell migration. However, GO classification is not ideal completely at present, because as it is different from traditional classification, the researchers can hardly find out tissue characteristics easily.

Prediction of gene function had enabled us to build an ORFome database on HFL22w and obtain 277 profiles. The built of ORFome database will promote genomic and proteomic researches on HFL22w, and the wide existence of 5 categories illustrate the importance of their functions.

The analysis of hierarchical cluster indicated that 5 tissue relations are consistent with their functions. Especially, HFL22w plays an important role in development of the hematopoietic system, but its most functions are similar to those of adult liver. This analysis will help us understand human hematopoietic mechanism and the relations of 5 tissues.

To sum up, we have built the world's biggest EST database on HFL22w, although it is still a preliminary attempt, the establishment of this database will promote our comprehensive research on human fetal liver.

*We especially thank Mr. MIMMI Brown from Genomics Institute of the Novartis Research Foundation (San Diego) for supplying microarray data. We also thank our colleagues LI JianQi, LI Dong, DU ChunJuan, DENG YangYang, WANG ZhongSheng for their helpful suggestions.*

1 Yu Y, Zhang C, Zhou G, et al. Gene expression profiling in human fetal liver and identification of tissue- and developmental-stage-specific genes through compiled expression profiles and efficient cloning of full-length cDNAs. Genome Res, 2001, 11(8): 1392－1403[DOI]

2 Ashburner M, Ball C A, Blake J A, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet, 2000, 25(1): 25－29[DOI]

3 Su A I, Wiltshire T, Batalov S, et al. A gene atlas of the mouse and human protein-encoding transcriptomes. Proc Natl Acad Sci USA, 2004, 101(16): 6062－6067[DOI]

4 MGC Project Team. The status, quality, and expansion of the NIH full-length cDNA project: the mammalian gene collection (MGC). Genome Res, 2004, 14: 2121－2127

5 Korf I, Flicek P, Duan D, et al. Integrating genomic homology into gene structure prediction. Bioinformatics, 2001, 17: S140－S148

6 Dennis G J, Sherman B T, Hosack D A, et al. DAVID: Database for annotation, visualization, and integrated discovery. Genome Biol, 2003, 4(5): 3[DOI]

7 Salamov A A, Nishikawa T, Swindells M B. Assessing protein coding region integrity in cDNA sequencing projects. Bioinformatics, 1998, 14(5): 384－390[DOI]

8 Hulo N, Sigrist C J A, Le Saux V, et al. Recent improvements to the PROSITE database. Nucleic Acids Res, 2004, 32: D134－D137[DOI]

9 Sauer I, Dunay I R, Weisgraber K, et al. An apolipoprotein E-derived peptide mediates uptake of sterically stabilized liposomes into brain capillary endothelial cells. Biochemistry, 2005, 44(6): 2021－2029[DOI]

10 Zhu Y, Hui D Y. Apolipoprotein E binding to low density lipoprotein receptor-related protein-1 inhibits cell migration via activation of cAMP-dependent protein kinase A. J Biol Chem, 2003, 278(38): 36257－36263[DOI]

11 Deuel T F, Senior R M, Chang D, et al. Platelet factor 4 is chemotactic for neutrophils and monocytes. Proc Natl Acad Sci USA, 1981, 78(7): 4584－4587[DOI]

12 Sironen R K, Karjalainen H M, Torronen K J, et al. Reticulon 4 in chondrocytic cells: Barosensitivity and intracellular localization. Int J Biochem Cell Biol, 2004, 36(8): 1521－1531[DOI]

13 Zermati Y, Garrido C, Amsellem S, et al. Caspase activation is required for terminal erythroid differentiation. J Exp Med, 2001, 193(2): 247－254[DOI]

14 Sordet O, Rebe C, Plenchette S, et al. Specific involvement of caspases in the differentiation of monocytes into macrophages. Blood, 2002, 100(13): 4446－4453[DOI]

15 Stifani S, Blaumueller C M, Redhead N J, et al. Human homologs of a *Drosophila* enhancer of split gene product define a novel family of nuclear proteins. Nat Genet, 1992, 2(2): 119-127. Erratum in: Nat Genet. 1992, 2(4): 343[DOI]

16 Kihm A J, Kong Y, Hong W, et al. An abundant erythroid protein that stabilizes free alpha-haemoglobin. Nature, 2002, 417(6890): 758－763[DOI]

17 Han Z C, Bellucci S, Tenza D, et al. Negative regulation of human megakaryocytopoiesis by human platelet factor 4 and beta thromboglobulin: Comparative analysis in bone marrow cultures from normal individuals and patients with essential thrombocythaemia and immune thrombocytopenic purpura. Br J Haematol, 1990, 74(4): 395－401[DOI]
18 Maione T E, Gray G S, Petro J, et al. Inhibition of angiogenesis by recombinant human platelet factor-4 and related peptides. Science, 1990, 247(4938): 77－79[DOI]
19 Zhang C G, Yu Y T, Zhang S W, et al. Characterization, chromosomal assignment, and tissue expression of a novel human gene belong to ARF GAP family. Genomics, 2000, 63(3): 400－408[DOI]
20 Liu X, Zhang C, Xing G, et al. Functional characterization of novel human ARFGAP3. FEBS Lett, 2001, 490(1-2): 79－83[DOI]
21 Qu X, Zhang C, Zhai Y, et al. Characterization and tissue expression of a novel human gene npdc1. Gene, 2001, 264(1): 37－44[DOI]
22 Zhang C G, Xing G C, Wei H D, et al. A new melanoma antigen-encoding gene subfamily in human chromosome X. Acta Genet Sin, 2001, 28(3): 197－203
23 Qu X H, Zhai Y, Wei H D, et al. Characterization and expression of three novel differentiation-related genes belong to the human NDRG gene family. Mol Cell Biochem, 2002, 229(1-2): 35－44[DOI]
24 Wang Z Q, Wei H D, He F C. Protein product encoded by a human novel gene E9730 enhances AP-1 activity through interacting with Jab1. Acta Biochim Biophys Sin (Shanghai), 2004, 36(1): 11－15
25 Wang Z, Wei H, Yu Y, et al. Characterization of Ceap-11 and Ceap-16, two novel splicing-variant-proteins, associated with centrosome, microtubule aggregation and cell proliferation. J Mol Biol, 2004, 343(1): 71－82[DOI]
26 Lu D, Searles M A, Klug A. Crystal structure of a zinc-finger-RNA complex reveals two modes of molecular recognition. Nature, 2003, 426(6962): 96－100[DOI]
27 Han Z G, Zhang Q H, Ye M, et al. Molecular cloning of six novel kruppel-like zinc finger genes from hematopoietic cells and identification of a novel transregulatory domain KRNB. J Biol Chem, 1999, 274 (50): 35741－35748[DOI]
28 Gilman A G. G proteins: Transducers of receptor-generated signals. Annu Rev Biochem, 1987, 56: 615－649[DOI]
29 Smith T F, Gaitatzes C, Saxena K, et al. The WD repeat: a common architecture for diverse functions. Trends Biochem Sci, 1999, 24: 181－185[DOI]
30 Neer E J, Schmidt C J, Nambudripad R, et al. The ancient regulatory-protein family of WD-repeat proteins. Nature, 1994, 371: 297－300[DOI]
31 Hershko A, Ciechanover A. The ubiquitin system. Annu Rev Biochem, 1998, 67: 425－479[DOI]
32 Gorina S, Pavletich N P. Structure of the p53 tumor suppressor bound to the ankyrin and SH3 domains of 53BP2. Science, 1996, 274: 1001－1005[DOI]
33 Luh F Y, Archer S J, Domaille P J, et al. Structure of the cyclin - dependent kinase inhibitor p19Ink4d. Nature, 1997, 389(6654): 999－1003[DOI]
34 Batchelor A H, Piper D E, de la Brousse F C, et al. The structure of GABP alpha/beta: an ETS domain-ankyrin repeat heterodimer bound to DNA. Science, 1998, 279(5353): 1037－1041[DOI]
35 Jacobs M D, Harrison S C. Structure of an IkappaBalpha/NF-kappaB complex. Cell, 1998, 95: 749－758[DOI]