Transgenic Research (2005) 14:179–192
DOI 10.1007/s11248-004-8147-8

Hypomorphic mutation in hnRNP U results in post-implantation lethality

Michael J. Roshon** & H. Earl Ruley*
Department of Microbiology and Immunology, Room AA4210 MCN, Vanderbilt University School of Medicine, 1161 21st Ave South, Nashville, TN 37232-2363, USA

Received 6 October 2004; accepted 21 December 2004

Key words: embryonic stem cell, hnRNP U, insertional mutagenesis, retrovirus gene trap

Abstract

The present study characterized an embryonic lethal mutation induced by insertion of the U3Neo gene trap retrovirus into an intron of the gene encoding heterogeneous ribonuclear protein U (*Hnrnpu*), which maps to the distal arm of mouse chromosome 1. Murine hnRNP U was found to be identical to the human protein at all but one of 341 amino acid residues. Embryos homozygous for the provirus showed obvious abnormalities after 6.5 days of development (E6.5) and were resorbed by E10.5. Expression of the inserted neomycin-resistance gene involved alternative splicing to a cryptic 3′ splice site located in the neomycin resistance gene resulting in a hypomorphic mutation. Homozygous mutant cell lines isolated from preimplantation blastocysts expressed hnRNP U transcripts at levels 2 to 5 times lower than wild-type cells, suggesting that nearly wild-type levels of hnRNP U are required for embryonic development.

Abbreviations: dNTP – deoxyribonucleoside triphosphate; HnRNP – heterogeneous ribonuclear protein; nt – nucleotide; ICM – inner cell mass; LTR – long terminal repeat; MEF – mouse embryonic fibroblast; NPT – neomycin phosphotransferase; PBS – phosphate buffered saline; RT PCR – reverse transcriptase polymerase chain reaction

Introduction

Transcripts synthesized by RNA polymerase II associate with RNA binding proteins to form heterogeneous ribonucleoprotein (hnRNP) complexes (Dreyfuss et al., 1993, 2002; Weighardt et al., 1996; Krecic and Swanson, 1999). Monoclonal antibodies raised against protein-RNA conjugates precipitate over 20 hnRNP proteins, named according to their sizes, from A1 (32 kDa) to U (120 kDa) (Pinol-Roma et al., 1988). The largest component in the complex, the 120kDa hnRNP U protein, interacts directly with RNA through a carboxy-terminal RGG sequence (Kiledjian and Dreyfuss, 1992). Since hnRNP U is entirely nuclear and does not shuttle between the nucleus and cytoplasm (Pinol-Roma and Dreyfuss, 1992; Pinol-Roma and Dreyfuss, 1993), remodeling of hnRNP complexes and/or dissociation of hnRNP U presumably precedes RNA transport from the nucleus.

Approximately 50% of the cellular hnRNP U is associated with the nuclear matrix, the insoluble material remaining after nuclei are treated with with DNase I digestion and high-salt extraction of protein (Fackelmayer et al.,

*Author for correspondence
E-mail: ruleye@ctrvax.vanderbilt.edu
**Present address: Department of Emergency Medicine, Carolinas Medical Center, Charlotte, NC 28232-2861, USA

1994; Gohring & Fackelmayer, 1997; Gohring et al., 1997). Chromosomal DNA binds to the nuclear matrix via interactions between DNA sequences known as scaffold attachment regions (SARs) or matrix associated regions (MARs) and specific DNA binding proteins (scaffold attachment factors, or SAFs) (Cockerill & Garrard, 1986; Gasser & Laemmli, 1986). A 120 kDa SAF protein characterized in human (Romig et al., 1992), chicken (von Kries et al., 1991), and rat (Tsutsui et al., 1993) cells has been identified as hnRNP U (Fackelmayer et al., 1994; von Kries et al., 1994). Binding of hnRNP U to scaffold/matrix attachment region (S/MAR) elements involves sequences in the amino-terminus that are independent of the RNA binding domain (Gohring & Fackelmayer, 1997; Gohring et al., 1997). The protein is methylated by Prmt1, the major mammalian arginine methyltransferase (Herrmann et al., 2004). However, the modification appears to be constitutive and is thus unlikely to regulate hnRNP U functions under steady-state growth conditions (Pawlak et al., 2002).

Recent studies suggest that HnRNP U may also interact with both specific and general transcription factors. For example, hnRNP U interacts with the glucocorticoid receptor and inhibits transcriptional activation by the hormone (Eggert et al., 1997). In addition, a subfraction of hnRNP U co-purifies with RNA polymerase II and inhibits transcriptional elongation, as is characteristic of the enzyme in the preinitiation complex (Kim & Nikodem, 1999). hnRNP U dissociates from the Pol II complex at an early stage in transcription, consistent with the failure of hnRNP U to co-localize with sites of transcription in vivo (Mattern et al., 1999). Finally, hnRNP U/SAF-A reportedly interacts with the p300 histone acetylase enzyme, and the two proteins co-localize on S/MAR sequences within the topoisomerase I gene prior to transcriptional activation (Martens et al., 2002).

As with other abundant, ubiquitous hnRNPs, hnRNP U is thought to function at one or more steps important for pre-mRNA synthesis, processing and transport. The available data support models in which hnRNP U functions as an adapter to localize proteins, RNA and DNA to the nuclear matrix and/or to one another. Such interactions could contribute to molecularly coupled events involved in transcription and mRNA biogenesis .

In the present study, we characterized mice and cell lines with a hypomorphic mutation in the hnRNP U gene, *Hnrnpu* (Mouse Genome Database (Blake et al., 2002) accession number, 1858195). The mutation was induced in murine embryonic stem cells following infection with the U3Neo gene trap retrovirus (von Melchner et al., 1992) and transmitted through the germline. Homozygous mutant mice were severely retarded in both growth and development indicating that hnRNP U is essential for embryonic development.

Materials and methods

Genotyping by PCR

Genomic DNA was extracted (Laird et al., 1991) from embryos following overnight incubation at 55°C in 10 to 50 µl of lysis buffer (100 mM Tris–HCl, pH 8.5, 5 mM EDTA, 0.2% SDS, 200 mM NaCl, 100 µg/ml Proteinase K). Samples were heated to 100°C for 10 min and 1 µl was used for PCR. DNA was precipitated in an equal volume of isopropanol and resuspended in 20–50 µl TE (10 mM Tris–HCl pH 8.0, 0.1 mM EDTA). Embryos were genotyped using a three primer PCR assay. A 400 nt fragment from the wild type allele was amplified by using a primer complementary to an upstream *Hnrnpu* exon (1B3 up; TCCTCAGCCACCTGTTGAAG) with a primer specific for genomic sequences downstream of the integration site (1B3 down; GCTCACTTAACATAGT TCCAC). A 500 nt fragment corresponding to the disrupted allele was amplified by using the 1B3 up primer together with a primer specific for the provirus (neoB; CGAATAGCCTCTCCACCCAA). PCR was performed in 25 µl reactions (10 mM Tris–HCl pH 8.3, 5 mM KCl, 1.5 mM $MgCl_2$, 200 µM of each deoxyribonucleoside triphosphate, 2.5 units of Amplitaq (Perkin–Elmer/Cetus), and each primer at 2 µM). Reactions proceeded through 35 cycles of denaturation (95°C for 1.0 min), primer annealing (55°C for 1.0 min), and primer extension (72°C for 2 min.).

Sectioning embryos

The decidua were dissected away from the uterus, rinsed in PBS, fixed in 4% paraformaldehyde for 1 h and dehydrated in a series of 20 min washes (three times in PBS, and in 25, 50, 75, 75, 95 and 100% ethanol). The decidua were infused with paraffin by a series of 30 min to 1 h washes (Ethanol:xylene, 1:1, twice in xylene, xylene:paraffin 1:1 at 60°C, and three times in paraffin at 60°C). The embryos were imbedded for sectioning. Seven micron sections were placed on polylysine coated slides, incubated on a 42°C slide warmer overnight and stained with hematoxylin and eosin.

Isolation of cell lines homozygous for the 1B3 provirus

Cell lines were established from the inner cell mass of blastocysts obtained by crossing 1B3 heterozygous mice (Robertson, 1987; Abbondanzo et al., 1993). The blastocysts were isolated initially from C57BL/6 animals; however better results were acheived after back crossing for 3 generations into a 129/Sv background. The cells were maintained on feeder layers of primary mouse embryonic fibroblasts (MEF). MEFs were plated and allowed to condition the media for at least 12 h before plating blastocyst-derived cells. Cells were cultured in DME supplemented with 15% preselected and heat inactivated FBS; 100 mM nonessential amino acids, 0.1 mM β-mercaptoethanol and 1000 U/ml leukemia inhibitory factor (Esgro®, GIBCO).

Blastocysts were flushed from the uterus 3.5 days p.c., rinsed in media and plated on MEFs. The cultures were incubated undisturbed for 2 days, allowing the blastocysts to attach and hatch from the zona pellucida, and cultured for two additional days with daily media changes. During this time the inner cell mass (ICM) forms a ball of cells on top of the attached trophoblast cells. Usually on the morning of the 5th day the ICM was dissected away from the trophoblast cells by using a drawn glass capillary. The dissected ICM was washed twice in PBS then trypsinized for 3 min with 0.25% trypsin EDTA. After 3 min, the ICM was broken into clumps of cells (not single cells) and the clumps were transferred to a single well of a 24 well dish containing a feeder layer. After 7 days with media changes every other day, the cells could be serially passaged every second or third day as needed and then cyropreserved.

Reverse transcriptase PCR

RT-PCR was performed as described (Kawasaki, 1990) starting with cellular RNA prepared by LiCl extraction. About 20 μg RNA was treated with 1 unit of RNAse free DNase (Gibco BRL) (20 mM Tris–HCl pH 8.4, 50 mM KCl, 2.5 mM $MgCl_2$) for 15 min at room temperature. First strand cDNA synthesis was performed at 42°C for 30 min in a 20 μl reaction containing: 5 μg RNA, 500 nM NEO A primer (5′-ATTGTCTGTTGTGCCCAGTCATA), 20 mM Tris-HCl (pH 8.4), 50 mM KCl 2.5 mM $MgCl_2$, 10 mM DTT 400 μM each dNTP, 8 units Super Script II reverse transcriptase. Two units of RNAse H was added and incubated for 10 min at 55°C. About 2 μl of the single strand cDNA was amplified through 35 cycles (95°C for 1.0 min, 55°C for 1.0 min and 72°C for 2 min.) in a 50 μl reaction containing: Neo (NeoA or NeoB; 5′-CGAATAGCCTCTCCACCCAA) and hnRNP U (1B3 Up) specific primers, 10 mM Tris·HCl, pH (8.3), 5 mM KCl, 1.5 mM $MgCl_2$, 200 μM of each deoxyribonucleoside triphosphate, 2.5 units of Amplitaq (Perkin-Elmer/Cetus), and each primer at 2 μM.

Gene mapping

The murine *Hnrnpu* gene was mapped by using the Jackson Laboratory community resource (C57Bl/6J × *Mus spretus*)F1 × *Mus spretus* interspecific backcross DNA panel. cDNA sequences upstream of the provirus integration site detected a restriction fragment length polymorphism when the DNA was digested with *Msp*I. Southern blots of *Msp*I digested DNAs from 94 backcross animals and both parental strains were probed with the upstream probe. Each sample was scored for the presence of the C57Bl/6J allele and the results were analyzed against the Jackson Laboratory database to determine the position on the existing map.

Results

Recessive embryonic lethal mutation associated with the 1B3 provirus

The 1B3 cell line was generated by infecting D3 ES cells with the U3Neo gene trap retrovirus and selecting for G418 resistant clones (von Melchner et al., 1992). 1B3 cells contain a single, intact provirus as assessed by Southern blot hybridization (Figure 1(a) and data not shown). A 45 nucleotide sequence extending to an *Mse*I site immediately upstream of the provirus was isolated by inverse PCR (von Melchner et al., 1992). This sequence hybridized to single-copy cellular DNA and thus distinguished the wild-type genomic sequence from the locus occupied by the provirus (Figure 1(b)). Of nearly 250 offspring analyzed, including progeny obtained after nine backcrosses into a C57Bl/6 background, none was homozygous for the 1B3 provirus (Table 1). Transmission of the 1B3 provirus therefore appeared to be tightly linked to a recessive embryonic lethal mutation. Heterozygous mice were somewhat under-represented with 152 observed as compared to 163 expected among 245 progeny, but the difference was not statistically significant ($\chi^2 = 0.12$). Moreover, heterozygous mice were indistinguishable from their wild-type littermates.

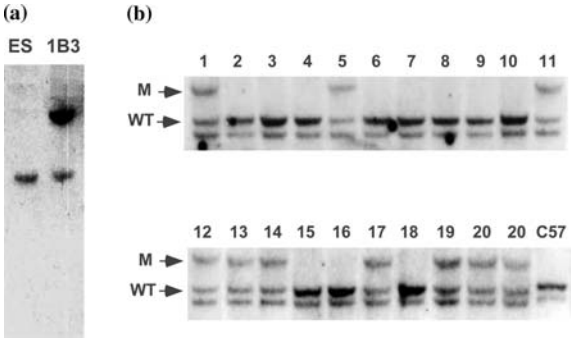

Figure 1. Failure to recover mice homozygous for the 1B3 provirus. (a) Southern blot analysis of the 1B3 provirus insert. DNA from the parental ES cell line and 1B3 cells was digested with *Hpa*I, and probed with a *Neo*-specific probe. (b) Southern blot analysis of progeny obtained by crossing 1B3 heterozygotes. About 10 μg of DNA was digested with *Eco*RI and probed with cDNA sequences upstream of the1B3 provirus. Twenty offspring were analyzed (1–20) and C57BL/6 DNA (C57) was included as a wild type control. Mutant (M) and wild-type (WT) alleles are indicated.

Table 1. Failure to recover mice homozygous for the 1B3 provirus

Backcross	Total	+/+	+/−	−/−
2–4	114	45	69	0
8–9	131	48	83	0
Total	245	93	152	0

Progeny of crosses between heterozygous mice after the indicated number of back crosses onto a C57B/6J background were genotyped as illustrated in Figure 1(b). The total number of progeny and numbers of wild type (+/+), heterozygous (+/−), and homozygous mutant (−/−) offspring are indicated.

The 1B3 provirus disrupts the hnRNP U gene

The 45-nt flanking sequence isolated by inverse PCR failed to detect transcripts by Northern blot analysis (not shown), but was used to isolate a 4 kb *Eco*RI fragment containing the viral integration site from a λ phage genomic DNA library. Further sequence analysis revealed a 74-nt region located 321 nt. upstream of the integration site that is identical to nucleotides 788 – 861 of a cDNA sequence reported for the human gene, *HNRNPU,* that encodes hnRNP U (Figure 2). The region of homology is flanked by consensus 3′ and 5′ splice sites, indicating that the 1B3 provirus integrated into an intron of the mouse *Hnrnpu* ortholog. The original 45-nt flanking sequence is also contained within the intron accounting for the failure of this probe to hybridize to cellular transcripts. Another exon, matching nucleotides 676–787 of the human hnRNP U cDNA, is located further upstream in the flanking DNA (Figure 2). Based on Southern blot analysis, using probes from either side of the provirus, insertion of the 1B3 provirus did not cause obvious deletions or rearrangements of adjacent cellular DNA (data not shown).

Isolation of the cDNAs encoding the mouse hnRNP U gene

The 4-kb genomic fragment, when used as a probe for Northern blot analysis, detected transcripts of approximately 2.9- and 3.7-kb in mouse testis RNA, and was therefore used to screen a mouse testis cDNA library (Stratagene). Four cDNAs, with inserts ranging in size from 1 to 2.6-kb, were isolated from among 5×10^5 clones screened. The complete sequence of the

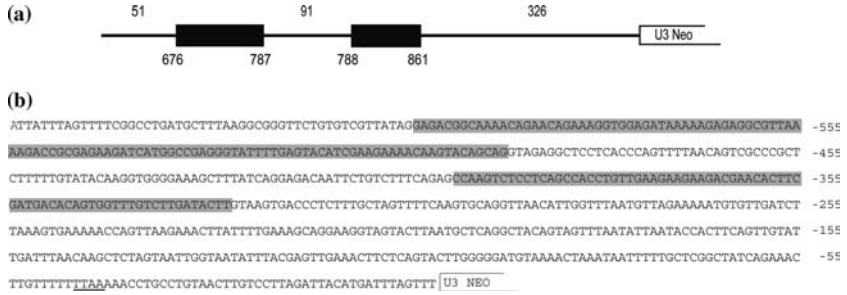

Figure 2. The 1B3 provirus disrupts the mouse *Hnrnpu* gene. (a) Schematic of the integration site. Exon and intron sequences are shown as solid boxes and lines, respectively. Distances beween exons and provirus are shown above and the corresponding nucleotides of the human cDNA are shown for the exon sequences (Kiledjian et al., 1991). (b) Genomic sequence upstream of the 1B3 provirus. Exon sequences are shaded and numbers to the right are relative to the integration site. The *Mse*I site involved in the amplification of 5′ flanking sequences by inverse PCR is underlined. This sequence was submitted to GenBank, accession number, AF073991.

2.6-kb cDNA (Figure 3), appears to contain the entire coding sequence for mouse hnRNP U. An open reading frame extending from nt 164 to 2560 begins with an ATG predicted to be in a favorable context for translational initiation (Kozak, 1991). The 3′ untranslated region contains a concensus polyadenylation signal (underlined) followed by a stretch of poly(A).

The encoded amino acid sequence of murine hnRNP U is nearly identical to both the rat sp120 protein and the human hnRNP U protein (Figure 3) and the canonical nucleic acid binding sites are identical in all three proteins. The protein also contains a potential nuclear localization signal (underlined) and an RGG sequence (shaded) (Kiledjian & Dreyfuss, 1992).

Phenotype of embryos homozygous for the 1B3 provirus

Failure to recover mice homozygous for the *Hnrnpu* mutation suggested that the gene is essential for embryonic development. To characterize embryonic defects associated with the mutation, 1B3 heterozygous mice were intercrossed and embryos were examined at different times of gestation. Blastocysts and embryos obtained after E6.5 were genotyped by PCR. Table 2 summarizes the data obtained from 214 embryos examined between E3.5 and E11.5. Resorbed embryos were first observed at E9.5. Assuming that the resorbed embryos were 1B3 homozygotes, two-thirds (12/18) of the homozygous mutant embryos were resorbed by E9.5. A similar ratio (8/11) was observed at E10.5.

At E3.5 (blastocyst stage) homozygous mutant embryos were indistinguishable from heterozygous or wild type littermates. Similarly at E6.5 it was not possible to distinguish mutants from normal littermates based on their gross morphological appearance. By E7.5, however, mutants could be easily distinguished from normal embryos (Figure 4(a)). Normal littermates (left) had undergone gastrulation and the headfold and neural groove were apparent. Mutant embryos lacked these features and exhibited a morphology similar to E6.5 embryos. At E8.5, normal embryos have completed gastrulation and begun organogenesis (Figure 4(b), left). Morphological landmarks such as the neural fold, primitive heart, and somites were present. In contrast, mutants showed little developmental progress (Figure 4(b), right). The embryonic portion of the embryo appeared essentially unchanged from the E7.5 mutant, while the extraembryonic portion appeared expanded. By E9.5, wild type embryos had progressed considerably (Figure 4(c), left). Mutant embryos on the other hand were severely delayed (Figure 4(c), right). At this stage 1B3 homozygotes displayed the size and development most closely resembling E8.5 embryos (Figure 4(e)). Mutants recovered at E10.5 showed very little developmental progress. The most advanced mutant observed is shown in Figure 4(d) (together with a wild type littermate) and enlarged in Figure 5(f). This embryo displayed neural folds, somites, and a beating heart. Besides the general developmental retardation, the most obvious defect is the large spherical allantois. This structure is seen on all mutants

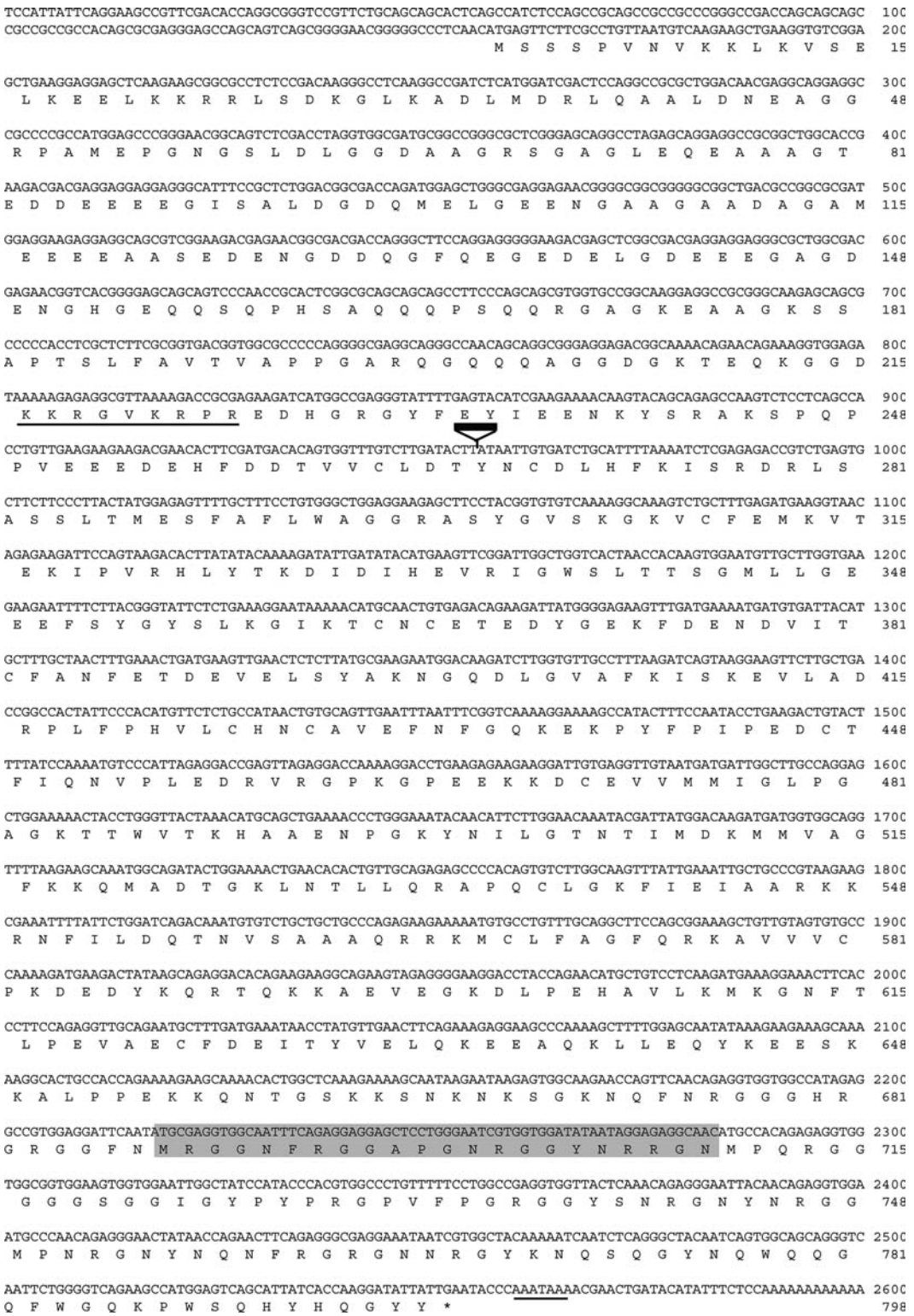

Figure 3. cDNA sequence of the mouse *Hnrnpu* gene. The cDNA and predicted amino acid sequences of the murine hnRNP U gene are shown as submitted to GenBank, accession number, AF073992. The putative nuclear localization and polyadenylation signals are underlined, and the RGG box is shaded. The position of the provirus integration site in relation to the coding sequence is shown.

isolated after E9.5 and is expected to arise due to continued growth of the allantois in the absence of fusion with the chorion. No homozygous mutant embryos were recovered after E11.5.

Phenotype of 1B3 homozygotes during early embryogenesis

To more closely characterize mutant phenotypes during early embryogenesis, histologic sections were prepared from decidua isolated from 1B3 heterozygote intercrosses from E7.5 to E8.5. At E7.5 (Figure 5(a)) the primitive streak and mesoderm layers are apparent in normal embryos. Presumptive mutants were readily distinguished from their normal littermates. The embryos appeared smaller, with markedly underdeveloped embryonic ectoderm. Extraembryonic regions, while less retarded in size, were disorganized and lacked identifiable extraembryonic cavities or membranes (Figure 5(b)).

By E8.5 normal embryos had begun organogenesis (Figure 5(c)) with neural tube, somites and heart being readily apparent. In contrast, development of mutant embryos was severely delayed (Figure 5(d)). The embryonic portion shows some disorganized progress since E7.5, with some mesoderm and primitive allantois apparent. The extraembryonic portion also shows some progress with some disorganized extraembryonic membranes and cavities. Approximately 1/3 of mutant embryos persisted until E10.5. These embryos were grossly abnormal but displayed structures characteristic of a neural axis, heart, amnion and yolk sac. All mutants observed at E10.5 had an abnormal allantois, presumably caused by failure to fuse with the chorion.

In summary, histological sections and whole mount morphology suggest that 1B3 heterozygotes develop normally until E6.5. Afterwards their development is greatly slowed such that by E8.5 the only sign of developmental progress in the embryonic portion is the presence of some embryonic mesoderm and the beginning of the allantois. In general, the defect seems to be less pronounced in the extraembryonic region so that by E8.5 there is formation of extraembryonic cavities and membranes, although they are considerably disorganized.

Effect of the 1B3 provirus on Hnrnpu gene expression

To determine if cells homozygous for the 1B3 provirus are viable, 1B3 heterozygotes were crossed, and the resulting blastocysts were cultured as described in Material and methods. Nineteen out of 33 blastocysts plated gave rise to stable cell lines; of which two were homozygous for the 1B3 provirus as assessed by Southern blot analysis (data not shown). Of the remaining cell lines 6 were wild-type and 11 were heterozygous. The viability of homozygous mutant cells could mean that the protein is not essential or that the 1B3 provirus did not completely disrupt *Hnrnpu* expression. To distinguish between these possibilities, levels of *Hnrnpu* transcripts were assessed by Northern blot analysis, probing with cDNA sequences downstream of the provirus integration site. The provirus inserts two poly(A) sites which, if utilized efficiently, will ablate transcription of

Table 2. Embryos homozygous for the 1B3 provirus die between E9.5 and E11.5

Embryonic age	Number	Wild type	Heterozygous	Homozygous mutant	Resorbed
3.5	8	3	3	2	0
6.5	15	4	7	4	0
7.5	15	3	6	6*	0
8.5	42	12	18	12*	0
9.5	67	15	3	6*	12
10.5	27	7	9	3*	8
11.5	30	7	12	0	11
Total	214	53	95	35	32

Embryos produced by crossing 1B3 heterozygotes were dissected from maternal tissue and genotyped by PCR using DNA extracted either from the entire embryo (E8.5), or from the yolk sac (E8.5). Embryos with obvious mutant phenotypes are marked (*).

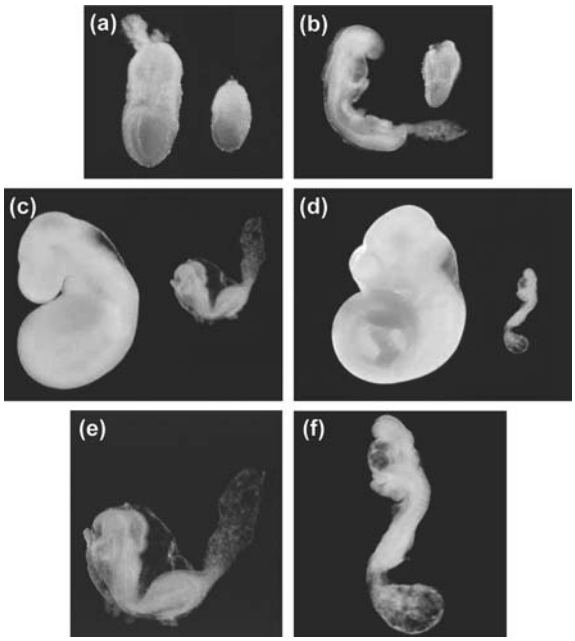

Figure 4. Wild-type and 1B3 homozygous mutant embryos. Wild type (left) and mutant (right) embryos at E7.5 (a), E8.5 (b), E9.5 (c), E10.5 (d) were photographed under dark field microscopy. Panels (e) and (f) are higher magnifications of mutants at E9.5 and E10.5, respectively. The genotypes of the embryos was confirmed by PCR.

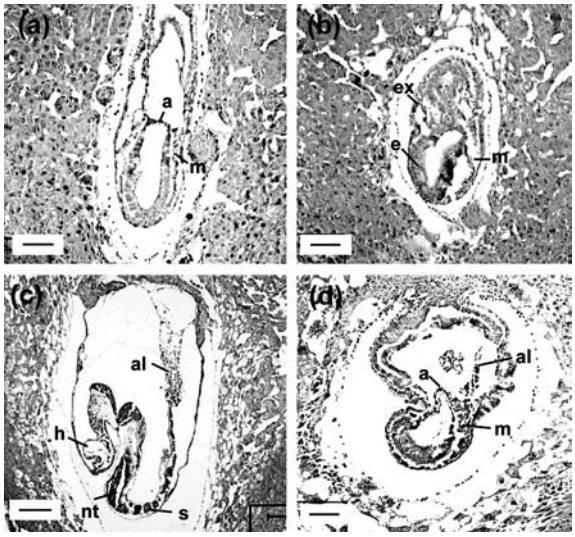

Figure 5. Histology of wild-type and mutant embryos. Embryos were fixed and sectioned within the decidua. (a) Normal and (b) mutant embryos at E7.5. (c) Normal and (d) mutant embryos at E8.5. Abbreviations: a, amnion; al, allantois; e, embryonic portion; ex, extraembryonic portion; h, heart; me, mesoderm; nt. neural tube; s, somite.

downstream hnRNP U sequences. Relative levels of hnRNP transcripts were quantified with a Fuji phosphorimager, with glyceraldehyde 3-phosphate dehydrogenase transcripts (GAPDH) serving as an internal standard. As shown in Figure 6, the *Hnrnpu* probe hybridized to two transcripts in both wild type and heteroyzygous cell lines. This pattern is similar to that observed in Hela cells, in which two transcripts are generated by alternative polyadenylation. Levels of the smaller and larger transcripts expressed in homozygous mutant cells were approximately two-fold (58% of wild-type) and 5-fold (18% of wild-type) lower than wild type cells, respectively. Similar results were obtained with the second homozygous mutant line (data not shown). Thus, insertion of 1B3 provirus resulted in a hypomorphic mutation.

Isolation of homozygous mutant cell lines, might select for cell variants that express higher levels of hnRNP U. Therefore, mutant embryos were also analyzed for *Hnrnpu* expression by in situ hybridization. Sections of wild type and morphologically mutant embryos were hybridized to ^{35}S-labeled probes prepared from sense and anti-sense cDNA sequences downstream of the provirus integration site. Both normal and mutant embryos showed expression throughout the embryo (data not shown), confirming that the phenotype associated with the 1B3 provirus did not result from a null mutation.

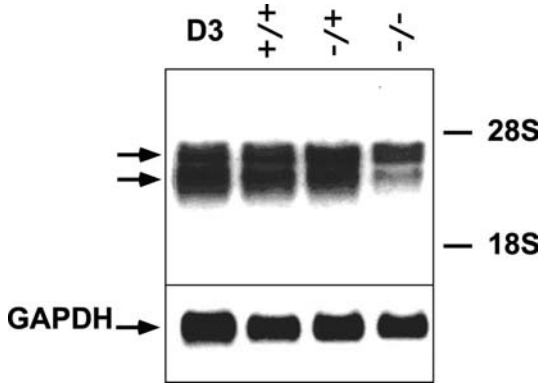

Figure 6. Disruption of *Hnrnpu* gene expression by the 1B3 provirus. Northern blot analysis of total RNA prepared from blastocyst derived cell lines. About 10 μg of total RNA for each cell line was hybridized to ^{32}P labeled probes derived from sequences downstream of the provirus integration site. Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) probes were used to normalize RNA loading. The migration of the ribosomal RNAs (18S and 28S) is indicated.

Splicing and polyadenylation of hnRNP U-Neo fusion transcripts

Each LTR of the U3Neo provirus contains sequences for 3′ processing and polyadenylation. Continued expression of *Hnrnpu* transcripts suggests that the viral poly(A) sites are not always utilized; instead, splicing removes provirus sequences inserted in the intron. The question then arises as to how hnRNP U-neo fusion transcripts are expressed. Previous Northern blot studies found high levels of a 1.5-kb *Neo* transcript in 1B3 cells (von Melchner et al., 1992), the approximate size expected for *Hnrnpu-Neo* fusion transcripts terminating in the 5′ LTR. One possibility is that fusion transcripts may combine the proximal upstream *Hnrnpu* exon, 5′ splice site, flanking intron and 5′ LTR into a single, terminal exon. However, this option contradicts current exon definition models in which exons in pre-mRNA are processed as autonomous units. Alternatively, the proximal *Hnrnpu* exon may maintain its autonomy and splice to a cryptic 3′ splice site, located either in *Neo* or in the adjacent intron as observed in most entrapment clones genenerated with U3Neo vectors (Osipovich et al., 2004).

To distinguish between these alternatives, *Hnrnpu-Neo* fusion transcripts expressed in ES cells were analyzed by reverse transcriptase PCR (RT-PCR). A primer specific for the hnRNP exon immediately upstream of the provirus and two different primers complementary the neomycin resistance gene were used to amplify transcripts extending from the *Hnrnpu* into the provirus (Figure 7(a)). Transcripts extending through the 5′ splice site, proximal intron and into the provirus would produce an RT-PCR product 503 nucleotides using the 1B3 up and neoB primers. As shown in Figure 7(b), the size of the major PCR product was significantly smaller (98 nt) than expected for transcripts colinear with the flanking DNA. Moreover, the major PCR product did not hybridize to a U3 probe (Figure 7(c)). Three independent RT-PCR products were cloned from separate amplification reactions and sequenced. As shown in Figure 7(d), all of these transcripts spliced from the proximal 5′ splice site in *Hnrnpu* to a cryptic 3′ splice site located in the *Neo* gene (Figure 7(d)). The same cryptic splice site, which is commonly used by U3 entrapment vectors containing wild-type *Neo* sequence (Roshon et al., 2003; Osipovich et al., 2004), contains a PyAG and a potential branch point sequence characteristic of 3′ splice sites, but lacks the upstream polypyrimidine stretch. Since the 3′ splice site is downstream of the neomycin phosphotransferase initiation codon, *Hnrnpu*-U3Neo transcripts are expected to encode a fusion protein consisting of the amino terminal 219 amino acids of hnRNP U fused to amino acid 10 of neomycin phosphotransferase (NPT).

The mouse hnRNP U gene maps to the distal end of chromosome 1

Early in the course of these studies, the mouse *Hnrnpu* gene was mapped by using the Jackson Laboratories (C57Bl/6J × *M. spretus*) F1 × *M. spretus* interspecific backcross DNA panel (see Materials and methods) (Rowe et al., 1994). This panel of DNAs was obtained from 94 N2 animals from the F1 × *M. spretus* backcross. The genetic maps of the DNA panel were anchored to published maps by mapping known loci, simple sequence length polymorphisms, and loci defined by endogenous retroviruses. This backcross panel allows accurate and efficient mapping with 1–5 cM resolution. The mouse hnRNP U gene mapped to the distal end of chromosome 1 (Figure 8). There are no known mutations mapping nearby for which hnRNP U would be a candidate.

Discussion

The present study characterized a hypomorphic mutation in the mouse *Hnrnpu* gene induced by insertion of the U3Neo gene trap retrovirus. Embryos homozygous for the 1B3 provirus experience retarded development of embryonic ectoderm at E6.5, severe growth retardation beginning at E7.5 and death between E9.5 and E11.5. None of the mutants advanced beyond the size or developmental stage of E8.5 embryos, but some formed neural plate, primitive heart, somites and extraembryonic tissues. This is the second report that an hnRNP gene is essential for mammalian development.

While the first observable defects coincided with mesoderm production, it is difficult to

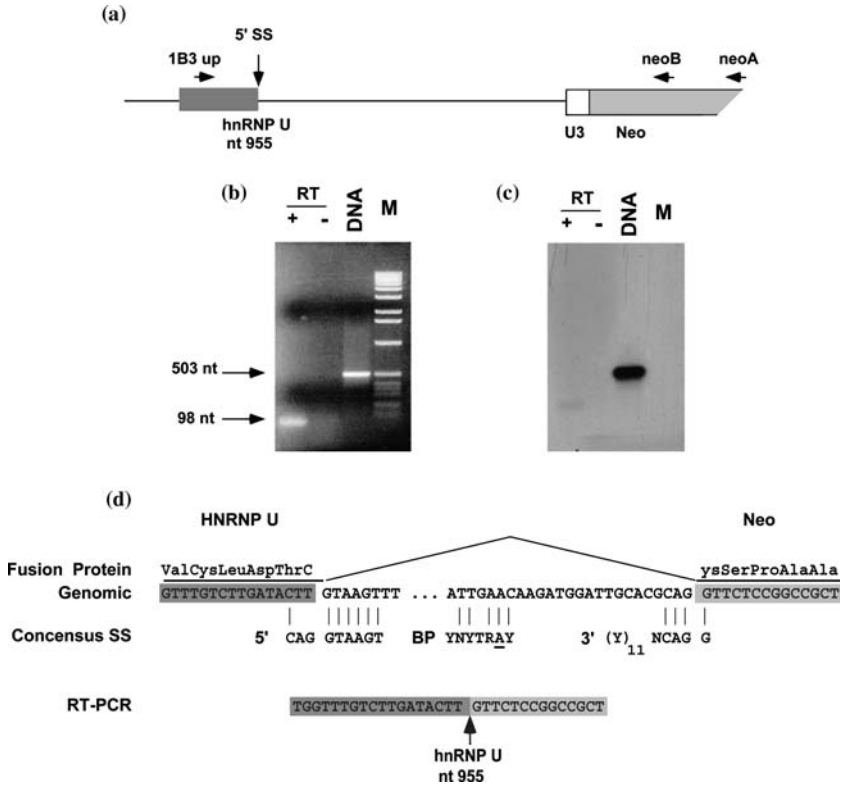

Figure 7. Analysis of *Hnrnpu*-U3Neo fusion transcripts. (a) Strategy used to amplify fusion transcripts. The provirus is shown downstream of the 5′ splice site (5′SS) of the *Hnrnpu* exon upstream and proximal to the 1B3 provirus (Figure 2). The positions of the primers used for reverse transcriptase PCR (RT-PCR) are indicated. (b) Ethidium bromide stained products of RT-PCR. PCRs were performed before (−RT) or after (+RT) first strand synthesis with the neoA primer. PCR reactions using DNA templates were performed to provide controls for products amplified from unspliced transcripts and products amplified from genomic DNA sequences that might contaminate RNA preparations. Sizes of the RT-PCR products are indicated. (c) PCR products from B were analyzed by Southern blot hybridization using a U3 oligo as a probe. (d) Sequence alignment between RT-PCR products and genomic sequences. Consensus sequences for 5′ splice sites (5′), 3′ splice sites (3′), and the branchpoint (BP) are shown below the genomic sequences. All RT PCR products were generated from transcripts that spliced from nt 955 of *Hnrnpu* to *Neo* as shown.

distinguish whether the phenotype reflects a specific function of the gene or is a secondary consequence of growth retardation. The phenotype is consistent with the partial loss of a widely expressed gene that is required for normal cell metabolism. Such mutations (assuming they do not affect cell viability) are expected to suppress embryonic growth beginning with gastrulation when the embryo experiences a dramatic increase in size and growth rate. For example, embryos with a hypomorphic mutation in DNA methyltransferase (Li et al., 1992) survive until E10.5, are stunted in their development and have similar defects in the allantois.

There is considerable variability in the severity of the mutant phenotype even within the same litter, which may reflect the number and proportion of mutant embryos and competition with nearby embryos. The reduction in hnRNP U could also contribute to the variation, assuming that the levels of gene expression are near the boundary of that required for function. The variation was also observed in mice back crossed for up to 9 generations and thus, does not appear to result from variation in genetic background. Mutants persisting beyond E9.5 consistently show defects in the allantois, as have been described for mutations affecting α5 integrin (Yang et al., 1993), DNA methyltransferase (Li et al., 1992), *Csk* (Imamoto & Soriano, 1993), *Brachyury* (Willison, 1990; Herrmann, 1992) and Hβ58 (Radice et al., 1991). Thus, defects in the chorioallantoic placenta, including failure of the allantois to fuse with the chorion, are frequently

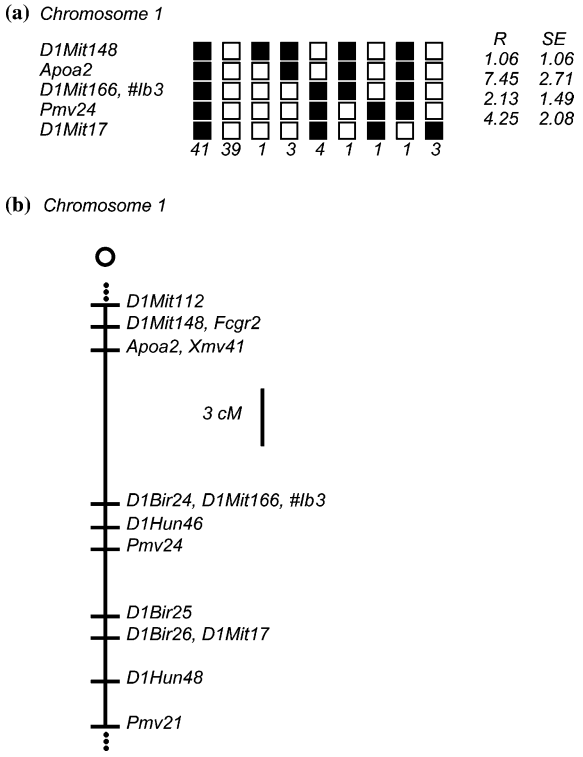

Figure 8. Mapping of the mouse *Hnrnpu* gene. (a) Haplotype from The Jackson BSS backcross showing part of Chromosome 1 with loci linked to #Ib3. Loci are listed in order with the most proximal at the top. The black boxes represent the C57BL6/Jei allele and the white boxes the SPRET/Ei allele. The number of animals with each haplotype is given at the bottom of each column of boxes. The percent recombination (R) between adjacent loci is given to the right of the figure, with the standard error (SE) for each R. (b) Map from The Jackson BSS backcross showing the distal end of Chromosome 1. The map is depicted with the centromere toward the top. A 3 cM scale bar is shown to the right of the figure. Loci mapping to the same position are listed in alphabetical order. Missing typings were inferred form surrounding data where assignment was unambiguous. Raw data from The Jackson Laboratory were obtained from the World Wide Web address, http://www.jax.org/resources/documents/cmdata.

associated with embryonic death during the E10 to E11 interval. Defects in the allantois are associated with insufficient mesoderm formation caused either by general growth retardation or by mesoderm-specific gene defects. For example, most BMP 4 null embryos arrest at gastrulation due to failure to produce mesoderm, whereas, death at later stages is associated with a failure to produce sufficient extraembryonic mesoderm including the yolk sac and allantois (Winnier et al., 1995).

U3 gene trap vectors were designed to disrupt cellular gene function by usurping the promoter of the occupied genes and by ablating transcription downstream of the two poly(A) sites (one in each LTR) carried by the provirus. Since poly(A) sites are not usually recognized when located in introns (Adami & Nevins, 1988; Levitt et al., 1989; Niwa et al., 1992), U3 gene traps were expected to select for clones in which the provirus had inserted into exons of transcriptionally active genes. However, our previous studies have demonstrated that U3Neo entrapment vectors are typically expressed following insertion into introns, from cellular transcripts that splice to cryptic 3′ splice sites in the neomycin phosphotransferase (NPT) coding sequence or in the adjacent intron (Roshon et al., 2003; Osipovich et al., 2004). The majority of *Hnrnpu-Neo* fusion transcripts utilized the same cryptic *Neo* splice site.

Insertion of U3 gene trap vectors into introns have had variable effects on the expression of the occupied genes, as illustrated by essentially null mutations in *Rangap1* (Ran GTPase activating protein) (DeGregori et al., 1994), *EphA2* (Eph receptor A2) (Chen et al., 1996), *Hnrpc* (Williamson et al., 2000), *Fus* (Hicks et al., 2000) and *Prmt1* (arginine methyltransferase 1) (Pawlak et al., 2000), as well as a hypomorphic allele involving *Hnrpa2b1* (hnRNP A2/B1) (Roshon et al., 2003). The severity of each mutation appears to depend on the efficiency of splicing between 5′ exons of the occupied gene and cryptic 3′ splice sites in or upstream of the entrapment vector. Thus, the hypomorphic mutation in *Hnrnpu*, like the previously described *Hnrpa2b1* mutation, results from transcripts that skip the relatively weak cryptic 3′ splice site in the U3Neo gene. By contrast, the null mutation in *Hnrpc* results from the efficient utilization of a cryptic 3′ splice site in the adjacent intron sequence.

These the expression of *Hnrnpu*-U3Neo fusion transcripts is consistent with an exon definition model in which splicing and polyadenylation require interactions between factors acting across exons (Robberson et al., 1990; Niwa et al., 1992). This model predicts that polyadenylation signals are not utilized unless they can be defined as part of a 3′ terminal exon (Adami & Nevins, 1988; Levitt et al., 1989; Niwa et al., 1990, 1992; Miller & Stoltzfus, 1992; Wasserman

& Steitz, 1993; Furth et al., 1994). We find that the proximal hnRNP U exon upstream of the provirus does not lose its identity; rather, the exon splices either to the *Neo* splice site or to the next hnRNP U exon. Moreover, the poly(A) site in the 5′ LTR appears to be used exclusively in conjunction with the *Neo* cyptic 3′ splice site.

cDNAs encoding murine hnRNP U were characterized as part of our analysis of the 1B3 mutation, providing an additional vertebrate hnRNP U sequence for interspecies comparisons (the chicken, rat and human sequences had been reported previously). Overall, murine and human hnRNP U differ at only a single residue, while functionally significant regions such as the RNA recognition motif and nuclear localization signal are completely conserved among all vertebrate species. Phylogenetic conservation of hnRNP U despite the presence of related genes, suggests that the protein has distinct functions in mRNA biogenesis or gene regulation.

Cell lines derived from homozygous mutant blastocysts expressed 2–5 times less hnRNP U transcripts and were more difficult to establish than wild type or heterozygous cell lines. We considered the possibility that in vitro propagation selected for cells that expressed higher levels of hnRNP U than are expressed in the corresponding mutant embryos. However, hnRNP U transcripts were easily detected in homozygous mutant embryos as assessed by in situ hybridization. While this analysis was not quantitative and some cell types may be more affected than others, the results suggest that the reduced levels of hnRNP U are not sufficient for embryonic development.

Without null cell lines, we were unable to determine if hnRNP U is required for cell viability. For example, cells deficient in hnRNPA1 (Ben-David et al., 1992) and lacking hnRNP C (Williamson et al., 2000) have been described. Even so, the availability of cell lines expressing reduced levels of hnRNP U transcripts, could assist cell-based studies of hnRNP U function.

Acknowledgements

We thank Drs Brigid Hogan and Lucy Liaw for help with in situ hybridization. This work was supported by Public Health Service Grants (R01NS043952, R01RR13166 and P01HL68744) and by a grant from the Kleberg Foundation. Additional support was provided by a Cancer Center Core grant (P30CA68485) to the Vanderbilt-Ingram Cancer Center. M.R. was supported by a Medical Scientist Training Program Grant (5T32-GM07347).

References

Abbondanzo SJ, Gadi I and Stewart CL (1993) Derivation of embryonic stem cell lines. *Methods Enzymol* **225**: 803–823.

Adami G and Nevins JR (1988) Splice site selection dominates over poly(A) choice in RNA production from complex adenovirus transcription units. *EMBO* J **7**: 2107–2116.

Ben-David Y, Bani MR, Chabot B, De Koven A and Bernstein A (1992) Retroviral insertions downstream of the heterogeneous nuclear ribonucleoprotein A1 gene in erythroleukemia cells: evidence that A1 is not essential for cell growth. *Mol Cell Biol* **12**: 4449–4455.

Blake JA, Richardson JE, Bult CJ, Kadin JA and Eppig JT (2002) The Mouse Genome Database (MGD): the model organism database for the laboratory mouse. *Nucleic Acids Res* **30**: 113–115.

Chen J, Nachabeh A, Scherer C, Ganju P, Reith A, Bronson R and Ruley HE (1996) Germline inactivation of the murine eck receptor tyrosine kinase by gene trap retroviral insertion. *Oncogene* **12**: 979–988.

Cockerill PN and Garrard WT (1986) Chromosomal loop anchorage of the Kappa immunoglobulin gene occurs next to the enhancer in a region containing Topoisomerase II sites. *Cell* **44**: 273–283.

DeGregori JV, Russ A, von Melchner H, Rayburn H, Priyaranjan P, Jenkins N, Copeland N and Ruley HE (1994) A murine homolog of the yeast RNA1 gene is required for post-implantation development. *Genes Dev* **8**: 265–276.

Dreyfuss G, Kim VN and Kataoka N (2002) Messenger-RNA-binding proteins and the messages they carry. *Nat Rev Mol Cell Biol* **3**: 195–205.

Dreyfuss G, Matunis MJ, Pinol-Roma S and Burd CG (1993) hnRNP proteins and the biogenesis of mRNA. *Annu Rev Biochem* **62**: 289–321.

Eggert M, Michel J, Schneider S, Bornfleth H, Baniahmad A, Fackelmayer FO, Schmidt S and Renkawitz R (1997) The glucocorticoid receptor is associated with the RNA-binding nuclear matrix protein hnRNP U. *J Biol Chem* **272**: 28471–28478.

Fackelmayer FO, Dahm K, Renz A, Ramsperger U and Richter A (1994) Nucleic-acid-binding properties of hnRNP-U/SAF-A, a nuclear-matrix protein which binds DNA and RNA in vivo and in vitro. *Eur J Biochem* **221**: 749–757.

Furth PA, Choe W-T, Rex JH, Byrne JC and Baker CC (1994) Sequences homologous to 5′ splice sites are required for the inhibitory activity of papillomavirus late 3′ untranslated regions. *Mol Cell Biol* **14**: 5278–5289.

Gasser SM and Laemmli UK (1986) Cohabitation of scaffold binding regions with upstream/enhancer elements of three developmentally regulated genes of D. melanogaster. *Cell* **46**: 521–530.

Gohring F and Fackelmayer FO (1997) The scaffold/matrix attachment region binding protein hnRNP-U (SAF-A) is directly bound to chromosomal DNA in vivo: a chemical cross-linking study. *Biochemistry* **36**: 8276–8283.

Gohring F, Schwab BL, Nicotera P, Leist M and Fackelmayer FO (1997) The novel SAR-binding domain of scaffold attachment factor A (SAF-A) is a target in apoptotic nuclear breakdown. *EMBO J* **16**: 7361–7371.

Herrmann BG (1992) Action of the Brachyury gene in mouse embryogenesis. *Ciba Foundation Symp* **165**: 78–86; discussion 86–91.

Herrmann F, Bossert M, Schwander A, Akgun E and Fackelmayer FO (2004) Arginine methylation of scaffold attachment factor A by heterogeneous nuclear ribonucleoprotein particle-associated PRMT1. *J Biol Chem* **279**: 48774–48779.

Hicks GG, Singh N, Nashabi A, Mai S, Bozek G, Klewes L, Arapovic D, White EK, Koury MJ, Oltz EM, et al. (2000) Fus deficiency in mice results in defective B-lymphocyte development and activation, high levels of chromosomal instability and perinatal death. *Nat Genet* **24**: 175–179.

Imamoto A and Soriano P (1993) Disruption of the csk gene, encoding a negative regulator of Src family tyrosine kinases, leads to neural tube defects and embryonic lethality in mice. *Cell* **73**: 1117–1124.

Kiledjian M and Dreyfuss G (1992) Primary structure and binding activity of the hnRNP U protein: binding RNA through RGG box. *EMBO J* **11**: 2655–2664.

Kim MK and Nikodem VM (1999) hnRNP U inhibits carboxy-terminal domain phosphorylation by TFIIH and represses RNA polymerase ii elongation. *Mol Cell Biol* **19**: 6833–6844.

Krecic AM and Swanson MS (1999) hnRNP complexes: composition, structure, and function. *Curr Opin Cell Biol* **11**: 363–371.

Laird PW, Zijderveld A, Linders K, Rudnicki MA, Jaenisch R and Berns A (1991) Simplified mammalian DNA isolation procedure. *Nucleic Acids Res* **19**: 4293.

Levitt N, Briggs D, Gil A and Proudfoot NJ (1989) Definition of an efficient synthetic poly (A) site. *Genes Dev* **3**: 1019–1025.

Li E, Bestor TH and Jaenisch R (1992) Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**: 915–926.

Martens JH, Verlaan M, Kalkhoven E, Dorsman JC and Zantema A (2002) Scaffold/matrix attachment region elements interact with a p300-scaffold attachment factor A complex and are bound by acetylated nucleosomes. *Mol Cell Biol* **22**: 2598–2606.

Mattern KA, van der Kraan I, Schul W, de Jong L and van Driel R (1999) Spatial organization of four hnRNP proteins in relation to sites of transcription, to nuclear speckles, and to each other in interphase nuclei and nuclear matrices of HeLa cells. *Exp Cell Res* **246**: 461–470.

Miller JT and Stoltzfus CM (1992) Two distant upstream regions containing cis-acting signals regulating splicing facilitate 3′ end processing of avian sarcoma virus RNA. *J Virol* **66**: 4242–4251.

Niwa M, MacDonald C and Berget S (1992) Are vertebrate exons scanned during splice-site selection? *Nature* **360**: 277–280.

Niwa M, Rose SD and Berget SM (1990) In vitro polyadenylation is stimulated by the presence of an upstream intron. *Genes Dev* **4**: 1552–1559.

Osipovich AB, White-Grindley EK, Hicks GG, Roshon MJ, Shaffer C, Moore JH and Ruley HE (2004) Activation of cryptic 3′ splice sites within introns of cellular genes following gene entrapment. *Nucleic Acids Res* **32**: 2912–2924.

Pawlak MR, Banik-Maiti S, Pietenpol JA and Ruley HE (2002) Protein arginine methyltransferase I: substrate specificity and role in hnRNP assembly. *J Cell Biochem* **87**: 394–407.

Pawlak MR, Scherer CA, Chen J, Roshon MJ and Ruley HE (2000) Arginine N-methyltransferase 1 is required for early postimplantation mouse development, but cells deficient in the enzyme are viable. *Mol Cell Biol* **20**: 4859–4869.

Pinol-Roma S, Choi YD, Matunis MJ and Dreyfuss G (1988) Immunopurification of heterogeneous nuclear ribonucleoprotein particles reveals an assortment of RNA-binding proteins. *Genes Dev* **2**: 215–217.

Pinol-Roma S and Dreyfuss G (1992) Shuttling of pre-mRNA binding proteins between nucleus and cytoplasm. *Nature* **355**: 730–732.

Pinol-Roma S and Dreyfuss G (1993) hnRNP proteins: localization and transport between the nucleus and cytoplasm. *Trends Cell Biol* **3**: 151–155.

Radice G, Lee JJ and Costantini F (1991) Hb58, an insertional mutation affecting early postimplantation development of the mouse embryo. *Development* **111**: 801–811.

Robberson BL, Cote GJ and Berget SM (1990) Exon definition may facilitate splice site selection in RNAs with multiple exons. *Mol Cell Biol* **10**: 84–94.

Robertson EJ (ed.) (1987) Teratocarcinomas and Embryonic Stem Cells, 1st Edn, Oxford, Washington DC, IRL Press.

Romig H, Fackelmayer FO, Renz A, Ramsperger U and Richter A (1992) Characterization of SAF-A, a novel nuclear DNA binding protein from HeLa cells with high affinity for nuclear matrix/scaffold attachment DNA elements. *EMBO J* **11**: 3431–3440.

Roshon M, DeGregori JV and Ruley HE (2003) Gene trap mutagenesis of hnRNP A2/B1: a cryptic 3′ splice site in the neomycin resistance gene allows continued expression of the disrupted cellular gene. *BMC Genomics* **4**: 2.

Tsutsui K, Okada S, Watarai S, Seki S, Yasuda T and Shohmori T (1993) Identification and characterization of a nuclear scaffold protein that binds the matrix attachment region DNA. *J Biol Chem* **268**: 12886–12894.

von Kries JP, Buck F and Stratling WH (1994) Chicken MAR binding proten p120 is identical to human heterogeneous nuclear ribonucleoprotein (hnRNP) U. *Nucleic Acids Res* **22**: 1215–1220.

von Kries JP, Buhrmester H and Stratling WH (1991) A matrix/scaffold attachment region binding protein: identification, purification, and mode of binding. *Cell* **64**: 123–135.

von Melchner H, DeGregori JV, Rayburn H, Reddy S, Friedel C and Ruley HE (1992) Selective disruption of genes expressed in totipotent embryonal stem cells. *Genes Dev* **6**: 919–927.

Wasserman KM and Steitz JA (1993) Association with terminal exons in pre-mRNAs: a new role for the U1 snRNP. *Genes Dev* **7**: 647–659.

Weighardt F, Biamonti G and Riva S (1996) The roles of heterogeneous nuclear ribonucleoproteins (hnRNP) in RNA metabolism. *Bioessays* **18**: 747–756.

Williamson DJ, Banik-Maiti S, DeGregori J and Ruley HE (2000) hnRNP C is required for postimplantation mouse development but is dispensable for cell viability. *Mol Cell Biol* **20**: 4094–4105.

Willison K (1990) The mouse Brachyury gene and mesoderm formation. *Trends Genet* **6**: 104–105.

Winnier G, Blessing M, Labosky PA and Hogan BL (1995) Bone morphogenetic protein-4 is required for mesoderm formation and patterning in the mouse. *Genes Dev* **9**: 2105–2116.

Yang JT, Rayburn H and Hynes RO (1993) Embryonic mesodermal defects in alpha 5 integrin-deficient mice. *Development* **119**: 1093–1105.