The murine H19 gene is activated during embryonic stem cell differentiation in vitro and at the time of implantation in the developing embryo

FRANCOISE POIRIER 1,2,*, C.-T. J. CHAN 1†, PAULA M. TIMMONS 1, ELIZABETH J. ROBERTSON 3,*, MARTIN J. EVANS 1,‡ and PETER W. J. RIGBY 1,2

1 Laboratory of Eukaryotic Molecular Genetics, MRC National Institute for Medical Research, The Ridgeway, Mill Hill, London, NW7 1AA, UK
2 Cancer Research Campaign, Eukaryotic Molecular Genetics Research Group, Department of Biochemistry, Imperial College of Science and Technology, London, SW7 2AZ, UK
3 Department of Genetics, University of Cambridge, Downing Street, Cambridge, CB2 3EH, UK

* Present address: Department of Genetics and Development, College of Physicians and Surgeons of Columbia University, 701 West 168th Street, New York, New York 10032, USA
† Present address: Mothercare Department of Pediatric Genetics, Institute of Child Health, University of London, 30 Guilford Street, London, WC1N 1EH, UK
‡ Present address: Wellcome/CRC Institute, University of Cambridge, Tennis Court Road, Cambridge, CB2 1QR, UK

Summary

The differentiation in vitro of murine embryonic stem cells to embryoid bodies mimicks events that occur in vivo shortly before and after embryonic implantation. We have used this system, together with differential cDNA cloning, to identify genes the expression of which is regulated during early embryogenesis. Here we describe the isolation of several such cDNA clones, one of which corresponds to the gene H19. This gene is activated in extraembryonic cell types at the time of implantation, suggesting that it may play a role at this stage of development, and is subsequently expressed in all of the cells of the mid-gestation embryo with the striking exception of most of those of the developing central and peripheral nervous systems. After birth, expression of this gene ceases or is dramatically reduced in all tissues.

Key words: embryonic stem cells, differentiation, implantation, H19 gene, transcription, in situ hybridisation, mouse development.

Introduction

A number of strategies have been used to identify genes involved in murine embryogenesis. Firstly, the murine homologues of genes known to regulate development in other species, most notably Drosophila, have been cloned and characterised (reviewed by Kessel and Gruss, 1990). Secondly, genes defined by spontaneous (Herrmann et al. 1990) or experimentally induced (reviewed by Gridley et al. 1987) mutations which result in developmental disorders have been analysed. More recently, this mutagenesis approach has been extended by using a gene trap strategy in embryonic stem cells (Gossler et al. 1989). Thirdly, systematic screening of tissue- or stage-specific cDNA libraries can be performed. This approach offers the unique advantage of allowing the direct isolation of genes specifically expressed in a defined cell population or at a predetermined stage of development. We have chosen this third strategy for isolating genes the expression of which is regulated in the peri-implantation mouse embryo.

Molecular analyses of events that occur during the early stages of mammalian development are hindered by the small size of the embryo, by the heterogeneity of cell types arising within the first few days after fertilisation and by the difficulty of access due to implantation. Very little is known about gene expression in the peri-implantation embryo although many important developmental decisions are taken at this time. Initially, cells become allocated to either the trophoderm lineage, which gives rise exclusively to extraembryonic tissues, or to the inner cell mass (ICM) lineage, from which the embryo proper is derived (Pedersen, 1986). Subsequently, extraembryonic lineages will differentiate (Rossant, 1986) while the
embryo will undergo gastrulation (Beddington, 1986). Our objective was to isolate genes involved in these early events of murine development.

The differentiation in vitro of murine embryonic stem (ES) cells to embryoid bodies mimicks events that occur in vivo shortly before and after embryonic implantation. ES cells, which are derived from in vitro cultures of blastocysts or isolated ICMs (Evans and Kaufman, 1981; Martin, 1981), correspond in developmental potential to cells of the early ICM and are able to contribute at high efficiency to all of the lineages of the embryo (Bradley et al. 1984; Beddington and Robertson, 1989). They will differentiate in vitro when cultured in suspension. Under these conditions they form small aggregates which, with time, undergo a reasonably ordered and synchronous programme of differentiation to form more complex structures. These are called embryoid bodies (EBs) because the process by which they are generated resembles, in part, that of the formation of the embryonic portion of the developing conceptus at the egg-cylinder stage (Martin and Evans, 1975; Martin et al. 1977; Doetschman et al. 1985). This in vitro system thus allows the generation of large quantities of cells similar to their normal embryonic counterparts present before and after implantation.

We have exploited this model system by using differential screening of cDNA libraries constructed from ES cells and from EBs to isolate genes regulated during this differentiation process in vitro. We present here a detailed analysis of one such gene using in situ hybridisation and show that it is indeed activated in a tissue-specific manner at the time of implantation, suggestive of a role during early embryogenesis. The pattern of expression at later stages of development and in the adult suggests that at these stages the gene is required prior to terminal cell differentiation. Nucleotide sequence determination shows that the gene is H19 which was originally isolated during a search for genes regulated by raf, a trans-acting locus involved in regulating the adult basal level of α-foetoprotein mRNA (Pachnis et al. 1984, 1988).

Materials and methods

Cell culture

The murine embryonic stem cell lines CP3 (Bradley and Robertson, 1986) and CCE (Robertson et al. 1986) were grown on feeder layers of mitomycin C-treated murine STO fibroblasts as previously described (Robertson, 1987). The cells were passaged every three to four days. Differentiation was induced by seeding lightly trypsinised cells into bacteriological Petri dishes and the resultant embryoid bodies were harvested 7 days later.

Construction and screening of cDNA libraries

Cytoplasmic RNA was prepared from ES cells and EBs by the NP40 lysis method of Favaloro et al. (1980). Poly(A)+ RNA was selected on oligo(dT)-cellulose as described by Aviv and Leder (1972). Two cDNA libraries were constructed in the bacteriophage vector λgt10 (Hyunh et al. 1985) using cDNA synthetised by the procedure of Watson and Jackson (1985). The first, derived from poly(A)+ cytoplasmic RNA from the CP3 line of ES cells contained 5×10^6 independent recombinants; the second, derived from poly(A)+ cytoplasmic RNA from CP3 7 day EBs, contained 2×10^7 independent recombinants. Unamplified libraries were plated at a density of 6000 plaques/10 cm plate and four replica filters were prepared from each plate. These were screened, according to the procedure of Benton and Davis (1977), with oligo(dT)-primed, 32P-labeled cDNA prepared from ES cell RNA (filters 1 and 3) or EB RNA (filters 2 and 4). Plaques showing a differential hybridisation signal on both filters 1 and 3 compared to filters 2 and 4 were purified through three cycles of comparative hybridisation with the same two cDNA probes. Plaques that survived this screen were grown in bulk culture for subsequent analysis.

We determined the sensitivity of our screening procedure using Aplysia bag cell mRNA, 50% of which is known to be mRNA for preproELH (Scheller et al. 1982). A λgt10 recombinant phage containing a cDNA insert corresponding to preproELH was mixed with vector phage at a ratio of 1:10, the phages were plated and replica filters were screened with 32P-labeled cDNA probes made from bag cell mRNA mixed with increasing amounts of unrelated mRNA (in this case mouse placental mRNA). We thus determined that we could detect plaques corresponding to mRNAs present at 1 part in 10^4.

Northern blot analyses

All RNA samples used for northern analysis were prepared by precipitation with LiCl (Auffray and Rougeon, 1980). RNA was electrophoresed in 1.5% (w/v) agarose gels containing 20 μM Mops pH 7.0, 5 mM sodium acetate, 1 mM EDTA and 0.7% (w/v) formaldehyde (Maniatis et al. 1982) and transferred to Amersham N-Bond nylon paper.

For oligolabelled cDNA probes (Feinberg and Vogelstein, 1983), hybridization was performed in 50% (v/v) formamide, 50 mM sodium phosphate pH 6.8, 2×Denhardt's solution, 0.5×SSC and 0.05% (w/v) SDS at 42°C. Filters were washed in 0.2×SSC, 0.2% (w/v) SDS at 65°C for 1 h. For antisense RNA probes, hybridization was carried out in 60% (v/v) formamide, 1 M NaCl, 50 mM Tris–HCl pH 7.5, 10×Denhardt's, 1% (w/v) SDS, 0.1% (w/v) sodium pyrophosphate, 100 μg ml⁻¹ yeast RNA, 10% (w/v) dextran sulphate at 65°C. Filters were washed in 0.2×SSC, 0.2% (w/v) SDS at 70°C for 1 h.

Nucleotide sequence determination

The cDNA clone LC10–8 was used to screen an 8.5 days post coitum (dpc) C57BL/6 embryo cDNA library (Fahrner et al. 1987) from which we isolated a much longer clone, LC10–8/1. The three BamHI fragments of LC10–8/1 were subcloned into Bluescript KS+ and each was then deleted unidirectionally using exonuclease III (Henikoff, 1984). Overlapping deletion clones were sequenced using the chain termination method (Sanger et al. 1977). The sequence was checked by unidirectionally deleting in the opposite orientation and sequencing the overlapping subclones. All of the nucleotides in the LC10–8/1 cDNA were clearly identified.

In situ hybridisation

The procedure of Wilkinson et al. (1987) was used throughout. Embryos or CCE-derived EBs were fixed overnight in 4% (w/v) paraformaldehyde, dehydrated and embedded in paraffin wax. 6 μm sections were cut and hybridized with 35S-labeled antisense RNA probes generated by in vitro
transcription of appropriate plasmid subclones with T7 RNA polymerase.

Antibody staining of embryoid bodies
7-day-old CP3 embryoid bodies were fixed in 4% (w/v) paraformaldehyde, embedded in OCT and 15μm frozen sections were stained with a monoclonal antibody raised against SSEA-1 (kindly provided by D. Solter; Solter and Knowles, 1978) or with a rabbit polyclonal antibody directed against Forssman antigen (a gift from K. Willison; Stinnakre et al. 1981). The results were revealed by using the horseradish peroxidase technique following the standard protocols described by Harlow and Lane (1988).

Results
Isolation of genes regulated during embryonic stem cell differentiation in vitro
In order to isolate genes which are either switched off or switched on during the in vitro differentiation of ES cells to EBs, we constructed cDNA libraries from each stage and differentially screened each library using cDNA probes from both cell populations. When ES cells are plated on bacteriological dishes, the three following morphological changes are observed sequentially: (i) formation of an outer layer of endoderm-like cells (24 h); (ii) cavitation (4 days) and (iii) appearance of an epithelium of columnar cells lining the cavities (4 to 6 days). These three steps parallel those observed in vivo around the time of implantation, namely: (i) formation of primitive endoderm (which gives rise to visceral and parietal endoderm); (ii) appearance of the proamniotic cavity and (iii) formation of primitive ectoderm.

CP3 ES cells were plated into bacteriological Petri dishes and the resultant EBs were allowed to differentiate for 7 days, at which time approximately 80% of them had a complex structure, being composed of an outer layer of endoderm-like cells surrounding a core in which some of the cells were organised into an epithelium lining one or two cavities (see Fig. 1B). The remaining 20% were classified as simple EBs, as differentiation had not proceeded beyond the stage of formation of an outer layer of endoderm. We monitored the loss of two markers of undifferentiated ES cells, SSEA-1 and Forssman antigens, in the EBs.

Fig. 1. Expression of three candidate clones in ES cells and EBs. (A) RNA blot analysis of LCI, LC25 and LC10-8 in ES cells and EBs. 10 μg of total RNA from undifferentiated CCE cells (lanes 1, 3, 5 and 7) and from day 7 EBs (lanes 2, 4, 6 and 8) were run on a northern gel as described in Materials and methods. After transfer, the filter was cut into four strips which were hybridized with random primed probes of LCI (lanes 1 and 2), LC25 (lanes 3 and 4), LC10-8 (lanes 5 and 6) and α-tubulin (lanes 7 and 8). (B) In situ hybridization of 35S-labeled antisense RNA probe of LC10-8 to sections of day 7 EBs. (1) Bright-field picture. After 7 days of differentiation, the embryoid bodies had developed a layer of endoderm and also generated a Reichert-like membrane. This membrane loosens the adhesive bonds between the endoderm layer and the core, resulting in cells of the outer layer detaching. (2) Dark-field picture. The signal is localised on the remaining cells of the outer layer. Hybridization with the sense strand probe showed no signal (data not shown).
These antigens correspond to glycosylated residues present on the cells of the ICM in vivo. After implantation, these antigens are present on the cells of the early primitive ectoderm but they later disappear from the embryonic portion of the conceptus. The cells of the primitive endoderm carry the Forssman but not the SSEA-1 antigen (Fox et al. 1981; Stinnakre et al. 1981). During this period of development these antigens can therefore be considered as markers for the differentiation of the ICM lineage. ES cells, like ICM cells, carry the SSEA-1 and Forssman antigens. In contrast, 80–90% of the cells in the core of day 7 EBs were negative for both markers while the cells in the outer layer were SSEA-1-negative and Forssman-positive. Both the morphological and antigenic properties of the day 7 EBs indicate that differentiation had followed a programme similar to that of ICM lineage cells in vivo.

We screened 20 000 plaques from each library, as described in Materials and methods, and isolated three clones that appeared to be differentially regulated. LC1 and LC25 derive from the screen of the ES cell library, while LC10–8 derives from the screen of the EB library. Fig. 1A shows that the cDNA inserts from LC1 and LC25 both hybridised, as expected, to transcripts present in ES cells but barely detectable in EBs. In contrast, LC10–8 hybridised to a 2.6 kb transcript present in ES cells but approximately threefold more abundant in EBs.

This limited screen thus allowed the identification of three genes that are regulated during ES cell differentiation. Subsequent analyses of these genes have shown that the expression of each is regulated in a unique fashion during mouse embryogenesis thus demonstrating the general utility of our experimental approach. In this report we describe the results obtained with LC10–8. LC1 encodes a β-galactoside-binding lectin of relative molecular mass 14×10^3; the expression pattern of this gene will be described elsewhere (F.P., P.M.T., C.-T.J.C., J.-L. Guenet and P.W.J.R., manuscript in preparation). LC25 encodes calcyclin, a putative calcium-binding protein of the S100 family which was originally isolated in a screen designed to identify genes regulated during the cell-cycle (Calabretta et al. 1986).

As EBs contain several cell types, we wished to determine in which cells LC10–8 transcripts accumulated. Fig. 1B shows the results obtained by hybridising an 35S-labeled, antisense RNA probe to sections of day 7 EBs. The cells of the outer layer hybridised strongly whereas the inner core cells were unlabeled. Wilkinson et al. (1988) used these same EBs to show, by hybridisation with α-foetoprotein (negative) and SPARC (positive) probes, that most of the outer cells are of parietal endoderm type. Only a small fraction contained AFP transcripts which are specific for visceral endoderm (Dziadek, 1978; Tilghman and Belayhew, 1982; Dziadek and Andrew, 1983). We therefore conclude that transcription of LC10–8 is activated when ES cells differentiate into a cell type(s) that is very similar to parietal endoderm.

**Nucleotide sequence analysis**

The sequence of the 1902 bp cDNA clone, LC10–8/1, isolated from the 8.5 dpc embryo library was determined. Computer-assisted searching of the nucleotide sequence data banks showed that LC10–8 corresponds to the same gene as H19, a cDNA clone isolated by Pachnis et al. (1984) because its expression, like that of α-foetoprotein, is down-regulated in liver soon after birth. We shall hereafter refer to LC10–8 as H19.

Comparison of our sequence with that of Pachnis et al. (1988), as modified by Brannan et al. (1990), indicates a number of discrepancies, which are detailed in the legend to Fig. 2. These differences, which could represent polymorphisms between the C57BL/6 and Balb/c mouse strains, do not affect any conclusions drawn regarding the overall structure of the gene, although they do affect details of the open reading frames (ORFs).

As was noted by Pachnis et al. (1988), a striking feature of their sequence is that it contains no long ORFs that span exon boundaries. The longest ORF in their Balb/c sequence (ORF5 in the nomenclature of Brannan et al. 1990) S' ORF in Fig. 2A) can encode a protein of only 132 amino acids (see Fig. 2B) and is entirely contained within exon 1. Our C57BL/6 sequence reveals another, longer, ORF (3' ORF in Fig. 2A) which spans exons II, III, IV and V and could encode a protein of 147 amino acids (see Fig. 2B). Both 5' and 3' ORFs have a Kozak consensus sequence (Kozak, 1986, 1987) for the initiation of translation at their first ATG. All of the other ORFs either lack an ATG codon, possess a poor Kozak consensus sequence or could encode only a small peptide. There are no transcriptional frameshifting sequences, of the sort found in retroviruses (Jacks et al. 1988) and coronaviruses (Brierley et al. 1989), that might mediate the synthesis of a single protein from the two longest ORFs. Analysis of the protein sequences derived by conceptual translation of the 5' and 3' ORFs did not reveal any known motifs.

**Pattern of H-19 expression during early embryogenesis**

Fig. 3 shows the results of an in situ hybridisation analysis of H-19 expression during the first 6.5 days of development. We did not detect expression in the fertilised egg (0.5 dpc), the morula (2.5 dpc; not shown), or the early blastocyst (3.5 dpc). However, in a late stage blastocyst (4.5 dpc) low levels of H19 RNA were detected in the cells of the trophectoderm. In the early postimplantation conceptus (5.5 dpc) a very strong signal was apparent in the ectoplacental cone, the trophoblastic giant cells, the extraembryonic endoderm and the extraembryonic ectoderm. Only the cells
of the embryonic ectoderm were devoid of \( H19 \) RNA. In the gastrulating (6.5 dpc) embryo, the distribution of transcripts observed a day earlier was maintained. We note that there was no detectable expression in the newly formed migrating mesoderm which derives from the embryonic ectoderm. \( H19 \) expression between fertilisation and gastrulation is thus characterised by the sudden accumulation of transcripts at the time of implantation in all of the cells that will contribute exclusively to the formation of extraembryonic tissues. \( H19 \) expression is maintained in most of these tissues; indeed, northern analysis of total embryo RNA showed that \( H19 \) transcripts were first detected in the embryo proper at 8.5 dpc. We note that there was no detectable expression in the parietal endoderm which then progressively enlarge. While \( H19 \) transcripts are initially found in the liver, cartilage, intestine, heart, tongue and mesenchyme (Fig. 4). Additional sites of \( H19 \) expression include the kidney and urogenital ridges (data not shown).

Within the positive organs, the distribution of grains was not uniform and it did not correlate with cell density. Rather there were clear quantitative differences in the levels of accumulated transcript between cells of endodermal and mesodermal origin, the signal being much stronger in the former. For example, this is apparent in sections of the intestine in which the grain density was higher in the epithelium (endodermal origin) than in the cells of the surrounding mesenchyme (mesodermal origin). This is a general feature of \( H19 \) expression at this stage of development, which is also observed in the stomach, lung and kidney.

An important feature of the pattern of \( H19 \) expression at this stage is that, in sharp contrast to the rest of the embryo, the majority of cells of neuroectodermal origin, namely the brain, spinal cord, dorsal root ganglia (see Fig. 4) and the trigeminal ganglion (data not shown) were devoid of detectable \( H19 \) transcripts. This result was confirmed when the exposure time was increased tenfold. It was also noticeable that blood cells, although of mesodermal origin, were negative.

Another aspect of the regulation of \( H19 \) expression can be seen in the developing long bones of the 14.5 dpc embryo. Ossification of the cartilaginous skeleton is initiated during the second half of gestation. This phenomenon starts in foci called ossification centres, which then progressively enlarge. While \( H19 \) is transcribed in cartilage, no transcripts were detected in the surrounding mesenchyme (Fig. 4). Within the positive organs, the distribution of grains was not uniform and it did not correlate with cell density. Rather there were clear quantitative differences in the levels of accumulated transcript between cells of endodermal and mesodermal origin, the signal being much stronger in the former. For example, this is apparent in sections of the intestine in which the grain density was higher in the epithelium (endodermal origin) than in the cells of the surrounding mesenchyme (mesodermal origin). This is a general feature of \( H19 \) expression at this stage of development, which is also observed in the stomach, lung and kidney.

An important feature of the pattern of \( H19 \) expression at this stage is that, in sharp contrast to the rest of the embryo, the majority of cells of neuroectodermal origin, namely the brain, spinal cord, dorsal root ganglia (see Fig. 4) and the trigeminal ganglion (data not shown) were devoid of detectable \( H19 \) transcripts. This result was confirmed when the exposure time was increased tenfold. It was also noticeable that blood cells, although of mesodermal origin, were negative.

Another aspect of the regulation of \( H19 \) expression can be seen in the developing long bones of the 14.5 dpc embryo. Ossification of the cartilaginous skeleton is initiated during the second half of gestation. This phenomenon starts in foci called ossification centres, which then progressively enlarge. While \( H19 \) is transcribed in cartilage, no transcripts were detected in the surrounding mesenchyme (data not shown), indicating that expression is shut off during this particular process of

Pattern of \( H19 \) expression in the midgestation embryo

\( H19 \) expression during murine embryogenesis

Fig. 2. Nucleotide sequence analysis. (A) Distribution of ORFs within the sequence of LC10-8/1. \( \uparrow \), initiation codons; \( \downarrow \), termination codons. The scale is in amino acid codons. (B) Conceptual sequences of proteins encoded by the two longest ORFs. The sequence of LC10-8/1 has been deposited in the EMBL database under the accession number X58196. It differs from the Balb/c sequence of Pachnis et al. (1988), as modified by Brannan et al. (1990), as follows: +G at 82; A not G at 133; C not G at 171; G not A at 449; G not A at 466; +T at 467; +G at 782; C not T at 819; T not C at 821; T not A at 1084; +C at 1291; –C between 1358 and 1359; –G between 1481 and 1482; C not A at 1509; –G between 1510 and 1511; A not T at 1805; +G at 1821.
In situ hybridization of $^{35}$S-labeled antisense RNA probe of H19 to sections of mouse embryos from fertilization to gastrulation. 0.5 dpc embryos were sectioned in the oviduct, 5.5 and 6.5 dpc embryos in their decidua. Early (3.5 dpc) and late (4.5 dpc) blastocysts were flushed out of the uterus, fixed in 4% (w/v) paraformaldehyde and introduced into the swollen ampulla of a female on the first day of gestation, which provided a convenient support tissue to process the blastocysts (Copp, 1978). The tissue around the embryo in panels 3 to 6 is therefore also oviduct. Even-numbered panels are dark-field views of the corresponding odd-numbered bright-field photographs. The bars represent 0.05 mm.

Autoradiographic exposure times were 15 days for panels 1–6 and 2 days for panels 7–10. (1, 2) 0.5 dpc; (3, 4) 3.5 dpc; (5, 6) 4.5 dpc; (7, 8) 5.5 dpc; (9, 10) 6.5 dpc. Note that two maternal tissues, the oviduct and the deciduum are negative. D, deciduum; EC, ectoplacental cone; G, trophoblast giant cell; ICM, inner cell mass; M, mesoderm; O, oviduct; PE, primitive (embryonic) ectoderm; PN, parietal endoderm; T, trophectoderm; XE, extraembryonic ectoderm.

differentiation. The cessation of H19 transcription in terminally differentiated cells was also observed in several other tissues (see below).

Expression of H19 in adult tissues

Fig. 5A shows a northern blot analysis of H19 expression in adult tissues. High levels of RNA were
Fig. 4. In situ hybridization of $^{35}$S-labeled antisense RNA of $H19$ to a sagittal section of a 14.5 dpc embryo, shown under bright-field (1) and dark-field (2) illumination. The bar represents 1 mm and exposure time was 2 days. Note the general absence of signal in the CNS (B, brain; SC, spinal cord; DRG, dorsal root ganglion; L, liver; I, intestine).

Fig. 5. RNA blot analysis of $H19$ mRNA in foetal, newborn and adult tissues. (A) Expression in adult tissues. 10 µg of total RNA were loaded per lane. The filter was hybridized with an actin probe as a control (lower panel). (1) brain; (2) salivary gland; (3) thymus; (4) liver; (5) kidney; (6) spleen; (7) heart; (8) skeletal muscle; (9) lung; (10) intestine; (11) stomach; (12) testis. (B) Expression in foetal and newborn kidney. 2 µg of poly(A)$^+$ RNA were loaded per lane. The filter was hybridized with an actin probe as a control (lower panel). (1) 16 dpc; (2) 18 dpc; (3) newborn day 1; (4) newborn day 3; (5) adult. (C) Expression in foetal and newborn liver. 2 µg of poly(A)$^+$ RNA were loaded per lane. The filter was hybridized with an albumin probe as a control (lower panel). (1) 16 dpc; (2) 18 dpc; (3) newborn day 1; (4) newborn day 3; (5) adult.
found in skeletal muscle and low levels in thymus, heart and lung; no RNA was detectable in brain, liver, kidney, spleen, intestine, stomach or testis. In Fig. 1, it can be seen that no \( H19 \) transcripts accumulate in the uterus or in the cells of the deciduum, which are maternally derived. The fact that \( H19 \) expression is limited to a subset of adult tissues contrasts sharply with its widespread expression in the developing embryo and indicates that transcription ceases during perinatal development.

We have analysed in more detail the time course of \( H19 \) expression during kidney (Fig. 5B) and liver (Fig. 5C) development. In the 3-day-old newborn animal \( H19 \) was still expressed in both organs at a level comparable to that seen in a late gestation embryo. There was a decrease in transcript abundance during the second week after birth but there were differences between the two organs, since in a 2-week-old animal \( H19 \) mRNA was no longer detectable in kidney while there was a residual amount in liver which only disappeared during the third week after birth (data not shown).

### Discussion

The \textit{in vitro} differentiation of embryonic stem cells to form embryoid bodies is accompanied by structural and morphological changes similar to those seen during the pre- and immediately postimplantation stages of mouse development. We have sought to exploit this \textit{in vitro} differentiation system in order to isolate genes the expression of which is regulated during this important but relatively inaccessible period of embryogenesis.

The \( H19 \) cDNA clone was isolated because the corresponding transcript is present at higher levels in EBs than in ES cells. Our \textit{in situ} hybridisation studies show that within the EBs \( H19 \) transcription is restricted to the outer layer of cells which are predominantly of parietal endoderm type. \textit{In vivo} the gene is activated at the time of implantation in the cells of the extraembryonic lineages but not in the cells of the embryonic ectoderm which will ultimately give rise to the embryo proper. There is thus a satisfying correlation between the distribution of \( H19 \) transcripts in the EBs and in the embryo, confirming that the similarities between the formation of EBs and the peri-implantation stages of development observed at the morphological level can be extended to the patterns of expression of certain genes. It should be possible to isolate additional genes regulated at the time of implantation by more extensive screening of our cDNA libraries and we suggest that this approach should be generally useful for the analysis of gene regulation at this particularly inaccessible stage of mammalian development. However, the similarities between ES cell differentiation \textit{in vitro} and embryogenesis must be interpreted with caution. Our approach is likely to suffer from the drawbacks intrinsic to any \textit{in vitro} model system. We have, for example, shown that the lectin and calcyclin genes, isolated because they are abundantly expressed in ES cells, are not expressed, at the sensitivity of detection of \textit{in situ} hybridisation, in the ICM cells of the blastocyst (F.P., P.M.T., C.-T.J.C., J.-L. Guenet and P.W.J.R., manuscripts in preparation). The calcyclin gene is known to be serum responsive (Calabretta \textit{et al.} 1986) and its expression in ES cells may therefore simply be a reflection of the fact that they are cultured in the presence of serum.

The pattern of \( H19 \) expression around the time of implantation suggests that the gene may play a role during this stage of embryogenesis. \( H19 \) transcripts are not detectable in the fertilised egg or during the first three days of development. Very low levels of transcript become detectable in the expanded blastocyst (4.5 dpc), specifically in the cells of the trophectoderm. After implantation (5.5 dpc) \( H19 \) transcripts accumulate to extremely high levels in all of the cells of the conceptus with the exception of the primitive ectoderm. This latter cell population is pluripotent and it is from it that the embryo proper is derived. The other cell types, which express \( H19 \), contribute exclusively to the formation of the extraembryonic tissues which, although only transiently required, play an important nutritive and protective role in ensuring the survival of the foetus (Rossant, 1986). Cells of the extraembryonic lineages are terminally allocated and display limited developmental plasticity. At this stage, it appears that accumulation of \( H19 \) transcripts coincides with the loss of pluripotency. Implantation also corresponds to the onset of a phase of rapid cell growth in all lineages (Snow, 1977) but expression of \( H19 \) does not correlate with the rate of cell division in general, since it is not detectable in the primitive ectoderm. The expression of only a very limited number of genes has been directly analysed at this stage (reviewed by Adamson, 1986). It is notable that both \( H19 \) and the gene encoding the \( \beta \)-galactoside-binding lectin (F.P., P.M.T., C.-T.J.C., J.-L. Guenet and P.W.J.R., manuscript in preparation) are abundantly transcribed in extraembryonic cell types but not detectably expressed in the primitive ectoderm.

The transcription of \( H19 \) is also strikingly regulated at later stages of development. While \( H19 \) transcripts are not detectable in the primitive ectoderm of the egg cylinder, they are found, by midgestation, in a wide range of tissues derived from it. The levels are very high in cells of endodermal origin but somewhat lower in most cells of mesodermal origin. In contrast, the cells of the central and peripheral nervous systems are generally devoid of \( H19 \) transcripts. However, detailed examination of the brain and spinal cord has revealed \( H19 \) transcripts in an extremely small subset of neurones. This gene therefore constitutes an early marker defining a new population of cells within the developing central nervous system (B. Feldman, J. Dodd and F. P., unpublished data).

\( H19 \) was originally isolated by Pachnis \textit{et al.} (1984) during a search for genes that are regulated by \textit{raf}, a trans-acting locus involved in regulating the adult basal level of \( \alpha \)-foetoprotein mRNA. It was also isolated by Davis \textit{et al.} (1987) during a search for genes activated during the differentiation of C3H10T1/2 cells to myoblasts; they called it \( \text{MyoH} \).
Our \textit{in situ} hybridisation studies confirm that the gene is indeed expressed in early foetal liver and in the muscle cell progenitors, i.e. the myotome at 9.5 dpc (data not shown). However, \textit{H19} is also expressed in many other lineages; for example, in the somites transcription is not confined to the myotomes. Although \textit{H19} is expressed abundantly in most cells of the embryo, transcripts are found in only a few tissues of the adult (this report; Pachnis \textit{et al.} 1984, 1988; Davis \textit{et al.} 1987). There are relatively high levels in skeletal muscle, although even these are 10- to 20-fold lower than in the embryo (Pachnis \textit{et al.} 1988), and lower levels in cardiac muscle, spleen and thymus. \textit{H19} transcription appears to cease, or to be greatly reduced, in all of the tissues in which it was expressed during embryogenesis. In both kidney and liver, down-regulation begins during the first week after birth and is complete by the third week. As \textit{H19} is not a member of a gene family (data not shown) it is unlikely that its function is performed in the adult by a closely related product and it can therefore be viewed as an embryonic gene.

\textit{H19} maps to the distal half of chromosome 7, 16±10 cM proximal to the albino locus (F.P. and J.-L. Guénet, unpublished data), a region of the mouse genome known to be imprinted (reviewed by Solter, 1988). It has now been demonstrated that the \textit{H19} gene is subject to parental imprinting, the result of which is that only the maternal allele is expressed (Bartolomei \textit{et al.} 1991). Moreover, the gene encoding the insulin-like growth factor II (\textit{Igf}-2), which maps close to \textit{H19}, is also imprinted but in this case transcripts are derived from the paternal allele (DeChiara \textit{et al.} 1991). It is intriguing that the pattern of expression of \textit{H19} during early embryogenesis shares many features with that of \textit{Igf}-2 (Lee \textit{et al.} 1990).

The function of the \textit{H19} gene remains obscure. Our sequence data, and those of Pachnis \textit{et al.} (1988) and of Brannan \textit{et al.} (1990), show that the murine cDNA contains only two ORFs of significant length. However, although the gene is conserved in other mammals and in chicken, neither of the two ORFs is conserved in the human sequence (Brannan \textit{et al.} 1990). Moreover, Brannan \textit{et al.} (1990) were unable to detect translation of the 5' ORF \textit{in vivo} and showed that the RNA is associated with a 28S cytoplasmic particle and not with ribosomes. Thus although the gene is transcribed by RNA polymerase II, spliced and polyadenylated, in the fashion generally taken to be characteristic of mRNAs, Brannan \textit{et al.} concluded that the final product of this gene may be an RNA and not a protein. Whether the 3' ORF can be translated \textit{in vivo} has not been directly tested and it could be functional. However, the fact that it is not conserved in the human sequence argues against this possibility. The evolutionary conservation of the gene and its highly regulated expression suggest that it does perform an important role during development. Moreover, Brunkowski and Tilghman (1991) have recently provided further support for this idea by showing that ectopic expression of the wild-type gene but not of mutated derivatives is lethal between 14 dpc and birth. Precise definition of the role played by \textit{H19} will require both genetic analysis of gain- and loss-of-function mutants and biochemical analysis of the cytoplasmic particle that appears to be the predominant location of \textit{H19} RNA within the cell.

We are grateful to Robin Lovell-Badge, David Wilkinson, Rosa Beddington and Vassilis Pachnis for helpful discussions and support throughout this work.

We thank Christine Watson for her help in the cDNA cloning, Jim Jackson and Rob Krumlauf for gifts of RNA, Davor Solter and Keith Willison for antibodies and K. Fahrner for her cDNA library.

During the initial phases of this work F.P. was supported by a Long Term Fellowship from the European Molecular Biology Organisation, while P.W.J.R. held a Cancer Research Campaign Career Development Award. C.-T.J.C. held a Medical Research Council Training Fellowship. This work was paid for by the Cancer Research Campaign and the Medical Research Council.

\section*{References}


