DNA Research **15**, 115–122, (2008) doi:10.1093/dnares/dsn005

Transcriptome Analysis of a cDNA Library from Adult Human Epididymis

Jian-Yuan Li[1,*], Hai-Yan Wang[1], Jie Liu[1], Qiang Liu[2], Jin-Song Zhang[2], Feng-Chun Wan[1], Fu-Jun Liu[1], Shao-Hua Jin[1], and Yong-Lian Zhang[2,*]

Shandong Research Center of Stem Cell Engineering/Centre Laboratory, Yu-Huang-Ding Hospital, Yan-Tai, Shan-Dong Province, Shandong 264000, People's Republic of China[1] and Shanghai Key Laboratory for Molecular Andrology, State Key Laboratory of Molecular Biology, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, People's Republic of China[2]

(Received 28 June 2007; accepted on 05 March 2008; published online 4 April 2008)

Abstract

Mammalian Gene Collection (MGC) verified over 9000 human full-ORF genes and FLJ Program reported 21 243 cDNAs of which 14 409 were unique ones and 5416 seemed to be protein-coding. The pity is that epididymis cDNA library was missing in their sequencing target list. Epididymis is a very important male accessory sex organ for sperm maturation and storage. Fully differentiated spermatozoa left from testis acquire their motility and capacity for fertilization via interactions with the epididymal epithelium duct lumen during passage through this convoluted duct. Here, we report that 20 000 clones from a healthy male epididymis cDNA library have been sequenced. The sequencing data provided 8234 known sequences and 650 unknown cDNA fragments. Hundred and six of 650 unknown cDNA clone inserts were randomly selected for fully sequencing. There were 25 unknown unique sequences and 19 released but unreported sequences came out. By northern blot analysis, four sequences randomly selected from the 19 released sequences with no known function showed positive mRNA signals in epididymis and testis. The signals for three of six from those unknown group showed as epididymis abundant in a region-specific manner but not in the testis and other tissues tested. All the sequencing data will be available on the website www.sdscli.com.

Key words: human epididymis cDNA library; transcriptomes for human epididymis; sperm maturation

1. Introduction

In 2002, Mammalian Gene Collection (MGC) sequenced and verified over 9000 human full-ORF genes and 7800 candidate full-ORF genes.[1] One year after, full-length Japan Program (FLJ) reported 21 243 cDNAs, of which 14 409 (10 897 clusters) were unique ones and half of them (5416) seemed to be protein-coding.[2] Ten months later, International Human Genome Sequencing Consortium announced on *Nature* that the euchromatic sequence of the human genome has been finished.[3] The current genome sequence (Build 35) contains 2.85 billion nucleotides interrupted by only 341 gaps. It covers 99% of the euchromatic genome and is accurate to an error rate of 1 event per 100 000 bases. It seems to encode only 20 000–25 000 protein-coding genes.[3] These advances should serve as a firm foundation for biomedical research in the decades ahead. Next step, people will be interested in profiling the temporal and special gene expression patterns in the individual human organs or cells on a genome-wide scale and comparing with some normal physiological processes and diseases

Edited by Osamu Ohara

* To whom correspondence should be addressed. Tel. + 86 21-54921263. Fax. + 86 21-54921011. E-mail: sdscli@126.com (J-Y.L.), ylzhang@sibs.ac.cn (Y-L.Z.)

status as well. However, we noticed that although MGC constructed >100 cDNA libraries, derived from a wide variety of tissues and cell lines (complete list at http://mgc.nci.nih.gov), and FLJ constructed 107 human cDNA libraries enriched for full-length cDNA clones, representing 61 tissues, 21 primary cell cultures, and 16 cell lines used to select the putative full-ORF clones, one reproductive organ, epididymis, which is beside the testis, connecting it with the vas deferens, has been neglected by both of the above groups. Mammalian spermatogonia undergo mitosis, meiosis, and some morphological changes to become a fully differentiated sperm in testis. But they are not mature, they pass through the epididymis, a convoluted duct, bathe in the duct lumen, and interact with some proteins in these particular microenvironments so that they acquire forward motility and fertility potential and become a so-called matured sperm stored in the end of this duct waiting for ejaculation. It was also reported that a great number of genes exclusively expressed in this organ. Besides, the epididymal epithelial cells in different regions of the duct, namely caput (the head part), corpus (the middle part), and cauda (the tail part), expressed different proteins secreted into the lumen and created an ever-changed microenvironment. For example, sperm gain their forward motility just in the caput region. Recently, the mouse epididymis has been divided into 10 segments anatomically by lobules bounded by connective tissue septa.[4] Analysis of differentially regulated genes in the mouse epididymis by DNA array demonstrates that more than 2168 genes, ~13% of the genes expressed in the epididymis, differ in their level of expression in different segments by at least fourfold. The large number of regulated genes within a grossly dissected single organ is unprecedented. In comparison, only 1186 genes are differentially regulated (fourfold, $P = 0.01$) between the kidney and the liver, two distinct organs with very different functions. This suggests that sperm maturation, transport, and storage in the epididymis are highly complex events. The more and more people pay attentions to this organ. Because understanding the molecular mechanisms of sperm maturation in epididymis will be of great helps not only in answering how the expression of region-specific programmatically expressed genes was controlled but also in male contraceptive drug design, personalized diagnosis, and treatment of infertility and sperm health evaluation. To this end, human genome U133 plus 2.0 microarray covered over 47 000 transcripts and variants representing 38 500 well-characterized human genes were applied by our lab to obtain a full profile of gene expression in the epididymis of a fertile young male.[5] Genes expressed exclusively or predominantly within one region could be searched from this data resources on the website http://www.scbit.org/human_epididymis_-transcriptomes) conveniently. However, those unknown epididymis-specific genes were still absent on the DNA chip. All of these urged us to complete this defect. Here, we partially sequenced a cDNA library for an adult human epididymis. We want to provide the following informations: (i) a list of the cDNAs expressed in the adult human epididymis; (ii) among them, which genes have been reported, which genes are just a known sequence but without any characterization data, and which genes are totally novel; (iii) are the novel ones human epididymis-specific and even region-specific in human epididymis? The data are not only useful for the epididymis community but also good for addendum the human transcriptome.

2. Materials and methods

2.1. Construction of cDNA library

Human epididymis was obtained from male of age 28 with one-child reproductive history and normal spermatogenesis without the other related diseases (the donor died of accident and were voluntary to donate), which was approved by the Ethics Committee of Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences. Samples were sliced into 1 mm^3 under low temperature, frozen by nitrogen, and stored at $-70^{\circ}C$. Human epididymis total RNA was extracted by Trizol (Invitrogen). The quality was determined by electrophoresis on a 1.1% formaldehyde degenerated agarose gel.

The cDNA library of human epididymis was constructed by Creator SMART cDNA Library Construction Kit (CLONTECH) according to the manufacturer's suggestions. Finally, the human epididymal cDNA was ligated into the vector pDNR-LIB and 5 μL of ligation product was transformed to 25 μL of electrocompetent cell XL1-Blue. The entire transformation mixture was transferred to the tubes containing 970 μL of LB broth. The method of consecutive dilution was used to titer the bacteria solution.[6] After incubating with shaking (225 rpm) for 1 h at 37°C, 1 μL of each transformation mixture was added into the appropriate tube containing 1 mL of LB broth and then 1 μL of dilution was added into 1 mL of LB broth. Combine 1 μL of final dilution with 50 μL of LB broth and spread it onto LB/Cm plates. Inverted the plates and incubated at 37°C overnight. The number of colonies was counted to determine the titer (cfu/mL). Calculated the titer according to the following formulas: colony $\times 10^3 \times 10^3$ cfu/mL.

2.2. Characterization of human epididymis cDNA library

2.2.1. Library amplification and PCR identification

The number of 150 mm plates was calculated based on the library titer. Library diluted by LB medium was spreaded onto LB/Cm plates which were inverted and incubated at 37°C overnight. Fourteen isolated colonies were picked and boiled as template for further PCR identification using M13 Sense and Antisense primer.

2.2.2. Sequencing and bioinformatics analysis

Thousand isolated colonies from the human epididymis cDNA library were first picked and sequenced by 5′-end one-pass using ABI 3100 sequencing system for quality evaluation of this cDNA library. Further, total of 20 000 colonies were sequenced by 5′-end one way. Then the sequencing results were blasted with Nr and EST database published on March 2004. The non-public cDNAs were identified by bilateral sequencing. All the sequencing was subjected to bioinformatics analysis.

Our sequencing results were compared with 11 monkey epididymis genes[7] and six human epididymis specific expression genes HE1-HE6[8] using clustalw software. The unmatched genes are amplified by PCR using cDNA library as template. The primers used for identification of each unmatched genes can be seen in supplementary Table S1.

2.3. Gene expression analysis using northern blot and RT–PCR

Some selected clones were used as probes for further northern analysis. Northern blot analysis was performed according to the procedure described previously.[9] The 10 μg total RNA of each sample was loaded in each lane. An 18S r-RNA hybridization signal was used as a loading control. Epididymis-specific expression of con16, 32, and 33 was further evaluated by RT–PCR using cDNA from caput, corpus, cauda epididymis, and other seven human tissues. The ESC42 and β-actin primers were applied as control.

3. Results and discussions

3.1. Evaluation of the human epididymis cDNA library

The quality of the total mRNAs prepared from an entire human epididymis was judged as A260/A280 = 1.85 and the 28S/18S = 2/1 shown in the Fig. 1A. The size of the resulted ds.cDNAs smeared from 4 kb to 200 bp on the agarose gel as shown in the Fig. 1B.

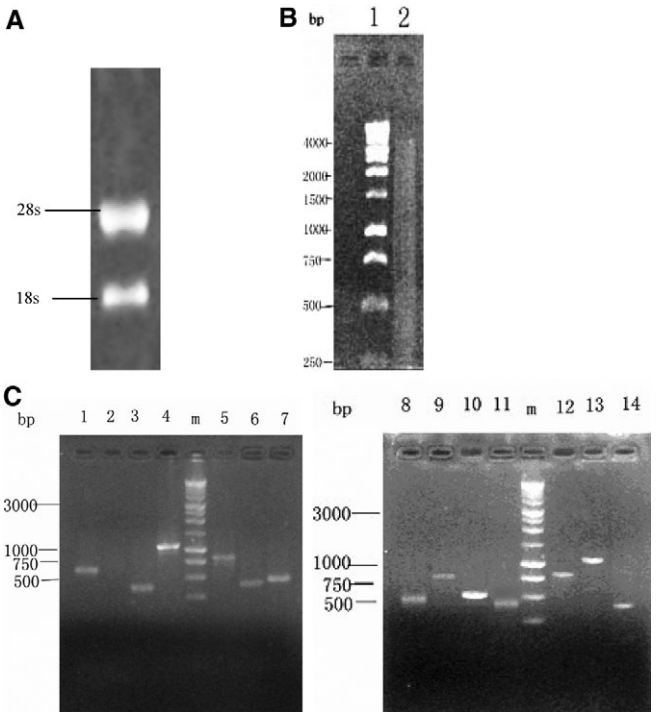

Figure 1. Quality analysis of the human epididymal cDNA library. (**A**) The integrity of the total RNAs of a human epididymis for constructing cDNA library analyzed by agarose gel. The loading amount of RNA is 20 μg; (**B**) LD-PCR products of human epididymis total RNAs. Two microlitter of the ss cDNA served as template for LD-PCR-based, second-strand synthesis using 15 thermal cycles. Lane 1, 1kb DNA ladder; lane 2, LD-PCR product; (**C**) evaluation of the transformation efficiency and the insert length of the cDNA library. Lane m, 1 kb DNA ladder; lane 1 ~ 14, randomly-selected cDNA clones from the library were served as template for PCR amplification.

Creator SMART cDNA Library Construction Kit (CLONTECH) was used to construct the library. It is the case that PCR employed in SMART kit can result in the clone frequencies bias. However, in this work, we have introduced the minimal PCR cycles (18 cycles) as recommended by the manufacturers which would reduce the bias significantly. The titer of this cDNA library is 2×10^7 cfu/mL. The 14 clones randomly selected from this library were examined to evaluate the quality of this library. Fig. 1C showed that 13 of 14 clones were positive. The distribution of the insert size was as follows: >1 kb for two clones, >700 bp for three clones, >500 bp for seven clones, and 400 bp for one clone. The quality of our library might be not very high because it contained a lot of relatively short inserts. However, it is true that the size of many epididymis-specific transcripts is quite small, less than 1 kb, such as at least 19 beta-defensins,[10] 7 lipocalins,[11] HE1, HE4, and HE5.[12]

It indicated that either the transformation efficiency (92.8%) or the insert length was high enough. For further testing the quality of the inserts,

1000 clones were sequenced at the 5′-teminal. The 450 bp sequencing data in average proved that 92.5% clones were positive. Among them, 92% of which could be found in the EST bank (December 2003) and only 8% could not. Seventy-five percentage of the insert sequence was protein coding, corresponding to 693 genes. To this end, we decided to sequence them all.

3.2. *Sequencing data for 20 000 clones and bioinformatics analysis*

As we know, the appearing frequency of the clones with same insert in the library was determined by their expression levels and in a direct proportion manner. So the number of the clones was not equal to the number of genes. And the insert of the clones could not be entirely the full-length cDNA. Owing to the financial limitation, we could not sequence all the clones in the library. According to the transcriptomes data from the human epididymis by DNA array, 26 893 qualifiers as 'present signal > 1' and 15 946 qualifiers as 'present signal > 50' were found in the whole epididymis, accounting for ~49 and 29% of the total on the array respectively were found in the whole epididymis. So, 49 and 29% of the gene number 25 000 which were announced by the International Human Genome Sequencing Consortium would be around 12 250 and 7250 respectively to be expressed in epididymis. The DNA array data also showed that the gene expression level in epididymis was not the same. The variations sometimes were very high. There were 0.4% genes with 'present signal > 5000', 6.8% genes with 'present signal > 500', 58.3% genes with 'present signal > 50'. It indicated that 41.7% of them (with 'present signal > 1') expressed in a very low level might be neglected. Therefore, we chose those genes with 'present signal > 50' holding 29% of the whole gene number 25 000 which will be only ~7250 genes expressed in human epididymis being sent for sequencing. Besides, both MGC and FLJ programs also sequenced 10 000 clones per each cDNA library only. So, we decided to sequence 20 000 clones. It might be enough to reveal those sequences for 29% of the total genes (~7250 genes) expressed in this library. The route map for analyzing these sequencing data was summarized in Fig. 2.

3.2.1. 7450 non-redundant sequences After removal of those recognizable contaminating sequences and the same sequence from different clones, the 5′-terminal one-pass sequences of the 20 000 clones were narrowed down to 7450. These sequences were blasted against NR and EST database published on the March 2003. Six thousand seven hundred and fifty cDNA sequences could be found and the rest 650 ones could not. After further analysis by using tblastx software against NR, EST database, GeneBank, and ProteinBank Swissprot, it was found that 5390 cDNAs had been reported with somewhat various biological functions using the software Genespring 7.0. Table 1 showed the classification of their functions and the numbers of genes belong to. Fig. 3 showed the chromosome localization of the 3394 cDNAs reported. Notably, quite low frequency appeared on chromosome 12 and Y for the epididymis transcripts. Although the rest 1360 cDNA sequences can be found on the database, they were only a predicted sequence resulted from bioinformatics. 650 ESTs could not be found in all the

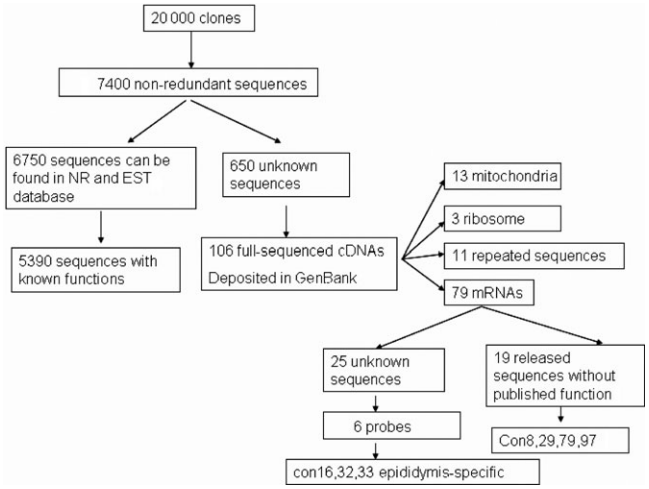

Figure 2. Schematic diagram of sequenced clones from human epididymal cDNA library.

Table 1. The statistical table of 5390 human epididymis sequences

Bio-activities	cDNA number
Apoptosis regulator	263
Cancer	167
Cell cycle	228
Chaperon	183
Enzyme	1191
Immunity protein	83
Motor	43
Nucleic acid binding	519
RNA	123
Signal transducer	173
Storage	152
Structure protein	291
Transport	557
Other group	450
Unclassified	918

databases and they were submitted to GenBank (EH041735-EH041927, FE192591-FE193047). The frequency of each of 7450 clones appearing in the sequenced 20 000 clones was also evaluated. Most of the clones are low-copy clones with the frequency no >10 among 20 000 clones. The contents of medium-copy (frequency 11–100) and high copy (frequency >100) clones are only 1.7% among the 7400 non-redundant clones.

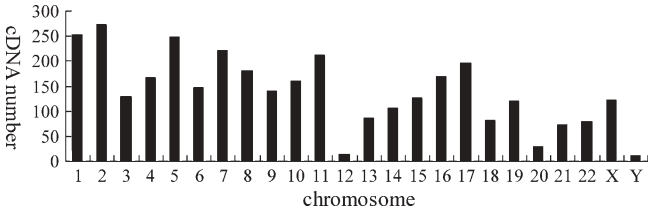

Figure 3. Chromosome location of 3394 cDNA sequences of human epididymis.

3.2.2. 650 EST For further understanding the nature of these so-called novel cDNA fragments, 106 of 650 clones were randomly selected for fully sequencing. The 106 sequences were submitted to GenBank (DV643899 ~ DV643998, DQ822205, DQ823637, DQ823638, EF426753, EF426754,

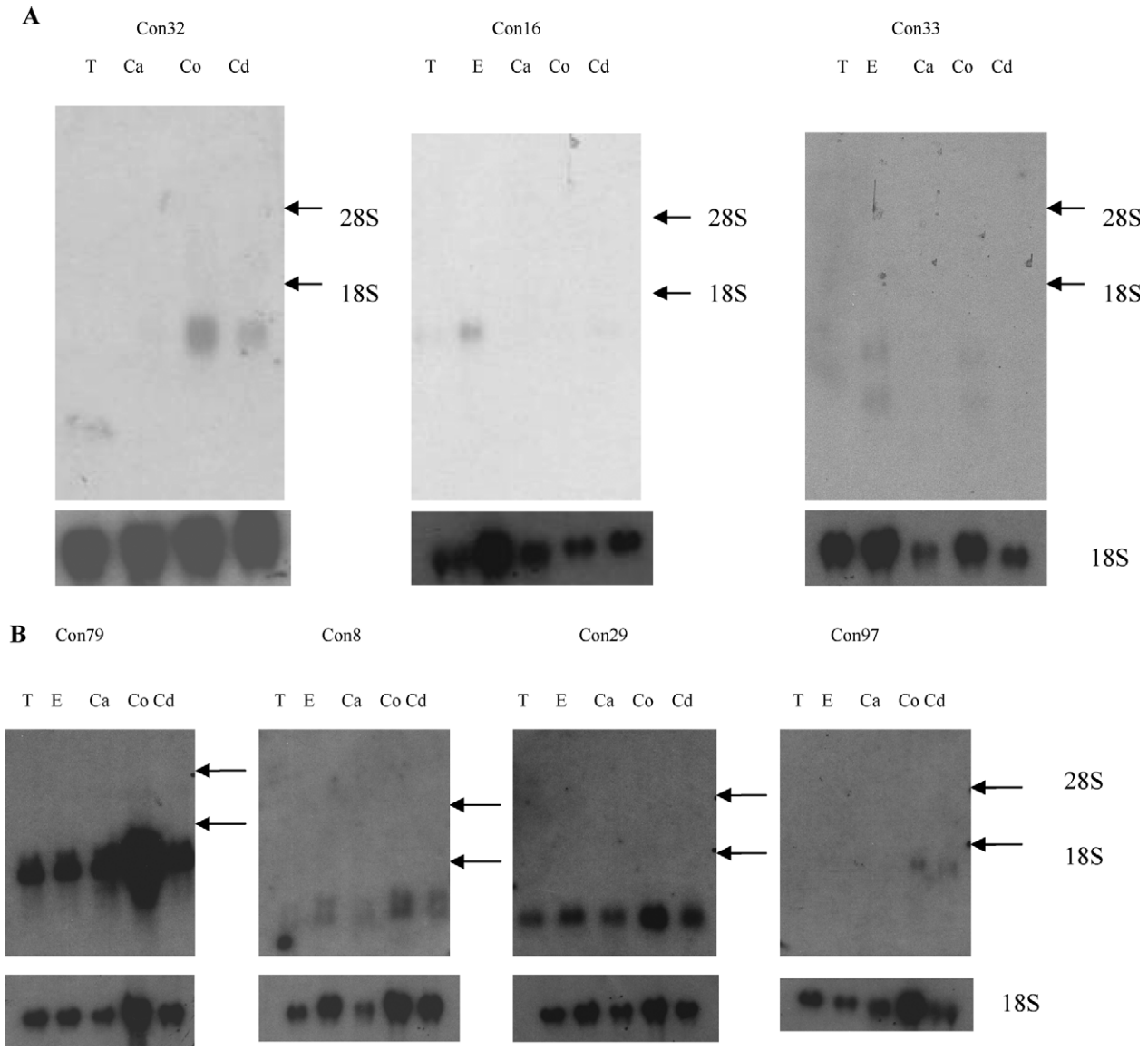

Figure 4. Northern blot validation of the mRNAs in human epididymis using corresponding ESTs as probes. Ten microgram of total RNA extracted from human testis (T), epididymis (E), caput epididymis (Ca), corpus epididymis (Co), cauda epididymis (Cd) were loaded. (**A**) Northern blot with con32, con16, con33, con39, con34, and con42 as probes which were unknown by NCBI NR database searching; (**B**) northern blot with con79, con8, con29, and con97 as probes which can be found in NR database but only full-length cDNA sequence were deposited in GenBank according to NCBI Blast searching.

EF426755). After getting rid of the contaminations or some repeated ones and being pieced together as long as possible, 79 independent unknown sequences were identified by Blast searching against GenBank NR database. These sequences were further searched against human genome database. Fifteen of 79 were found with significant redundant sequence in human genome and seemingly non-coding-transcripts; 8 of 79 clones were surprisingly not found in human genome; 19 of them have already been released during the course of this project but their function is still totally unknown; 25 sequences seemed to be totally unknown and unique in human genome (Fig. 2).

For further checking if these novel genes expressed in human epididymis, 6 of 25 novel clones were selected for northern blot analysis. Con16, 32, and 33 did express in human epididymis even in a region-specific manner and not in testis (Fig. 4A and Table 2). Con32 and 33 were abundant in corpus region and con16 was abundant in cauda epididymis. Their tissue specificity was further tested by RT–PCR

Table 2. Summary of seven clones tested by northern blot assay

Contig name	Length (bp)	Description	mRNA size by northern blot	Region-specific expression in epididymis
con16	587	Unknown	About 1.4 kb	Cauda
con32	553	Unknown	About 1.0 kb	Corpus
con33	198	Unknown	Double mRNA about 1.2 and 0.9 kb, respectively	Corpus
con8	530	Px19-lik protein/cGI-106 1-530	About 1.3 kb	
con29	362	ORF52/ FLJ20451	About 0.7 kb	
con79	421	HSPC337	About 1.2 kb	
con97	495	50-495 in TCTP	About 1.3 Kb	

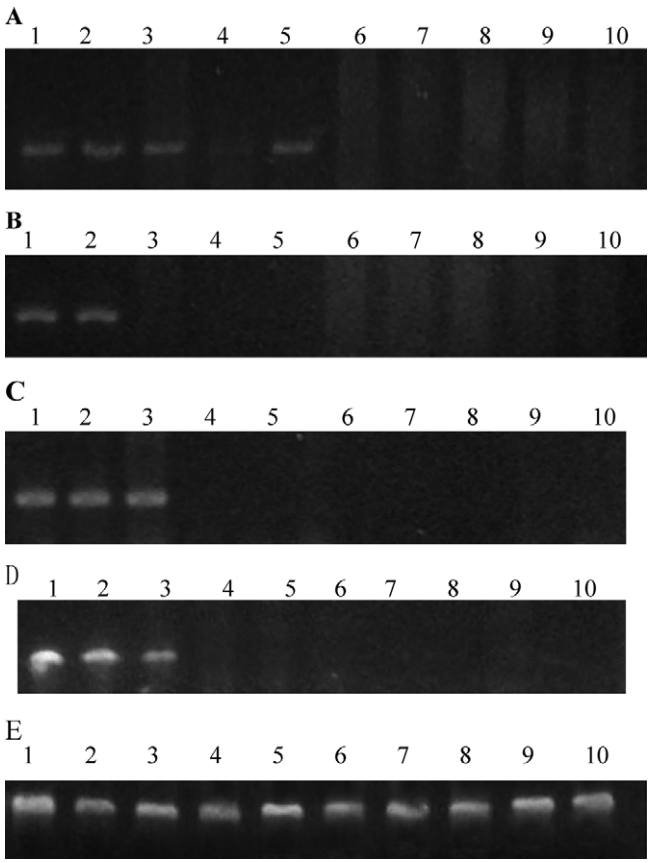

Figure 5. Epididymis-specific expression for con16, con32, con33 using RT–PCR with multiple human tissue cDNA as template. (**A**) Con16; (**B**) con32; (**C**) con33; (**D**) ESC42; (**E**) β-actin. 1, caput epididymis; 2, corpus epididymis; 3, cauda epididymis; 4, heart; 5, lung; 6, liver; 7, spleen; 8, kidney; 9, stomach; 10, testis.

Table 3. Summary of epididymis-specific genes identified in the cDNA library

Gene symbol	Relative signal from cDNA array			Appearance in 20 000 sequenced clones	Not found in 20 000 but can be identified by PCR from the library
	Caput	Corpus	Cauda		
HE1	6396	7178	5338	No	Yes
HE2	6062	5429	268	Yes	
HE3	1038	5433	2988	Yes	
HE4	766	494	3206	No	Yes
HE5	5056	7963	6698	Yes	
HE6	5269	3965	626	Yes	
AEGL1	1037	7011	5105	Yes	
CRES	2	5	3	Yes	
ADAM7	3914	4984	1740	Yes	
ESP13.2	2123	5144	3874	Yes	
GPX5	1071.6	94	35	Yes	
ESC342	3065	835	19	Yes	
LCN6	511	104	39	Yes	
CST11	4944	191	9	Yes	
ESC42	3605	835	19	Yes	
ESC6	**531**	**1922**	**723**	**Yes**	
ESC9	**131**	**791**	**125**	**No**	**Yes**
ESC177	**30**	**44**	**18**	**No**	**Yes**
ESC384	**1694**	**210**	**25**	**No**	**Yes**
ESC615	**57**	**1543**	**496**	**Yes**	
ESC112	**ND**	**ND**	**ND**	**No**	**Yes**
ESC461	**ND**	**ND**	**ND**	**Yes**	
ESC513	**ND**	**ND**	**ND**	**No**	**Yes**

The bold letters indicate the eight novel genes.

(Fig. 5). Epididymis-specific gene ESC42 was used as positive control which has the highest expression in caput epididymis.[7] Con32 and 33 were exclusively expressed in human epididymis and con16 was expressed in human epididymis except lung. These three sequences have been extended by electronic cloning (EST overlapping and extension) and 5′race and have been deposited in GenBank. Con16 belongs to the defensin family and con32 belongs to the colipase family. They might play the important roles in sperm maturation and innate immunity in male reproductive tract. Northern blot analysis was also applied for testing if 4 of 19 released but unknown sequences were present in human epididymis. Although all of them (Con8, 29, 79, 97) did express in human epididymis, it looked neither epididymis-specific nor region-specific (Fig. 4B and Table 2). Since the released sequences were found in the cDNA library from non-epididymis tissues, it is not surprised to find these four clones were not epididymis-specific and also expressed in testis. The information dealing with these clones are summarized in Table 2.

Next, we used BLAST search to compare the sequences for the reminder 544 of 650 clones again with relevant data sets published on the March of 2006. 342 of 544 EST sequence could be found in the NR and EST database during the course of this project. The sequences for the rest 202 clones were regarded as unknown ESTs and 193 of them were submitted to the EST database (EH041735-EH041927).

3.2.3. Qualification of the library

Fifteen reported human epididymis-specific genes and eight monkey epididymis-specific genes newly discovered by our lab were checked in this cDNA library for qualifying its representative. It is happy to see that all the above cDNAs could be detected in this library shown in Table 3 and Fig. 6. However, only 69.6% of them could be found in our sequencing data. This means that sequencing 20 000 clones is not enough to cover all the expressing transcripts in this library. In another words, due to the financial limitation, at least one of three genes in this library had not been sequenced. Nevertheless, Table 3 showed that there is 87% of the 15 reported epididymis-specific genes could be found in our sequencing data, but only 37.5% for the eight newly discovered ones. Notably, it seems that those genes appearing in our sequencing data are basically with higher expression level. It makes sense, because as we mentioned above that the probability for the cDNA clone appearing in the library was determined by their expression levels and in a direct proportion manner. From this point of view, the sequencing data presented in this report might include the majority of the human epididymal transcriptome with relatively higher expression level. Further works based on the novel cDNAs resource will be good for revealing the whole transcriptome for an organ—human epididymis.

In summary, the provided data possess the following significances. (i) Providing a list of the reported genes which are also expressed in epididymis. With the aid of human epididymis cDNA array data on our website, people can know the expression level and region localization of the particular genes in epididymis. (ii) Providing a number of unknown ESTs which are expressed in human epididymis. It is the source making the novel epididymis expressed gene discovery much more convenient. (iii) Providing a number of epididymis-specific transcripts which make up for the deficiency of the updated human transcriptome.

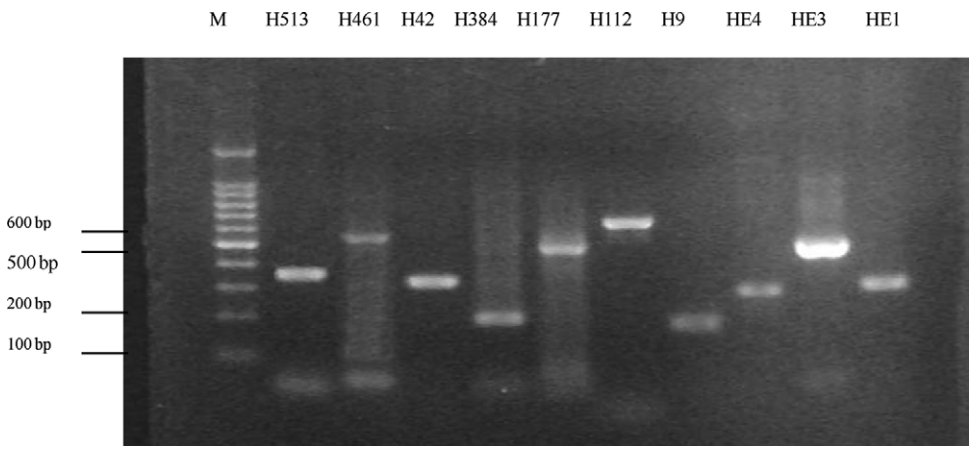

Figure 6. PCR identification of previous identified human epididymis-specific clones which could not found in the 20 000 sequencing data with the human cDNA library as template.

Funding

This work is supported by Shandong Province Science & Technology Key Program (032050102), The 973 Program (2006CB504002) & (2006CB944002), The National Natural Sciences Foundation of China (30230190), Shanghai Science and Technology Funding (05DZ22103), The CAS Knowledge Creative Program (KSCX1-YW-R-54), State 863 High Technology R & D project of China 2004AA221120.

Supplementary Data: Supplementary data are available online at www.dnaresearch.oxfordjournals.org.

References

1. Strausberg, R. L., Feingold, E. A., Grouse, L. H., Derge, J. G., Klausner, R. D., Collins, F. S., Wagner, L., Shenmen, C. M., Schuler, G. D., Altschul, S. F., Zeeberg, B., Buetow, K. H., et al. 2002, Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences, *Proc. Natl. Acad. Sci. USA*, **99**, 16899–16903.
2. Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H., et al. 2004, Complete sequencing and characterization of 21,243 full-length human cDNAs, *Nat. Genet.*, **36**, 40–45.
3. International Human Genome Sequencing Consortium. 2004, Finishing the euchromatic sequence of the human genome, *Nature*, **431**, 931–945.
4. Johnston, D. S., Jelinsky, S. A., Bang, H. J., DiCandeloro, P., Wilson, E., Kopf, G. S. and Turner, T. T. 2005, The mouse epididymal transcriptome: transcriptional profiling of segmental gene expression in the epididymis, *Biol. Reprod.*, **73**, 404–413.
5. Zhang, J. S., Liu, Q., Li, Y. M., Hall, S. H., French, F. S. and Zhang, Y. L. 2006, Genome-wide profiling of segmental-regulated transcriptomes in human epididymis using oligo microarray, *Mol. Cell Endocrinol.*, **250**, 169–177.
6. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. and Struhl, K. 1994, *Current Protocols in Molecular Biology*, John Wiley and Sons: New York.
7. Liu, Q., Hamil, K. G., Sivashanmugam, P., Grossman, G., Soundararajan, R., Rao, A. J., Richardson, R. T., Zhang, Y. L., O'Rand, M. G., Petrusz, P., French, F. S. and Hall, S. H. 2001, Primate epididymis-specific proteins: characterization of ESC42, a novel protein containing a trefoil-like motif in monkey and human, *Endocrinology*, **142**, 4529–4539.
8. Kirchhoff, C. 1999, Molecular characterization of epididymal proteins, *Rev. Reprod.*, **3**, 86–95.
9. Li, P., Chan, H. C., He, B., So, S. C., Chung, Y. W., Shang, Q., Zhang, Y. D. and Zhang, Y. L. 2001, An antimicrobial peptide gene found in the male reproductive system of rats, *Science*, **291**, 1783–1785.
10. Patil, A. A., Cai, Y., Sang, Y., Blecha, F. and Zhang, G. 2005, Cross-species analysis of the mammalian beta-defensin gene family: presence of syntenic gene clusters and preferential expression in the male reproductive tract, *Physiol. Genom.*, **23**, 5–17.
11. Suzuki, K., Lareyre, J. J., Sánchez, D., Gutierrez, G., Araki, Y., Matusik, R. J. and Orgebin-Crist, M. C. 2004, Molecular evolution of epididymal lipocalin genes localized on mouse chromosome 2, *Gene*, **339**, 49–59.
12. Kirchhoff, C., Osterhoff, C., Pera, I. and Schröter, S. 1998, Function of human epididymal proteins in sperm maturation, *Andrologia*, **30**, 225–232.