The EMBO Journal vol.3 no.10 pp.2279–2285, 1984

A gene sequence expressed only in undifferentiated EC, EK cells and testes

Alexander J.Stacey and Martin J.Evans[1]

Whitehead Institute for Biomedical Research, 9 Cambridge Centre, Cambridge, MA 02139, USA, and [1]University of Cambridge, Department of Genetics, Downing Site, Cambridge, UK

Communicated by M.Ashburner

A clone, EC1, has been isolated from a cDNA library prepared from 4-day embryoid bodies formed by suspension culture of PSMB EC cells. This clone has been used to screen a variety of RNA sourcecs including adult tissues, embryonal carcinoma (EC), and endoderm cell lines. A 3-kb poly(A)$^+$ RNA species was found to be present only in undifferentiated EC cells and adult mouse testes. This species was significantly reduced in testes of W/W^v mice compared with wild-type at this locus. Germ cells and their progeny are therefore implicated as the source of the RNA in testes. Hybrid-selected RNA from PSMB could be translated *in vitro* into a 35-kd protein, but no translatable message was evident in either PYS-2 (parietal), or PSA5-E (visceral) endoderm cell lines. DNA sequencing of the EC1 insert revealed that it is 744 bp in length, the 3′ 460 bp of which are in open reading frame. Comparison with known sequences have shown no significant homology. EC1 subclones in M13 have been used to generate single-stranded probes for hybridisation to RNA *in situ* in tissue sections. Hybridisation of the strand complementary to RNA produces a signal limited to the central regions of embryoid bodies formed on suspension culture of embryo-derived EK cells coinciding with the presence of undiffferentiated cells. Probing of a mouse genomic library and Southern blots of liver DNA with EC1 reveals that the gene is present as a single copy.

Key words: EC cells/EK cells/cDNA clone/gene expression

Introduction

Understanding the first stages of embryogenesis at a molecular level requires the isolation of probes for gene products which are differentially expressed during the formation of early tissues. The first overt differentiation of the totipotent embryonic stem cells results in the formation of trophectoderm on day 3.5 post-coitum. One day later, the pluripotent cells of the inner cell mass (ICM) form primitive endoderm, which further differentiates into visceral and parietal endoderm (Gardner and Papaioannou, 1975; Gardner, 1982). The study of these events has been greatly facilitated by the use of cultured cells, particularly embryonal carcinoma (EC) cells (reviewed by Graham, 1977; Martin, 1980). The initial differentiation of EC cells *in vitro* has many features in common with the formation of endoderm by the ICM (reviewed by Hogan *et al.*, 1983).

In recent years, several groups have generated recombinant DNA clones complementary to RNAs which are transcribed following differentiation of embryonic stem cells. A cDNA clone for an intermediate filament component specific to trophectoderm has been isolated from a trophoblastoma cDNA library (Brûlet and Jacob, 1982). Three independently isolated cDNA clones are now available for two components of the basal lamina produced by parietal endoderm; type IV procollagen from the parietal endoderm of day 13.5 embryos (Kurkinen *et al.*, 1983), and both type IV procollagen and laminin from differentiated F9 cells (Wang and Gudas, 1983). A marker for visceral endoderm, alpha-fetoprotein, was also cloned by Law *et al.* (1980) from yolk sac mRNA.

As is indicated by the source tissues of some of these cloned cDNAs, these probes are homologous to RNAs which are present in the first differentiated cells of the embryo and also in tissues at much later stages. We were particularly interested in obtaining probes specific to the pluripotent stem cells, as proteins characteristic of this cell type in addition to providing the converse class of regulated sequences, could also be involved in regulating the mechanism of differentiation itself.

There has been a report of a clone from an undifferentiated PCC3/A/1 cDNA library which is complementary to an RNA specific to undifferentiated EC cells (Brûlet *et al.*, 1983). This RNA proves, however, to be transcribed from retroviral elements present in a few thousand copies spread throughout the genome. We describe here the isolation and preliminary characterisation of a cDNA clone complementary to RNA specific to undifferentiated EC and embryo-derived EK cells (Evans and Kaufmann, 1981). It appears to be a single copy gene, the transcript of which is potentially capable of translation into a protein. An homologous RNA of the same size is also present in adult mouse testes, most probably in the germ cells.

Results

Isolation of EC1

A cDNA library was prepared in plasmid pAT153 from RNA isolated from PSMB EC cells which had formed simple embryoid bodies by growth as clumps in suspension for 96 h (Lovell-Badge and Evans, 1980). Such a library could be expected to contain clones representative of RNA species present in the two tissue types evident at this stage, the core of undifferentiated EC cells and the outer monolayer of endoderm.

Colony hybridisation of the library with ^{32}P-labelled cDNA from RNA extracted from undifferentiated PSMB cells, the presumptive visceral endoderm cell line PSA5-E and the presumptive parietal endoderm cell lines ICME-6 and PYS-2 indicated a number of putative differentially expressed clones. Reciprocal hybridisation of nick-translated cloned cDNA to dot blots of PSMB, PSA5-E, PYS-2, ICME-6 and STO poly(A)$^+$ RNA eliminated all but one of these clones. A single cDNA clone gave a reproducible signal when hybridised to undifferentiated PSMB RNA, but failed to hybridise to any other. This clone was named EC1.

Identification of RNA homologous to EC1 in various tissues

A variety of tissues and cell lines were screened for possession

of homologous RNAs by Northern blotting. The finding was that out of these sources tested (Table I) only undifferentiated EC cells, EK cells, embryoid bodies from either, and testes possess homologous RNA. Figure 1a shows such a blot which indicates the presence of a 3-kb species in the polyadenylated RNA present in F9 EC cells and testes and also a 1.5-kb species in F9. The larger species was found in blots of other EC cells but the 1.5-kb band was only observed in F9 RNA. The nature of this species in F9 is unknown. The abundance of the 3-kb band was estimated by comparison of the signal intensity with that of a known quantity of pBR322 blotted from the same gel. It represents ~0.005% of total RNA and 0.15% of poly(A)$^+$ RNA.

Table I. Survey of RNA from various sources for the presence of sequences homologous to EC1

RNA source	Poly(A)$^+$ RNA species detected homologous to EC1
Brain	–
Liver	–
Spleen	–
Kidney	–
Heart	–
Skeletal muscle	–
+/+ Testes	3 kb
W/W^v Testes	–
Whole day 9 embryo[a]	–
10 embryo[a]	–
11 embryo[a]	–
12 embryo[a]	–
Undifferentiated PSMB	3 kb
F9	3 kb + 1.5 kb
B2B2	3 kb
PSMB embryoid bodies	3 kb
PYS-2 parietal endoderm	–
ICME-6 parietal endoderm	–
PSA5-E visceral endoderm	–

[a]A prepared blot (a kind gift of A.Schnieke) was screened.

The finding of homologous RNA in preparations from whole testes led us to examine further the probable cell type responsible. Mice carrying the combination of alleles W/W, W/W^v and W^v/W^v at the W-locus (dominant spotting) are all known to have severe germ cell deficiencies resulting in anaemia and sterility (reviewed by Silvers, 1979). The histology of testes from such mice has shown that while some germ cells and the various stages of spermatogenesis are greatly reduced, interstitial cells and Sertoli cells of the seminiferous tubules are all present (Coulombre and Russell, 1954). Therefore a comparison of RNA from mice +/+ at this locus and one of these mutants should reveal if the germ cells are the likely source of the signal. The relative severity of the germ cell defect in the three strains is as follows: W/W > W/W^v > W^v/W^v. Atlhough theoretically the most suitable for the comparison, W/W homozygotes die around birth. W/W^v mice often survive to adulthood and so were used as an expedient source of tissue.

Figure 1b shows a blot of +/+ and W/W^v testes

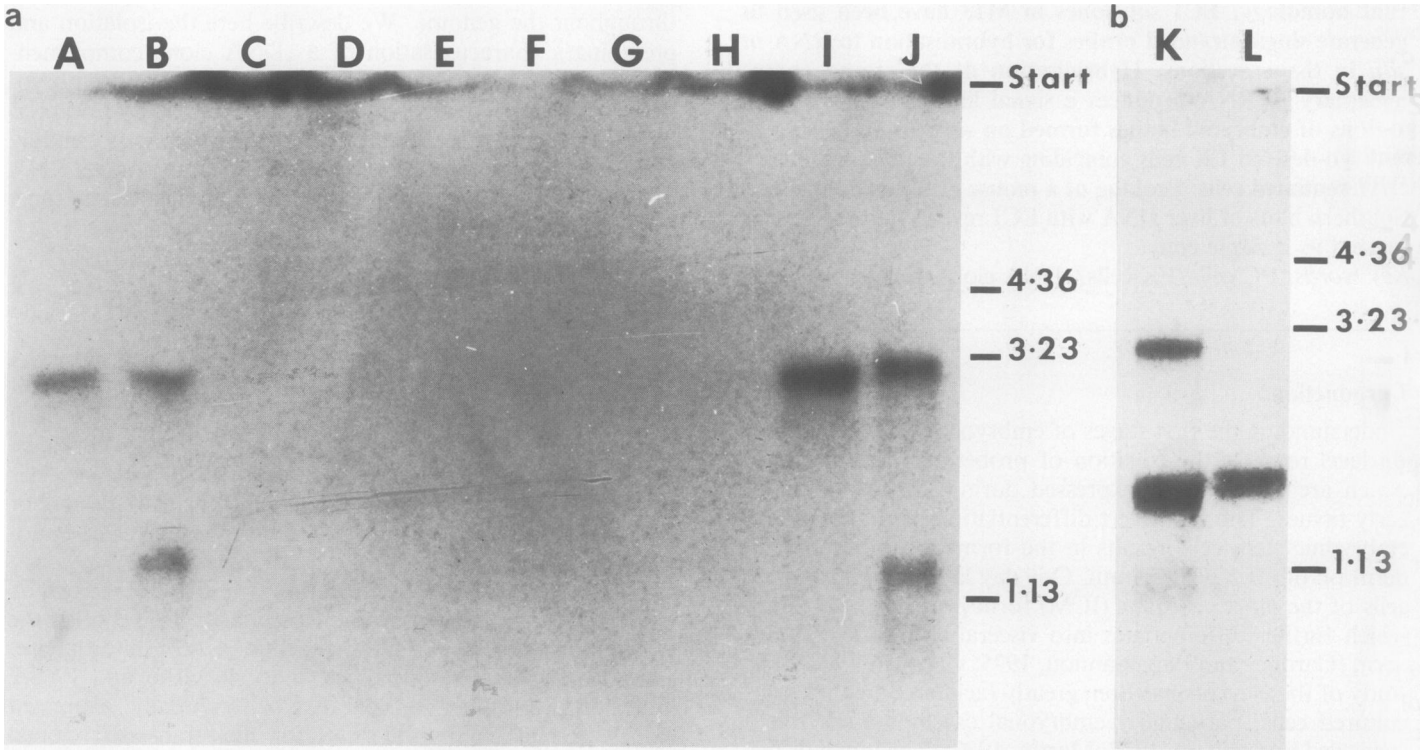

Fig. 1(a) Northern blot of F9 and various organ RNAs. RNA was isolated from undifferentiated F9 cells and adult mouse organs, glyoxalated, separated on a 1.5% agarose gel, transferred to 'genescreen' and hybridised with ^{32}P-labelled EC1 plasmid DNA ^{32}P-labelled to a specific activity of 9 x 10^7 c.p.m./μg at a concentration of 1.5 x 10^6 c.p.m./ml. **Lanes A–H** each contain 20 μg total RNA from **(A)** testes, **(B)** F9, **(C)** heart, **(D)** skeletal muscle, **(E)** brain, **(F)** kidney, **(G)**liver, *(H)* spleen. **I–J** each contain 5 μg poly(A)$^+$ RNA from **(I)** testes and **(J)** F9. Autoradiographic exposure was for 7 days. **(b)** RNA was isolated and blotted as in **(a)** from testes of +/+ and W/W^v mice of the same age. Both genotypes were the F1 offspring of a C57BL/6J-W^v/+ x WB/Re a/a-W/+ cross. The blot was hybridised with a mixture of ^{32}P-labelled EC1 plasmid DNA at 1 x 10^6 c.p.m./ml and a ^{32}P-labelled rat α-tubulin probe, pL 1 (Lemischka *et al.*, 1981), which detects a 1.5-kb species, at 0.5 x 10^6 c.p.m./ml. **Track K** contains 5 μg poly(A)$^+$ RNA from + + testes and **L** contains 5 μg poly(A)$^+$ RNA from W/W^v testes.

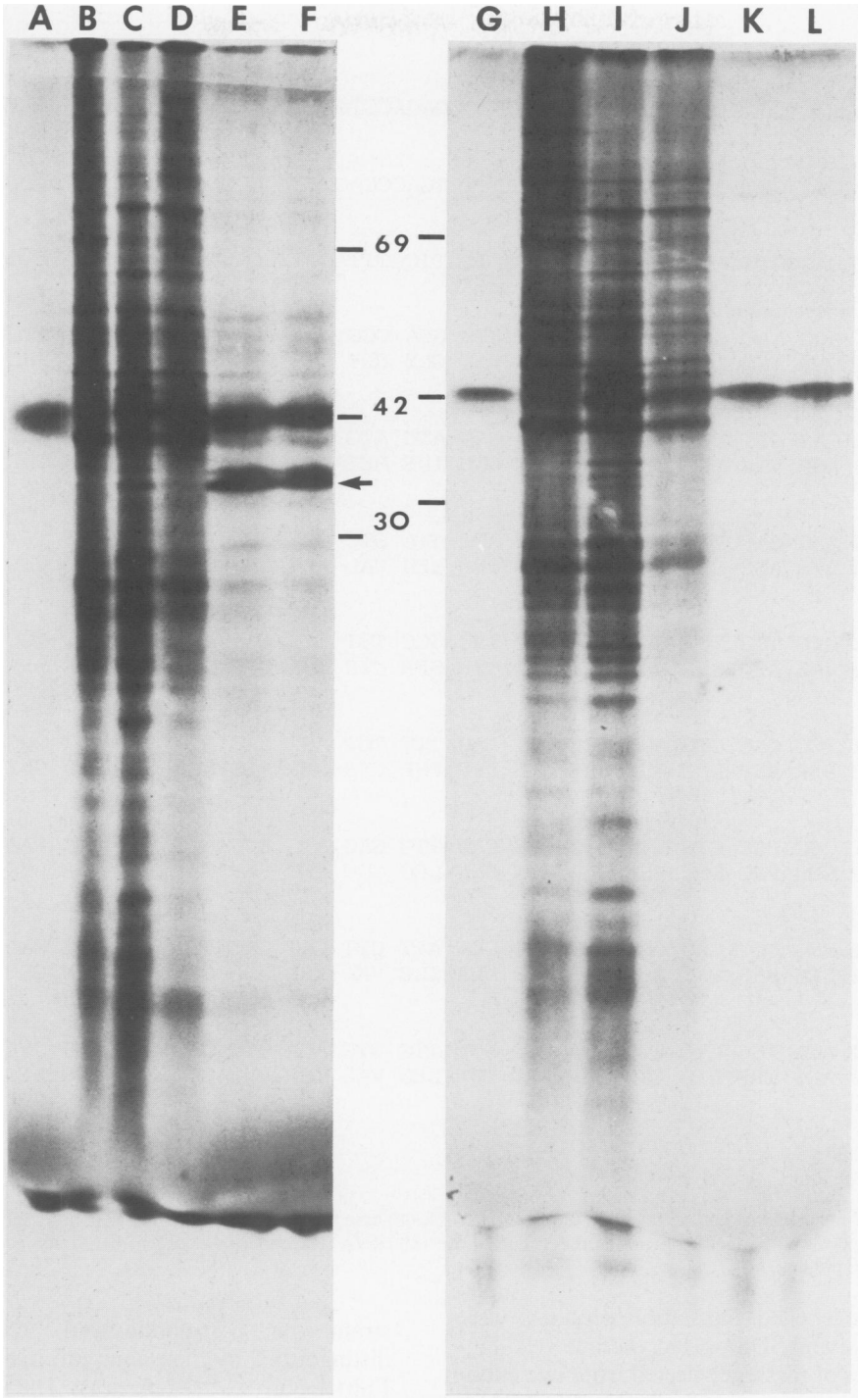

Fig. 2. *In vitro* translation of poly(A)$^+$ and hybrid-selected RNAs from EC and endoderm cells. 15 μg of EC1 plasmid bound to nitrocellulose was used to select message from 50 μg poly(A)$^+$ RNA. Message eluted was applied to a rabbit reticulocyte lysate for *in vitro* translation in the presence of [^{35}S]methionine. 3 μg each of poly(A)$^+$ RNA was similarly translated. The products of translation were separated on a 12% polyacrylamide SDS gel and products visualised by autofluorography. Translates of poly(A)$^+$ RNA were loaded at 100 000 TCA precipitable c.p.m. per slot; hybrid-selected translates and no RNA controls had 40 000 TCA precipitable c.p.m. per slot. Translates were: **(a)** no RNA, **(b)** PSA5-E poly(A)$^+$, **(c)** PSMB poly(A)$^+$, **(d)** PYS-2 poly(A)$^+$, **(e** and **f)** PSMB hybrid selected, two different RNA preparations, **(g)** no RNA, **(h)** PSA5-E poly(A)$^+$, **(i)** B2B2 poly(A)$^+$, **(j)** PYS-2 poly(A)$^+$, **(k)** PYS-2 hybrid selected, **(l)** PSA5-E hybrid selected. Autofluorographic exposure was for 10 days.

poly(A)$^+$ RNA. The 1.8-kb tubulin band which is an internal control for the quantity of RNA present is slightly stronger in the +/+ track. The difference probably represents tubulin mRNA involved in sperm tail synthesis. It is apparent that the 3-kb species homologous to EC1 is absent in W/W^v testes. This is evidence in favour of the germ cells, or certain of their progeny, specifically expressing the 3-kb RNA.

In vitro translation of hybrid-selected message

If RNA species are to make some contribution to cellular phenotype, they must almost certainly encode a protein *in vivo*. We wished to determine whether RNA homologous to EC1 could function as a message. RNA eluted after hybridisation to EC1 immobilised on nitrocellulose was applied to a rabbit reticulocyte message-dependent lysate for translation

```
                5'  GTGGACGTGTCGTCGAGTCCTGGTGAGTCACCACCGAGGTTGGGTGAGTCACCGGACGC

GACCACCCTCGACGACTGGTTAGGAGTCCACCGACCGACACGTGAGGTCAAGAAGGCGGTAGCACGCGCGATAAT

TATGTCTCCGGAGACGTTACTACTACACAAAAATTCTGACCCACTCAAGGCGCCTTCGTTTCGTCTCCTGTCTAG

TCTCGTAAGAGGGTGCAAATAAAGTCCTTTCGCTCGTTCGTTTACGTTCCACTTAAGGCGGTCCTAAAATCGTGA

ATG AAA ACG AAC GAC ACC ACC GGC GTG GGA GGG AAA AGA AGC CAC CCC TTC CCA AGT
MET LYS THR ASN ASP THR THR GLY VAL GLY GLY LYS ARG SER HIS PRO PHE PRO SER

AAT GAA GAC GTT TGG CTG TCT ATG GAG ATC ATG TTG TCG TGG AGT CAA ACC TTA GGT
ASN GLU ASP VAL TRP LEU SER MET GLU ILE MET LEU SER TRP SER GLN THR LEU GLY

GGA CAG AAG GAC AGT TTC ATC CCT ACC TTA GTC CGT CTA TGG TGT TTT TGT TCG TGT
GLY GLN LYS ASP SER PHE ILE PRO THR LEU VAL ARG LEU TRP CYS PHE CYS SER CYS

CTT GGA GTG TTG TCC GGG GGA GTT TAC TCC TGT TTC AAA AAA GGT TTT TCC CGG ACT
LEU GLY VAL LEU SER GLY GLY VAL TYR SER CYS PHE LYS LYS GLY PHE SER ARG THR

GAC TTT TCC GAT CTG ACA TTT TGT AAG TTT TGT AGT CAC ACA GGT TCT GGG TCG AAC
ASP PHE SER ASP LEU THR PHE CYS LYS PHE CYS SER HIS THR GLY SER GLY SER ASN

ATG ACT CTC CCC CTG TTC GTC CAG CAA TTG GAG ATG TCG ACT AAG AGC AAA GAG ATC
MET THR LEU PRO LEU PHE VAL GLN GLN LEU GLU MET SER THR LYS SER LYS GLU ILE

GTA TGT CTG TCT GTC GAT GAC GAT GCA ATT GTT GAT GAC TGT ATG AAC GAA CCA GGA
VAL CYS LEU SER VAL ASP ASP ASP ALA ILE VAL ASP ASP CYS MET ASN GLU PRO GLY

AAT TCT AGA TGG CTG TGT ACA CCG ACT TTG GTG ACA TTC GAG AAG TGT CTA CCA CTG
ASN SER ARG TRP LEU CYS THR PRO THR LEU VAL THR PHE GLU LYS CYS LEU PRO LEU

CGA C  3'
ARG
```

Fig. 3. Nucleotide sequence of the cDNA strand representing the RNA (+). The 10-base sequence underlined has nine bases in common with a consensus sequence found to be homologous to a purine-rich region at the 3′ end of the 18S rRNA (Hagenbüchle *et al.*, 1978). Amino acids have been assigned to the single long open reading frame present.

in vivo. [^{35}S]methionine-labelled translation products were separated on one-dimensional SDS-polyacrylamide gels. Figure 2 shows the translates of message selected from two endoderm and one EC line compared with those of poly(A)$^+$ RNA from the same sources and also EK B2B2 cells. It is apparent from the figure that a message which encodes a 35-kd protein *in vitro* (arrowed) has been selected by EC1 from PSMB RNA, but none is apparent from the endoderm cell lines. This represents strong evidence for the differential expression of this message between EC and endoderm.

There is evidence of a strong 35-kd band in the translates of PSMB and EK B2B2 poly(A)$^+$ RNA (tracks C and I) which is weaker or absent from the endodermal RNA transcripts. Bands in one-dimensional protein gels, however, often represent a number of polypeptide species.

Nucleotide sequence of EC1

The DNA sequence of EC1 was determined by the method of Sanger *et al.* (1977). Figure 3 shows the sequence of the + strand which corresponds to the RNA. The two strands were distinguished by their ability to hybridise to undifferentiated PSMB poly(A)$^+$ RNA dots. The sequence shows two features. (i) A single long reading frame which extends from base 284 off the 3′ end. The predicted amino acid sequence is shown below this reading frame. Examination of the amino acids shows no notable regions of hydrophobicity apart from two groups in each of which seven of nine residues are hydrophobic and one group where six of seven are. (ii) A 10-base sequence (underlined in Figure 3), of which nine bases conform to a consensus sequence present on many eukaryotic mRNAs and shown by Hagenbüchle *et al.* (1978) as homologous to a region at the 3′ end of the 18S rRNA. The probability of this occurring anywhere in the sequence is 0.0028, and its presence here directly upstream from the long open reading frame suggests that translation may occur.

Computer searches of the EMBL DNA database and the Genebank DNA database revealed no significant homology

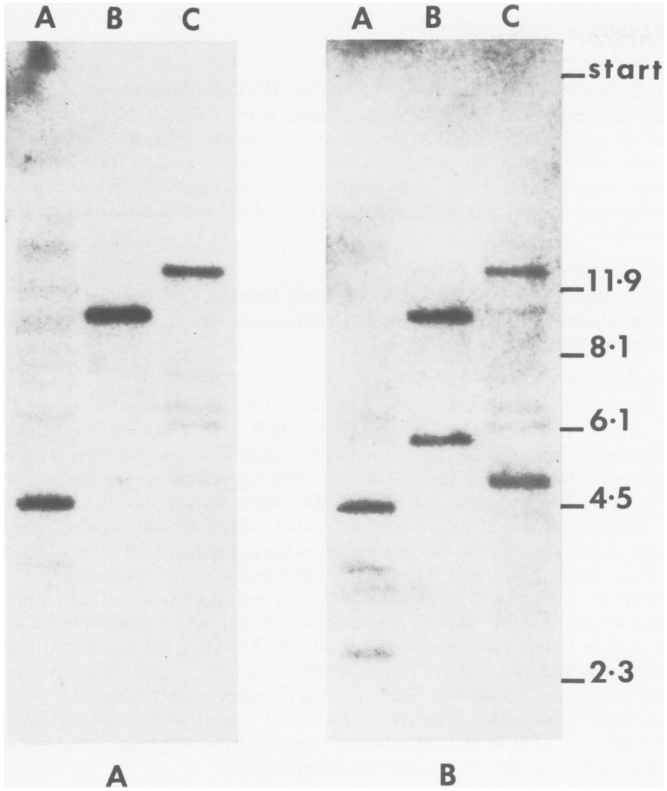

Fig. 4. Genomic blots showing fragments homologous to EC1. 20 μg each of DBA/2J mouse liver was digested to completion with **(A)** *Bam*HI, **(B)** *Eco*RI, **(C)** *Xba*I. Samples were separated on 0.8% agarose, transferred to 'genescreen' and hybridised in **part A** with EC1 plasmid DNA labelled with ^{32}P to a specific activity of 5 x 10^7 c.p.m./μg at a concentration of 1.5 x 10^6 c.p.m./ml. Autoradiographic exposure was for 6 days. **Part B** shows a re-hybridisation of the same blot with a mixture of EC1 and HF677 (Chu *et al.*, 1982) labelled to specific activities of 4 x 10^6 c.p.m./μg and 4.5 x 10^6 c.p.m./μg, respectively, both at concentrations of 0.75 c.p.m./ml. Autoradiographic exposure was for 4 days.

with any known DNA sequences. A search for the putative protein sequence in the Doolittle protein sequence databank likewise revealed no homologies.

Genomic regions homologous to EC1

The copy number of EC1 homologous sequences in the mouse genome was estimated. Screening of a Ch4A BALB/c library (prepared by L.Hood and a kind gift of J.Rogers) showed positive signals from five out of ~4 x 10^5 plaques. This indicates a low copy number. Figure 4 shows two autoradiographs of a single Southern blot of DBA/2J mouse liver DNA digested to completion with three restriction enzymes. In part A the blot was hybridised to EC1 and in part B to an equal mixture of EC1 and Hf 677 (Chu *et al.*, 1982) both labelled to the same specific activity. Hf 677 is a cDNA clone from the human α-1(I) collagen gene which is homologous to a single copy sequence at the 3′ end of the mouse α-1(I) collagen gene (Schnieke *et al.*, 1983). Part A shows that EC1 is homologous to single restriction fragments of the enzymes used, the sizes of which are: *Bam*HI, 4.6 kb; *Eco*RI, 10 kb; *Xba*I, 13 kb. Part B shows that the intensities of these bands are very similar to those of the α-1(I) collagen gene when allowance is made for the relative sizes of the restriction fragments. α-1(I) collagen bands are evident as a 2.6-kb *Bam*HI fragment, a 5.8-kb *Eco*RI fragment and a 5-kb *Xba*I fragment. These data provide conclusive evidence that EC1 defines a gene present in the mouse genome as a single copy.

Hybridisation of EC1 to RNA in situ in B2B2 embryoid bodies

In order to illustrate directly the pattern of expression of homologous RNA in EK and EK-derived endoderm cells, EC1 probes were hybridised with RNA in frozen sections of 4-day embryoid bodies. M13 subclones carrying the EC1 sequence in opposite orientations allowed the preparation of separate − and + single stranded probes which could be ^3H labelled to high specific activities, isolated and applied to slides in the absence of any vector DNA.

When embryoid bodies were hybridised with − and + stranded probes, respectively, both treatments showed a general background of grains over the sections due to non-specific binding of the probes. But superimposed on this background is a strong signal from the − stranded probe which was limited to the inner portion of each embryoid body. This signal was particularly strong where cells were clustered. We suggest that the difference in signal intensity between the inner portion and the outer endoderm layer is due to the presence of an RNA, the expression of which ceases or is significantly reduced on differentiation of EK cells into endoderm.

Discussion

These results show that EC1 is a probe for a single copy gene, the transcription of which is limited to undifferentiated EC and EK cells and some cells present in the testes. The transcript is potentially capable of translation into a protein, as indicated by the nucleotide sequence of the cDNA and also by the translation *in vitro* of hybrid-selected message to a 35-kd protein.

The expression of a transcript common to both EC cells and testes is particularly intriguing and reminiscent of the considerable specific cell-surface homologies shared by these two cell types. EK B2B2, which also shows homologous RNA, was derived directly from an embryo (Evans and Kaufmann, 1981). These three sources of homologous RNA have one prominent feature in common, the presence of pluripotential stem cells. On this basis, we would suggest that the RNA we detect is a characteristic of those stem cells. This is consistent with the drastic reduction in quantity of the 3-kb RNA in germ cell-deficient W/W^v testes. However, as all stages of spermatogenesis are reduced in such testes, one cannot conclude anything about the pattern of EC1 RNA expression during the production of male gametes from the primary stem cells. Work aiming to pinpoint the precise stages expressing EC1 RNA by *in situ* hybridisation to tissue sections is in progress.

One of the many questions raised by three closely associated papers from Rigby and co-workers (Brickell *et al.*, 1983; Murphy *et al.*, 1983; Scott *et al.*, 1983) is the significance of repetitive elements to gene regulation. As mentioned in the Introduction, a cDNA clone which detects a 6-kb transcript specific to undifferentiated EC cells has been isolated by Brûlet *et al.* (1983). The genomic DNA sequence corresponding to this takes the form of a transposon-like or retroviral element scattered in one to a few thousand copies throughout the genome. Particles associated with endogenous retroviruses have been known for some time to be expressed during early development (Chase and Pikó, 1973) and often

show stage-specific expression (Yotsuyanagi and Szöllösi, 1981).

It may be argued that the transcription of such elements does not indicate any active role in differentiation. This idea is discussed further in Jaenisch *et al.* (1981) and Jaenisch (1983). However, the findings of Rigby and co-workers show that a whole set of transcripts (Set 1), which hold a repetitive element in common, were regulated in concert on differentiation of EC and EK cells. This could indicate a signal function for the repeated sequence. Clearly more than a single class of elements are involved and both viewpoints are not incompatible. Probing of the mouse genome with EC1 indicates that this cDNA fragment contains no evidence of the Set 1, or any other repeat. This is confirmed by the nucleotide sequence.

Although our data suggest that EC1 RNA could be translated *in vivo*, a direct demonstration of the presence of a protein has yet to be made. However, it is possible that some characteristic could be ascribed to EC1 if it were found to be related to a known gene or protein. As has been mentioned, no homologies have emerged so far from searches of sequence data. A check of 14 cloned oncogenes, *fos, abl,* Ki-*ras,* Ha-*ras, mos, fms, fes, rel, yes, myb, myc, erb* (total), *src* and *sis* (A.Stacey, unpublished data, oncogene dot blots a kind gift of A.Schnieke), also showed no homology. It is also possible that EC1 may represent one of the several genes known to affect embryonic development defined by mutants (reviewed by McLaren, 1976; by Magnuson and Epstein, 1981). Determination of the chromosomal location of the EC1 gene would be a first step towards this.

A comparison of the present data with a survey of proteins present in EC monolayers, embryoid bodies and other cell lines carried out by Lovell-Badge and Evans (1980) shows a possibly interesting correlation. Out of 32 polypeptides found to be differentially expressed on the formation of embryoid bodies, four fit the criteria to be expected of a protein product of EC1, that is, presence in EC monolayers and embryoid bodies and absence in the two differentiated lines checked, PSA5-E and STO. These species have mol. wt. estimated as 33, 36, 45 and 58 kd. Taking into account possible post-translational modifications *in vivo*, either of the two smallest proteins could correspond to the putative EC1 protein. It should be feasible to compare the 35-kd translation product and these cellular proteins by proteolytic mapping of each species eluted from preparative gels. As translates of the mRNA preparations from EC cells as used here show a high degree of complexity on two-dimensional gels (Evans *et al.*, 1983a) these could be used for further resolution and identification of the hybrid-selected message translation products. Antisera could also be raised if sufficient quantities of individual spots were isolated. This could provide the key to a full characterisation of EC1.

EC1 clearly defines a gene of some interest. It is a structural gene which is specifically expressed at a moderate level in undifferentiated EC cells and whose expression is down-regulated at differentiation. As EC cells and more especially EK cells are essentially embryonic epiblast cells (Evans *et al.*, 1983b) it is to be expected that EC1 will define a gene whose transcription is regulated during early mouse embryogenesis. The fact that this regulation is the converse of that seen for the other structural genes for which probes are available will facilitate studies of the control of gene expression in early mammalian development.

Materials and methods

Cell culture

Cell lines used were as follows: PSMB and PSA5-E (Adamson *et al.*, 1977), F9 (Bernstine *et al.*, 1973), PYS-2 (Lehman *et al.*, 1974), ICME-6 (Adamson *et al.*, 1977), STO, EK B2B2 (Evans and Kaufmann, 1981). EC and EK cells were maintained in an undifferentiated state and induced to form embryoid bodies as described for EC cells (Martin and Evans, 1975a,1975b). Endoderm lines were grown in Dulbecco's modified Eagle's medium containing 10% newborn calf serum.

RNA preparation

The method used was that of Auffray and Rougeon (1980). Poly(A)$^+$ RNA was selected by oligo(dT)-cellulose (Collaborative Research) chromatography as described by Aviv and Leder (1972).

cDNA synthesis and cloning

Single-stranded cDNA was generated from poly(A)$^+$ RNA by incubation for 1 h at 42°C in a mixture comprising 50 mM Tris-HCl, pH 8.3, 6 mM magnesium acetate, 5 mM DTT, 1 mM each nucleotide triphosphate, 5 μg/ml oligo(dT) (PL Biochemicals), 25 U/ml human placental ribonuclease inhibitor (Enzo Biochemicals), 250 U/ml AMV reverse transcriptase (a kind gift of J.Beard) and 10–25 μg/ml poly(A)$^+$ RNA. When labelled single-stranded cDNA was required for use as a probe, in the screening of the cDNA library, dATP was replaced with [^{32}P]dATP. Before second strand synthesis the RNA template was removed by alkaline hydrolysis and single-stranded cDNA purified by gel filtration. Second stranding was carried out by incubation for 4–5 h at 46°C in a mixture comprising 50 mM Tris-HCl, pH 8.3, 9 mM magnesium acetate, 20 mM DTT, 1 mM each nucleotide triphosphate, 2–5 μg/ml single-stranded cDNA, 500 V/ml AMV reverse transcriptase. cDNA was rendered blunt-ended with S1 nuclease and cloned with combinations of *Eco*RI and *Xba*I linkers into pAT153 and a derivative plasmid carrying an *Xba*I linker within two regenerated *Bam*HI sites.

Screening cDNA clones

Initial screening was by differential colony hybridisation (Grunstein and Hogness, 1975) using ^{32}P-labelled single-stranded cDNA probes. Reciprocal hybridisation of ^{32}P-labelled plasmid DNA from putative differentially expressed clones was to RNA dots on aminothiophenol paper [prepared as described by Seed (1982)], Northern blots onto aminothiophenol paper were carried out with poly(A)$^+$ RNA from undifferentiated PSMB, PSMB embryoid bodies, undifferentiated EK B2B2, PYS-2, ICME-6 and PSA5-E. Northern blots onto 'genescreen' (New England Nuclear) were carried out by a method modified in two respects from that described for nitrocellulose by Thomas (1980). Transfer was carried out in 10 x SSC rather than 20 x, and 1% SDS rather than 0.1% was included in the hybridisation solution. By this method, RNA from adult mouse brain, liver, heart, kidney, spleen, skeletal muscle, wild-type testes, W/W^v testes and undifferentiated F9 cells was blotted and screened. A prepared filter carrying total RNA from day 9, 10, 11, 12 embryos was also screened (a kind gift from A.Schnieke).

Hybrid selection and in vitro translation of messages

Hybrid selection was carried out by a method which is described by Parnes *et al.* (1981). The rabbit reticulocyte lysate was purchased from R.Jackson (Cambridge) and was used essentially as described by Pelham and Jackson (1976). [^{35}S]methionine-labelled translation products were separated on 12% polyacrylamide SDS gels as described by Laemmli (1970). Preparation for autofluorography was carried out using 'enhance' fixative and scintillant (New England Nuclear).

DNA sequencing

DNA sequencing of the EC1 insert in subclones in M13 mp8 was carried out by the dideoxynucleotide chain termination method (Sanger *et al.*, 1977) as described, with the minor modification being that reactions were carried out in Eppendorf tubes. The optimal dideoxy:deoxynucleotide ratio for each individual reaction was predetermined in trial experiments. The sequence given was derived from sequencing the insert in both orientations in the *Eco*RI site of M13 mp8. Selected regions, notably the center, were sequenced from subcloned *Hae*III fragments inserted into the *Sma*I site of M13 mp8.

Single-stranded probe preparation for in situ hybridisation

^{3}H-Labelled single-stranded probes were generated by primer extension on M13 templates under reaction conditions close to those used for sequencing, omitting dideoxynucleotides. The method used was that originally devised by J.F.Burke (in preparation) for probes to be used for S1 mapping. Here it has been adapted for use in *in situ* hybridisation. Akam (1983) has similarly used the same method. The probes used were prepared as follows. The EC1 fragment was subcloned into the *Eco*RI site of M13 mp9 and subclones carrying the insert in opposite orientations picked by the ability of their single-stranded DNA to form double-stranded hybrids over the insert region.

Approximately 1 pmol of single-stranded template and 5 pmol M13 universal pentadecamer primer (New England Biolabs) in 50 μl, 20 mM Tris-HCl, pH 8.3, 20 mM $MgCl_2$, 1 mM DTT was denatured by boiling for 1 min and annealed by cooling gradually to room temperature over 15 min. This was transferred to a tube in which [^{3}H]dCTP (50 Ci/mmol), [^{3}H]dATP (50 Ci/mmol) and [^{3}H]dTTP (97 Ci/mmol) had been dried down to give a final concentration of 20 μM each. Cold dGTP was added to the same concentration. The primer extension reaction commenced with the addition of 12 units Klenow polymerase (Boehringer). Incubation continued at room temperature for 30 min. A chase of cold nucleotide triphosphates was added to 100 mM each and the reaction continued for a further 30 min. The reaction was heat killed at 70°C for 10 min, carrier DNA added to 0.1 mg/l and DNA ethanol precipitated. Following *Hind*III digestion NaOH was added to 0.1 M, the solution was heated to 70°C for 30 s and loaded immediately onto a 1.5% agarose gel. Labelled fragments were collected by electrophoresis onto NA45 membrane (Schleicher and Schuell). The size of the single-stranded eluted fragment was reduced by sonication. The estimated specific activity of the probe was 4 x 10^8 d.p.m./μg.

In situ hybridisation to tissue sections

Embryoid bodies were placed in OCT cutting compound (Miles), frozen in liquid nitrogen and 10 μm frozen sections cut in a cryostat and collected on subbed slides. Preparation of slides and hybridisation conditions were essentially as described by Brahic and Haase (1978) with the modification that the initial fixation of sections was carried out with 4% paraformaldehyde in 1 x PBS (Hafen *et al.*, 1983). Hybridisation was carried out at 37°C for 18 h in volumes of 5 μl under polythene coverslips. The hybridisation solution comprised 50% formamide, 0.6 M NaCl, 10 mM Tris-HCl, pH 7.5, 1 mM EDTA, 1 x Denhardt's solution with an additional 0.8 mg/ml BSA, 1 mg/ml carrier DNA and 3 x 10^6 c.p.m./ml probe. Autoradiography was carried out using Kodak NTB-2 emulsion.

Acknowledgements

A.S. would like to thank Liz Robertson and Gregg Magrane for advice on growing EC cells, Michael Kuehn for advice on *in situ* hybridisation, Robin Lovell-Badge for helpful criticisms, Wolfram Ostertag and Colin Stewart for discussion. During part of this work A.S. held a research studentship from the Science and Engineering Research Council and also acknowledges the financial support of the Stiftung Volkswagenwerk. Work in the University of Cambridge, Genetics Department, was supported by a project grant from the Science and Engineering Research Council.

References

Adamson,E.D., Evans,M.J. and Magrane,G.G. (1977) *Eur. J. Biochem.*, **79**, 607-615.

Akam,M.E. (1983) *EMBO J.*, **2**, 2075-2084.

Auffray,C. and Rougeon,F.J. (1980) *Eur. J. Biochem.*, **107**, 303-314.

Aviv,H. and Leder,P. (1972) *Proc. Natl. Acad. Sci. USA*, **69**, 1408-1412.

Bernstine,E.G., Hooper,M.L., Grandchamp,F. and Ephrusfi,B. (1973) *Proc. Natl. Acad. Sci. USA.*, **70**, 3899-3903.

Brahic,M. and Haase,A.T. (1978) *Proc. Natl. Acad. Sci. USA*, **75**, 6125-6129.

Brickell,P.M., Latchman,D.S., Murphy,D., Willison,K. and Rigby,P.W.J. (1983) *Nature*, **306**, 756-760.

Brûlet,P. and Jacob,F. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 2328-2332.

Brûlet,P., Kaghad,M., Xu,Y-S., Croissant,O. and Jacob,F. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 5641-5645.

Chase,D.G. and Piko,L. (1973) *J. Nat. Cancer Inst.*, **51**, 1971-1975.

Chu,M.L., Myers,J.C., Bernard,M.P., Ding,J.F. and Ramirez,F. (1982) *Nucleic Acids Res.*, **10**, 5952-5934.

Coulombre,J.L. and Russell,E.S. (1954) *J. Exp. Zool.*, **126**, 277-296.

Evans,M.J. and Kaufmann,M.H. (1981) *Nature*, **292**, 154-156.

Evans,M.J., Lovell-Badge,R.H., Latchman,D., Stacey,A. and Brzeski,H. (1983a) in Celis,J.E. and Bravo,R. (eds.), *Gene Expression in Normal and Transformed Cells*, Plenum Press, NY, pp. 87-99.

Evans,M.E., Robertson,E., Bradley,A. and Kaufmann,M.H. (1983b) in McLaren,A. and Wylie,C.C. (eds.), *Current Problems in Germ Cell Differentiation*, Symposium of British Society for Developmental Biology, Cambridge University Press, pp. 139-155.

Gardner,R.L. (1982) *J. Embryol. Exp. Morphol.*, **68**, 175-198.

Gardner,R.L. and Papaioannou,V.E. (1975) in Balls,M. and Wild,A.E. (eds.), *Early Development of Mammals*, Cambridge University Press, pp. 107-132.

Graham,C.F. (1977) in Sherman,M.I. (ed.), *Concepts in Mammalian Embryogenesis*, MIT Press, Cambridge, MA, pp. 315-394.

Grunstein,M. and Hogness,D. (1975) *Proc. Natl. Acad. Sci. USA*, **72**, 3961-3965.

Hafen,E., Levine,M., Garber,R.L. and Gehring,W.J. (1983) *EMBO J.*, **2**, 617-623.

Hagenbüchle,O., Santer,M. and Argetsinger-Steitz,J. (1978) *Cell*, **13**, 551-563.

Hogan,B.L.M., Barlow,D.P. and Tilly,R. (1983) in Gardner,R.L. (ed.), *Embryonic and Germ Cell Tumours in Man and Animals*, Oxford University Press Cancer Surveys, Vol. **2**, pp. 117-140.

Jaenisch,R. (1983) *Cell*, **32**, 5-6.

Jaenisch,R., Jahner,D., Nobis,P., Simon,I., Lohler,J., Harbers,K. and Grotkopp,D. (1981) *Cell*, **24**, 519-529.

Kurkinen,M., Barlow,D.P., Helfman,D.M., Williams,J.G. and Hogan, B.L.M. (1983) *Nucleic Acids Res.*, **11**, 6201-6209.

Laemmli,U.K. (1970) *Nature*, **227**, 680-685.

Law,S., Tamoaki,T., Kreuzaler,F. and Dugaiczyk,A. (1980) *Gene*, **10**, 54-61.

Lehman,J.M., Speers,W.C., Swartzendruber,D.E. and Pierce,G.B. (1974) *J. Cell. Physiol.*, **84**, 13-28.

Lemischka,I.R., Farmer,S., Racaniello,V.R. and Sharp,P.A. (1981) *J. Mol. Biol.*, **151**, 101-120.

Lovell-Badge,R.H. and Evans,M.J. (1980) *J. Embryol. Exp. Morphol.*, **59**, 187-206.

Magnuson,T. and Epstein,C.J. (1981) *Biol. Rev.*, **56**, 396-408.

Martin,G.R. (1980) *Science (Wash.)*, **209**, 768-776.

Martin,G.R. and Evans,M.J. (1975a) in Sherman,M.I. and Solter,D. (eds.), *Teratomas and Differentiation*, Academic Press, London/NY, pp. 169-187.

Martin,G.R. and Evans,M.J. (1975b) *Proc. Natl. Acad. Sci. USA*, **72**, 1441-1445.

McLaren,A. (1976) *Annu. Rev. Genet.*, **10**, 361-368.

Murphy,D., Brickell,P.M., Latchman,D.S., Willison,K. and Rigby,P.W.J. (1983) *Cell*, **35**, 865-871.

Parnes,J.R., Velon,B., Felsenfeld,A., Ramanathan,L., Ferrini,V., Appella, E. and Seidman,J.G. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 2253-2257.

Pelham,H.R.B. and Jackson,R.J. (1976) *Eur. J. Biochem.*, **67**, 247-256.

Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463-5467.

Schnieke,A., Horbers,K. and Jaenisch,R. (1983) *Nature*, **304**, 315-320.

Scott,M.R.D., Westphal,K.-H. and Rigby,P.W.J. (1983) *Cell*, **34**, 557-567.

Seed,B. (1982) *Nucleic Acids Res.*, **10**, 1799-1810.

Silvers,W.K. (1979) *The Coat Colours of Mice*, published by Springer-Verlag, NY/Heidelberg/Berlin, pp. 206-241.

Thomas,P.S. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 5201-5205.

Wang,S-Y. and Gudas,L.J. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 5880-5884.

Yotsuyanagi,Y. and Szöllösi,D. (1981) *J. Natl. Cancer Inst.*, **67**, 677-683.

Received on 4 June 1984; revised on 10 July 1984