MOLECULAR AND CELLULAR BIOLOGY, July 1997, p. 4080–4086 Vol. 17, No. 7
0270-7306/97/$04.00+0

Proviral Inactivation of the *Npat* Gene of Mpv 20 Mice Results in Early Embryonic Arrest

MARCO DI FRUSCIO,[1,2] HANS WEIHER,[3] BARBARA C. VANDERHYDEN,[1,4] TAKASHI IMAI,[5] TADAHIRO SHIOMI,[5] TADA-AKI HORI,[5] RUDOLF JAENISCH,[6] AND DOUGLAS A. GRAY[1,4]*

Ottawa Regional Cancer Centre, Ottawa, Ontario, Canada K1H 8L6[1]; Department of Biochemistry[2], and Department of Medicine,[4] University of Ottawa, Ottawa, Ontario, Canada K1H 8M5; Forschungszentrum Karlsruhe, Institute of Genetics, 76021 Karlsruhe, Germany[3]; Genome Research Group, National Institute of Radiological Sciences, Inage-ku, Chiba-shi 263, Japan[5]; and Whitehead Institute for Biomedical Research, Cambridge, Massachusetts 02142[6]

Received 13 January 1997/Returned for modification 18 February 1997/Accepted 8 April 1997

The Mpv 20 transgenic mouse strain was created by infection of embryos with a defective retrovirus. When Mpv 20 heterozygous animals were crossed, no homozygous neonatal mice or midgestation embryos were identified. When embryos from heterozygous crosses were cultured in vitro, approximately one quarter arrested as uncompacted eight-cell embryos, indicating that proviral insertion resulted in a recessive lethal defect whose phenotype was manifest very early in development. Molecular cloning of the Mpv 20 insertion site revealed that the provirus had disrupted the *Npat* gene, a gene of unknown function, resulting in the production of a truncated *Npat* mRNA. Expression of the closely linked *Atm* gene was found to be unaffected by the provirus.

The insertion of retroviral proviruses into the host genome can have mutagenic consequences, either by the activation of host genes through juxtaposition of positive viral regulatory elements (7) or by the transcriptional silencing of adjacent host sequences that occurs as murine type C viruses infect germ line cells (15, 16). The former phenomenon has been exploited with great success for the discovery of oncogenes; the latter has been useful for identifying mammalian developmental mutations (8, 21). Through retroviral insertional mutagenesis screens, genes whose inactivation caused recessive lethal defects manifest at the midgestation (25), neonatal (22), and adult stages (29) have been identified. The use of defective viruses incorporating selectable markers has further enhanced the utility of the strategy by allowing the infection and selection of ES cell clones in which genes have been trapped (3, 5). There are two great advantages in using retroviruses as mutagenic agents: they generate discrete, localized insertion mutations (as opposed to the large deletions that frequently accompany the insertion of DNA introduced by pronuclear microinjection), and the provirus provides the probe with which the site of insertion can be cloned.

We have used the defective murine type C virus myeloproliferative sarcoma virus (MPSV)-neo (23) to infect early mouse embryos in a screen which previously gave rise to the adult kidney-specific mutant strain Mpv 17 (29). In this paper, we describe Mpv 20, a mouse strain in which insertional mutagenesis has resulted in a recessive lethal phenotype manifest at a very early stage of embryonic development. We also report the identification of the mutated gene in Mpv 20 as *Npat*, a recently described gene closely linked to the *Atm* locus (13). No function has been reported for *Npat*, and its predicted amino acid sequence is not sufficiently similar to known sequences to allow predictions about the function of the Npat protein. The data presented herein suggest that whatever its function, the *Npat* gene protein is absolutely required for development beyond the eight-cell stage.

* Corresponding author. Mailing address: Ottawa Regional Cancer Centre, 501 Smyth Rd., Ottawa, Ontario, Canada K1H 8L6.

MATERIALS AND METHODS

Generation of transgenic mouse lines. Retroviral infection of 4- to 16-cell CFW mouse embryos was accomplished as described previously (15, 29). Briefly, Rat 1 cells infected with MPSV-neo and Friend murine leukemia virus (kindly provided by W. Ostertag) as a helper virus were used as producer cells. Embryos were cocultured with producer cells and then transferred to pseudopregnant C57BL/6J × CBA females. Progeny mice were genotyped by Southern analysis of tail biopsy DNA with a *neo*-specific probe. Of the 42 mice examined, 11 contained one or more copies of the *neo* gene (a total of 18 proviruses). Each provirus was segregated by breeding founder animals to wild-type CFW mice, resulting in 18 transgenic strains.

In vitro culture of embryos. Five-week-old Mpv 20 heterozygous females were superovulated by intraperitoneal injection of 5 IU of pregnant mares' serum gonadotropin followed 44 h later by 5 IU of human chorionic gonadotropin. Following the second injection, the females were paired with either heterozygous or normal CFW male mice. The following morning, the females were checked for vaginal plugs. Mated females were killed 48 to 64 h after human chorionic gonadotropin injection, and their oviducts and uteri were gently flushed with CZ-B embryo culture medium (2) supplemented with 875 mg of taurine per liter (CZ-B/T). Embryos were collected, washed, and transferred to a solution containing 0.8 ml of CZ-B/T supplemented with 1 mg of glucose per ml in embryo culture dishes (Falcon; Becton Dickinson) for subsequent culture. Embryo development was assessed at 12-h intervals.

Molecular cloning. The Mpv 20 integration site was cloned by established methods as described previously (29). Briefly, *Bgl*II-digested liver DNA from a heterozygous animal was fractionated on a 0.7% agarose gel. DNA with a molecular size of 10 to 20 kb was purified from a gel slice with Geneclean (Bio 101) and ligated into the phage vector λEMBL3 (Stratagene). Ligated DNA was packaged with Gigapack reagents (Stratagene). Phage were plated onto *Escherichia coli* K802. Filter lifts were hybridized to a radiolabelled 754-bp *Nhe*I-to-*Spe*I fragment derived from the MPSV long terminal repeat (LTR). Filters were hybridized in a standard solution (5× SSC [1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate]–50% formamide at 42°C) and washed under high-stringency conditions (0.1× SSC, 0.1% sodium dodecyl sulfate), under which the probe does not detect endogenous proviral sequences on Southern blots. Five positive clones, one of which was λBg4.1, were identified from approximately 10^6 phage. Additional clones were obtained with a 2.5-kb genomic *Bam*HI fragment from λBg4.1 to screen a phage library constructed from DNA partially digested with *Mbo*I, fractionated into 15- to 20-kb fractions by sucrose gradient centrifugation, and ligated into λEMBL3 arms as described above.

RNA isolation and Northern blotting. Tissues were flash frozen in liquid nitrogen and processed for RNA extraction by the LiCl-urea method as described previously (29). RNA samples were resuspended in water, and optical densities were determined spectrophotometrically. Equal quantities of Mpv 20

TABLE 1. Results of tail DNA biopsies of test crosses

Stage of development	No. of progeny from Mpv 20/wt[a] × Mpv 20/wt matings with:			
	Total	wt/wt	Mpv 20/wt	Mpv 20/Mpv 20
Adult	199	67	132	0
Day 11 embryo	55	21	34	0

[a] wt, wild type.

TABLE 2. Development of embryos to blastocysts in vitro

Cross	No. of blastocysts/ no. of embryos	%
wt[a] × wt	49/54	91
wt × Mpv 20/wt	43/50	86
Mpv 20/wt × Mpv 20/wt	110/155	71[b]

[a] wt, wild type.
[b] Significantly different from wt × wt cross (Z test; $\alpha = 0.0006$).

heterozygous and wild-type RNAs were pipetted into a loading dye containing ethidium bromide and were separated on a 1% formaldehyde–agarose gel. The integrity and quantity of the RNA samples were confirmed by UV transillumination. RNAs were transferred to Hybond N (Amersham) by capillary transfer. RNAs were then cross-linked to the membrane with a UV cross-linker (Bio-Rad) and probed with cDNA fragments labelled with [^{32}P]dCTP by using a random oligonucleotide system (Boehringer Mannheim).

RNA quantitation. Northern blots were probed with *Npat* or *ATM* probes and exposed to PhosphorImager screens for 12 h and were then imaged on a Molecular Dynamics model SI PhosphorImager. The blots were stripped and rehybridized with a ubiquitous, abundant control probe from the *Unp* gene (10). The volumes of RNA bands were determined by using ImageQuaNT software (Molecular Dynamics). Values for *Npat* or *Atm* bands were divided by the corresponding *Unp* values to correct for RNA loading.

Exon trapping. Phage spanning the integration site were digested with *Sal*I to release inserts. Inserts were isolated by using Geneclean (Bio 101) and subjected to either a partial *Mbo*I digest or a complete *Bam*HI digest. DNA was cloned into the *Bam*HI site of the exon-trapping vector pSPL1 (Gibco BRL).

Transfections were carried out by the calcium phosphate method (6) into African green monkey Cos-1 cells. Briefly, cells were plated at a density of 8×10^5 per 100-mm-diameter tissue culture dish. When cells reached 60% confluence, 10 μg of plasmid DNA was transfected and RNA was isolated 24 h later.

cDNA was generated with SuperScript II reverse transcriptase according to the manufacturer's instructions (Gibco BRL). First-strand synthesis was primed with the SA4 oligonucleotide.

DNA sequencing. Sequencing was performed either manually or with an automated system. Manual DNA sequencing was performed by using a T7 dideoxy sequencing kit (Pharmacia). Automated sequencing was performed on an ABI 377 DNA sequencer with dye-terminator chemistry.

Probes. The *Npat* reverse transcription (RT)-PCR probe was amplified from a murine testis cDNA library with the following oligonucleotide primers: 5mnpat, 5′ TTTAATGTCTCCTGGTAGACG 3′, and 3mnpat, 5′ GGCCTAGCAGTTC ATCAATTG 3′. Cycling conditions were 40 cycles at 94°C for 30 s, 55°C for 30 s, and 72°C for 1 min, followed by a 10-min cycle at 72°C.

A 2.4-kb *ATM* probe was generated by *Eco*RI digestion of the mouse mATM 6 cDNA (T. Shiomi), followed by purification by low-melting-point agarose.

RACE cloning. 5′ rapid cloning of cDNA ends (RACE) was performed with a random-primed and oligo(dT)-primed mouse lung cDNA library (Stratagene) as the source of the template. PCR was performed with oligonucleotides 3mnpat and KS (5′ CGAGGTCGACGGTATCG 3′). Cycling conditions were 40 cycles of 94°C for 30 s, 58°C for 30 s, and 72°C for 90 s followed by a final soak at 72°C for 10 min.

RESULTS

Generation and identification of a mutant transgenic mouse line. Mpv 20 transgenic mice were generated during the same screen for recessive lethal mutations that generated the Mpv 17 strain, whose adult kidney phenotype has been described previously (29). Of the 18 transgenic mouse lines tested, only Mpv 17 and Mpv 20 were found to harbor recessive lethal phenotypes. In the Mpv 20 strain, no animals that were homozygous for the MPSV-neo provirus were identified among the progeny of heterozygous crosses (the results of tail DNA biopsies of test crosses are summarized in Table 1). In addition, no homozygotes were among day 11 midgestation embryos from similar crosses (Table 1). These data suggested the influence of an early recessive lethal phenotype, possibly prior to implantation. To examine this possibility, we cultured preimplantation embryos from heterozygous × heterozygous or heterozygous × wild-type crosses. Approximately one quarter of the embryos from the crosses of heterozygous and heterozygous animals failed to proceed from the morula to the blastocyst stage (Table 2). Many of the arrested embryos had the appearance of uncompacted eight-cell embryos (Fig. 1).

Molecular cloning of the Mpv 20 proviral insertion site. In order to identify the gene or genes affected by the insertion of proviral sequences, we cloned the region of the insertion in bacteriophage. A proviral LTR fragment was used to screen a genomic library constructed from Mpv 20 heterozygous liver DNA, yielding clone Bg4.1 (Fig. 2). Additional overlapping phage clones were obtained from a wild-type genomic library, and a restriction map of the locus was deduced by single and multiple digests with several enzymes (Fig. 2). To confirm that these phage clones were in fact derived from the insertion site, a Southern blot of DNA from wild-type and Mpv 20 heterozygous mice was hybridized with a probe derived from phage M2041, which was predicted to contain the wild-type "preinsertion" site. As expected, the insertion of proviral sequences resulted in the appearance of additional bands in the Mpv 20 heterozygous lanes that were in complete agreement with the previously deduced restriction map of the locus (Fig. 3). Southern blotting analyses with additional enzymes confirmed that proviral insertion in Mpv 20 was not accompanied by any gross genomic rearrangements near the integration site (data not shown).

Identification of a transcriptional unit at the Mpv 20 locus. An exon-trapping strategy was used to identify putative exons in the cloned phage DNA. Subcloned *Mbo*I fragments from phage M2017 as well as the 2.5-kb *Bam*HI fragment from phage M2041 (in both orientations) were used to transfect Cos-1 cells. One exon was identified in this screen and was used to probe a Southern blot of cloned phage DNAs digested

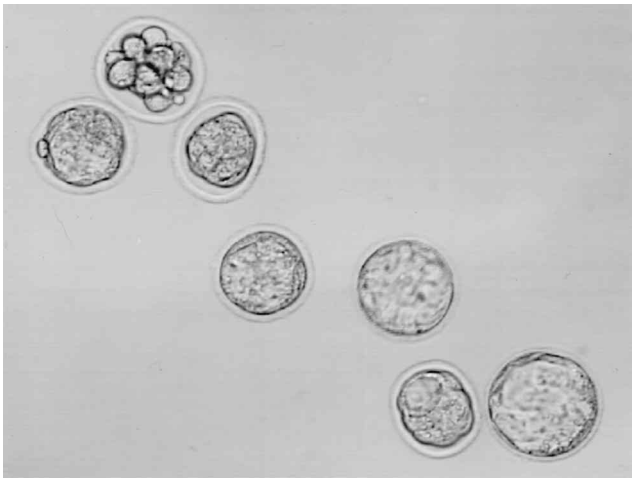

FIG. 1. Cultured embryos from an Mpv 20 heterozygous cross (photographed by phase-contrast microscopy at 3.5 days postcoitum). The uppermost embryo arrested at the uncompacted eight-cell stage, while the rest compacted and, in some cases, formed blastocysts.

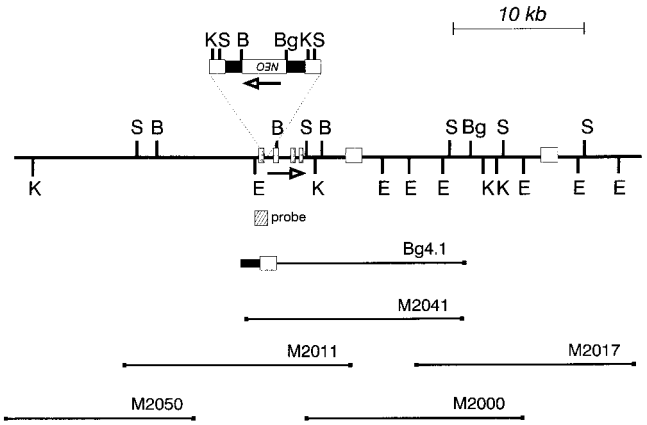

FIG. 2. Map of the Mpv 20 locus. The MPSV-neo provirus is indicated at the top (in the transcriptional orientation opposite to that of the cellular exons), with its site of integration indicated by broken lines. Sequenced exons (Fig. 4a) are indicated by shaded boxes, and exons whose approximate positions were inferred from hybridization analyses are indicated by open boxes. The box marked "probe" indicates the position of a genomic fragment used for the Southern blotting analysis (Fig. 3). The positions of overlapping phage clones are depicted below the genomic restriction map. B, *Bam*HI; Bg, *Bgl*II; E, *Eco*RI; K, *Kpn*I; S, *Sac*I (*Sst*I).

with various enzymes. The trapped exon sequences hybridized to the 2.5-kb *Bam*HI fragment of M2041 (data not shown).

Identification of *Npat* as the disrupted gene. In order to confirm that the trapped exon was derived from the integration site, an *Sst*I fragment from clone Bg4.1 that contained part of the proviral LTR as well as the genomic flanking sequence and a 9-kb *Eco*RI fragment from clone M2041 that included the site of proviral insertion were subcloned into the plasmid vector pGEM-4 (Promega). The sequence of the trapped exon was identified within the genomic sequence from the insertion site. The predicted translation products of the trapped exon and of possible exons in the sequence near the insertion site were compared to entries in the translated GenBank database. This search resulted in the identification of the human *NPAT* gene product as one very similar to the predicted translation product of putative exons near the proviral insertion site (Fig. 4). Surprisingly, the trapped exon was aligned to the translation product of its supposedly antisense strand (it appears to have been fortuitously trapped as a consequence of elements in the DNA which can function as splice acceptors and donors in the orientation opposite to that of *Npat*). Mapping data from somatic cell hybrids had placed the Mpv 20 provirus on chromosome 9 (5a), which is also the site of the murine *Npat* gene (18). Based on the sequencing and mapping data, we concluded that the MPSV-neo provirus in Mpv 20 mice had inserted into the *Npat* gene. The sequencing data suggest that the provirus was situated in an intron of *Npat* in the transcriptional orientation opposite to that of the *Npat* gene.

Confirmation of the relationship of Mpv 20 and *Npat* by Southern analysis. To confirm that Mpv 20 was an insertion mutation of the *Npat* gene, we isolated an *Npat* cDNA by RT-PCR. Oligonucleotides were chosen corresponding to a region of predicted mouse protein product that had 100% identity to the human NPAT protein (13). These primers were used to amplify a partial mouse cDNA from a testis cDNA library. An RT-PCR fragment which spans a region corresponding to amino acids 281 to 302 of the human protein was obtained, and its identity was confirmed by hybridization to the human NPAT cDNA (13) and by sequencing (data not shown). This probe was subsequently used in both Northern and Southern blot analyses of Mpv 20 and wild-type samples. Southern blots of DNAs digested with a series of enzymes were probed with the above-mentioned fragment, which identified both the wild-type as well as the disrupted alleles of *Npat* (Fig. 5). These data confirm that the provirus in Mpv 20 had inserted into the *Npat* gene.

Analysis of *Npat* and *ATM* RNA expression. Northern blots of total RNA (Fig. 6a) or poly(A)-enriched RNA (data not shown) were probed with the RT-PCR *Npat* fragment. In samples from brain, lung, and kidney, there was a demonstrable difference in *Npat* expression (Fig. 6a). Decreased expression of *Npat* was noted in Mpv 20 heterozygotes in which one *Npat* allele was expected to be inactivated by proviral insertion. Because the *Npat* cDNA probe in this analysis was derived from exons 3′ to the provirus (with respect to the *Npat* gene), it could not distinguish between the transcriptional inactivation of the disrupted *Npat* allele and the transcriptional termination upstream of the sequences contained within the probe. To investigate the mechanism of inactivation further, we obtained cDNA sequences from *Npat* 5′ exons by RACE. When Northern blots were probed with the 5′ RACE cDNA, an RNA species of approximately 1 kb was detected in tissues from Mpv 20 heterozygous mice but not in samples from wild-type mice (Fig. 6b). An RNA band of similar mobility was detected in the same lanes with an MPSV LTR probe (Fig. 6c), suggesting that the disrupted *Npat* allele was transcribed but that the transcripts did not include the *Npat* exons immediately 3′ to the provirus.

Northern blots were stripped and reprobed with a murine *ATM* cDNA fragment to determine if the expression of this closely linked gene (13) was also affected by the integration event. There was no obvious difference in *Atm* expression between Mpv 20 heterozygous and wild-type tissue RNAs (Fig. 6d). To quantify levels of *Npat* and *Atm* expression in RNA samples, Northern blots were analyzed with a PhosphorImager. *Npat* mRNA levels were 1.5- to 1.8-fold higher in wild-type animals than in Mpv 20 heterozygous mice (Table 3). *Atm* mRNA levels were not substantially different between wild-type and heterozygous mice.

DISCUSSION

The Mpv 20 mouse strain was one of two recessive lethal mutant strains (the other being Mpv 17) obtained from a screen of 18 transgenic lines obtained by retroviral infection of embryos (29). If these data are grouped with data from the

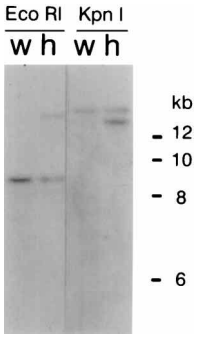

FIG. 3. Southern blot analysis of heterozygous (h) versus wild-type (w) CFW DNA. A genomic fragment adjacent to the insertion site (from phage clone M2041 [Fig. 2]) was used to probe DNA samples digested completely with the enzymes indicated. In each case, a second band, whose size was consistent with the deduced restriction map of the locus, was evident in heterozygous DNA. These data confirm that clone M2041 is derived from the proviral insertion site.

FIG. 4. (a) Sequence of genomic DNA at the Mpv 20 proviral insertion site. The lowercase sequences did not align with *NPAT* or *Npat* cDNA sequences and are therefore intronic. The uppercase sequences represent exons whose predicted translation products matched the translated *NPAT* entry (exons correspond to positions 551 to 906 of the human *NPAT* cDNA [GenBank accession no. D83243]). Putative introns have appropriate splice donor and acceptor sites. The position of the proviral insertion is indicated by a filled triangle (its position was determined by sequencing off the proviral LTR of clone Bg4.1 into the flanking genomic sequence). The *Eco*RI and *Sac*I sites adjacent to the provirus (Fig. 2) are in italics and in bold and italics, respectively; the *Bam*HI site is double underlined. The underlined regions represent a perfectly conserved direct repeat of unknown function. The sequence represents a contig of multiple sequence reactions of both DNA strands. (b) Alignment of predicted translation products of the *Npat* exons flanking the Mpv 20 insertion site and the published *NPAT* translation product (12). Identical residues are indicated by a vertical line. There is 82% identity between the predicted murine Npat and human NPAT amino acid sequences within this region. The intron harboring the Mpv 20 provirus separates exons encoding the residues indicated by triangles.

Mov strains (reviewed in reference 8), there have been four mutant strains out of 70 transgenic lines, giving a mutation frequency of approximately 5 to 6%. Although the sample size is obviously quite small, it is intriguing that the proviral insertions in these strains appear to favor noncoding regions. In the Mov 13 strain, proviral insertion occurred in the α(1)I collagen gene (11, 14, 22), resulting in the inactivation of a *cis*-acting transcriptional element (1). In the Mov 34 strain, proviral insertion into the 5′ region of a gene encoding a proteasomal subunit (9, 25, 26) was found to inactivate RNA expression from the gene. Recently, Harbers et al. (12) have reported that insertion of a provirus into the first intron of a gene encoding a ubiquitin-conjugating enzyme caused a 70% reduction in the steady-state levels of mRNA from the gene, resulting in a placental defect and embryonic lethality. In the present work, we show that in the Mpv 20 strain, the provirus inserted into an

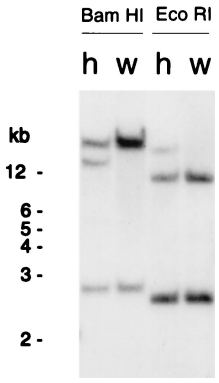

FIG. 5. Confirmation that the *Npat* RT-PCR probe is derived from the Mpv 20 locus. A Southern blot of Mpv 20 heterozygous (h) and CFW wild-type littermate (w) DNA was probed with the RT-PCR fragment. The RT-PCR fragment detected an extra band in heterozygous DNA and must therefore be derived from the locus. The sizes of the bands detected indicate that the exons from which the probe was derived must lie entirely between the *Eco*RI sites flanking the insertion site (Fig. 2).

intron of the *Npat* gene. Retroviral insertion can occur at many sites in the mammalian genome, but there is some evidence of preferential insertion into active chromatin (20, 28). It may be that the noncoding sequences that are the sites of proviral insertion in the Mov and Mpv strains have a conformation that is particularly accessible to the viral integrase or that disruptions of coding sequences are merely underrepresented in the current small sample size and would be detected if more mutant strains were generated. Although in the Mov 13 strain the mechanism of inactivation involves the disruption of a *cis*-acting element (1), in other strains the mechanism of gene inactivation is not fully clear. Upon passage through embryonic cells, type C proviruses can initiate a general inactivation by a mechanism that involves the methylation (15, 16) of a genomic expanse that includes not only proviral sequences but also flanking genomic regions. We have not been able to detect proviral expression in Mpv 20 mice using a *neo* probe (30). If there was more extensive inactivation in the Mpv 20 strain, one might expect regional differences in methylation that could be detected by Southern analysis of DNAs digested with methylation-sensitive restriction enzymes and probed with *Npat* cDNAs, but we detected no such differences between Mpv 20 and wild-type DNAs (data not shown). Further, one might expect the adjacent *Atm* gene to be affected by any regional inactivation (particularly given the proximity of the *Npat* and *ATM* transcriptional start sites), but this does not seem to be the case (Fig. 6d). Rather, our data are best explained by a model of altered transcription leading to a nonfunctional *Npat* transcript. In one scenario, we speculate that in the reverse orientation there remain sequence elements in the MPSV-neo provirus that could provide functional polyadenylation and transcriptional termination signals (although these elements are not obvious from an examination of the MPSV DNA sequence). The detection of a 1-kb hybrid *Npat*-LTR transcript containing a 5′ but not a 3′ *Npat* exonic sequence is consistent with this model, which would imply that approximately 600 bases of the mRNA are derived from *Npat*, the remainder being comprised of MPSV-neo sequences and the poly(A) tail [the truncated hybrid mRNA was detected in poly(A)$^+$ RNA samples, indicating that polyadenylation had occurred]. An alternative explanation for the 1-kb hybrid mRNAs involves an unusual splicing event, resulting in the inclusion of reversed proviral sequences and the exclusion of the proximal 3′ *Npat* exons but not of the exon containing the *Npat* polyadenylation signal. In either case, the reading frame of *Npat* would be lost or become nonsensical at the same position. If the reading frame of *Npat* is similar to that of the human *NPAT* gene, we predict that the truncated mRNA encodes a truncated protein

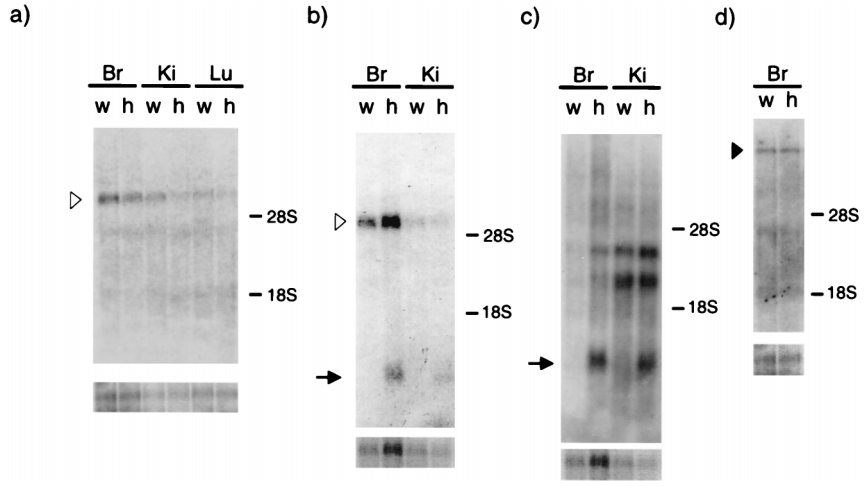

FIG. 6. (a) Northern blot analysis of tissue RNAs from a heterozygous Mpv 20 animal and a wild-type littermate. Both animals were female. The upper panel depicts the pattern detected with an *Npat* RT-PCR cDNA probe. The position of the *Npat* mRNA is indicated by an open triangle. The lower panel represents the same blot hybridized with a control human transferrin receptor cDNA probe (obtained from the American Type Culture Collection). In each matched tissue pair, the *Npat* band was of greater intensity in the wild-type sample, consistent with the inactivation of one *Npat* allele in Mpv 20 heterozygous mice. The positions of ribosomal markers (as determined by UV transillumination) are indicated. (b) Northern analysis with a 5′ RACE product probe. The open triangle indicates the position of *Npat* mRNA. A truncated mRNA (arrow) was detected in samples from Mpv 20 heterozygous mice but not in samples of wild-type littermate RNA. The control hybridization (lower panel) was to transferrin receptor cDNA, as in panel a. (c) Northern analysis with a probe derived from the MPSV LTR. A transcript with a mobility similar to that detected in panel b (arrow) was detected in heterozygous but not in wild-type poly(A)$^+$ RNA. Other bands were detected in wild-type as well as in Mpv 20 heterozygous RNA and probably are derived from endogenous type C proviruses. The control hybridization (lower panel) was to transferrin receptor cDNA. (d) Analysis of *Atm* RNA expression. The Northern blot from panel a was stripped and rehybridized with a mouse *Atm* cDNA probe. No difference was detected between *Atm* RNA levels (closed triangle) in heterozygous Mpv 20 and wild-type littermate brain samples. The loading control hybridization (lower panel) was to transferrin receptor cDNA. h, heterozygous; w, wild type; Br, brain; Ki, kidney; Lu, lung.

TABLE 3. mRNA expression in Mpv 20 heterozygous versus wild-type mice

Probe	Tissue	PhosphorImager counts (adjusted to loading controls)[a]		Ratio
		wt/wt	Mpv 20/wt	
Npat	Brain	0.027	0.017	1.6
Npat	Kidney	0.015	0.010	1.5
Npat	Lung	0.016	0.009	1.8
ATM	Brain	0.036	0.032	1.1

[a] wt, wild type.

of 203 amino acids with a molecular mass of approximately 23 kDa. Until the function of the gene product is known, it will be impossible to predict whether such a truncated Npat protein can retain any functional activity. Clearly, if a truncated protein does exist, it is lacking some activity that is essential for early embryogenesis.

The relative levels of *Npat* mRNA in Mpv 20 heterozygous and wild-type mice showed less than the expected twofold difference, varying from 1.5- to 1.8-fold in the various tissues (Table 3). Harbers et al. (12) have also reported that proviral insertion into an intron can result in less than the expected reduction in mRNA levels from the mutated gene. Our data can be interpreted in the context of either of the previously described models of inactivation. The data may mean that at some frequency the transcriptional machinery is able to proceed through the inverted provirus without termination or that the splicing machinery is occasionally able to correctly excise the provirus-containing intron. Another explanation for the less-than-twofold difference is that there is elevated, perhaps compensatory, transcription from the wild-type *Npat* gene in the heterozygous animals. These models make different predictions with respect to the expression of *Npat* in homozygous embryos. If the mechanism of inactivation is "leaky," there might be a small amount of functional Npat protein of zygotic origin even in homozygous embryos (although the amount must be insufficient for viability). If, on the other hand, there is elevated transcription of the wild-type gene in heterozygotes, this could not occur in homozygotes and there would be no Npat protein of zygotic origin. In the absence of any information regarding the role of Npat, it is difficult to predict whether small amounts would have any functional importance.

The lack of homozygous Mpv 20 midgestation embryos and the preimplantation arrest of one quarter of the embryos from the Mpv 20 heterozygous crosses suggest a recessive lethal mutation whose phenotype is among the earliest known. There are few reported examples of mutations affecting the development of the preimplantation mouse embryo. The earliest of these, the c^{25H} strain, carries a 5-centimorgan deletion spanning the albino locus on chromosome 7 (17), and embryonic arrest occurs at the two- to six-cell stage. In a second example, insertion of exogenous DNA into and/or the deletion of 2 centimorgans of chromosome 1 in the βS12 transgenic line resulted in the generation of a morula decompaction (*mdn*) phenotype in which embryos were observed to compact and then to decompact and developmentally arrest (4). We have not observed the compaction-decompaction phenomenon in Mpv 20 embryos. Because of the large genomic rearrangements that have occurred in the c25H and βS12 strains the gene or genes responsible for the early embryonic arrest have not been identified in either case. With respect to targeted (as opposed to random) mutations, there is a surprising paucity of descriptions of preimplantation phenotypes in the literature. Intuitively, most mutations that have consequences that are lethal to cells might be expected to have associated early embryonic lethality, but if this is the case, the mutations have not been reported. One exception is the targeted mutation of *Rad51* (the mammalian homolog of the yeast *RAD51* and *E. coli recA* genes, required for DNA repair). *Rad51* is apparently essential for cell viability in that no ES cell lines that were homozygous null for the gene could be obtained. Among embryos from a heterozygous cross, very few $Rad51^{-/-}$ embryos were found to have progressed to the morula stage (27). A second example of early embryonic lethality is provided by the mutation of E-cadherin, which is known to be required for the compaction process (24) and whose absence could therefore be expected to have deleterious effects on development. Inactivation of E-cadherin by homologous recombination caused the dissociation of blastomeres and the resulting arrest of homozygous mutant embryos (19).

The function of *Npat* is not currently known. Based on the presence of protein motifs in the predicted translation product, it has been speculated that the human *NPAT* gene encodes a nuclear protein that may be a substrate of cyclin-dependent kinases (13). There are also sequence similarities between *NPAT* and transcription factors such as *Oct-1*, although no subcellular localization or other data that would support a nuclear role for the *Npat* protein product have been reported. Until more information is available, it will be difficult to explain how the disruption of the *Npat* gene leads to embryonic arrest at or around the uncompacted eight-cell stage. Two possibilities must be considered: either the lack of zygotic *Npat* is of no consequence until maternal stores are depleted (which must occur at or around the eight-cell stage) or the *Npat* gene product is not required by the embryo until the eight-cell stage. The former scenario implies that the *Npat* gene product is involved in any number of activities related to the structure, metabolism, and cycle of the cells, whereas the latter scenario implies that the *Npat* product has a direct or indirect role in events occurring at the eight-cell stage (notably, cell polarization or compaction). In the future, we will manipulate the zygotic levels of *Npat* by microinjection to distinguish between these models.

ACKNOWLEDGMENTS

Aspects of this work were supported by National Institutes of Health research grant R35CA44339 to R.J. and by National Cancer Institute of Canada grant 2251 to D.G. T.I. is supported in part by grants from the Ministry of Education, Science and Culture and by a research grant on aging and health from the Ministry of Health and Welfare of Japan. A Fonds pour la Formation de Chercheurs et l'Aide à la Recherche studentship was awarded to M.D.

REFERENCES

1. **Barker, D. A., H. Wu, S. Hartung, M. Briendl, and R. Jaenisch.** 1991. Retrovirus-induced insertional mutagenesis: mechanism of collagen mutation in Mov 13 mice. Mol. Cell. Biol. **11:**5154–5163.
2. **Chatot, C. L., C. A. Ziomek, B. D. Bavister, J. L. Lewis, and I. Torres.** 1989. An improved culture medium supports development of random-bred 1-cell mouse embryos *in vitro*. J. Reprod. Fertil. **86:**679–688.
3. **Chen, Z., G. A. Friedrich, and P. Soriano.** 1994. Transcriptional enhancer factor 1 disruption by a retroviral gene trap leads to heart defects and embryonic lethality in mice. Genes Dev. **8:**2293–2301.
4. **Cheng, S. S., and F. Costantini.** 1993. Morula decompaction (*mdn*), a preimplantation recessive lethal defect in a transgenic mouse line. Dev. Biol. **156:**265–277.
5. **DeGregori, J., A. Russ, H. von Melchner, H. Rayburn, P. Priyaranjan, N. A. Jenkins, N. G. Copeland, and H. E. Ruley.** 1994. A murine homolog of the yeast RNA1 gene is required for postimplantation development. Genes Dev. **8:**265–276.

5a. **Francke, U.** Unpublished data.

6. **Graham, F. L., and A. J. van der Eb.** 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. Virology **52:**456–467.
7. **Gray, D. A.** 1991. Insertional mutagenesis: neoplasia arising from retroviral integration. Cancer Invest. **9:**295–304.
8. **Gridley, T., P. Soriano, and R. Jaenisch.** 1987. Insertional mutagenesis in mice. Trends Genet. **3:**162–166.
9. **Gridley, T., D. A. Gray, T. Orr-Weaver, P. Soriano, D. E. Barton, U. Francke, and R. Jaenisch.** 1990. Molecular analysis of the Mov 34 mutation: transcript disrupted by proviral integration in mice is conserved in Drosophila. Development **109:**235–242.
10. **Gupta, K., M. Chevrette, and D. A. Gray.** 1994. The Unp proto-oncogene encodes a nuclear protein. Oncogene **9:**1729–1731.
11. **Harbers, K., M. Kuehn, H. Delius, and R. Jaenisch.** 1984. Insertion of retrovirus into the first intron of α1 (I) collagen gene leads to embryonic lethal mutation in mice. Proc. Natl. Acad. Sci. USA **81:**1504–1508.
12. **Harbers, K., U. Müller, A. Grams, E. Li, R. Jaenisch, and T. Franz.** 1996. Provirus integration into a gene encoding a ubiquitin-conjugating enzyme results in a placental defect and embryonic lethality. Proc. Natl. Acad. Sci. USA **93:**12412–12417.
13. **Imai, T., M. Yamauchi, N. Seki, T. Sugawara, T. Saito, Y. Matsuda, H. Ito, T. Nagase, N. Nomura, and T. Hori.** 1996. Identification and characterization of a new gene physically linked to the ATM gene. Genome Res. **6:**439–447.
14. **Jaenisch, R., K. Harbers, A. Schnieke, J. Löhler, I. Chumakov, D. Jähner, D. Grotkopp, and E. Hoffmann.** 1983. Germline integration of Moloney murine leukemia virus at Mov 13 locus leads to recessive lethal mutation and early embryonic death. Cell **32:**209–216.
15. **Jähner, D., and R. Jaenisch.** 1980. Integration of Moloney leukemia virus into the germ line of mice: correlation between genotype and virus activation. Nature **287:**456–458.
16. **Jähner, D., and R. Jaenisch.** 1985. Retrovirus induced de novo methylation of flanking host sequences correlates with gene inactivity. Nature **315:**594–597.
17. **Lewis, S. E.** 1978. Developmental analysis of lethal effects of homozygosity for the c25H deletion in the mouse. Dev. Biol. **65:**553–557.
18. **Matsuda, Y., T. Imai, T. Shiomi, T. Saito, M. Yamauchi, T. Fukao, Y. Akao, N. Seki, H. Ito, and T. Hori.** 1996. Comparative genome mapping of the ataxia telangiectasia region in mouse, rat, and syrian hamster. Genomics **34:**347–353.
19. **Riethmacher, D., V. Brinkmann, and C. Birchmeier.** 1995. A targeted mutation in the mouse E-cadherin gene results in defective preimplantation development. Proc. Natl. Acad. Sci. USA **92:**855–859.
20. **Rohdewohld, H. M., H. Weiher, W. Reik, R. Jaenisch, and M. Breindl.** 1987. Retrovirus integration and chromatin structure: Moloney murine leukemia proviral integration sites map near DNase I-hypersensitive sites. J. Virol. **61:**336–343.
21. **Rudnicki, M. A., and R. Jaenisch.** 1991. Insertional mutagenesis in the mouse, p. 13–39. *In* S. Tilghman (ed.), Genome analysis, vol. II. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
22. **Schnieke, A., K. Harbers, and R. Jaenisch.** 1983. Embryonic lethal mutation in mice induced by retrovirus insertion into the α(I) collagen gene. Nature **304:**315–320.
23. **Seliger, B., R. Kollek, C. Stocking, T. Franz, and W. Ostertag.** 1986. Viral transfer, transcription, and rescue of a selectable myeloproliferative sarcoma virus in embryonal cell lines: expression of the *mos* oncogene. Mol. Cell. Biol. **6:**286–293.
24. **Shirayoshi, Y., T. S. Okada, and M. Takeichi.** 1983. The calcium-dependent cell-cell adhesion system regulates inner cell mass formation and cell surface polarization in early mouse development. Cell **35:**631–638.
25. **Soriano, P., T. Gridley, and R. Jaenisch.** 1987. Retroviruses and insertional mutagenesis: proviral integration at the Mov 34 locus leads to early embryonic death. Genes Dev. **1:**366–375.
26. **Tsurumi, C., G. N. DeMartino, C. A. Slaughter, and K. Tanaka.** 1995. cDNA cloning of p40, a regulatory subunit of the human 26S proteasome, and a homolog of the Mov-34 gene product. Biochem. Biophys. Res. Commun. **210:**600–608.
27. **Tsuzuki, T., Y. Fijii, K. Sakumi, Y. Tominaga, K. Nakao, M. Sekiguchi, A. Matsushiro, Y. Yoshimura, and T. Morita.** 1996. Targeted disruption of the Rad51 gene leads to lethality in embryonic mice. Proc. Natl. Acad. Sci. USA **93:**6236–6240.
28. **Vijaya, S., D. L. Steffen, and H. L. Robinson.** 1986. Acceptor sites for retroviral integrations map near DNase I-hypersensitive sites in chromatin. J. Virol. **60:**683–692.
29. **Weiher, H., T. Noda, D. A. Gray, A. H. Sharp, and R. Jaenisch.** 1990. Transgenic mouse model of kidney disease: insertional inactivation of ubiquitously expressed gene leads to nephrotic syndrome. Cell **62:**425–434.
30. **Weiher, H.** Unpublished data.