The EMBO Journal vol.14 no.10 pp.2326–2332, 1995

Retroviral sequences located within an intron of the *dilute* gene alter *dilute* expression in a tissue-specific manner

Peter K.Seperack, John A.Mercer[1], Marjorie C.Strobel[2], Neal G.Copeland[2] and Nancy A.Jenkins[2,3]

Institute for Arthritis and Autoimmunity, 400 Morgan Lane, West Haven, CT 06516, [1]Department of Physiology, University of Texas Southwestern Medical Center, Dallas, TX 75235 and [2]Mammalian Genetics Laboratory, ABL–Basic Research Program, NCI–Frederick Cancer Research and Development Center, Frederick, MD 21702, USA

[3]Corresponding author

The murine *dilute* coat color locus encodes an unconventional myosin heavy chain that is thought to be required for the elaboration or maintenance of dendrites or organelle transport in melanocytes and neurons. In previous studies we showed that the *d* mutation carried by many inbred strains of mice (now referred to as *dilute viral*, d^v), is caused by the integration of an ecotropic murine leukemia virus (*Emv-3*) into the *dilute* gene and that phenotypic revertants of d^v (termed d^+) result from viral excision; a solo viral long terminal repeat (LTR) is all that remains in revertant DNA. In the studies described here we show that *Emv-3* sequences are located within an intron of the *dilute* gene in a region of the C-terminal tail that is differentially spliced. We also show that these *Emv-3* sequences result in the production of shortened and abnormally spliced *dilute* transcripts and that the level of this effect varies among tissues. This tissue-specific effect on *dilute* expression likely accounts for the absence of neurological abnormalities observed in d^v mice. Surprisingly, we also found that the solo viral LTR present in revertant d^+ DNA produces a tissue-specific effect on *dilute* expression, although this effect is less dramatic than with the full-length provirus and produces no obvious mutant phenotype. These findings have important implications for understanding the effects of viral sequences on mammalian gene expression.

Key words: *dilute*/myosin heavy chain/retroviral insertional mutagenesis/solo LTRs/tissue-specific splicing

Introduction

More than 70 loci have been identified in the mouse that affect coat color (Silvers, 1979). Among the best studied of these mutations is the recessive autosomal mutation *dilute*, which has been used extensively as a phenotypic marker for assessing mutagenesis frequencies in the mammalian germline (Russell, 1951). Over the years, hundreds of *dilute* (*d*) alleles, both spontaneous and induced, have been identified and characterized genetically (Russell, 1971). The discovery that the original *dilute* mutation carried by many inbred strains (now called *dilute viral* or d^v) was caused by the integration of an ecotropic murine leukemia virus (*Emv-3*) into the *dilute* locus (Jenkins *et al.*, 1981) has allowed the cloning of genomic DNA (Hutchison *et al.*, 1984; Rinchik *et al.*, 1986; Jenkins *et al.*, 1989; Strobel *et al.*, 1990) and corresponding cDNA (Mercer *et al.*, 1991) from the *dilute* gene.

The nucleotide and deduced amino acid sequence of *dilute* indicated that it encodes a novel class of myosin heavy chain (Mercer *et al.*, 1991). Myosin heavy chains (MHCs) contain a globular N-terminal, or head, region that binds actin and produces mechanochemical force via ATP hydrolysis. The C-terminal, or tail, region can vary in sequence and overall structure and it is the tail region that largely determines the functional specificity of each myosin heavy chain. The N-terminal head region of *dilute* shows extensive homology with the head regions of all known MHCs (Mercer *et al.*, 1991). Following the head region is an imperfect amino acid repeat that is similar to the calmodulin binding repeats identified in brush border myosin I (Mercer *et al.*, 1991). Cheney and Mooseker (1992) have termed this repeat the 'IQ motif'. The tail region of *dilute* contains both coiled coil and globular domains. Since our initial characterization of *dilute*, other members of its class have been described, including chicken p190 (Espreafico *et al.*, 1992), yeast *MYO2* (Johnston *et al.*, 1991) and yeast *MYO4* (Pascolo *et al.*, 1992; Haarer *et al.*, 1994).

Skin melanocytes have dendritic processes that are filled with pigment granules called melanosomes. The melanosomes are transported through these dendrites into the developing hair shaft (Silvers, 1979). In contrast, *dilute* melanocytes lack pigment-filled dendritic processes, resulting in an uneven distribution of pigment within the hair shaft and a lightening of coat color (Russell, 1949). The level of pigment synthesis in a *dilute* mouse is normal. The apparent lack of melanocyte processes observed in *dilute* mutants suggests that *dilute* function is required either for the elaboration or maintenance of dendritic processes or for the transport of melanosomes within them.

Most *dilute* alleles (termed *dilute lethal*, d^l) also produce a neurological defect, characterized by opisthotonos beginning at 8–10 days of age, continuing until death of the animal at 2–3 weeks of age (Searle, 1952). d^l mutations represent the null phenotype at the *d* locus (Moore *et al.*, 1988; Strobel *et al.*, 1990). The exact cause of this neurological defect and the cell types affected in d^l mice remain to be determined. Transcripts from the *dilute* locus are expressed in a number of embryonic and adult tissues and are very abundant in neurons of the central nervous system, cephalic ganglia and spinal ganglia (Mercer *et al.*, 1991). Collectively, these results suggest that *dilute* may also function in the elaboration or maintenance of neuronal cell processes and/or organelle transport within neurons.

The d^v allele, which was so useful in cloning the *dilute* gene, differs from all other *d* alleles in that it does not produce any obvious neurological defects and is phenotypically unstable (Schlager and Dickie, 1971). In a survey of 1 115 818 mice screened for germline and somatic reversion at *d*, 10 germline d^+ revertants (4.5×10^{-6} reversions/gamete) and one somatic d^+ revertant (9×10^{-7} somatic reversions/animal analyzed) were identified (Seperack *et al.*, 1988). All germline d^+ revertants resulted from virus excision (Copeland *et al.*, 1983; Seperack *et al.*, 1988). Analysis of two somatic d^+ revertants (Seperack *et al.*, 1988; our unpublished results) suggested that reversion results from intrachromosomal homologous recombination between LTRs of one provirus to generate an unintegrated circular viral DNA molecule containing one LTR, with a single LTR remaining at the revertant site. Homologous recombination between LTRs has been reported for retroviral-like elements carried by other species, including *Ty* elements of yeast and *copia*-like elements of *Drosophila melanogaster* (Fink *et al.*, 1981; Zachar *et al.*, 1985) and appears to be a common cause for the reported instability of mutations carried by a number of different species.

In the studies described here we show that the viral integration that caused the d^v mutation occurred within an intron of the *dilute* gene, in a region of the C-terminal tail that is differentially spliced. The effect of viral sequences (both the full-length provirus and the solo viral LTR) on the processing of *dilute* transcripts is also examined. These studies provide a model to explain the absence of a neurological phenotype for the d^v allele and provide new insights into the disruption of mammalian gene expression by retroviral insertion.

Results and discussion

The Emv-3 integration site is located in an intron of the dilute tail

Hybridization of *dilute* cDNA clones to genomic lambda clones spanning the *Emv-3* integration site (obtained from a wild-type C57BL/6J mouse genomic DNA library) and limited sequencing of genomic regions containing *dilute* exon sequences (Strobel *et al.*, 1990) indicate that the *Emv-3* provirus is located within an intron of the *dilute* gene between amino acids 1319 and 1320 of the *dilute* brain cDNA sequence (Mercer *et al.*, 1991; Figure 1). *Emv-3* is oriented in the same transcriptional direction as *dilute*. The *dilute* tail region spans amino acids 1041–1852; thus *Emv-3* is located within the *dilute* tail. These findings are consistent with genetic data indicating that *dilute* revertants are phenotypically normal, although a single 523 bp LTR is retained in revertant DNA (Hutchison *et al.*, 1984).

The dilute tail is alternatively spliced

Evidence for alternative splicing within the *dilute* tail was originally obtained by comparing the sequences of a series of *dilute* cDNA clones isolated from a mouse B16 melanoma library wild-type at *d* (Strobel *et al.*, 1990; P.K.Seperack, unpublished observations) with a set of brain cDNA clones isolated from a C57BL/6J mouse (Mercer *et al.*, 1991). All of the alternative splicing that we have detected in *dilute* to date is located in a region of the *dilute* tail containing the *Emv-3* integration site. In the studies described here we have further characterized the nature of this alternative splicing, with the hope that this may provide clues as to why viral integration induces a coat color phenotype in the absence of a neurological phenotype. Since all of the alternatively spliced products were located between *dilute* exons A and E (Figure 1), we used oligodeoxynucleotide primers from these two exons (primers 41 and 43 in Figure 1) to amplify first-strand cDNA from brain and skin of adult C57BL/6J mice and two wild-type cell lines, the B16 melanoma and melan-c melanocytes (Bennett *et al.*, 1989) (Figure 2). Three different bands were detected, with lengths of 370, 298 and 289 bp. Only the 298 bp band was detected in brain cDNA. In skin the 370 bp band was most prevalent, while a lesser amount of the 289 bp band and a minor amount of the 298 bp band (not visible in Figure 2) was also detected. In the B16 melanoma and melan-c melanocytes virtually all of the PCR product migrated at 370 bp, however, small amounts of the 289 bp band were also detected.

Gel purification and direct sequencing was used to determine the structure of each PCR product and limited sequencing of *dilute* genomic DNA was used to confirm predicted intron/exon boundaries (Figure 1). The 370 bp PCR product was found to contain exons A, C, D and E, the 298 bp product exons A, B, C and E and the 289 bp product exons A, C and E. Exons A, C and E are thus common to all three alternative splice forms, with the difference in each splice form being due to the presence or absence of exons B and D.

Published sequences of cDNAs closely related to *dilute* (Espreafico *et al.*, 1992; Sanders *et al.*, 1992) lack another exon we call exon F. Since our 3′ primer was in exon E, this exon would not have been amplified. PCR analysis of cDNA using a 5′ primer in exon E and a 3′ primer located 3′ of exon F indicated that nearly all brain transcripts lack exon F and nearly all melanoma and melanocyte transcripts contain exon F (data not shown). Exons B, D and F are 9, 81 and 75 bases in length respectively and their inclusion or exclusion in *dilute* transcripts would therefore not disrupt the *dilute* reading frame. The nucleotide and derived amino acid sequences of each of these *dilute* exons are shown in Figure 3.

Functional implications of alternative splicing in the dilute tail

The proximal region of the dilute tail has an α-helical coiled coil region, which in a chicken homolog of *dilute* (p190) has been shown to promote dimerization (Cheney *et al.*, 1993). Just distal to the long coiled coil region are two putative shorter α-helical coiled coil regions (Espreafico *et al.*, 1992). Rotary shadowed electron micrographs of purified p190 (Cheney *et al.*, 1993) are consistent with this hypothesis, although they do not definitely prove the existence of such regions, since small coiled coil regions are difficult to resolve by electron microscopy. If these additional small coiled coil regions do in fact exist, it is important to note that the most distal one would be interrupted by the presence of exon F in the *dilute* protein expressed in melanocytes, but not in the *dilute* protein expressed in brain. The presence or absence of exon F could therefore dramatically alter the structure of the

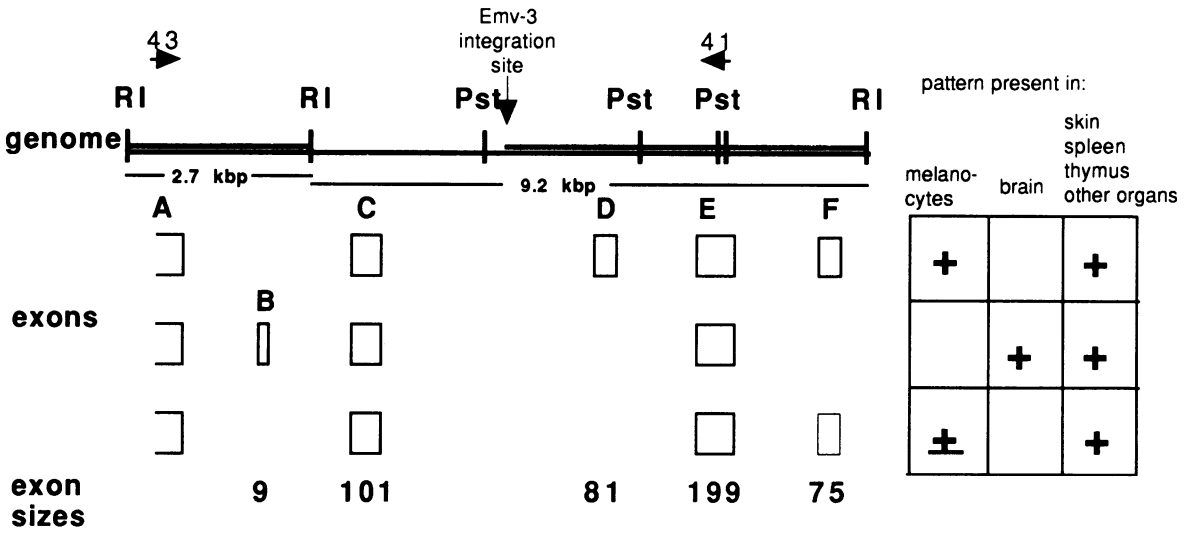

Fig. 1. Genomic organization of a small region of the wild-type allele of the *d* locus. The diagram of *dilute* genomic DNA in the region of *Emv-3* integration shows restriction sites (R, *Eco*RI; Pst, *Pst*I) and tissue-specific alternative splicing of the *dilute* transcripts. The direction of transcription for both the *dilute* gene and the *Emv-3* provirus is left to right. Exons are shown as boxes; the sizes of the exons are exaggerated for clarity. The precise size of each exon (in bp) is indicated. Oligodeoxynucleotide primers used for PCR amplification are shown as arrows. Genomic regions that have been sequenced (Strobel *et al.*, 1990; P.K.Seperack, unpublished results) are shown as double lines.

dilute tail. Perhaps the alternative splicing present in the *dilute* tail allows dilute to interact with different structures in melanocytes and neurons. Such a functional difference for *dilute* in melanocytes and neurons is consistent with genetic data obtained for *dilute suppressor* (*dsu*), an unlinked phenotypic suppressor of *dilute* that suppresses the coat color but not the neurological defects associated with d^l alleles (Moore *et al.*, 1988).

The Emv-3 provirus disrupts dilute expression

The effects of *Emv-3* integration on *d* gene expression were determined by Northern analysis of RNAs isolated from B16 (+/+) and S91 (d^v/d^v) melanoma cells and from tissues of wild-type, homozygous d^v and homozygous d^+ mice (Figure 4). Three wild-type *dilute* transcripts, 12, 8 and 7 kb in size, were observed in all wild-type tissues and cells when a *dilute* cDNA probe located 5′ of the *Emv-3* integration site was used as a hybridization probe (Figure 4A). The 7 and 8 kb transcripts are known to result from the alternative use of two different poly(A) sites (Mercer *et al.*, 1991); the structure of the 12 kb transcript has not yet been confirmed, but is also likely to result from alternative poly(A) site usage (J.A.Mercer, unpublished results). The poly(A) sites for the 7 and 8 kb messages (and probably the 12 kb message) are located in 3′ untranslated sequences, indicating that alternative poly(A) site usage would not affect the overall structure of the dilute protein. In some cases, the 7 and 8 kb messages each migrate as doublets (Figure 4A and C, lanes 3 and 7). The 7 and 8 kb doublets may reflect alternative splicing that we have not yet detected or alternative 5′- or 3′-ends.

In S91 (d^v/d^v) melanoma cells and skin from d^v mice the normal messages were either not detected or detected only after extremely long exposures. Instead, an abnormal 7.7 kb transcript was observed (Figure 4A, lanes 2 and 6). In d^v/d^v brain the three wild-type *dilute* transcripts were also observed, albeit at reduced levels relative to wild-type brain (Figure 4C, lanes 3 and 5). Finally, in skin, but not in brain, of homozygous d^+ mice a new 5.5 kb transcript was observed (compare Figure 4A, lanes 7 and 8). Increased levels of the 12 kb message were also observed in d^+ brain compared with wild-type brain (Figure 4A, lane 7). The origin of these abnormal transcripts is discussed below.

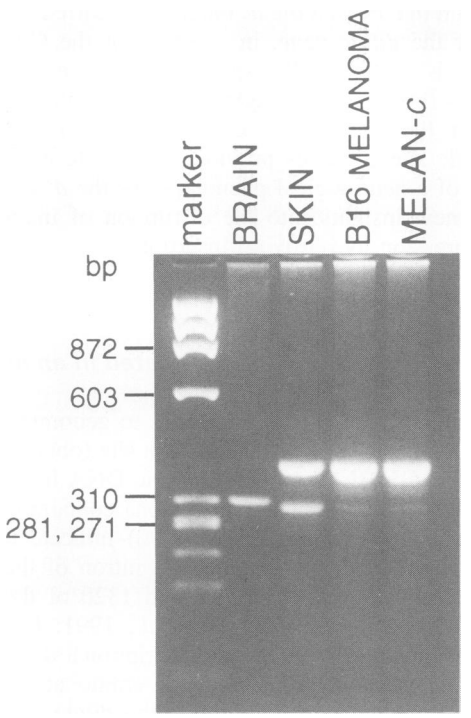

Fig. 2. Alternative splicing of *dilute* transcripts demonstrated by PCR amplification of C57BL/6J tissue and cultured cell line cDNAs using primers 43 and 41 (Figure 1). Marker lane is *Hae*III-cleaved ϕX174 DNA. Ten microliters of a 50 μl reaction was loaded for all but the brain. Only 2 μl of the brain reaction was loaded on the gel because a much larger amount of product was obtained.

Fig. 3. Nucleotide sequence of *dilute* exons (top) with derived amino acid sequence and sequence of a portion of the 2.7 kb genomic *Eco*RI DNA fragment containing exons A and B (bottom).

The 7.7 kb d^v transcript represents a chimeric viral–dilute message

Since *Emv-3* is located within an intron of the *dilute* gene it was possible that the 7.7 kb message identified in d^v melanocyte and skin RNA represents a chimeric message composed of *Emv-3* and *dilute* sequences. To test this possibility, we stripped the Northern blot shown in Figure 4A and rehybridized the blot with a probe from the *Emv-3 env* gene (Chattopadhyay *et al.*, 1980). While the probe hybridized very strongly to multiple transcripts present in B16 and S91 RNA (data not shown), presumably reflecting multiple infections of these cell lines by ecotropic viruses during their long passage *in vivo* and *in vitro*, the probe hybridized only to the 7.7 kb transcript in d^v tissue RNA (Figure 4B, lanes 5 and 6). The smaller transcript seen in all lanes represents background hybridization to 28S ribosomal RNA. The 7.7 kb transcript was not detected in wild-type and revertant tissues, confirming that it contains *Emv-3* sequences.

The 7.7 kb transcript is not large enough to represent the fusion of full-length *Emv-3* and *dilute* transcripts. A more likely origin for this transcript is shown in Figure 5. According to this model, the *dilute* exon C splice donor is joined with the splice acceptor of the *Emv-3 env* gene, with subsequent transcriptional termination in the *Emv-3* 3′ LTR. This model is consistent with all of the data that follow and with the predicted size of a *dilute* message containing the *Emv-3 env* gene and terminating in the *Emv-3* LTR. To determine if this splice actually occurs, we PCR amplified d^v skin cDNA using a primer located in exon A (primer 43, Figure 1) and a primer located in the *Emv-3 env* gene. The resulting amplified fragment was gel purified and sequenced. The amplified fragment was found to contain the exon C-*Emv-3 env* splice, as proposed in Figure 5 (data not shown).

To provide further evidence that the 7.7 kb transcript does not extend beyond the *Emv-3* LTR, the filter shown in Figure 4B was stripped and rehybridized with a *dilute* cDNA probe containing sequences located immediately 3′ of the *Emv-3* integration site. The 7.7 kb d^v transcript was not detected with this probe (Figure 4C, lanes 2 and 6), consistent with the predicted termination of this message in the 3′ *Emv-3* LTR (Figure 5). Additional studies are, however, needed to precisely define the transcription termination site within the viral LTR.

Effects similar to *Emv-3* on gene expression have also been reported to occur when an early transposable element (ETn) integrated into intron 2 of the mouse *Fas antigen* gene (Adachi *et al.*, 1993). As with d^v, the ETn element integrated in the same transcriptional orientation as the *Fas antigen* gene and abnormal splicing and premature transcriptional termination within the ETn LTR was observed. Interestingly, when the ETn element was inserted into an intron of the human *CSF3* gene in a plasmid that carries exon 1, intron 1 and exon 2, *CSF3* expression was greatly reduced only when the ETn element was inserted in the same transcriptional orientation as the *CSF3* gene. Insertion of the ETn element in the reverse orientation had almost no effect on *CSF3* expression (Adachi *et al.*, 1993).

Basis for the normal neurological phenotype of d^v mice

The d^v mutation is one of only a few *d* alleles to be identified that produces a coat color defect in the absence of a neurological defect. An explanation for this phenotype can be obtained from the Northern blot results shown in Figure 4. In homozygous d^v skin and S91 d^v melanoma RNA almost no normal *dilute* transcripts were identified (Figure 4A and C, lanes 2 and 6), while some normal *dilute* transcripts could be identified in homozygous d^v brain RNA (Figure 4A and C, lanes 5). The effects of *Emv-3* on *dilute* expression thus appear to be tissue-

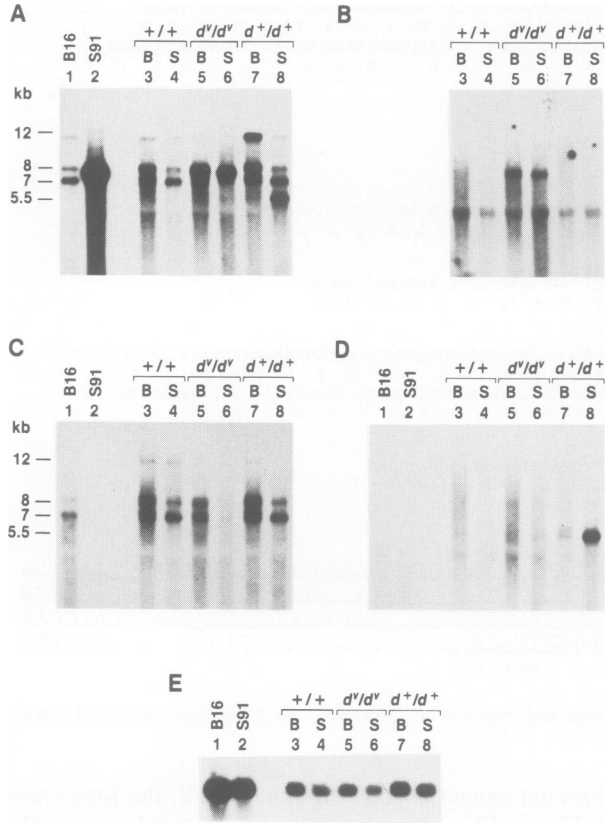

Fig. 4. Northern blot of poly(A)$^+$ RNA from: lane 1, B16 melanoma cells, wild-type at *d*; lane 2, S91 melanoma cells, homozygous for d^v; lane 3, 10-day-old C57BL/6J +/+ brain; lane 4, C57BL/6J +/+ skin; lane 5, 11-day-old DBA/2J d^v/d^v brain; lane 6, DBA/2J d^v/d^v skin; lane 7, 13-day-old DBA/2J d^{+18J}/d^{+18J} brain; lane 8, DBA/2J d^{+18J}/d^{+18J} skin. The same blot was hybridized with: (**A**) a *dilute* cDNA probe from the head region, 5′ of the *Emv-3* integration site; (**B**) an ecotropic MuLV-specific probe from the *env* region (Chattopadhyay *et al.*, 1980); (**C**) a *dilute* cDNA probe from the tail region, 3′ of the *Emv-3* integration site; (**D**) a probe from the intron 5′ of the *Emv-3* integration site; (**E**) a glyceraldehyde-3-phosphate dehydrogenase cDNA probe (Tso *et al.*, 1985) to detect variation in RNA loading. The probes used in (A–D) are diagrammed in Figure 5 and described in detail in Materials and methods. Autoradiographic exposure times: (A) 48; (B) 18; (C) 48; (D) 72; (E) 12 h.

specific. Presumably, enough normal *dilute* expression occurs in d^v brain so that normal neurological function is observed.

A characteristic of retroviral RNA processing is the presence of read-through transcripts that extend into neighboring cellular sequences. Such read-through transcripts have been shown to constitute up to 15% of viral RNA in avian leukosis virus-infected cells (Swain and Coffin, 1993). This inefficiency in 3′-end processing by viral LTRs may explain why some normal *dilute* transcripts are detected in d^v brain RNA. It is not clear, however, why this 3′-end processing should be more efficient in skin (melanocytes) than in brain.

One other ecotropic provirus, *Mov13*, has been shown to have a tissue-specific effect on gene expression (Kratochwill *et al.*, 1989). However, the mechanism of this tissue-specific effect seems to be quite different from that observed for *dilute*. In this case *Mov13* is integrated in the first intron of the *α1(I) collagen* gene, in the opposite transcriptional orientation. In some tissues collagen transcripts are never detected, while in other tissues wild-type levels of an apparently normal sized collagen RNA are produced. It is hypothesized that collagen expression in different tissues requires the use of different *cis* regulatory sequences, which in turn are affected to a different degree by *Mov13*. Thus in some tissues collagen transcription is never initiated, while in others collagen is expressed normally and the provirus is subsequently removed by splicing.

Origin of the 5.5 kb d^+ message

In d^+ skin the 5.5 kb *dilute* transcript was detected with a *dilute* cDNA probe located 5′ of *Emv-3* (Figure 4A, lane 8), but not with a *dilute* cDNA probe located 3′ of *Emv-3* (Figure 4C, lane 8). The size of this transcript is consistent with transcriptional read-through from *dilute* exon C into the *Emv-3* LTR and subsequent transcriptional termination in the *Emv-3* LTR (Figure 5). If this were the case, the probe from the *dilute* intron located immediately 5′ of *Emv-3* (Figure 5, top) should hybridize to the 5.5 kb message detected in d^+ skin RNA. As can be seen in Figure 4D lane 8, the 5.5 kb transcript was detected with the *Emv-3* intron probe.

It is interesting to note that we have not identified any

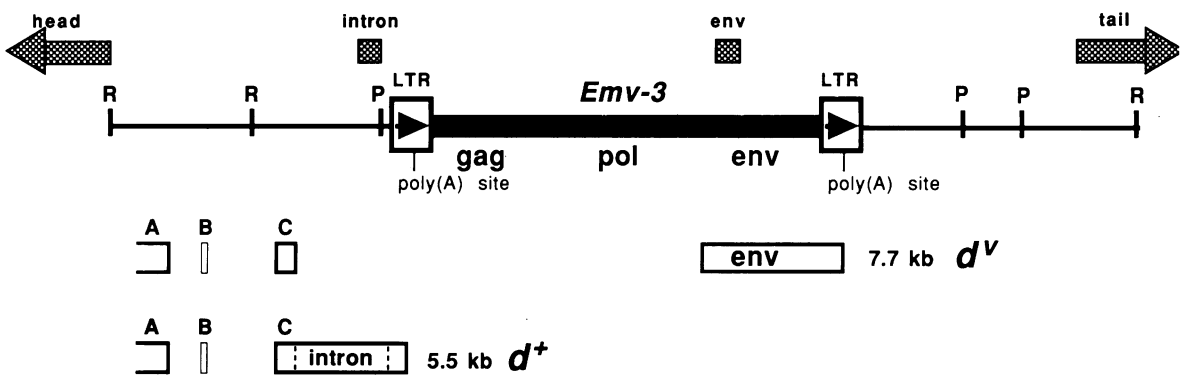

Fig. 5. Genomic organization of a small region of the d^v allele. The diagram (similar to Figure 1) shows an 18 kb genomic region of the d^v allele containing the proviral integration site with proposed abnormal transcript structures for the d^v and d^+ alleles. There is ~4 kb of *dilute* cDNA sequence located 5′ of the region shown. The locations of probes discussed in the text that were hybridized to the Northern blots are shown above the genomic restriction map. Restriction sites: R, *Eco*RI; P, *Pst*I.

dilute transcripts that terminate in the 5′ *Emv-3* LTR nor have we identified any *dilute* transcripts that initiate in the 3′ *Emv-3* LTR, despite the fact that both LTRs have an identical sequence. This is consistent with previous studies indicating that 5′ and 3′ LTRs are functionally non-equivalent (Herman and Coffin, 1986).

The Emv-3 LTR present in d^+ revertant DNA also produces a tissue-specific effect on dilute expression

Perhaps the most interesting and unexpected finding of this study is the tissue-specific effect on *dilute* expression we observed for the solo *Emv-3* LTR present in d^+ revertant DNA. This effect can be seen in the high levels of the truncated 5.5 kb *dilute* transcript expressed in d^+ skin RNA compared with the very low levels of this transcript expressed in d^+ brain RNA (compare Figure 4A, lanes 7 and 8). Unlike in d^v skin, however, enough normal *dilute* transcripts are still made in d^+ skin to allow for a normal coat color phenotype. The mouse genome is thought to contain a number of solo viral LTRs that are likely to be remnants of previous viral excision events. It is easy to predict that these solo LTRs would have a profound effect on gene expression when they are located in exon sequences. However, when a solo viral LTR is located in an intron sequence, it might be assumed that it will have little or no effect on gene expression. The results presented here indicate that this may not always be the case and that the effects of these solo viral LTRs on gene expression may be tissue-specific. The effects of proviral integration and solo viral LTRs on gene expression can therefore be ascertained only after several different tissues are examined.

Materials and methods

Mice

The C57BL/6J, DBA/2J and DBA/2J-d^{+18J}/d^{+18J} mice are maintained at the NCI–Frederick Cancer Research and Development Center.

PCR amplification and sequencing of dilute cDNA

First-strand cDNA was synthesized from 1 μg total RNA with random oligonucleotide primers and AMV reverse transcriptase (Boehringer Mannheim). The reaction was diluted 10-fold with 10 mM Tris–HCl, pH 8, 1 mM EDTA and heated to 100°C for 10 min. PCR amplification was performed on a 0.1 μl aliquot of the cDNA (~0.5 ng) using reagents from the GeneAmp kit (Perkin-Elmer-Cetus) and primers 43 and 41 (shown in Figure 1), with temperature steps of 94°C denaturation, 45°C annealing and 72°C extension for 1 min each for 30 cycles. PCR products were fractionated on 3% NuSieve/1% SeaKem agarose (FMC) gels in TBE buffer.

PCR fragments to be sequenced were excised from the gel. The gel slice was melted and 1 μl of the gel was reamplified under the conditions described above, except that the concentration of primer 41 was reduced 50-fold, allowing for production of an excess of single-stranded cDNA (Gyllensten and Erlich, 1988). The products were purified by spin dialysis (Centricon 30 columns; Amicon) and dideoxy sequencing was performed with the Sequenase (USB) kit with ^{32}P-end-labeled oligonucleotide 41 (Figure 1) as primer.

RNA blots

Total tissue RNA was prepared from mice of the genotypes described in Figure 4 or from B16 and S91 melanoma cells using RNAzol (Cinna/Biotecx) (Chomczynski and Sacchi, 1987) and poly(A) RNA was purified with oligo(dT) spin columns (mRNA Purification Kit; Pharmacia). Two micrograms/lane was fractionated on 1% agarose–formaldehyde gels and transferred to nylon membranes (Cuno). The blots were probed with *dilute* cDNA sequences (Mercer *et al.*, 1991) from the 5′ head region (nucleotides 526–993), the tail region 3′ of the *Emv-3* integration site (nucleotides 4110–5635), a 340 bp fragment of the *dilute* intron 5′ of the *Emv-3* integration site (Hutchison *et al.*, 1984), an ecotropic murine leukemia virus *env* probe (Chattopadhyay *et al.*, 1980) and a glyceraldehyde-3-phosphate dehydrogenase cDNA probe (Tso *et al.*, 1985) to detect variation in loading. Probes were labeled with ^{32}P (Feinberg and Vogelstein, 1983) and hybridized to the membranes by the method of Church and Gilbert (1984).

Acknowledgements

We thank Linda Cleveland, Debbie Swing, and Janet Koslovsky for technical assistance. This work was supported in part by the National Cancer Institute, DHHS, under contract NO1-CO-46000 with ABL. J.A.M. was supported by a Leukemia Society of America Fellowship and NIH grant R01 NS30848.

References

Adachi,M., Watanabe-Fukunaga,R. and Nagata,S. (1993) Aberrant transcription caused by the insertion of an early transposable element in an intron of the Fas antigen gene of *lpr*. *Proc. Natl Acad. Sci. USA*, **90**, 1756–1760.

Bennett,D.C., Cooper, P.J., Dexter,T.J., Devlin,L.M., Heasman,J. and Nester,B. (1989) Cloned mouse melanocyte lines carrying the germline mutations albino and brown. *Development*, **105**, 379–385.

Chattopadhyay,S.K., Lander,M.R., Rands,E. and Lowy,D.R. (1980) Structure of endogenous murine leukemia virus DNA in mouse genomes. *Proc. Natl Acad. Sci. USA*, **77**, 5774–5778.

Cheney,R.E. and Mooseker,M.S. (1992) Unconventional myosins. *Curr. Opin. Cell Biol.*, **4**, 27–35.

Cheney,R.E., O'Shea,M.K., Heuser,J.E., Coelho,M.V., Wolenski,J.S., Espreafico,E.M., Forscher,P., Larson,R.E. and Mooseker,M.S. (1993) Brain myosin-V is a two-headed unconventional myosin with motor activity. *Cell*, **75**, 13–23.

Chomczynski,P. and Sacchi,N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction. *Anal. Biochem.*, **162**, 156–159.

Church,G.M. and Gilbert,W. (1984) Genomic sequencing. *Proc. Natl Acad. Sci.USA*, **81**, 1991–1995.

Copeland,N.G., Hutchison,K.W. and Jenkins,N.A. (1983) Excision of the DBA ecotropic provirus in *dilute* coat-color revertants of mice occurs by homologous recombination involving the viral LTRs. *Cell*, **33**, 379–387.

Espreafico,E.M., Cheney,R.E., Matteoli,M., Nascimento,A.A., De Camilli,P.V., Larson,R.E. and Mooseker,M.S. (1992) Primary structure and cellular localization of chicken brain myosin-V (p190), an unconventional myosin with calmodulin light chains. *J. Cell Biol.*, **119**, 1541–1557.

Feinberg,A.P. and Vogelstein,B. (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.*, **132**, 6–13.

Fink,G., Farabaugh,P., Roeder,G. and Chaleff,D. (1981) Transposable elements (Ty) in yeast. *Cold Spring Harbor Symp. Quant. Biol.*, **45**, 575–580.

Gyllensten,U.B. and Erlich,H.A. (1988) Generation of single-stranded DNA by the polymerase chain reaction and its application to direct sequencing of the HLA-DQA locus. *Proc. Natl Acad. Sci. USA*, **85**, 7652–7656.

Haarer,B.K., Petzold,A., Lillie,S.H. and Brown,S.S. (1994) Identification of MYO4, a second class V myosin gene in yeast. *J. Cell Sci.*, **107**, 1055–1064.

Herman,S.A. and Coffin,J.M. (1986) Differential transcription from the long terminal repeats of integrated avian leukosis virus DNA. *J. Virol.*, **60**, 497–505.

Hutchison,K.W., Copeland,N.G. and Jenkins,N.A. (1984) Dilute-coat-color locus of mice: nucleotide sequence analysis of the d^{+2J} and d^{+Ha} revertant alleles. *Mol. Cell. Biol.*, **4**, 2899–2904.

Jenkins,N.A., Copeland,N.G., Taylor,B.A. and Lee,B.K. (1981) Dilute (*d*) coat colour mutation of DBA/2J mice is associated with the site of integration of an ecotropic MuLV genome. *Nature*, **293**, 370–374.

Jenkins,N.A., Strobel,M.C., Seperack,P.K., Kingsley,D.M., Moore,K.J., Mercer,J.A., Russell,L.B. and Copeland,N.G. (1989) A retroviral insertion in the dilute (*d*) locus provides molecular access to this region of mouse chromosome 9. *Prog. Nucleic Acid Res. Mol. Biol.*, **36**, 207–220

Johnston,G.C., Prendergast,J.A. and Singer,R.A. (1991) The *Saccharomyces cerevisiae* MYO2 gene encodes an essential myosin for vectorial transport of vesicles. *J. Cell Biol.*, **113**, 539–551.

Kratochwill,K., von der Mark,K., Kollar,E.J., Jaenisch,R., Mooslehner, K., Schwarz,M., Haase,K., Gmachl,I. and Harbers,K. (1989) Retrovirus-induced insertional mutation in Mov13 mice affects collagen I expression in a tissue-specific manner. *Cell*, **57**, 807–816.

Mercer,J.A., Seperack,P.K., Strobel,M.C., Copeland,N.G. and Jenkins, N.A. (1991) Novel myosin heavy chain encoded by murine *dilute* coat colour locus. *Nature*, **349**, 709–713.

Moore,K.J., Seperack,P.K., Strobel,M.C., Swing,D.A., Copeland,N.G. and Jenkins,N.A. (1988) Dilute suppressor *dsu* acts semidominantly to suppress the coat color phenotype of a deletion mutation, d^{l20J}, of the murine dilute locus. *Proc. Natl Acad. Sci. USA*, **85**, 8131–8135.

Pascolo,S., Ghazvini,M., Boyer,J., Colleaux,L., Thierry,A. and Dujon,B. (1992) The sequence of a 9.3 kb segment located on the left arm of the yeast chromosome XI reveals five open reading frames including the CCE1 gene and putative products related to MYO2 and to the ribosomal protein L10. *Yeast*, **8**, 987–995.

Rinchik,E.M., Russell,L.B., Copeland,N.G. and Jenkins,N.A. (1986) Molecular genetic analysis of the dilute–short ear (*d–se*) region of the mouse. *Genetics*, **112**, 321–342.

Russell,E.S. (1949) A quantitative histological study of the pigment found in the coat-color mutants of the mouse. IV. The nature of the effects of genic substitution in five major allelic series. *Genetics*, **34**, 146–166.

Russell,L.B. (1971) Definition of functional units in a small chromosomal segment of the mouse and its use in interpreting the nature of radiation-induced mutations. *Mutat. Res.*, **11**, 107–123.

Russell,W.L. (1951) X-ray-induced mutations in mice. *Cold Spring Harbor Symp. Quant. Biol.*, **16**, 327–336.

Sanders,G., Lichte,B., Meyer,H.E. and Kilimann,M.W. (1992) cDNA encoding the chicken ortholog of the mouse dilute gene product. Sequence comparison reveals a myosin I subfamily with conserved C-terminal tail. *FEBS Lett.*, **26**, 295–298.

Schlager,G. and Dickie,M.M. (1971) Natural mutation rates in the house mouse. *Mutat. Res.*, **11**, 89–96.

Searle,A.G. (1952) A lethal allele of dilute in the house mouse. *Heredity*, **6**, 395–401.

Seperack,P.K., Strobel,M.C., Corrow,D.J., Jenkins,N.A. and Copeland, N.G. (1988) Somatic and germ-line reverse mutation rates of the retrovirus-induced dilute coat-color mutation of DBA mice. *Proc. Natl Acad. Sci. USA*, **85**, 189–192.

Silvers,W.K. (1979) *Coat Colors of Mice*. Springer-Verlag, New York, NY, pp. 1–379.

Strobel,M.C., Seperack,P.K., Copeland,N.G. and Jenkins,N.A. (1990) Molecular analysis of two mouse dilute locus deletion mutations: spontaneous dilute lethal 20J and radiation-induced dilute prenatal lethal Aa2 alleles. *Mol. Cell. Biol.*, **10**, 501–509.

Swain,A. and Coffin,J.M. (1993) Influence of sequences in the long terminal repeat and flanking cell DNA on polyadenylation of retroviral transcripts. *J. Virol.*, **67**, 6265–6269.

Tso,J.Y., Sun,X.H., Kao,T.H., Reece,K.S. and Wu,R. (1985) Isolation and characterization of rat and human glyceraldehyde-3-phosphate cDNA: genomic complexity and molecular evolution of the gene. *Nucleic Acids Res.*, **13**, 2485–2502.

Zachar,Z., Davison,D., Garza,D. and Bingham,P.M. (1985) A detailed developmental and structural study of the transcriptional effects of insertion of the Copia transposon into the white locus of *Drosophila melanogaster*. *Genetics*, **111**, 495–515.

Received on January 30, 1995; revised on March 1, 1995