The Construction of cDNA Libraries from Human Single Preimplantation Embryos and Their Use in the Study of Gene Expression During Development

JAMES ADJAYE,^{1,2} ROB DANIELS,¹ and MARILYN MONK¹

Submitted: January 14, 1998

Accepted: January 14, 1998

Purpose: The construction and application of polymerase chain reaction (PCR)-based cDNA libraries from unfertilized human oocytes and single preimplantation-stage embryos are described. The purpose of these studies is to provide a readily available resource for the study of gene expression during human preimplantation development.

Methods: Rapid, reproducible, and efficient procedures for the construction of PCR-based cDNA libraries from fewer than 10 cells were first developed in small populations of fibroblast cells. We then constructed cDNA libraries from eight unfertilized oocytes and single two-cell, -4-cell, -7-cell, and blastocyst-stage embryos. Differential display PCR using the libraries as template allows the analysis of stage-specific expression of embryonic genes. Genomic libraries are also prepared from parental samples (cumulus cells and sperm) corresponding to the individual embryo generating the cDNA library.

Results: The complexities (between 10^5 and 10^6 clones) of the human embryo libraries indicate that they may represent the entire active gene population at these early stages of human development. Nucleotide sequence analyses of random clones showed the presence of a variety of transcripts, such as the human transposable element, LINE-1, Alu repeat sequences, housekeeping genes, and tissue-specific genes, (e.g., α -globin, FMR-1, and interleukin-10). Also present at the expected frequency are the ubiquitous cytoskeletal elements, β -actin, keratin-18, and α -tubulin. In addition to cDNAs corresponding to known expressed sequence tags (ESTs) in the GenBank and dbEST databases, a high proportion of novel sequences was also detected. Several cDNAs were detected only at specific stages of preimplantation development by the differential display analysis.

Conclusions: PCR-based cDNA libraries from single human preimplantation embryos provide a new and important

resource for the identification and study of novel genes or gene families. As such, they will increase our basic understanding of the molecular control of human development. In the clinical context, the libraries identify the time of onset of specific genes, and hence the diseases resulting from mutation of these genes, and provide information about new methods of preimplantation diagnosis. The molecular analysis of early gene transcription in human embryogenesis may be expected to lead to advances in contraception, assisted reproduction, and preimplantation genetic diagnosis.

KEY WORDS: human preimplantation embryos; cDNA libraries; gene expression; genomic imprinting; zygotic gene activation.

INTRODUCTION

Analysis of the molecular basis of human development will lead to a greater understanding of abnormal function leading to congenital abnormality and inherited genetic disease. So far, detailed studies of gene expression in preimplantation human embryos have been limited to previously characterized genes using reverse transcription-polymerase chain reaction (RT-PCR). Such analyses have been hampered by the scarcity and small size of human embryos, and there has been little opportunity to screen for the many human gene sequences that are rapidly becoming available (1). Furthermore, a major limitation of current procedures is that novel genes cannot be identified and characterized. To circumvent these difficulties, we have constructed stage-specific cDNA libraries for early human development (2). Subtractive hybridization and differential display PCR can now be used to identify key regulatory genes participating in the early developmental program.

All procedures for the construction of the human embryo cDNA libraries must be ultrasensitive. To

¹ Molecular Embryology Unit, Institute of Child Health, 30 Guilford Street, London WC1N 1EH, UK.

² To whom correspondence should be addressed.

achieve this sensitivity and to generate the large number of clones required to span the gene products of a given embryonic cell, it is necessary to use PCR procedures (3) to generate a high enough concentration of cDNAs for the construction of the libraries. Such methods are now being exploited to answer various biological questions such as the mechanisms underlying cell differentiation in postimplantation mouse embryos (4), in preimplantation and postimplantation human embryos (2,5), in the isolation of novel genes expressed during mouse preimplantation embryogenesis (6–8), and in quail embryo neural crest cells (9).

This paper describes the construction, characterization, and applications of human oocyte, two-cell-, four-cell-, seven-cell-, and blastocyst-stage-embryo libraries with complexities between 1.0×10^5 and 2.5×10^6 clones. Applications of the libraries to several areas of interest, such as expression of CpG island-containing “tissue-specific” genes, developmental genes expressed in a stage-specific manner, and a search for monoallelic expression of imprinted genes are described. The libraries are a valuable resource for the study of gene expression during human preimplantation development and, thus, obviate the need for research on the human embryos themselves.

MATERIALS AND METHODS

Preimplantation Embryos

Anonymized embryos derived by *in vitro* fertilization or intracytoplasmic sperm injection, and spare to requirement for infertility treatment, were donated by patients attending the Assisted Conception Unit (ACU) at King's College Hospital (KCH), London. For each embryo used, the corresponding parental samples (sperm and cumulus cells) were obtained. All embryos were lysed at the ACU before transfer to ICH. The research was carried out under license from the HFEA (Human Fertilisation and Embryo Authority).

Construction and Qualitative Analysis of the Libraries

The protocol for cDNA synthesis and library construction is essentially that described by Beliaevsky *et al.* (10), with some modifications to increase efficiency. For full details see Adjaye *et al.* (2).

Differential Display PCR Analysis of cDNA Libraries

Lambda DNA was obtained by the plate lysate method (11) from approximately 1.0×10^5 plaque forming units of the four-cell and seven-cell libraries. PCR was performed as described (2,12) with modifications involving the use of vector specific primers KS and SK (2), end-labeled using T₄ polynucleotide kinase and $\gamma^{32}\text{P}$ -dATP (Amersham). The $\gamma^{32}\text{P}$ -labeled PCR products were resolved on 6% polyacrylamide, 8 M urea sequencing gels. Dried gels were exposed to X-ray film for at least 18 hr.

RESULTS

Qualitative Analysis of the Libraries

The measured complexities of the oocyte, two-cell, four-cell, seven-cell, and blastocyst libraries were 2.0×10^6 , 1.0×10^5 , 5.0×10^5 , 3.5×10^6 , and 2.5×10^6 clones, respectively. cDNA libraries based on PCR procedures could show a disturbance in relative gene frequencies due to variable PCR efficiencies. However, estimated clone frequencies of the cytoskeletal elements β -actin, α -tubulin, and keratin-18, and of the repetitive sequence, LINE-1, in the embryo libraries show the proportional representation of transcribed genes expected in human cells (Table I). The tissue-specific α -globin gene is also present among the expressed sequences in the four-cell cDNA library (frequency, 0.003%). This is in accordance with our earlier observation of the presence of this transcript in fresh human embryonic material (13).

To isolate and identify other transcripts present in the oocyte, two-cell-, four-cell-, seven-cell-, and blastocyst-stage libraries, randomly picked clones ranging in size up to 400 bp were sequenced. Sequence homologies were analyzed with the use of the Basic Local Alignment Search Tool (BLAST) algorithm (14) to search GenBank, EMBL and the dbEST databases. Some of the interesting findings emerging from the analysis of the sequenced clones include evidence for expressed repetitive sequences, an abundance of transcripts also found in brain and skeletal muscle, sequences with homologies to oncogenes, transcription factors, cytokines, homeobox, and stem-cell genes, and novel transcripts with no homology to sequences in the databases. For more details see Adjaye *et al.* (2). Matching cDNAs show 60 to 99% nucleotide sequence identity with known genes in databases.

Table I. Abundance of cDNA Clones for Various Ubiquitous Cytoskeletal Elements^a

	β -Actin	α -Tubulin	Keratin-18	LINE-1	rRNA	α -Globin
Clones screened	1.5×10^4	3.0×10^4	3.0×10^4	1.5×10^4	1.4×10^4	3.0×10^4
Clone frequency (%)	0.05	0.006	0.006	0.6	0.03	0.003

^a The clone frequencies were determined by colony/plaque hybridization using probes described by Adjaye *et al.* (2).

Identification of Novel Embryonic Stage-Specific Transcripts Using Differential Display PCR

To identify candidate genes that may play a role in embryonic genome activation, and therefore preimplantation development, differential display PCR analysis, using cDNAs from staged embryo cDNA libraries as template, was employed to indicate genes expressed at one or the other stage (Fig. 1). Several modifications of the method of Liang and Purdee (12) and Zimmermann and Schultz (15) were introduced to overcome the limitations imposed by the extremely low number of preimplantation embryos available for our research purposes and the minute amount (picograms) of mRNA present in single embryos. Figure 1 shows several bands (highlighted by asterisks) amplified from the four-cell library but not the seven-cell library, and vice versa. For further characterization of these bands, the DNA is eluted, reamplified, cloned, and sequenced.

Similar stage-specific cDNAs have been demonstrated in oocyte, two-cell, and blastocyst-stage libraries (Hurley, Adjaye, and Monk, unpublished). Confirmation of stage-specific expression involves the use of gene-specific primers on the libraries themselves and RT-PCR on fresh embryos. Northern blots using a variety of fetal and adult human tissues to provide mRNA are performed to determine whether the expression of the embryo stage-specific genes is confined to the embryo or whether expression also occurs at later developmental stages and in the adult. These experiments are in progress.

DISCUSSION

The construction of PCR-based cDNA libraries prepared from eight human oocytes and single two-cell-, four-cell-, seven-cell-, and blastocyst-stage embryos is described. The complexities of the oocyte, two-cell, four-cell, seven-cell, and blastocyst libraries were between 1.0×10^5 and 2.5×10^6 clones, and we may expect that the libraries are representative given that the human genome is estimated to consist of 50,000 to 100,000 genes (1). This reproducibly high

yield of clones is made possible by the single-tube procedure described by Adjaye *et al.* (2).

In-depth nucleotide sequence analyses of a number of transcripts that bear homologies of between 60 and 86% with entries in the databases (2) suggest that these embryonic transcripts represent new members of gene families. These genes may be similar, but not necessarily identical, to genes of known function in other human tissues and therefore may turn out to be novel embryonic genes.

Applications of the Human Embryo cDNA Libraries

Rapid progress in the study of gene expression during preimplantation development in humans has been hindered by the inherent limitation in the amount of biological material present in embryos. Now, with the availability of these embryonic cDNA libraries, research into early human development can be initiated in many areas. Here, we consider three areas: zygotic gene activation, regulation of CpG island gene expression, and genomic imprinting.

Zygotic Gene Activation. One of the major goals of developmental biology is to identify mechanisms that regulate the onset of zygotic gene transcription/translation and thus the initiation of the developmental program. During oogenesis, the growing mammalian oocyte synthesizes and accumulates organelles, proteins, and mRNAs that constitute the maternal contribution toward early development. As a result of oocyte maturation and fertilization, this contribution both supports and directs early development, at least until the time of activation of the embryonic genome (10,17). Despite the fundamental importance of zygotic gene activation, very little is known about the molecular basis of this in mammals (especially humans) and the repertoire and biochemical functions of genes that are activated. To isolate genes specifically expressed from the embryonic genome, three approaches can be used.

Differential Display PCR Analysis. This technique, developed by Liang and Pardee (12), is an effective method that can be used to identify and separate

Fig. 1. A section of an autoradiogram of PCR-amplified cDNA fragments obtained from differential display PCR using the four-cell and seven-cell cDNA libraries as template. The primers used are KS and SK, which flank the unique *Eco*R1 site present in λ -ZAPII and pBluescript (2). The samples in the 4c and 7c lanes at the left were resolved for 2 hr and identical samples at the right for 5 hr. Exposure was for 18 hr at room temperature. The asterisks point to specific cDNAs (transcripts) exhibiting differential stage-specific expression. Regions of higher intensity may be due to an accumulation of a greater number of cDNAs of a similar size range.

cDNAs that are differentially expressed between various cell types or, in this case, specific embryonic stages using the libraries as templates for the reaction. The profiles of genes expressed at each embryonic stage

during preimplantation development can be displayed and visualized on a single gel as demonstrated in Fig. 1.

Construction of Subtraction Libraries. Stage-specific subtraction libraries enriched for transcripts specific for a particular embryonic stage can be generated using individual pools of cDNAs synthesized from embryos at that stage and using the new and highly effective method, termed suppression subtractive hybridization (SSH). This is a powerful technique developed for the generation of subtracted cDNA libraries based primarily on a recently described technique called suppression PCR and combines normalization and subtraction in a single procedure (18).

Aspects of Regulation of Embryonic Gene Expression—The Role of CpG Islands. Daniels *et al.* (13,19) previously showed that myotonin protein kinase (a predominantly muscle-specific gene) and α -globin (a red blood cell-specific gene) are expressed during early human development and proposed that CpG-island genes may be expressed at these early stages. They hypothesized that this expression may prevent methylation and, hence, mutation [deamination of methylated CpG to TpG (20)] of these sequences. In keeping with this hypothesis, transcripts of keratin-18 (an epithelial cell-specific intermediate filament protein), *SNRPN* (an imprinted, CpG island-containing gene expressed predominantly in brain and heart (21, 22), and several other CpG island-containing tissue-specific genes [see Adjaye *et al.* (2)] have been identified in our embryonic libraries.

These results, clearly backed by direct research on the human embryos themselves, demonstrates the usefulness of the embryo libraries for use in studies on the regulation of embryonic gene expression.

Genomic Imprinting. Imprinted genes are expressed differentially according to whether they are inherited from the egg or sperm (23). The genes have identical nucleic acid sequence but their function is determined by the parent of origin. It follows that imprinted genes are differentially modified in egg and sperm.

An interesting question is whether both parental alleles of an imprinted gene are expressed before the selective silencing of one or the other allele. For the investigation of this question, the embryonic libraries, together with genomic libraries prepared from parental samples (cumulus cells and sperm), corresponding to the individual embryo generating the cDNA library, can be utilized. Polymorphism(s) within the 3' end of the coding sequence or 3'-UTR of the imprinted gene must be identified in the parental genomic DNA libraries for use as a marker(s) to distinguish between

mRNAs derived from each parental copy of the gene in the embryonic libraries. Huntriss *et al.* (submitted) have recently demonstrated monoallelic paternal expression of the imprinted *SNRPN* gene in human preimplantation embryos. *SNRPN* expression has also been detected in the cDNA libraries (Adjaye, Ribugent, and Monk, unpublished), thus confirming the feasibility of studying the expression of imprinted genes using the embryo libraries.

CONCLUDING REMARKS

We have demonstrated that the construction of single human embryo cDNA libraries is feasible and that these libraries provide a limitless resource for the identification and study of novel and previously characterized genes. The libraries are also an important new resource for the investigation of the molecular programming of early mammalian development.

ACCESS TO LIBRARIES

Queries regarding the use of these libraries should be directed to Dr. James Adjaye.

ACKNOWLEDGMENTS

The use of embryos is covered by HFEA license number R0063 to Dr. Virginia Bolton. This work was supported by the Medical Research Council.

REFERENCES

- Adams MD, Kelley JM, Gocayne JD, Dubnick M, Polymeropoulos MH, Xiao H, Merrill CR, Wu A, Olde B, Moreno RF, Kerlavage AR, McCombie WR, Venter JC: Complementary DNA sequencing: Expressed sequence tags and human genome project. *Science* 1991;252:1651–1656
- Adjaye J, Daniels R, Bolton V, Monk M: cDNA libraries from single human preimplantation embryos. *Genomics* 1997;46:337–344
- Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB, Erlich HA: Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 1988;239:487–491
- Smith DE, Gridley T: Differential screening of a PCR-generated mouse embryo cDNA library: Glucose transporters are differentially expressed in early postimplantation mouse embryos. *Development* 1992;116:555–561
- Buraczynska MJ, Van Keuren ML, Buraczynska KM, Chang YS, Crombez E, Kurnit DM: Construction of human embryonic cDNA libraries: HD, PKD1 and BRAC1 are transcribed widely during embryogenesis. *Cytogenet Cell Genet* 1995;71:197–202
- Rothstein JL, Johnson D, DeLoia JA, Skowronski J, Solter D, Knowles B: Gene expression during preimplantation mouse development. *Genes Dev* 1992;6:1190–1201
- Temeles GL, Ram PT, Rothstein JL, Schultz RM: Expression patterns of novel genes during mouse preimplantation embryogenesis. *Mol Reprod Dev* 1994;37:121–129
- Weng DE, Morgan RA, Gearhart JD: Estimates of mRNA abundance in the mouse blastocyst based on cDNA library analysis. *Mol Reprod Dev* 1989;1:233–241
- Bevan SG, Southey MC, Armes JE, Venter DJ, Newgreen DF: Spatiotemporally exact cDNA libraries from Quail embryos: A resource for studying neural crest development and neuro-cristopathies. *Genomics* 1996;38:206–214
- Beliaevsky A, Vinogradova T, Rajewsky K: PCR-based cDNA library construction: General cDNA libraries at the level of a few cells. *Nucleic Acids Res* 1989;17:2919–2932
- Sambrook J, Fritsch EF, Maniatis T: *Molecular Cloning: A Laboratory Manual*, 2nd ed. Cold Spring Harbor Laboratories Press, Cold Spring Harbor, NY, 1989
- Liang P, Pardee AB: Differential display of eucaryotic messenger RNA by means of the polymerase chain reaction. *Science* 1992;257:967–971
- Daniels R, Lowell S, Bolton V, Monk M: Transcription of tissue-specific genes in human preimplantation embryos. *Hum Reprod* 1997;12:2251–2256
- Altschul SF, Gish W, Miller W, Myers EW, Lipman, DJ: Basic local alignment search tool. *J Mol Biol* 1990;215:403–410
- Zimmermann JW, Schultz RM: Analysis of gene expression in the preimplantation mouse embryo: Use of mRNA differential display. *Proc Natl Acad Sci USA* 1994;91:5456–5460
- Nothias J-Y, Majumder S, Kaneko KJ, DePamphilis ML: Regulation of gene expression at the beginning of mammalian development. *J Biol Chem* 1995;270:22077–22080
- Schultz RM: Regulation of zygotic gene activation in the mouse. *BioEssays* 1993;15:531–538
- Diatchenko L, Lau Y-F C, Campbell AP, Chenchik A, Moqadam F, Huang B, Lukyanov S, Lukyanov K, Gurskaya N, Sverdlov ED, Siebert, PD: Suppression subtractive hybridisation: A method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proc Natl Acad Sci USA* 1996;93:6025–6030
- Daniels R, Kinis T, Serhal P, Monk M: Expression of the myotonin protein kinase gene in preimplantation human embryos. *Hum Mol Genet* 1995;4:389–393
- Cross SH, Bird AP: CpG islands and genes. *Curr Opin Genet Dev* 1995;5:309–314
- Reed ML, Leff SE: Maternal imprinting of human *SNRPN*, a gene deleted in Prader-Willi syndrome. *Nature Genet* 1994;6:163–167
- Mc Allister G, Amara SG, Lerner MR: Tissue-specific expression and cDNA cloning of small nuclear ribonucleoprotein-associated polypeptide N. *Proc Natl Acad Sci USA* 1988;85:5296–5300
- Barlow DP: Gametic imprinting in mammals. *Science* 1995;270:1610–1613