Sequence Analysis of Libraries from Individual Human Blastocysts

G. MOROZOV,¹ O. VERLINSKY,¹ S. RECHITSKY,¹ V. KUKHARENKO,^{1,2} E. GOLTSMAN,¹
V. IVAKHNENKO,¹ V. GINDILIS,¹ C. STROM,¹ A. KULIEV,¹ and Y. VERLINSKY^{1,3}

Submitted: January 7, 1998

Accepted: January 7, 1998

Purpose: It has recently become possible to construct cDNA libraries from individual human blastocysts to investigate the expression of embryonic genes in human preimplantation development. We have previously reported the expression of β -actin, CD-59, and homeobox OCT-3 and identified almost-complete homology of sequences to human histone 3.1 and human ribosome protein S25. In the present paper, our further sequencing analysis of cDNA libraries from single human blastocysts is described.

Methods: cDNA libraries were constructed from 13 blastocysts. Sequence analysis was performed in 120 clones from one of these cDNA libraries with fragments of 50 to 1000 bp. Their sequence identity was analyzed using the expressed sequence tag (EST) database.

Results: The presence of two housekeeping genes, hexokinase 1 and serine/threonine phosphorylase, and four other ESTs was demonstrated, the identity of which, with particular gene expression in preimplantation development, has not yet been established.

Conclusions: The data demonstrate the usefulness of constructing cDNA libraries from individual human blastocysts and their value in the analysis of genetic expression in human preimplantation development.

KEY WORDS: human blastocyst; cDNA library; housekeeping genes; expressed sequence tags.

INTRODUCTION

It has recently become possible to construct cDNA libraries from individual human blastocysts to investi-

gate the expression of embryonic genes in human preimplantation development (1,2). Using these libraries, we have previously reported the expression of β -actin, CD-59, and homeobox OCT-3 and also identified almost-complete identity of the cloned sequences to human histone 3.1 and ribosome protein S25 (2). As the information on genetic expression in human preimplantation development is still limited to particular genes or gene families (3–5), we present here our further results on a systematic study of the expression of genes involved in preimplantation development through the construction of cDNA libraries from individual human blastocysts. The present paper describes the results of our study on the identification of six expressed sequence tags (ESTs), two of them corresponding to housekeeping genes, hexokinase 1 and serine/threonine phosphorylase, and four corresponding to ESTs found by other authors in fetal brain and pancreatic tumor cells, whose actual identities have not yet been established.

MATERIALS AND METHODS

Spare embryos were obtained from the IVF program at Illinois Masonic Medical Center. These were spare embryos or those that would have been discarded otherwise because of an aneuploidy discovered by preimplantation genetic analysis of the oocytes. The study protocol was approved by the Institutional Review Board at Illinois Masonic Medical Center.

The cDNA libraries were constructed by the amplification method using poly(dT) primer and dG-tailing as an adaptor, as described elsewhere (2). The principal steps in a library construction are shown in Fig. 1.

Primers for reverse transcription (RT) polymerase chain reaction (PCR) were synthesized using an Applied Biosystems 381A DNA synthesizer (Camden, NJ). The primer design optimized their compatibility

¹ Reproductive Genetics Institute, 836 West Wellington Avenue, Chicago, Illinois 60657.

² National Research Center for Medical Genetics, RAMS, Moscow, Russia.

³ To whom correspondence should be addressed.

Fig. 1. Principal steps in constructing cDNA libraries.

and minimized potential hairpin formation and primer-dimer formation using the OLIGO program software (6). Primer sequences and conditions of PCR are pre-

sented in Table I. The amplified cDNAs were enriched to a high molecular fraction by chromatography on a minicolumn with Sephacryl S-400 (USB Corp. (7).

Table I. Designed Primers and PCR Conditions

Primer	Sequence	Gene	PCR conditions		
			T _{anneal} (°C)	[MgCl ₂] (mM)	t _{elongation} (min)
Primers for 1st-round PCR					
Lyb-1	GCGCCGCGGCCGC(T) ₁₈		62	6	3
Lyb-2	GCCGTCGACGCGT(C) ₁₅				
Primers for 2nd-round PCR					
Lyb-3	GCGCCGCGGCCGCTT		62	6	3
Lyb-4	CGGAATTCGCCGTGGACGCGTCC				
act-up	TCCAGGGCGACGTAGCACAGCTT	β-actin	60	2.0	1
act-dn	CGACGAGGCCAGAGCAAGAGA				
cd59-up	CTGCTGACTGCAAAACAGCC	CD-59	49	2.0	1
cd59-dn	CCACCATTTCGAAGCTGTTC				
mlh-up	AGCCTCAGTAAAGAATGCG	MLH1	50	2.0	1
mlh-dn	ATAAAGGAATACTATCAGAAGGC				
hlag-up	TTCCCAGAGCAGTCTTCC	HLA-G	51	2.0	1
hlag-dn	GGGATCAGCCTCCTTCG				
VH4	GGTGTACCCCTGGATGAAG	HOXA4	47	2.0	1
DH2	CCGGTTCTGGAACCAGAT				
pou-up	CCAAGCTCCTGAAGCAGAAGAGG	OCT-3	49	2.0	1
pou-dn	CTTCTGGCGCCGTTGCAGAAC				

Table II. Gene Expression in Individual Human Blastocysts Identified by PCR Analysis of DNA from cDNA Libraries and RT-PCR Products [Data Were Partly Reported Elsewhere (2)]

Library No.	Blastocyst No.	Presence of sequences corresponding to					
		β-Actin	CD59	Oct-3	HoxA4	Hla-G	Mlh-1
cDNA libraries of 3' tags							
1	1	+	—	+	—	—	—
2	3	+	+	+	—	—	—
3	4	+	+	+	—	—	—
4	5	+	+	—	—	—	—
5	11	+	+	+	—	—	—
6	13	—	+	—	—	—	—
7	15	+	—	+	ND	ND	ND
8	18	+	—	—	ND	ND	ND
9	21	—	—	+	ND	ND	ND
10	28	—	+	—	ND	ND	ND
11	30	+	+	+	ND	ND	ND
12	32	—	—	+	ND	ND	ND
13	34	+	—	—	ND	ND	ND
RT-PCR of individual blastocysts							
	22	+	+	+	—	—	—
	23	+	—	+	—	—	—
	24	+	+	+	—	—	—
	25	+	+	+	—	—	—

Fig. 2. PCR products after second-round PCR obtained from different individual human blastocysts. Lane: (1) 123 ladder; (2-5) different blastocysts; (6) RT-PCR of water (negative control); (7) PCR of water (negative control).

The cDNAs amplified with primer set Lyb3/Lyb4 were digested using the restriction endonucleases *NotI*-*EcoRI*, and ligated with the λ gt11 Sfi-Not vector. Approximately 1.2×10^6 plaque forming units (pfu) was obtained. Restriction nuclease digestion and ligation were performed as described by Sambrook *et al.* (8).

To propagate the library, 50 ml of Packagene extract (Packagene System; Promega) was added to 10 ml of ligated mixture, followed by incubation at room temperature for 3 hr. The mixture of the assembled phage particles was evaluated for quantity and propagated in BNN102 strain (C600Hfl), and the total phage DNA of each library was isolated and purified by

Fig. 3. Repartition of ESTs according to the sequence comparison with gene databases. Each category is represented as a percentage of the total number of analyzed clones.

Table III. Expressed Sequence Tags from One of the Libraries Showed Identity with Those Deposited in Genetic Databases

No. of ESTs	Correspondence to ESTs from GenBank	Accession No.
14	Human hexokinase I	M75126
26	EST overexpressed in human pancreatic cancer cells	U46224
49	EST expressed in human fetal brain	X65376
67	Human Ser/Thr phosphorylase	M29074
88	EST expressed in human fetal brain	X65389
114	EST overexpressed in human pancreatic cancer cells	U46454

immunoprecipitation (7). Plasmid DNA was propagated by transforming the bacterial strain INV-alphaF' (9) and isolated by boiling (10). The PCR inserts were sequenced by the dideoxy Sanger method (11), modified for double-stranded DNA (12) using bacteriophage T7 DNA polymerase (13).

RESULTS AND DISCUSSION

Thirteen cDNA libraries were constructed from individual human blastocysts at different stages of development. The results of PCR analysis of the RT-PCR products and the library cDNAs are summarized in Table II and Fig. 2. cDNA libraries from nine embryos showed the presence of mRNA for β -actin; seven, for CD-59; and eight, for OCT-3 (2). Altogether these three genes (β -actin, CD-59, and OCT-3) were present in only six blastocysts. In contrast, specially designed intron-spanning primer sets failed to reveal mRNA for the HOXA4, HLA-G, and mLH-1 genes, which were expected to be present based on previous studies in the human and mouse (14,15).

As reported previously, we found a sequence identical to human histone 3.1 in one of the cDNA libraries, as well as a sequence identical to human ribosome protein S25 in the other (2). The sequencing analysis of 120 clones from the third library, with fragments of 50 to 1000 bp, is shown in Fig. 3. Figure 3 shows that 5% of the clones contained sequences identical to known human sequences, while 8% were "junky" clones, and 87% contained sequences that showed only partial identity to the known human sequences. It should be mentioned that no clones were found that contained sequences other than the human ones. Of those six clones with sequences identical to known human sequences, two contained housekeeping genes, hexokinase I and Ser/Thr phosphorylase, and four other ESTs (Table III), which had been detected previously in fetal brain and pancreatic tumor cells (16).

The data demonstrate the usefulness of constructing cDNA libraries from individual human preimplanta-

tion embryos and the value of the analysis of single-pass sequencing reactions for library evaluation. The availability of a publicly accessible database for ESTs from cDNA libraries constructed from different human tissues at different stages of development (17) makes the EST analysis of cDNA libraries from individual blastocysts an important part in the analysis of gene expression in human preimplantation development. The four unidentified ESTs in human blastocysts may represent developmentally important genes. Work is under way to obtain full-length clones for these sequences to identify their function.

REFERENCES

1. Adjaye J, Daniels R, Bolton V, Monk M: cDNA libraries from single human preimplantation embryos. *Genomics* 1997;46:337-344
2. Verlinsky Y, Morozov G, Verlinsky O, *et al.*: Isolation of cDNA libraries from individual human preimplantation embryos. *Mol Hum Reprod* 1998 (in press)
3. Campbell S, Swann HR, Seif MW, *et al.*: Cell adhesion molecules in the oocyte and preimplantation human embryo. *Hum Reprod* 1995;10:1571-1578
4. Jay P, Diriong S, Taviaux S, *et al.*: Isolation and regional mapping of cDNAs expressed during early human development. *Genomics* 1997;39:104-108
5. Taylor CT, Johnson PM: Complement-binding proteins are strongly expressed by human preimplantation blastocysts and cumulus cells as well as gametes. *Mol Hum Reprod* 1996; 2:52-59
6. Rychlik W: OLIGO 4.1 Primer Analysis Software
7. Titus D: Promega Protocols and Application Guide. Promega Co., 1995
8. Sambrook J, Fritsch EF, Maniatis T: *Molecular Cloning: A Laboratory Manual*, 2nd ed. Cold Spring Harbor, NY, Cold Spring Harbor Laboratory Press, 1989
9. Hanahan D: Studies on transformation of *Escherichia coli* plasmids. *J Mol Biol* 1983;166:557-580
10. Holmes DS, Quigley M: A rapid boiling method for the preparation of bacterial plasmids. *Anal Biochem* 1982;114:193-197
11. Sanger F, Nicklen S, Coulson AR: DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* 1977; 74:5463-5467
12. Chen EY, Seeburg PH: Supercoil sequencing: A fast and simple method for sequencing plasmid DNA. *DNA* 1985;4:165-170

13. Tabor S, Richardson CC: DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. *Proc Natl Acad Sci USA* 1987;84:4767-14771
14. Verlinsky Y, Morozov G, Gindilis V, *et al.*: Homeobox gene expression in human oocytes and preembryos. *Mol Reprod Dev* 1995;41:127-132
15. Xu Y, Jin P, Mellor AL, Warner CM: Identification of the *Ped* gene at the molecular level: The Q9 MHC class I transgene converts the *Ped Slow* to the *Ped Fast* phenotypes. *Biol Reprod* 1994;51:695-699
16. ENTREZ, Document Retrieval System, Release 23.0. National Center for Biotechnology Information, 1996
17. Hillier LD, Lennon G, Becker M, *et al.*: Generation and analysis of 280,000 human expressed sequence tags. *Genet Res* 1996;6:807-828