Ectopic expression of the H19 gene in mice causes prenatal lethality

Mary E. Brunkow¹ and Shirley M. Tilghman

Howard Hughes Medical Institute and Department of Molecular Biology, Princeton University, Princeton, New Jersey 08544 USA

The mouse H19 gene is expressed in a broad array of tissues of both endoderm and mesoderm origin in the developing mouse embryo. Its expression is repressed in all tissues except skeletal muscle shortly after birth. This gene is unusual in that it may not encode a protein, despite its conservation in mammals. The RNA product is found as a spliced and polyadenylated RNA in a cytoplasmic particle. To probe whether this unusual gene is functional, excess copies were introduced into mouse zygotes. Transgenic progeny were obtained at a very low frequency, but in no instance was the transgene expressed. That the gene itself was deleterious to embryos was established by introducing into zygotes a mutant of the structural gene in which its most conserved segment was deleted. Transgenic founders were obtained at a higher frequency, and these expressed the altered transgene at high rates in a subset of the tissues that express the endogenous H19 gene. The lethal effects are manifested late in gestation, between day 14 and birth.

[Key Words: H19 gene; mouse embryo; ectopic expression; transgenic expression]

Received January 11, 1991; revised version accepted March 5, 1991.

The mouse H19 gene encodes one of the most abundant RNA polymerase II products in the mouse embryo, in cells of both endoderm and mesoderm lineages. This high-level expression is repressed in all tissues except skeletal muscle shortly after birth. The gene is strongly conserved at the nucleotide level in humans and rodents. However, despite the fact that the mature 2.5-kb RNA is spliced and polyadenylated, it does not contain an open reading frame characteristic of mRNAs (Pachnis et al. 1988; Brannan et al. 1990). The localization of the H19 RNA to a cytoplasmic ribonucleoprotein (RNP) particle led us to the conclusion that this RNA may not encode a protein (Brannan et al. 1990).

The H19 gene was originally identified as a fetal liver-specific cDNA whose repression after birth paralleled that of the α -fetoprotein (AFP) gene (Pachnis et al. 1984). In that tissue both genes are influenced by two *trans*-acting loci, *raf* and *Rif*. *raf* affects the adult basal level of AFP and H19 mRNAs while *Rif* determines the degree of their inducibility during liver regeneration (Olsson et al. 1977; Belayew and Tilghman 1982). H19 has since been cloned by at least three independent groups: first by Davis et al. (1987) during their screen for genes that are activated during the differentiation of C3H10T1/2 cells (referred to in that report as myoH) and, more recently, by F. Poirier and colleagues (pers. comm.) and Wiles (1988) (referred to in that report as SO5A) during screens for genes that are activated during embryonic stem cell

and embryonal carcinoma cell differentiation, respectively. In each instance it was identified by virtue of its activation in differentiated cells, rather than by virtue of its function.

As a first step in understanding the function of this unusual RNA, extra copies of the gene have been introduced into the mouse germ line in an effort to generate possible "gain-of-function" mutations. It was hoped that the phenotype of such mutations might illuminate the normal function of the gene. These experiments demonstrated that the transgene acts as a dominant lethal mutation. The deleterious effects are first manifested in late gestation, between embryonic day 14 (day e14) and birth.

Results

Microinjection of H19 gene into mouse zygotes

To achieve overproduction of the H19 gene in vivo, we utilized a segment of DNA that spanned from 4 kb upstream of the transcriptional start site to 7.5 kb 3' of the gene (Yoo-Warren et al. 1988). This included the promoter region and two 3' enhancers as illustrated in Figure 1 for construct RRSR, named after the restriction fragments that define its borders. This segment of DNA is sufficient to direct high-level transcription in hepatoma cells (Yoo-Warren et al. 1988). In all of these experiments, the structural gene was modified by the insertion of two *Xba*I linkers into a *Sac*I site at +580 bp in exon 1. The *Sac*I site was chosen because it encompasses one of the least conserved regions of the gene (Brannan et al. 1990). By using this variable site, we hoped that the

¹Present address: Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, Ontario M5G 1X5 Canada.

Figure 1. H19 gene constructs used for microinjection into mouse embryos. The five-exon H19 gene body (~2.8 kb) is represented by the stippled boxes, the 3' enhancers at +8 and +9.5 kb, relative to the transcription initiation site, by hatched ovals. The restriction sites, whose positions are used to name the constructs are *Bam*HI (B), *Eco*RI (R), and *Sal*I (S). Δ Bam refers to an intragenic deletion of a 0.9-kb *Bam*HI fragment. All of the transgenes have been marked by the insertion of two *Xba*I linkers at a *Sac*I site in the first exon at position +580 bp. The position of a cluster of 3' distal DNase I hypersensitive sites is indicated by a vertical arrow. The numbers refer to kilobase pairs of DNA, with the transcriptional start site defining +1 bp.

linker would not interfere with the normal function of the gene. This marker was used to detect the transgene by Southern blot analysis of *Xba*I-digested genomic DNA, as well as the transgene transcript using an RNase protection assay that distinguishes it from the endogenous transcript, as described below.

The 14.5-kb RRSR fragment was purified away from vector sequences and microinjected into fertilized mouse eggs derived from (C57BL/6 $\times$ SJL) F₁ intercrosses, and live pups were analyzed for the presence of the transgene. Two founder animals, each carrying three to four copies of the transgene, were obtained at a very low frequency (2 of 569 embryos transferred to the oviducts of pseudopregnant females, or 0.35%). Surprisingly, neither of these lines expressed the transgene in any tissue examined (data not shown). Southern blot analysis revealed no gross rearrangements of the H19 transgenes that might explain the absence of any expression.

Identification of a DNase I hypersensitive site 3' of the H19 gene

Two possibilities to explain the absence of transgene expression suggested themselves. It was possible that the

transgene itself was deleterious to the developing embryo, and, therefore, the only live progeny that could be obtained were those in which the transgene had integrated into a region of the genome that inhibited expression. On the other hand, the low frequency could have been a technical problem, and the lack of expression could have been the consequence of the absence of important regulatory elements that are necessary to reconstitute an active H19 gene locus in vivo but not in hepatoma cells. To address the latter possibility, chromosomal regions 5' and 3' farther from the endogenous H19 gene were assayed for DNase I hypersensitivity, on the premise that regions in chromatin that are functionally important for transcriptional activation and tissue specificity often show hypersensitivity to the nonspecific endonuclease DNase I (Emerson et al. 1985; Plumb et al. 1986; Grosveld et al. 1987; Lui et al. 1988).

Nuclei were isolated from liver of neonatal mice at a time when the H19 gene was actively transcribed (Pachnis et al. 1984; Brannan et al. 1990) and treated with increasing concentrations of DNase I. The DNA was purified, digested to completion with *Bam*HI, and analyzed by Southern blotting using a DNA probe derived from a region 3' of the enhancers. As shown in Figure 2, the 11.4-kb *Bam*HI fragment was cleaved at multiple sites

Figure 2. Mapping DNase I hypersensitivity sites 3' to the H19 gene. Nuclei were isolated from neonatal liver and treated with 0–50 $\mu\text{g/ml}$ of DNase I for 3 min at 25°C. Genomic DNA was purified, digested to completion with *Bam*HI, and analyzed by blotting. The autoradiograph shows the specific cleavage products of an 11.4-kb *Bam*HI restriction fragment that spans the 3' end of the gene. In the diagram, the H19 gene body is represented by the stippled box; the enhancers at +8 and +9.5 kb are represented by hatched ovals. The restriction sites included are *Bam*HI (B), *Bgl*II (Bg), *Eco*RI (R), *Hind*III (H), *Sal*I (S), and *Xba*I (X). The heavy horizontal line represents the sequences included in the RRSR transgene. The probe used for hybridization was a 1.5-kb *Hind*III–*Bam*HI fragment, as indicated. The prominent cluster of DNase I hypersensitive sites that correspond to the 1.5-kb cleavage product is indicated by the heaviest vertical arrow. Other hypersensitive sites are shown on the map with arrows representing the relative intensities of bands on the autoradiograph.

by DNase I. The digestion products migrating at 6.0 and 4.5 kb mapped to the two liver-specific enhancers, as expected. In addition, a very strong cluster of hypersensitive sites generated a band at 1.5 kb, and these mapped at +11.5 kb, 1 kb farther downstream of the RRSR transgene. No equivalent hypersensitive sites were detected 5' of the gene (data not shown).

To test whether the very strong hypersensitive site was indeed a regulatory element necessary for H19 gene expression *in vivo*, a 2.5-kb *Eco*RI–*Bam*HI fragment containing it was cloned onto the 3' end of RRSR to create the construct RRSB (Fig. 1). Microinjection of purified RRSB DNA into the pronuclei of mouse eggs resulted in the generation of no founders from 423 embryos implanted.

Microinjection of a deletion mutant of the H19 gene

The failure to generate transgenic mice that carried RRSB in the germ line forced us to consider the possibility that the H19 gene product, if inappropriately ex-

pressed, was deleterious to the developing embryo. To confirm that it was the gene itself, a deletion mutant was generated, by removing an internal 0.9-kb *Bam*HI fragment containing the most highly conserved domain between the mouse and human genes (Brannan et al. 1990). We reasoned that its deletion, which removes approximately one-half of the first exon, all of the second, and a portion of the third exon, would result in a nonfunctional gene.

The constructs RRSR Δ Bam and RRSB Δ Bam (Fig. 1) were generated and tested by transfection into Hep3B cells to confirm that the deletion did not affect either transcriptional competence or stability of the transcript (data not shown). They were then microinjected into mouse embryos. Eight founder animals were obtained (8/1252 eggs implanted, or 0.64%), and six of the seven analyzed in detail expressed the transgene at high levels (see below). This strongly suggested that it was the intact gene itself that was deleterious to mouse embryos.

To confirm that there was a significant difference in the ability of the wild-type and internally deleted trans-

genes to be incorporated stably into the mouse germ line in a transcriptionally active state, eggs were injected with either RRSB or RRSB Δ Bam, and equal numbers were coimplanted into foster mothers. This procedure would eliminate any differences due to the harvesting, injection, or transfer of the eggs at different times. Two RRSB Δ Bam transgenic founders were obtained from 452 eggs transferred, and one that survived for analysis expressed the transgene (see below).

From 434 eggs injected with RRSB, two founders were obtained, both of which were mosaic for the transgene, based on the observation that the ratio of intensities of the endogenous DNA to transgene DNA on Southern blots was greater than two (data not shown). One founder died before it could be analyzed; the second one failed to transmit its transgene to progeny. A postmortem Southern blot analysis of the transgene distribution in this mouse revealed that the transgene, represented by the lower *Xba*I band in Fig. 3, was under-represented in the brain, liver, and testes, as compared with most other tissues. The absence of the transgene in both testes and sperm DNA was consistent with its absence in progeny derived from this mouse.

Tissue specificity of H19 transgenes in fetal and neonatal mice

Seven transgenic lines were established by mating founders to CD1 mice, four with RRSR Δ Bam and three with RRSB Δ Bam (see Table 1), and these were analyzed in detail for their patterns of expression. Transgene RNA was detected in total RNA from tissues with an RNase protection assay, using a probe that spans the two *Xba*I linkers of the transgene (Fig. 4A). Because the probe includes the two *Xba*I linkers, the transgene transcript protects a 115-base fragment, whereas the endogenous H19

Table 1. Summary of transgenic lines obtained by microinjection

Transgene	Lines obtained ^a			Transcriptional activity
	germ line	mosaic	other	
RRSR	2 ^b	0		none
RRSB	0	2 ^c		ND ^d
RRSR Δ Bam	4	0	1	yes
RRSB Δ Bam	2	1	2	yes

^aThe distinction between germ line, mosaic lines, and other is based on both gene dosage measurements of Southern gels and progeny testing. Three additional founders that carried RRSR Δ Bam or RRSB Δ Bam died at birth or as adults before progeny could be obtained, and are listed as other.

^bOne of these was sacrificed at birth to assay transgenic expression; the other was bred to establish a line. Neither expressed the transgene.

^cOne mosaic RRSB founder did not transmit the transgene to 70 progeny analyzed; the other died as an adult before producing offspring.

^d[ND] Not determined.

transcript protects 100 bases of the probe, up to the *Xba*I insertion.

All of the transgenic lines expressed the transgene at high levels in the fetal and neonatal liver and gut (Fig. 4B). There was a striking absence of expression in the yolk sac and skeletal muscle, which express the endogenous gene at levels comparable to liver and gut in the neonate. In the high-copy-number transgenics, expression of the transgene was also apparent in other tissues, including the brain where normally the endogenous gene is silent (RRSB Δ Bam line 908 and RRSR Δ Bam line 382). Neither of these aberrant patterns of expression (i.e., absence in yolk sac and muscle and expression in the brain) could be attributed to the 3' distal DNase I hypersensitive site, which had no appreciable effect on the tissue-specific pattern of transgene expression.

A fourth RRSR Δ Bam transgenic line, 367, did not express the transgene in any tissue at any time (data not shown). This line carries only one or two copies of the transgene, as determined by analysis of a slot blot.

Temporal regulation of the H19 transgenes

The repression of transcription of the H19 gene after birth is highly tissue dependent. For example, the levels of H19 RNA in the liver decrease several hundred-fold during the third week after birth, while in the gut the levels fall below detection by 1 week of age. In contrast, the extent of repression in skeletal muscle is only 10-fold, resulting in the maintenance of high levels of H19 RNA in the adult (Pachnis et al. 1988; see Fig. 5). In all other tissues examined, where the concentration of H19 RNA starts out significantly lower than in liver, gut, and skeletal muscle, no RNA is detected in the adult animal.

For the low-copy-number lines (840 and 109), no expression of the transgene was detected in adult tissues (Fig. 5), with the exception of very low levels in the heart

Figure 3. Distribution of RRSB DNA in a transgenic founder. DNA was isolated from the tissues indicated and digested with *Xba*I. The DNA was separated on a 1% agarose gel, blotted onto nitrocellulose, and hybridized to a labeled fragment that detects the endogenous H19 gene (upper band) and the transgene DNA (lower band). L and R refer to the left and right testes, respectively, and the sperm derived from them.

Figure 4. Analysis of H19 transgene expression in neonates. (A) (1) Exon 1 of the H19 gene is represented by the stippled box; the site of the *Xba*I linker insertion and transcription initiation site (horizontal arrow) are indicated. (2) A 137-bp *Eco*RI (R)–*Bam*HI (B) fragment containing an internal *Sma*I deletion was cloned into the vector SP64. (3) The vector was linearized with *Pvu*II (P) to generate a 318-base antisense RNA probe. This probe will protect 115 bases of the transgene transcript and 100 bases of the endogenous transcript. RNA was extracted from animals carrying either the RRSRΔBam (B) or RRSBΔBam (C) transgene. The probe was hybridized to 3 μg of total RNA from yolk sac of day 16–18 fetuses (Y) and liver (L), gut (G), skeletal muscle (M), heart (H), kidney (K), and brain (B) from day-1 neonates. The independent lines analyzed are indicated by a # sign. Marker (m) is *Msp*I-digested pBR322, end-labeled by the Klenow fragment of DNA polymerase I. The full-length 318-base probe (P) is shown, and the 115-base transgenic (TRANS) and 100-base endogenous (ENDOG) products are indicated. The apparent variability in the levels of endogenous H19 RNA in this figure and in Figs. 5 and 6 are the consequence of transgene RNA, which was in some cases present at levels sufficient to saturate the labeled RNA probe. The copy numbers for the transgenic lines described here and in Figs. 5 and 6, are 109, 4 copies; 382, 35 copies; 699, 60 copies; 840, 4 copies; 908, 35 copies; and 9, ~200 copies. These numbers were determined by slot blot analysis of transgenic vs. nontransgenic genomic DNA.

in line 840 (seen only with long exposures). In several high-copy-number transgenics (382 and 699) persistent expression in liver, gut, kidney, and heart was noted occasionally, although the absence in skeletal muscle was maintained. In each case, however, the adult levels were below those observed in the neonate.

To characterize further the response of the RRSBΔBam and RRSRΔBam transgenes to developmental cues, transgene expression was monitored in tissues at several stages from before birth through adulthood. In Figure 6, the temporal modulation of five lines is illustrated for liver and gut. In liver, both the transgene RNA and endogenous H19 RNA declined at the same time, primarily between 1 and 2 weeks of age; although in one of the high-copy-number animals (908), persistent expression of the transgene was evident even at 4 weeks. Apparent differences in timing reflect the fact that the RNA is declining rapidly during this period, and a difference of

from 1 to 2 days can result in marked differences in H19 RNA concentration. The repression of both endogenous and transgene H19 RNA in the gut was rapid after birth, although persistent levels were seen once again in the high-copy-number transgenics (908 and 9) (Fig. 6). In the high-copy-number lines, it was difficult to detect the expression of the endogenous H19 RNA, due to the vast overexpression of the transgene (see the liver and gut samples for day 1 of 699). By using Northern blot analysis, we confirmed that the endogenous RNA is appropriately expressed at that time (data not shown).

These experiments measured steady-state levels of RNA. Therefore, any differences in timing between transgenic and endogenous expression patterns could be caused by either prolonged transcription of the transgene or increased stability of the transgene transcript. The half-lives of the H19 RNA and the significantly shorter RRSBΔBam and RRSRΔBam transgenic transcripts have

Figure 5. Analysis of H19 transgene expression in adults. An aliquot of 30 μ g of total RNA from 4-week-old animals carrying either the RRSR Δ Bam (A) or RRSB Δ Bam (B) transgene was analyzed by the RNase protection assay as in Fig. 3. Tissues analyzed were liver (L), gut (G), skeletal muscle (M), heart (H), and kidney (K). The 115-base transgenic (TRANS) and 100-base endogenous (ENDOG) products are indicated.

not been determined. The apparent persistent, ubiquitous expression of the transgene in the high-copy-number transgenics is likely the consequence of very low expression per gene copy that is amplified by the high copy number.

Defining the period of embryonic lethality of the H19 gene

The generation of transgenic progeny that expressed the mutated form of the H19 gene, coupled with the absence of germ-line transgenic progeny that expressed the wild-type gene, suggested that overexpression of the latter was deleterious to embryos. In an effort to define the period of lethality, one-cell zygotes were microinjected with either RRSB or RRSB Δ Bam and transferred into the oviducts of pseudopregnant foster mothers. Embryos were then sacrificed at different times during gestation and scored for the presence and expression of the transgenes.

Of 193 embryos that had received RRSB, 39 were still alive by day e11, and of these, 11 were transgenic (>25%; Table 2). More importantly, when 6 of these 11 were tested for RNA expression, all expressed the transgene at high levels (Fig. 7A). Similar results were obtained at day e13 and day e14, that is, embryos that expressed the transgene at high levels were readily detected. Although fewer embryos carrying the control transgene RRSB Δ Bam were examined, ~20% were transgenic.

There was no difference in the number of resorption

sites in RRSB and RRSB Δ Bam foster mothers at this stage of development (data not shown). All sites were tested for the presence of the transgene by isolating genomic DNA. In general, they contained only maternal DNA, as determined using a restriction fragment-length polymorphism that distinguished the C57BL/6J and SJL/J alleles of the AFP gene from the CD1 allele in the foster mother (Belayew and Tilghman 1982; data not shown). Therefore, they could not be scored for the presence of the transgene.

At day e15 and day e16, no live transgenic animals that expressed the wild-type transgene were obtained from 118 eggs that had been microinjected with RRSB. Three animals, which appeared to be highly mosaic for the transgene, based on the relative intensity of the diagnostic 1.4-kb *Xba*I fragment (Fig. 7B, lanes 1–3), were recovered but no transgene RNA was detected in their tissues (data not shown). A resorption site recovered at day e15 was shown by DNA blotting to contain both fetal DNA and the transgene (lane 4), suggesting that the fetus had recently died. Finally, one fetus, which had just died, was recovered at day e16 and was shown by Southern blotting to have been transgenic (Fig. 7B, lane 5).

Discussion

Lethality of the H19 gene in transgenic mice

The H19 gene encodes a conserved 2.5-kb RNA that is not apparently translated but is present in an abundant RNP particle (Brannan et al. 1990). Its nonuniform distribution in mammals and its striking temporal regulation distinguish it from the constitutively expressed RNP particle RNAs found, for example, in small nuclear ribonucleoproteins (snRNPs) (Maniatis and Reed 1987) and the signal recognition particle (Walter and Blobel 1982). Our goal in these studies was to introduce excess copies of the mouse H19 gene into the mouse germ line, in an effort to gain insight into its possible function in vivo. The gene was introduced under its own transcriptional control elements, with the hope that classical gain-of-function mutations would be generated. The overexpression of this gene perturbed embryonic development, which was manifested late in gestation by a failure to survive. Only four live transgenic founders were obtained from >1420 eggs implanted. In addition, the transgene in these four animals was either silent, in the case of two RRSR transgenics, or mosaic for the transgene, in the case of two RRSB animals. These are precisely the kinds of animals that would be expected if the gene was having a deleterious effect. That this deleterious effect was due to the gene itself, and not due to the flanking DNA, was demonstrated by our ability to obtain transgenic progeny with the internally mutated H19 gene. These progeny expressed the transgene at high levels in a subset of the tissues where the gene is normally expressed.

The mouse embryo can tolerate expression of the wild-type H19 transgene until at least day e14. Apparently normal transgenic fetuses could be recovered up to that

Figure 6. Temporal modulation of H19 transgenes in liver and gut. Levels of H19 transgene expression were determined in liver and gut, relative to levels of endogenous H19 gene expression, in two RRSR Δ Bam (A) and three RRSB Δ Bam (B) lines. Aliquots of 3 μ g of total RNA from fetal, 1-day-old, 1-week-old, and 2-week-old tissues, and 30 μ g of total RNA from 3- and 4-week-old tissues were analyzed by RNase protection as described in Materials and methods and the legend to Fig. 3. The upper panels in A and B illustrate the temporal modulation of gene expression in liver; the lower panels show pattern of expression in the gut. The sizes of the transgene and endogenous protected products are indicated, as described previously. Fetal gut RNA sample from RRSB Δ Bam line 908 is absent in this experiment.

time, and the majority of these expressed the transgene at high levels. From day e15 on, no live animals that expressed the transgene were recovered. The high frequency of detecting transgenic day e11–e14 embryos expressing the wild-type gene (>20%) contrasts with the complete absence of transgene expression in the two RRSR postnatal founders. It reinforces the argument that the failure of the latter to express the transgene is not the consequence of a lack of regulatory sequences, or a repressor element in the 0.9-kb *Bam*HI fragment deleted in

RRSR Δ Bam and RRSB Δ Bam. Rather, the animals that survive to birth do so because the transgene is integrated in a silent configuration.

Three reasons for the lethal effects can be proposed. It is possible that the overexpression of the transgene in the appropriate tissues, presumably liver and/or gut, is deleterious to the embryo. Certainly we observe levels of transgene expression in prenatal animals that either approach or exceed endogenous levels. If transgene expression in liver and/or gut is deleterious, it suggests that the gene dosage of the H19 gene is under strict control. The H19 gene has been mapped to the distal region of mouse chromosome 7 (Pachnis et al. 1984; T. Glaser and D. Housman, pers. comm.), within a region that undergoes both maternal and paternal genetic imprinting (Cattanach 1986; Solter 1988). Thus, it is conceivable that the prenatal lethality is the consequence of a disruption in tightly controlled H19 gene dosage.

On the other hand, it may be that expression in inappropriate tissues, such as that observed in the brain of high-copy-number lines, caused the lethal effects. In that case, however, we might have expected to obtain some low-copy-number lines that expressed the wild-type gene in an appropriate fashion. Finally, it is formally possible that the two *Xba*I linkers in the first exon, which were used to mark the transgene for detection, created a dominant lethal gene.

The nonuniform distribution of the transgene in the mosaic animal that was analyzed in detail (Fig. 3) is consistent with either of the first two models. Three tissues exhibited a very low distribution of RRSB DNA: the

Table 2. Analysis of H19 transgenics in utero

Transgene	Days of gestation	Number of eggs transferred	Number alive	Number transgenic (%) ^a	Transgene expression
RRSB	e8–10	56	19	3 (16)	ND ^b
	e11	193	39	11 (28)	yes
	e12–14	149	22	5 (23)	yes
	e15–16	118	27	5 (18) ^c	no ^c
RRSB Δ Bam	e9–10	78	15	3 (20)	ND ^b
	e11	107	20	4 (20)	yes
	e12–16	83	7	0	—

^aPercent refers to the ratio of transgenic fetuses obtained to the number of live fetuses observed at the time of dissection.

^b[ND] Not determined.

^cThree live transgenic fetuses obtained at e15–16 were mosaic for the transgene based on Southern blot analysis, but none expressed the transgene. Two other transgenic progeny were detected, one that appeared to have died immediately prior to dissection and one resorption site (see Fig. 7B).

Figure 7. Analysis of founder transgenic embryos. Either RRSB or RRSB Δ Bam was injected into zygotes and transferred into the oviducts of CD1 foster mothers. The mothers were sacrificed, and RNA and DNA were extracted from the live embryos. (A) RNA from whole embryos (e11–e13) or separated body (B) and head (H) (e14) were analyzed by RNase protection. (N) Non-transgenic littermates. Note the absence of transgene expression in only 1 of 11 transgenic embryos, at day e12. (B) DNAs from three mosaic fetuses (lanes 1–3), one resorption site (lane 4), and one dead embryo (lane 5) obtained at e15 and e16 were digested with *Xba*I, subjected to gel electrophoresis, and transferred to nitrocellulose. The filter was hybridized with a 1.4-kb *Xba*I fragment of the H19 gene to detect the diagnostic 1.4-kb transgene band.

liver, which presumably would have expressed it at high levels; and the brain and testes, which would not. At the present time, we are generating mosaic animals with embryonic stem cells transfected with wild-type H19 genes to address the mechanism underlying the late prenatal lethality.

There is a striking difference in the frequency of obtaining transgenic progeny before birth, as opposed to after birth for both RRSR and RRSB (4.7% vs. 0.28%, respectively), a result that is predicted if the gene is deleterious before birth. Rather surprisingly, there is also a difference for the “mutant” RRSB Δ Bam and RRSR Δ Bam constructs, of 2.6% for prenatal versus 0.6% for postnatal transgenic animals. Although some of this difference may be explained by differences in efficiencies of generating transgenic animals during the course of this study, it may also suggest that the highly expressed minigene is

not entirely innocuous. In fact, transgenic animals carrying the deletion mutant were obtained at only twice the frequency of wild-type transgenic animals after birth. However, unlike the wild-type lines, the majority of the mutant transgenes were expressed at high levels, at least in a subset of the tissues that express the endogenous gene. It remains to be determined whether the survival of these lines was a consequence of some common aspect of their pattern of expression, such as the absence of expression in the yolk sac.

There are few precedents for genes transcribed by RNA polymerase II that do not encode functional mRNAs other than a subset of the small nuclear RNAs (snRNAs); the *Hsrw* gene in *Drosophila melanogaster*, which has been suggested to function as a translational regulator (Fini et al. 1989); and an antisense RNA in *Xenopus laevis*, which may regulate the concentration of fibroblast growth factor mRNA (Kimelman and Kirschner 1989). None of these examples is consistent with our observations regarding H19 RNA: Its structure is very different from that of the short RNAs associated with the splicing machinery, we cannot detect its presence associated with the translation machinery, and no “sense” transcript has been detected to date.

Transcriptional control of the H19 gene in transgenic mice

The pattern of expression of the RRSR Δ Bam and RRSB Δ Bam transgenes suggests that the *cis*-acting elements that control the liver and gut-specific transcription of the H19 gene have been separated from those that control the mesoderm specificity. The 17 kb of mouse genomic DNA analyzed in this study had been shown previously to direct high-level transcription in hepatoma cells (Yoo-Warren et al. 1988). The transgenes did indeed exhibit expression in the liver, as well as the intestine, another endodermally derived tissue, at levels that approximated those of the endogenous gene. This high-level expression was characteristic of most, but not all, integration sites and appeared to follow some degree of copy-number dependence, in that the higher-copy-number lines expressed the transgene at higher levels.

The constellation of gene products in the visceral endoderm of the yolk sac is very similar to that in liver and gut (Meehan et al. 1984). The striking absence of transgene expression in the yolk sac was surprising, especially because we had demonstrated previously enhancer-dependent expression of the H19 gene in PC13 embryonal carcinoma cells, after differentiation into visceral endoderm (Yoo-Warren et al. 1988). It is possible that the yolk sac expression observed *in vivo* derives not primarily from the endoderm layer but from the underlying mesodermal layer. In that case, the absence of transgene expression is not surprising, as no mesodermal tissue expresses the transgene to any appreciable extent. Otherwise, it must be proposed that the visceral endoderm utilizes a distinct set of regulatory elements from those in liver and gut. This is not entirely without precedent, as each of the three AFP gene enhancers has been shown

to have overlapping specificities for liver, gut, and visceral endoderm (Hammer et al. 1987).

It is less surprising that the transgenes were uniformly not expressed in cells of mesodermal origin, despite the fact that there is high-level endogenous H19 gene expression in muscle of both fetuses and adults. Examination by transient transfection of DNA extending 15 kb 5' and 19 kb 3' of the gene RRSR has failed to reveal regulatory elements that could serve as muscle-specific enhancers in several muscle-derived cell lines (M.E. Brunkow, unpubl.). Rather, it is likely that the H19 transcriptional control region is very large and that the endodermal and mesodermal transcriptional control elements are separately encoded.

Several examples exist in *Drosophila* where independent elements, spanning many kilobases of DNA, direct differential gene expression. The segmentation gene, *fushi tarazu*, is expressed in a pattern of seven stripes at the blastoderm stage, and later during neurogenesis, in a specific set of neuronal precursors (Hafen et al. 1984; Carroll and Scott 1985; Hiromi et al. 1985). Three separable upstream sequences have been identified that are important for these different patterns of expression. Likewise, the upstream sequences of the small heat shock gene, *hsp26*, contain two independently acting *cis* elements; the more promoter-proximal one confers heat shock inducibility, while the second element, located farther upstream, directs expression of *hsp26* in the ovary (Cohen and Meselson 1985). The yolk protein genes *yp1* and *yp2* likewise have two separate 5' elements that direct expression in fat bodies and ovaries (Garabedian et al. 1985).

Although we favor the conclusion that the H19 gene control region is encoded in separable and widely dispersed elements, only a subset of which have been included in the transgenes in this study, this has not yet been rigorously demonstrated. Additional experiments are required to rule out the possibility that the six RRSRΔBam and RRSBΔBam lines that we have characterized were selected precisely because their pattern of transgene expression was compatible with survival through embryogenesis. This cautionary note is based on our data, which suggest that the internally deleted transgenes may have deleterious effects, albeit less pronounced ones than exhibited by the wild-type gene, which is never expressed.

Materials and methods

Transgenic mice and vector constructions

The plasmids used for generating transgenic mice include 4 kb of DNA 5' to the H19 structural gene, and either 8 kb (RRSR) or 10.5 kb (RRSB) of DNA 3' to the gene. The constructs RRSRΔBam and RRSBΔBam have a deletion of a 0.9-kb *Bam*HI fragment internal to the 2.8-kb gene body. This deletion removes half of exon 1, all of exon 2, and part of exon 3 (see Pachnis et al. 1988). All transgenes were marked by the insertion of two synthetic *Xba*I linkers 580 bp downstream of the transcription initiation site (Yoo-Warren et al. 1988).

The H19 gene fragments were cloned into the *Eco*RI (for RRSR and RRSRΔBam) or *Eco*RI–*Bam*HI (for RRSB and RRSBΔBam) sites of a pUC18 derivative. This derivative was engineered by synthetic oligonucleotide cloning to contain *Cl*aI sites at either end of the polylinker. DNA was prepared for pronuclear injection by digesting the plasmids with *Cl*aI, separating the H19 gene inserts (containing 6–70 bp of polylinker DNA on the 5' and 3' ends) from the vector by agarose gel electrophoresis, and purifying the DNA by elution from NA45 membrane (Schleicher & Schuell; Sambrook et al. 1989) and cesium chloride centrifugation (Hogan et al. 1986). The lengths of the RRSR, RRSB, RRSRΔBam, and RRSBΔBam inserts are as follows: 15, 17.5, 14.1, and 16.6 kb, respectively.

The concentration of these DNA fragments was estimated by ethidium bromide fluorescence relative to a marker DNA of known concentration and adjusted to 4–8 ng/μl. DNA was microinjected into one of the pronuclei of fertilized one-cell mouse eggs derived from (C57BL/6 × SJL) F₁ intercrosses (Hogan et al. 1986). Injected embryos were transferred to the oviducts of pseudopregnant CD-1 females and allowed to develop to term. Founder animals were identified by blot analysis of tail DNA (Southern 1975) and subsequently bred to maintain transgenic lines.

DNA and RNA analysis

Typically, 10 μg of genomic DNA was digested with a sixfold excess of the appropriate restriction endonuclease, electrophoresed on 1.0% agarose in 1× TBE running buffer, and transferred to nitrocellulose (Millipore) by the method of Southern (1975). To identify transgenic animals, tail DNA was digested with *Xba*I, and filters were probed with a nick-translated (Maniatis et al. 1975) 1.4-kb *Xba*I fragment that hybridizes to a 9.0-kb endogenous *Xba*I fragment and a 1.4-kb *Xba*I fragment that results from restriction at the artificial *Xba*I linkers in the first exon the H19 transgene.

Expression of the transgene was analyzed by an RNase protection assay (Melton et al. 1984). The probe included a 135-bp *Sma*I–*Bam*HI fragment (+566 to +682) that spans the two *Xba*I linkers inserted at +580 bp relative to the start site of H19 gene transcription, cloned into SP64 (Promega). The vector was linearized at the unique *Pvu*II site of SP64, purified by agarose gel electrophoresis, and used as a template in an in vitro transcription reaction utilizing [³²P]CTP and SP6 polymerase. From 3 to 30 μg of total RNA was hybridized to 1 × 10⁵ to 5 × 10⁵ cpm radiolabeled probe overnight at 55°C or 60°C in a buffer containing 80% formamide, 40 mM PIPES (pH 6.7), 0.4 M NaCl, and 1.0 mM EDTA. Single-stranded RNA was then digested in buffer containing 10 mM Tris-HCl (pH 7.5), 300 mM NaCl, 5 mM EDTA, 40 μg/ml of RNase A, and 2 μg/ml of RNase T1 (Sigma Chemical Co.), and incubated for 1 hr at 25°C. The digestion was inactivated with SDS and proteinase K, extracted once with phenol/chloroform/isoamyl alcohol (PCI; 25 : 24 : 1), and precipitated with 2 volumes of ethanol, and the products were analyzed on gels of 10% acrylamide, 7 M urea, and 1× TBE. Gels were exposed to Kodak XAR-5 x-ray film with intensifying screens at –70°C.

Acknowledgments

We thank the members of the laboratory for encouragement and many helpful discussions. This work was supported by grants from the National Institutes of Health (CA44976) and The National Foundation of the March of Dimes. S.M.T. is an Investigator of the Howard Hughes Medical Institute.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

- Belayew, A. and S.M. Tilghman. 1982. Genetic analysis of α -fetoprotein synthesis in mice. *Mol. Cell. Biol.* **2**: 1427–1435.
- Brannan, C.I., E.C. Dees, R.S. Ingram, and S.M. Tilghman. 1990. The product of the H19 gene may function as an RNA. *Mol. Cell. Biol.* **10**: 28–36.
- Carroll, S.B. and M.P. Scott. 1985. Localization of the *fushi tarazu* protein during *Drosophila* embryogenesis. *Cell* **43**: 47–57.
- Cattanach, B.M. 1986. Parental origin effects in mice. *J. Embryol. Exp. Morphol.* (suppl.) **97**: 137–150.
- Cohen, R.S. and M. Meselson. 1985. Separate regulatory elements for the heat-inducible and ovarian expression of the *Drosophila hsp26* gene. *Cell* **43**: 737–746.
- Davis, R.L., H. Weintraub, and A.B. Lassar. 1987. Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* **51**: 987–1000.
- Emerson, B.M., C.D. Lewis, and G. Felsenfeld. 1985. Interaction of specific nuclear factors with the nuclease-hypersensitive region of the chicken adult β -globin gene: Nature of the binding domain. *Cell* **41**: 21–30.
- Fini, M.E., W.G. Bendena, and M.L. Pardue. 1989. Unusual behavior of the cytoplasmic transcript of *hsr omega*: An abundant, stress-inducible RNA that is translated but yields no detectable protein product. *J. Cell Biol.* **108**: 2045–2057.
- Garabedian, M.J., M. Hung, and P.C. Wensink. 1985. Independent control elements that determine yolk protein gene expression in alternative *Drosophila* tissues. *Proc. Natl. Acad. Sci.* **82**: 1396–1400.
- Grosveld, F., G. Blom van Assendelft, D.R. Greaves, and G. Kollias. 1987. Position-independent, high-level expression of the human β -globin gene in transgenic mice. *Cell* **51**: 975–985.
- Hafen, E., A. Kuroiwa, and W.J. Gehring. 1984. Spatial distribution of transcripts from the segmentation gene *fushi tarazu* during *Drosophila* embryonic development. *Cell* **37**: 833–841.
- Hammer, R.E., R. Krumlauf, S.A. Camper, R.L. Brinster, and S.M. Tilghman. 1987. Diversity of alpha-fetoprotein gene expression in mice is generated by a combination of separate enhancer elements. *Science* **235**: 53–58.
- Hiromi, U., A. Kuroiwa, and W.J. Gehring. 1985. Control elements of the *Drosophila* segmentation gene *fushi tarazu*. *Cell* **43**: 603–613.
- Hogan, B., F. Costantini, and E. Lacy. 1986. *Manipulating the mouse embryo: A laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- Kimelman, D. and M.W. Kirschner. 1989. An antisense mRNA directs the covalent modification of the transcript encoding fibroblast growth factor in *Xenopus* oocytes. *Cell* **59**: 697–696.
- Lui, J., T. Bergman, and K.S. Zaret. 1988. The mouse albumin promoter and a distal upstream site are simultaneously DNase I hypersensitive in liver chromatin and bind similar liver-abundant factors in vitro. *Genes & Dev.* **2**: 528–541.
- Maniatis, T. and R. Reed. 1987. The role of small nuclear ribonucleoprotein particles in pre-mRNA splicing. *Nature* **325**: 673–678.
- Maniatis, T., A. Jeffrey, and D.G. Kleid. 1975. Nucleotide sequence of the rightward operator of phage λ . *Proc. Natl. Acad. Sci.* **72**: 1184–1188.
- Meehan, R.R., D.P. Barlow, R.E. Hill, B.L. Hogan, and N.D. Hastie. 1984. Pattern of serum protein gene expression in mouse visceral yolk sac and foetal liver. *EMBO J.* **3**: 1881–1885.
- Melton, D.A., P.A. Krieg, M.R. Rebagliati, T. Maniatis, K. Zinn, and M.R. Green. 1984. Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. *Nucleic Acids Res.* **12**: 7035–7056.
- Olsson, M., G. Lindahl, and E. Ruoslahti. 1977. Genetic control of alpha-fetoprotein synthesis in the mouse. *J. Exp. Med.* **145**: 819–827.
- Pachnis, V., A. Belayew, and S.M. Tilghman. 1984. Locus unlinked to α -fetoprotein under the control of the murine *raf* and *Rif* genes. *Proc. Natl. Acad. Sci.* **81**: 5523–5527.
- Pachnis, V., C.I. Brannan, and S.M. Tilghman. 1988. The structure and expression of a novel gene activated in early mouse embryogenesis. *EMBO J.* **7**: 673–681.
- Plumb, M.A., V.V. Lobanenko, R.H. Nicolas, C.A. Wright, S. Zavou, and G.H. Goodwins. 1986. Characterisation of chicken erythroid nuclear proteins which bind to the nuclease hypersensitive regions upstream of the β^A - and β^H -globin genes. *Nucleic Acids Res.* **14**: 7675–7693.
- Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: A laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Solter, D. 1988. Differential imprinting and expression of maternal and paternal genomes. *Annu. Rev. Genet.* **22**: 127–146.
- Southern, E.M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**: 503–517.
- Walter, P. and G. Blobel. 1982. 7SL small cytoplasmic RNA is an integral component of the signal recognition particle. *Nature* **299**: 691–698.
- Wiles, M.V. 1988. Isolation of differentially expressed human cDNA clones: Similarities between mouse and human embryonal carcinoma cell differentiation. *Development* **104**: 403–413.
- Yoo-Warren, H., V. Pachnis, R.S. Ingram, and S.M. Tilghman. 1988. Two regulatory domains flank the mouse H19 gene. *Mol. Cell. Biol.* **8**: 4707–4715.

Ectopic expression of the H19 gene in mice causes prenatal lethality.

M E Brunkow and S M Tilghman

Genes Dev. 1991, **5**:

Access the most recent version at doi:[10.1101/gad.5.6.1092](https://doi.org/10.1101/gad.5.6.1092)

References

This article cites 27 articles, 11 of which can be accessed free at:
<http://genesdev.cshlp.org/content/5/6/1092.full.html#ref-list-1>

License

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

Focused on your science.

USC
SCIENTIFIC
SIDE BY SCIENCE