Dsi-1, a Region with Frequent Proviral Insertions in Moloney Murine Leukemia Virus-Induced Rat Thymomas

S. VIJAYA,^{1†} D. L. STEFFEN,^{1*} C. KOZAK,² AND H. L. ROBINSON¹

Worcester Foundation for Experimental Biology, Shrewsbury, Massachusetts 01545,¹ and National Institute of Allergy and Infectious Diseases, Bethesda, Maryland 20892²

Received 15 August 1986/Accepted 16 December 1986

Dsi-1 is a region of chromosomal DNA that underwent proviral insertion in 3 of 24 Moloney murine leukemia virus-induced rat thymomas. In one of these tumors, a provirus is also integrated adjacent to the proto-oncogene *c-myc*. The proviruses in *Dsi-1* have been characterized and appear to be complete. The proviruses were located within a 2-kilobase region that contained four prominent DNase I-hypersensitive sites. These hypersensitive sites were observed in Moloney murine leukemia virus-induced thymomas but not in NRK cells. The region of *Dsi-1* immediately 3' to the insertions cross-hybridized with human and chicken DNA, indicating that it contains highly conserved sequences. No evidence could be found for the expression of this highly conserved region. *Dsi-1* was mapped to mouse chromosome 4. This location demonstrates that *Dsi-1* is different from 16 of the known proto-oncogenes (*c-abl*, *c-erbA*, *c-erbB*, *c-ets-1*, *c-ets-2*, *c-fes*, *c-fos*, *c-myb*, *c-myc*, *c-raf*, *A-raf*, *c-Ha-ras*, *c-Ki-ras*, *N-ras*, *c-sis*, and *c-src*) and 12 cellular regions of tumor-associated integrations in retrovirus-induced tumors (*c-erbB*, *Fis-1*, *int-1*, *int-2*, *Mis-1/pvt-1*, *Mlvi-1*, *Mlvi-2*, *c-mos*, *c-myb*, *c-myc*, *Pim-1*, and *c-Ha-ras*). Hybridization experiments indicated that *Dsi-1* is probably different from five additional proto-oncogenes (*c-fgr*, *c-fms*, *c-mos*, *neu*, and *c-yes*) and from two additional frequent integration regions (*lck* and *Mlvi-3*).

Moloney murine leukemia virus (MoMLV) is a nonacute retrovirus which induces thymic lymphomas in rats and mice. Nonacute retroviruses lack transduced oncogenes and induce tumors only after a latent period of several months. Some tumors induced by nonacute viruses contain a provirus within or adjacent to a cellular proto-oncogene (7, 12, 19, 30, 32, 34-37, 41, 45, 52, 63, 64). In these cases, the insertion is believed to activate the oncogenic potential of the proto-oncogene. In other cases, tumor-associated proviral insertions have occurred in regions not associated with a proto-oncogene (9, 13, 16, 17, 28, 32, 39, 44, 49, 51, 55-57, 61). These regions, termed frequent integration regions, are associated with tumorigenesis because tumors contain insertions in these regions much more frequently than would be expected by chance. One region, *int-1*, initially identified as a frequent integration region, was subsequently shown to be a proto-oncogene (4). Whether the remaining frequent integration regions are also proto-oncogenes remains to be determined. Prior to this study, six proto-oncogenes and 12 frequent integration regions had been identified as sites of tumor-associated integrations. We report here the identification and characterization of *Dsi-1*, yet another frequent integration region, which we observed in thymic lymphomas induced in rats by MoMLV.

MATERIALS AND METHODS

Cell lines and virus. The RCC2-1M line of T cells was established from an MoMLV-induced rat thymoma (60). RCC2-1M cells were grown in RPMI 1640 medium supplemented with 10% heat-inactivated fetal calf serum and 5×10^{-5} M 2-mercaptoethanol. NRK, a continuous line of rat kidney fibroblast cells, was grown in Dulbecco modified Eagle medium supplemented with 10% calf serum. The

hamster-mouse hybrid somatic cell lines used for chromosomal mapping have been described (23).

Tumors. Lymphomas were induced by inoculating Fisher rats with MoMLV as described previously (52).

Molecular cloning of *Dsi-1*. The locus *Dsi-1* (David Steffen integration 1) was molecularly cloned as two *Hind*III fragments in lambda Charon 4a from RT10-2, an MoMLV-induced thymoma which had five integrated proviruses (60). *Dsi-1* was mapped for restriction endonuclease sites and assigned the same orientation as the provirus in this region in the tumor RT10-2. Repetitive sequences in *Dsi-1* were identified by hybridizing blots of restricted DNA with nick-translated rat DNA (specific activity, $>10^8$ cpm/ μ g of DNA). *Hind*III to *Kpn*I (ds1-HK) and *Sma*I to *Eco*RI (ds1-SR) fragments which were free of repeat sequences were subcloned into pUC18 or the Bluescribe vector (Vector Cloning Systems).

Southern blots. Isolation of DNA from tumors, restriction enzyme digestions, separation of digested DNA on agarose gels, transfer to nitrocellulose filters (Schleicher & Schuell BA 85), and hybridization to nick-translated probes were done as described (33). Fragment sizes were estimated by using *Hind*III-digested lambda DNA or 32 P-end-labeled 1-kilobase (kb) ladder fragments (Bethesda Research Laboratories) as size markers.

Northern blots. Northern analyses were performed on cytoplasmic RNA prepared from cells and tissues as follows. Cells were harvested by trypsinization and washed once with cold phosphate-buffered saline containing 2 mM vanadyl ribonucleoside complex (VRC) (Bethesda Research Laboratories). The cells were then suspended in cold hypotonic buffer containing 10 mM Tris, pH 7.5, 10 mM MgCl₂, 10 mM NaCl, and 20 mM VRC, and 10% Nonidet P-40 was added to a final concentration of 0.5 to 1%. As soon as the suspension cleared, nuclei were pelleted by centrifugation for 3 min at $700 \times g$. The supernatant was poured into a tube containing sufficient *para*-aminosalicylate (PAS), so-

* Corresponding author.

† Present address: National Institute of Allergy and Infectious Diseases, Bethesda, MD 20892.

dium dodecyl sulfate (SDS), and EDTA to bring the final solution to 6% PAS, 1% SDS, and 1 mM EDTA. The solution was gently warmed to dissolve the PAS and extracted twice with phenol-chloroform-isoamyl alcohol (25:24:1) and once with chloroform-isoamyl alcohol (24:1). Sodium acetate (pH 6.0) was added to a final concentration of 0.25 M, followed by 2.5 volumes of 100% ethanol. RNA was allowed to precipitate for at least 2 h at -70°C and collected by centrifugation. Tissues were homogenized in phosphate-buffered saline containing 2 mM VRC, washed once with the same solution, suspended in hypotonic buffer containing 20 mM VRC, and lysed by the addition of Nonidet P-40 as described above. Polyadenylated [poly(A)⁺] RNA was prepared by passing the total RNA over an oligo(dT) column (Collaborative Research) as recommended by the manufacturer. Poly(A)⁺ RNA was ethanol precipitated, collected by centrifugation, and stored as an ethanol suspension. A portion of this suspension was dissolved in water, and the amount of RNA was estimated by determining the optical density at 260 nm. RNA to be used in Northern analyses was pelleted from the ethanol and redissolved in MOPS-acetate buffer (20 mM MOPS [3-(N-morpholino)propanesulfonic acid], 5 mM sodium acetate, and 1 mM EDTA, pH 7.0) containing 50% deionized formamide and 2.2 M formaldehyde.

Northern blot analyses were carried out in 1% agarose gels containing 2.2 M formaldehyde. Electrophoresis was done in recirculated MOPS-acetate buffer at 50 V for approximately 8 h (until the bromophenol blue dye moved 12 to 16 cm). The RNA was transferred to nitrocellulose filters by blotting overnight in the presence of $20\times$ SSC ($1\times$ SSC is 0.15 M NaCl, 0.015 M sodium citrate, pH 7.0). Filters were baked under vacuum for 2 h at 80°C . Prehybridization was performed for 4 h at 60°C in 50% formamide–50 mM sodium phosphate (pH 6.5)–0.8 M NaCl–1 mM EDTA–0.05% bovine serum albumin–0.05% Ficoll type 400–0.05% polyvinylpyrrolidone 360–250 μg of denatured salmon sperm DNA per ml–500 μg of yeast RNA per ml. Hybridization was done for 16 to 18 h at 60 to 65°C in prehybridization buffer plus approximately 10^6 cpm of [^{32}P]RNA probe per ml. Riboprobes were generated by transcription from Bluescribe vectors with T3 or T7 polymerases by the protocol in Promega Biotec Technical Bulletin 002. Filters were washed two to four times for 20 min each in 50 mM NaCl–20 mM sodium phosphate (pH 6.5)–1 mM EDTA–0.1% SDS at the same temperature used for hybridization. RNAs hybridizing with the probe were visualized by autoradiography with or without intensifying screens. The specificity of hybridization bands was tested by incubating filters with 2 μg of RNase A per ml in $2\times$ SSC for 15 min at room temperature, followed by washing the filters in $2\times$ SSC–0.1% SDS for 30 min at 50°C . Rat rRNAs (28S, 4,700 nucleotides; 18S, 1,900 nucleotides) were used as size markers.

RNase protection. Poly(A)⁺ RNA (5 μg) from RCC2-1M or NRK cells was hybridized to 3×10^5 cpm of [^{32}P]RNA transcripts from each strand of the ds1-SR fragment. Hybridization was carried out in a total volume of 30 μl for at least 8 h at either high stringency (40 mM PIPES [piperazine-N,N'-bis(2-ethanesulfonic acid)]-acetate, pH 6.7, 0.4 M sodium chloride, 1 mM EDTA, 80% formamide) at 45°C or low stringency (10 mM MOPS-acetate, pH 7.0, 0.4 M NaCl, 1 mM EDTA, 50% formamide) at 37°C . Hybridization was performed in a total volume of 30 μl for at least 8 h. Following hybridization, 300 μl of RNase digestion buffer (10 mM Tris, pH 7.5, 5 mM EDTA, 300 mM NaCl) containing 40 μg of RNase A and 2 of RNase T₁ per ml was added

FIG. 1. Mapping of thymoma-associated insertions in *Dsi-1* in thymomas RT2-2 and RT13-1. (A) Schematic indicating restriction endonuclease sites in *Dsi-1* and MoMLV proviral DNA, sites of insertion of MoMLV in *Dsi-1* (Y), hypersensitive sites in *Dsi-1* ($\oplus$), and probes for *Dsi-1* (heavy black lines). (B) Autoradiographs of Southern blots of normal and tumor-derived rat DNAs. All DNAs were digested with *Eco*RI and then with the indicated second enzyme. Hybridization was with the ds1-SR probe. Size markers are given (in kilobases) for *Hind*III-digested lambda DNA. H, *Hind*III; K, *Kpn*I; L, *Sal*I; R, *Eco*RI; S, *Sma*I; X, *Xho*I.

to the reaction mixture and incubated at 30°C for 1 h. RNase digestion was terminated by adding 20 μl of 10% SDS and 50 μg of proteinase K, followed by additional incubation at 37°C for 15 min. The reaction mixture was then extracted once with phenol-chloroform and precipitated by the addition of 1 ml of ethanol. The precipitate was dissolved in loading buffer and electrophoresed in 5% acrylamide gels containing 8 M urea for approximately 90 min at 200 V in 89 mM Tris base–89 mM boric acid–2 mM EDTA, pH 8.0. The gels were exposed to Kodak XAR-5 film for 4 to 12 h.

Tests for homology between *Dsi-1* and proto-oncogenes or frequent integration regions. To test for homology between two regions, the molecular clone of one region was digested with the restriction endonuclease required to release it from its vector, fractionated by gel electrophoresis, and hybridized to a radiolabeled probe for the other region. For *c-mos*, *Mis-1*, and *Fis-1*, the probes pMS1 (40), pMo-1c (28), and p1.8 (49) were hybridized to clones containing the 13.5 kb of *Dsi-1* sequences bounded by *Hind*III sites on either side of the proviral integrations (Fig. 1). Lambda clones containing *Mlvi-1*, *Mlvi-2*, and *Mlvi-3* (55–57) were hybridized with the *Dsi-1* probe, ds1-SR. For *Pim-1*, a clone containing the sequences of mouse *Pim-1* extending from -9.5 to $+1.3$ kb on the *Pim-1* map and the 1.1-kb PrA subclone (9) were used as probes for hybridization with the 13.5 kb of cloned rat *Dsi-1* sequences. The *Pim-1* sequences were also hybridized to the ds1-SR probe. To test for *src*-related sequences, the 558-base-pair *Bam*HI fragment encoding the kinase domain of *v-erbB* (59) was hybridized to the 13.5 kb of cloned rat *Dsi-1* sequences.

The homologies between the *v-erbB* probe and *lck* or *Dsrc4* were determined by comparing the published sequences of these genes (32, 62, 65). This comparison was

done with the ALIGN command of the IFIND program on the BIONET national computer resource. Default settings (window = 20, word size = 3, density = less, gap penalty = 4) were used.

RESULTS

Most MoMLV-induced thymic lymphomas in Fischer rats contain proviruses in at least 10 to 20 different cellular regions. In a series of 27 lymphomas, the tumor with the smallest number of insertions, RT10-2, contained five proviruses. We molecularly cloned these five proviruses along with flanking cellular sequences (60). We have previously shown that one of these proviruses is inserted near the proto-oncogene *c-myc* (52). The remaining four proviruses are inserted in regions which have not been correlated with known genes. These regions have been designated *Dsi-1*, *dsi-2*, *dsi-3*, and *dsi-4* (60). The purpose of the present study was to determine whether *Dsi-1* represents a frequent integration region in MoMLV-induced lymphomas.

Identification of proviral insertions in *Dsi-1* in two additional thymomas. To determine whether insertions in *Dsi-1* occurred in other MoMLV-induced thymomas, DNAs from 24 thymomas were digested with the restriction endonuclease *Hind*III and analyzed by the Southern technique for sequences hybridizing with a probe for *Dsi-1* sequences, designated ds1-SR (Fig. 1). All 24 of the tumor DNAs as well as control DNA from the liver of a normal rat revealed the expected 13.5-kb *Hind*III fragment of *Dsi-1*. Two tumors, designated RT2-2 and RT13-1, revealed additional fragments of roughly 8.4 and 9.6 kb, respectively. These novel fragments were determined to result from proviral insertions by examining the sizes of fragments generated by double digestions with an enzyme that does not cleave within the viral genome (*Eco*RI) and enzymes which cleave once within the MoMLV genome but not between the 3' *Eco*RI site in *Dsi-1* and the sites of insertion. Southern blots of the products of these digestions were hybridized with the ds1-SR probe for sequences immediately upstream of the 3' *Eco*RI site in *Dsi-1* (Fig. 1). For example, in tumor RT2-2, the sizes of novel fragments positioned proviral *Xho*I, *Sal*I, and *Hind*III sites 8.2, 6.1, and 4.85 kb upstream of the 3' *Eco*RI site in

TABLE 1. Localization of *Dsi-1* to mouse chromosome 4

Mouse chromosome	No. of somatic cell hybrids ^a (chromosome/ <i>Dsi-1</i>)				% Discordant ^b
	+ / +	- / -	+ / -	- / +	
1	4	4	2	3	39
2	3	6	2	1	25
3	2	4	1	2	27
4	6	7	0	0	0
5	2	6	5	0	38
6	4	4	2	3	38
7	7	3	0	4	29
8	3	6	3	1	31
9	4	6	3	1	29
10	0	6	7	1	57
11	0	6	4	0	40
12	2	1	1	4	63
13	2	4	1	1	25
14	1	5	5	2	54
15	3	0	0	5	63
16	2	3	1	2	38
17	2	1	1	5	67
18	3	3	1	2	33
19	4	5	3	2	46
X	4	3	2	4	46

^a The total number of somatic-cell hybrids is different for different mouse chromosomes because for some hybrids, the presence of mouse chromosomes was determined by using marker loci (which reveals only some chromosomes) rather than by cytogenetics (which reveals all chromosomes).

^b Calculated as the percentage of the total number of hybrids examined for a chromosome which either have *Dsi-1* and lack the indicated chromosome or lack *Dsi-1* and have the chromosome. Such discordancies are evidence that *Dsi-1* is not located on that chromosome.

Dsi-1. These positions indicate that this insertion lies 1.4 kb upstream of the *Eco*RI site in *Dsi-1* and in the same orientation as the insertion in tumor RT10-2.

The results of these experiments confirmed that the perturbations in *Dsi-1* in RT2-2 and RT13-1 were due to proviral insertions. In addition, these results revealed that the transcriptional orientation of the proviruses in RT2-2 and RT13-1 was the same as that of the molecularly cloned provirus in *Dsi-1* in RT10-2. Finally, the results indicate that the three thymoma-associated insertions reside within a 2-kb region of *Dsi-1*.

Evolutionary conservation of *Dsi-1*. To determine whether the nucleotide sequence of the *Dsi-1* region has been evolutionarily conserved, the ds1-HK and ds1-SR probes for unique *Dsi-1* sequences were hybridized with Southern blots of *Hind*III-digested rat, mouse, hamster, human, and chicken DNAs. Hybridization with the ds1-HK probe did not reveal easily detected bands with human and chicken DNAs (Fig. 2A; data not shown). In contrast, hybridization with the ds1-SR probe revealed strongly hybridizing fragments in hamster and human DNAs (Fig. 2B) and two weaker yet clearly detectable fragments in chicken DNA (data not shown). Since the hybridizations and washes were carried out under conditions of moderately high stringency (67°C, 3× to 6× SSC), detection of hybridization indicates a significant extent of homology. Thus, these results indicate a relatively strong evolutionary conservation of the ds1-SR region of *Dsi-1*, but less evolutionary conservation of the ds1-HK region.

Location of *Dsi-1* on mouse chromosome 4. Since the hamster and mouse *Hind*III fragments of *Dsi-1* differed in size (Fig. 2B), we were able to identify the mouse chromo-

FIG. 2. *Dsi-1* is evolutionarily conserved. Hybridization of Southern blots of *Hind*III-digested DNAs from the indicated species with the ds1-HK probe (A) and the ds1-SR probe (B). Hybridization and washes were done at 67°C in 3× to 6× SSC. Sizes are indicated (in kilobases).

TABLE 2. Relationship between *Dsi-1*, proto-oncogenes, and frequent integration regions

Sequence ^a	Cross-hybridization with <i>Dsi-1</i> ^b	Mouse chromosomal location	Identity with <i>Dsi-1</i> ^c	Chromosomal location reference
<i>c-abl</i>	(-)	2	-	15
<i>Dsi-1</i> *	+	4	+	Table 1
<i>c-erbA</i>	nt	11	-	66
<i>c-erbB</i> *	(-)	11	-	50
<i>c-ets-1</i>	nt	9	-	63
<i>c-ets-2</i>	nt	16	-	63
<i>c-fes/c-fps</i>	(-)	7	-	3
<i>c-fgr</i>	(-)		(-)	
<i>fim-1</i> *	nt		?	
<i>fim-2</i> *	nt		?	
<i>Fis-1</i> *	-	7	-	49
<i>c-fms</i>	(-)		(-)	
<i>c-fos</i>	nt	12	-	10
<i>int-1</i> *	nt	15	-	38
<i>int-2</i> *	nt	7	-	42
<i>int41</i> *	nt		?	
<i>int-H</i> *	nt		?	
<i>c-kit</i>	nt		?	
<i>lck</i> *	(-)	4	(-)	31
<i>mas</i>	nt		?	
<i>met</i>	nt		?	
<i>Mis-1/pvt-1</i> *	-	15	-	16, 22
<i>MLvi-1</i> *	-	15	-	26
<i>MLvi-2</i> *	-	15	-	57
<i>MLvi-3</i> *	-		(-)	
<i>c-mos</i> *	-	4	(-)	54
<i>c-myb</i> *	nt	10	-	48
<i>c-myc</i> *	nt	15	-	8
<i>N-myc</i>	nt		?	
<i>neu</i>	(-)		(-)	
<i>Pim-1</i> *	-	17	-	20
<i>c-raf/c-mht/c-mil</i>	nt	6	-	24
<i>A-raf</i>	nt	X	-	21
<i>c-Ha-ras</i> *	nt	7	-	48
<i>c-Ki-ras</i>	nt	6	-	5
<i>N-ras</i>	nt	3	-	47
<i>c-rel</i>	nt		?	
<i>c-ros</i>	(-)		?	
<i>c-sis</i>	nt	15	-	25
<i>c-ski</i>	nt		?	
<i>c-src</i>	(-)	2	-	3
<i>c-yes</i>	(-)		(-)	

^a Regions where proviral integrations have been observed are marked (*).^b Symbols: +, hybridization; -, no hybridization; (-), hybridization to *v-erbB* tested; nt, not tested.^c Symbols: +, identical; - and (-), not identical, based on chromosomal location [-] or on cross-hybridization results only [(-)]; ?, identity uncertain.

some that carries *Dsi-1* by examining a series of hamster-mouse somatic cell hybrids containing different subsets of mouse chromosomes. DNAs from the hybrids were digested with *Hind*III and analyzed by the Southern technique for fragments hybridizing with the ds1-SR probe. For each of the 20 mouse chromosomes, we were then able to ask whether hybrids containing or lacking a specific mouse chromosome contained mouse *Dsi-1* sequences. These analyses demonstrated that *Dsi-1* is located on mouse chromosome 4 (Table 1).

Relationship between *Dsi-1*, proto-oncogenes, and other frequent integration regions. Because *Dsi-1* is located on mouse chromosome 4, *Dsi-1* must be different from any proto-oncogenes or frequent integration regions which are

located on other mouse chromosomes. To date, at least 17 of the known proto-oncogenes have been localized to specific mouse chromosomes (Table 2). Of these, only *c-mos* is on mouse chromosome 4. Of the 18 tumor-associated integration regions, 13 have been localized to a specific mouse chromosome. Of these, only *lck* is on mouse chromosome 4 (Table 2).

To test whether *Dsi-1* was related to *c-mos*, a probe for the mouse *c-mos* sequences was hybridized to the 13.5 kb of cloned *Dsi-1* sequences. No homology was observed. To test whether *Dsi-1* was related to *MLvi-3*, the *Dsi-1* probe ds1-SR was hybridized to molecularly cloned chromosomal sequences for *MLvi-3*. No homology was observed. To test whether *Dsi-1* was related to *lck*, we took advantage of the homology between *lck* and the *src* family of oncogenes (32, 61) and hybridized a probe for chicken *v-erbB* sequences, another member of the *src* family, to the cloned *Dsi-1* sequences. Because homology between the *v-erbB* probe and mouse *lck* would be expected to be relatively weak, we included as a positive control *Dsrc4*, a molecularly cloned *src*-related sequence from *Drosophila melanogaster* (62). Both *lck* and *Dsrc4* had about 60% homology with a 200-base-pair region of the *v-erbB* probe. No hybridization of the *v-erbB* probe to *Dsi-1* was observed, whereas hybridization to *Dsrc4* was clearly visible on the same filter. This result suggests that *Dsi-1* is not related to *lck* or other members of the *src* gene family.

The results of the chromosomal localization studies provide unambiguous evidence that *Dsi-1* must be different from at least 16 of the known oncogenes and 7 of the 12 previously identified frequent integration regions. The results of the cross-hybridization experiments indicate that *Dsi-1* is not identical to two additional frequent integration regions, is not *c-mos*, and does not contain *src*-related sequences. Clearly, our data do not exclude the possibility that *Dsi-1*

FIG. 3. Failure to detect transcripts of *Dsi-1* in Northern blots of cytoplasmic poly(A)⁺ RNAs. Hybridization of Northern blots of 5 to 10 μ g of cytoplasmic poly(A)⁺ RNAs of RCC2-1M, NRK, and chicken embryo fibroblast cells with riboprobes for the indicated sequences. Size markers indicate the position of RNAs present in poly(A)⁺ RNAs electrophoresed on the same gel. Filters were (RNase +) or were not (RNase -) treated with RNase A to reduce background hybridization.

represents sequences outside the kinase domain of a member of the *src* gene family, nor can we exclude the possibility that *Dsi-1* is related to one of the proto-oncogenes or frequent integration regions whose mouse chromosomal location and cross-hybridization have not been determined. However, the evidence to date indicates that *Dsi-1* is different from previously identified tumor-related sequences.

Transcriptional activity of *Dsi-1*. Analyses of *Dsi-1* for DNase I-hypersensitive sites had revealed two clusters of prominent hypersensitive sites in a line of T cells cultured from an MoMLV-induced thymoma (60). This thymoma did not have a provirus integrated in *Dsi-1*. The hypersensitive sites were not present in the chromatin of a cell line cultured from NRK cells. Since cell-type-specific hypersensitive sites are frequently associated with tissue-specific mRNAs, cytoplasmic poly(A)⁺ RNAs from RCC2-1M cells and NRK cells were analyzed for transcripts of *Dsi-1*. Ds1-SR sequences were used as probes for these transcripts because the evolutionary conservation of this region of *Dsi-1* suggested that it might contain coding sequences (Fig. 2).

Cytoplasmic poly(A)⁺ RNAs were prepared from RCC2-1M and NRK cells and analyzed on Northern blots under stringent conditions for hybridization (50% formamide, 65°C) with RNA probes for both strands of ds1-SR. Since preliminary analyses for transcription of *Dsi-1* had indicated that transcripts from this region were not abundant, 5 to 10 µg of poly(A)⁺ RNA was loaded per lane. Both RNA strand probes hybridized to a number of bands in RCC2-1M and NRK cells (Fig. 3). Some of the bands were detected in RNA from both cell types, while others were unique to RCC2-1M or NRK RNA.

To test whether the bands represented specific hybridization, the nitrocellulose filters were digested with RNase A and then reanalyzed by autoradiography. All of the bands were completely susceptible to RNase digestion. In contrast, bands representing specific hybridization to the mRNA for the kinase domain of the chicken epidermal growth factor receptor were completely resistant to repeated RNase digestions (Fig. 3C). Cytoplasmic poly(A)⁺ RNA hybridized in solution with riboprobes for both strands of Dsi-SR also did not reveal RNase-resistant RNAs. Thus, we could detect no transcription of *Dsi-1*-specific sequences in RCC2-1M or NRK cells. Since RNA was not available for a tumor with an insertion in *Dsi-1*, we could not test whether insertions in *Dsi-1* alter *Dsi-1* expression.

DISCUSSION

Characterization of *Dsi-1*. *Dsi-1* is a 13.5-kb region of chromosomal DNA that contains proviral insertions in 10 to 20% of MoMLV-induced thymomas in rats. Three tumor-associated insertions in *Dsi-1* have been characterized (Fig. 1). Each is in the same transcriptional orientation. None appears to have deletions or rearrangements in *Dsi-1* or proviral sequences. Each of the tumor-associated insertions occurs in a 2-kb region in *Dsi-1* which is associated with a cluster of DNase I-hypersensitive sites (60). These hypersensitive sites may be T-cell-specific, since they were observed in MoMLV-induced thymomas but not in NRK cells.

Regions of *Dsi-1* are evolutionarily conserved (Fig. 2), a finding which allowed us to map this region to mouse chromosome 4 (Table 1). Interestingly, the evolutionarily conserved region of *Dsi-1*, despite its residence between two clusters of hypersensitive sites in the RCC2-1M line of T cells, does not appear to be expressed (Fig. 3). Thus, these hypersensitive sites may mark a region that was expressed at

an earlier stage in T-cell development (18) or a conserved noncoding region such as that containing enhancer elements in the major intron of rearranged immunoglobulin heavy-chain genes (2, 11, 14).

Dsi-1 resides on the same chromosome as the proto-oncogene *c-mos* and the apparent proto-oncogene *lck*. Cross-hybridization studies indicate that the *Dsi-1* sequences we have cloned contain neither *c-mos* nor *lck*. However, we cannot exclude the possibility that tumor-promoting integrations at *Dsi-1* affect the expression of *c-mos* or *lck*, either because *Dsi-1* represents a distal portion of one of these genes or because *Dsi-1* has a longer-range effect on their expression. A precedent for such exists in integrations at *pvt-1* which are more than 70 kb away from *c-myc* (16) that affect *c-myc* expression. Unfortunately, the absence of RNA from a tumor with an integration in *Dsi-1* prevented us from testing these hypotheses.

MLV-induced thymomas are associated with insertions in at least eight different frequent integration regions. Avian leukemia virus-induced B-cell lymphomas have been associated with only one integration region, the *c-myc* proto-oncogene, and more than 95% of such tumors have proviruses integrated in or near *c-myc* (45). In contrast, MLV-induced thymomas have been associated with insertions in eight different chromosomal regions (7, 9, 16, 28, 30, 32, 52, 55–57, 61). The incidence of insertions into a particular one of these regions is variable and presumably depends on the inducing MLV, the rodent inoculated, and the protocol being used for tumor induction. Typically, only a minority of tumors will have insertions into any one region. The reason for these striking differences remains to be elucidated.

Role of frequent integration regions in tumorigenesis. Our prejudice is that most if not all regions with tumor-associated insertions affect cell growth. Three lines of reasoning support this bias. First, at least 6 and perhaps as many as 9 of the 19 regions with tumor-associated integrations are related to or affect the expression of known oncogenes. Second, insertions into at least five of these regions result in increased levels of expression of the target region. This would be consistent with expression of genes in these regions affecting tumorigenesis. And third, in the absence of selection, MLV provirus would not be expected to be identified multiple times in a given region because MLV proviruses insert into many different regions of chromosomal DNA (1, 53). Recently we have shown that integration is nonrandom, with proviruses preferentially integrating near DNase I-hypersensitive sites (60). However, there are many DNase I-hypersensitive sites in chromosomal DNA (29, 33), far more than have been found or might be expected to be found to be frequent integration regions.

Many MLV-induced tumors contain insertions in several frequent integration regions. Interestingly, many MLV-induced thymomas have insertions in more than one frequent integration region. The tumor RT10-2 from which the insertion in *Dsi-1* was molecularly cloned also has a provirus adjacent to *c-myc* (52). Many thymomas have insertions in *Myi-1* and *Myi-2* (58). Similarly, many mouse mammary tumor virus-induced tumors have insertions in *int-1* and *int-2* (43). Several lines of evidence indicate that tumor induction is the result of several genetic events (6, 27, 46). Most MLV-induced tumors have 10 to 20 proviruses per tumor. Consequently, proviral insertion may be one of the more frequent mutagenic events in the infected host, and it would not be unreasonable to expect that more than one of the mutations leading to tumor formation might involve insertional mutagenesis.

Identification of frequent integration regions may identify proto-oncogenes with subtle effects on cell growth. Of the 19 currently identified tumor-associated integration regions, only 6 represent currently known proto-oncogenes. If all frequent integration regions are involved in tumor development, the finding that only a third of these are related to known proto-oncogenes suggests that the known proto-oncogenes include only a fraction of the genes that control cell growth. Known proto-oncogenes have been identified as homologs of viral genes which confer or enhance transforming potential or as genes in transfected DNA able to alter the ordered growth of certain established cell lines. In contrast, frequent integration regions have been identified by virtue of their containing a provirus. Consequently, frequent integration regions may identify genes with subtle as well as dramatic effects on cell growth.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grants CA-30674 to D. Steffen, and CA-27223 to H. Robinson and core grant P30-12708 to the Worcester Foundation for Experimental Biology from the National Institutes of Health (NIH). Computer resources used to carry out our studies were provided by the NIH-sponsored BIONET National Computer Resource for Molecular Biology (grant 1 U41 RR-01685-03).

We are indebted to P. Tschlis for providing molecularly cloned *Mlvi-1*, *Mlvi-2*, and *Mlvi-3* sequences, A. Berns for a unique sequence probe for *Pim-1*, P. Jolicoeur for a unique sequence probe for *Mis-1*, and Jonathan Silver for a unique sequence probe for *Fis-1*.

LITERATURE CITED

- Bacheler, L. T., and H. Fan. 1979. Multiple integration sites for Moloney murine leukemia virus in productively infected mouse fibroblasts. *J. Virol.* **30**:657-667.
- Banerji, J., L. Olsen, and W. Schaffner. 1983. A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. *Cell* **33**:729-740.
- Blatt, C., M. E. Harper, G. Franchini, M. N. Nesbitt, and M. I. Simon. 1984. Chromosomal mapping of murine *c-fes* and *c-src* genes. *Mol. Cell. Biol.* **4**:978-981.
- Brown, A. M. C., R. S. Wildin, T. J. Prendergast, and H. E. Varmus. 1986. A retrovirus vector expressing the putative mammary oncogene *int-1* causes partial transformation of a mammary epithelial cell line. *Cell* **46**:1001-1009.
- Cahilly, L. A., and D. L. George. 1985. Regional mapping of c-Ki-ras proto-oncogene on mouse chromosome 6 by in situ hybridization. *Cytogenet. Cell Genet.* **39**:140-144.
- Cairns, J. 1975. Mutation, selection, and the natural history of cancer. *Nature (London)* **255**:197-200.
- Corcoran, L. M., J. M. Adams, A. R. Dunn, and S. Cory. 1984. Murine T lymphomas in which the cellular *myc* oncogene has been activated by retroviral insertions. *Cell* **37**:113-122.
- Crews, S., R. Barth, L. Hood, J. Prehn, and K. Calame. 1982. Mouse *c-myc* oncogene is located on chromosome 15 and translocated to chromosome 12 in plasmacytomas. *Science* **218**:1319-1321.
- Cuyppers, H. T., G. Selten, W. Quint, M. Zijlstra, E. R. Moandag, W. Boelens, P. van Wezenbeek, C. Melief, and A. Berns. 1984. Murine leukemia virus-induced T-cell lymphomagenesis: integration of proviruses in a distinct chromosomal region. *Cell* **37**:141-150.
- D'Eustachio, P. 1984. A genetic map of mouse chromosome 12 composed of polymorphic DNA fragments. *J. Exp. Med.* **160**:827-838.
- Emorine, L., M. Kuehl, L. Weir, P. Leder, and E. E. Max. 1983. A conserved sequence in the immunoglobulin J_K-C_K intron: possible enhancer element. *Nature (London)* **304**:447-449.
- Fung, Y. K., W. G. Lewis, L. B. Crittenden, and J. H. Kung. 1983. Activation of the cellular oncogene *c-erbB* by LTR insertion: molecular basis for induction of erythroblastosis by avian leukosis virus. *Cell* **33**:357-368.
- Garcia, M., R. Wellinger, A. Vellaz, and H. Doggelmann. 1986. A new site of integration for mouse mammary tumor virus proviral DNA common to BALB/c(C3H) summary and kidney adenocarcinomas. *EMBO J.* **5**:127-134.
- Gillies, S. D., S. L. Morrison, V. T. Oi, and S. Tonegawa. 1983. A tissue-specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. *Cell* **33**:717-728.
- Goff, S. P., P. D'Eustachio, F. H. Ruddle, and D. Baltimore. 1982. Chromosomal assignment of the endogenous proto-oncogene *c-abl*. *Science* **218**:1317-1319.
- Graham, M., J. M. Adams, and S. Cory. 1985. Murine T lymphomas with retroviral inserts in the chromosomal 15 locus for plasmacytoma variant translocations. *Nature (London)* **314**:740-743.
- Gray, D. A., C. M. McGrath, R. F. Jones, and V. L. Morris. 1986. A common mouse mammary tumor virus integration site in chemically induced precancerous mammary hyperplasias. *Virology* **148**:360-368.
- Groudine, M., and H. Weintraub. 1982. Propagation of globin DNAase I-hypersensitive sites in absence of factors required for induction: a possible mechanism for determination. *Cell* **30**:131-139.
- Hayward, W. S., B. G. Neel, and S. M. Astrin. 1981. Activation of a cellular onc gene by promoter insertion in ALV-induced lymphoid leukemia. *Nature (London)* **290**:475-480.
- Hilkens, J., H. T. Cuyppers, G. Selten, V. Kroexen, J. Hilgers, and A. Berns. 1986. Genetic mapping of *Pim-1* putative oncogene to mouse chromosome 17. *Somat. Cell Mol. Genet.* **12**:81-88.
- Huebner, K., A. ar-Rushdi, C. A. Griffin, M. Isobe, C. Kozak, B. S. Emanuel, L. Nagarajan, J. L. Cleveland, T. I. Bonner, M. D. Goldsborough, C. M. Croce, and U. Rapp. 1986. Actively transcribed genes in the *raf* oncogene group, located on the X chromosome in mouse and human. *Proc. Natl. Acad. Sci. USA* **83**:3934-3938.
- Jolicoeur, P., L. Villeneuve, E. Rassart, and C. Kozak. 1985. Mouse chromosomal mapping of a murine leukemia virus integration region (*Mis-1*) first identified in rat thymic leukemia. *J. Virol.* **56**:1045-1048.
- Kozak, C. A. 1983. Genetic mapping of a mouse chromosomal locus required for mink cell focus-forming virus replication. *J. Virol.* **48**:300-303.
- Kozak, C. A., M. A. Gunnell, and U. R. Rapp. 1984. A new oncogene, *c-raf*, is located on mouse chromosome 6. *J. Virol.* **49**:297-299.
- Kozak, C. A., J. F. Sears, and M. D. Hoggan. 1983. Genetic mapping of the mouse proto-oncogene *c-sis* to chromosome 15. *Science* **221**:867-869.
- Kozak, C. A., P. G. Strauss, and P. N. Tschlis. 1985. Genetic mapping of a cellular DNA region involved in induction of thymic lymphomas (*Mlvi-1*) to mouse chromosome 15. *Mol. Cell. Biol.* **5**:894-897.
- Land, H., L. F. Parada, and R. A. Weinberg. 1983. Tumorigenic conversion of primary embryo fibroblasts requires at least two cooperating oncogenes. *Nature (London)* **304**:596-602.
- Lemay, G., and P. Jolicoeur. 1984. Rearrangement of a DNA sequence homologous to a cell-virus junction fragment in several Moloney murine leukemia virus-induced rat thymomas. *Proc. Natl. Acad. Sci. USA* **81**:38-42.
- Lewin, B. 1980. Complexity of mRNA population, p. 694-727. *In* Gene expression, vol. 2. John Wiley & Sons, Inc., New York.
- Li, Y., C. A. Holland, J. W. Hartley, and N. Hopkins. 1984. Viral integration near *c-myc* in 10-20% of MCF 247 induced AKR lymphomas. *Proc. Natl. Acad. Sci. USA* **81**:6801-6811.
- Marth, J. D., C. Distech, D. Pravtcheva, F. Ruddle, E. G. Krebs, and R. M. Perlmutter. 1986. Localization of a lymphocyte-specific protein tyrosine kinase gene (*lck*) at a site of frequent chromosomal abnormalities in human lymphomas.

- Proc. Natl. Acad. Sci. USA 83:7400-7404.
32. Marth, J. D., R. Peet, E. G. Krebs, and R. M. Perlmutter. 1985. A lymphocyte-specific protein-tyrosine kinase gene is rearranged and overexpressed in the murine T cell lymphoma LSTRA. *Cell* 43:393-404.
 33. Miles, B. D., and H. L. Robinson. 1985. High-frequency transduction of *c-erbB* in avian leukosis virus-induced erythroblastosis. *J. Virol.* 54:295-303.
 34. Mushinski, J. R., M. Potter, S. R. Bauer, and E. P. Reddy. 1983. DNA rearrangement and altered RNA expression of the *c-myc* oncogene in mouse plasmacytoid lymphosarcomas. *Science* 220:795-798.
 35. Nagarajan, L., E. Louie, Y. Tsujimoto, A. ar-Rushdi, K. Huebner, and C. M. Croce. 1986. Localization of the human *pim* oncogene (PIM) to a region of chromosome 6. *Proc. Natl. Acad. Sci. USA* 83:2556-2560.
 36. Neel, B. G., W. S. Hayward, H. L. Robinson, J. Fang, and S. M. Astrin. 1981. Avian leukosis virus-induced tumors have common proviral integration sites and synthesize discrete new RNAs: oncogenesis by promoter insertion. *Cell* 23:323-334.
 37. Noori-Daloui, M. R., R. A. Swift, H. J. Kung, L. B. Crittenden, and R. L. Witter. 1981. Specific integration of REV proviruses in avian bursal lymphomas. *Nature (London)* 294:574-576.
 38. Nusse, R., A. van Ooyen, D. Cox, Y.-K. T. Foy, and H. E. Varmus. 1984. Mode of proviral activation of a putative mammary oncogene (*int-1*) on mouse chromosome 15. *Nature (London)* 307:131-136.
 39. Nusse, R., and H. E. Varmus. 1982. Many tumors induced by the mouse mammary tumor virus contain a provirus integrated in the same region of the host genome. *Cell* 31:99-109.
 40. Oskarsson, M., W. L. McClements, D. G. Blair, J. V. Maizel, and G. F. Vande Woude. 1980. Properties of a normal mouse cell DNA sequence (*sarc*) homologous to the *src* sequence of Moloney sarcoma virus. *Science* 207:1222-1224.
 41. Payne, G. S., J. M. Bishop, and H. E. Varmus. 1982. Multiple arrangements of viral DNA and an activated host oncogene in bursal lymphomas. *Nature (London)* 295:209-214.
 42. Peters, G., C. Kozak, and C. Dickson. 1984. Mouse mammary tumor virus integration regions *int-1* and *int-2* map on different mouse chromosomes. *Mol. Cell. Biol.* 4:375-378.
 43. Peters, G., A. E. Lee, and C. Dickson. 1986. Concerted activation of two potential proto-oncogenes in carcinomas induced by mouse mammary tumour virus. *Nature (London)* 320:628-631.
 44. Peters, G., S. Brookes, R. Smith, and C. Dickson. 1983. Tumorigenesis by mouse mammary tumor virus: evidence for a common region for provirus integration in mammary tumors. *Cell* 33:364-377.
 45. Robinson, H. L., and G. C. Gagnon. 1986. Patterns of proviral insertion and deletion in avian leukosis virus-induced lymphomas. *J. Virol.* 57:28-36.
 46. Ruley, H. E. 1983. Adenovirus early region 1A enables viral and cellular transforming genes to transform primary cells in culture. *Nature (London)* 304:602-606.
 47. Ryan, J., C. P. Hart, and F. H. Ruddle. 1984. Molecular cloning and chromosome assignment of murine *N-ras*. *Nucleic Acids Res.* 12:6063-6072.
 48. Sakaguchi, A. Y., P. A. Lalley, B. U. Zabel, R. W. Ellis, E. M. Scolnick, and S. L. Naylor. 1984. Chromosome assignments of four mouse cellular homologs of sarcoma and leukemia virus oncogenes. *Proc. Natl. Acad. Sci. USA* 81:525-529.
 49. Silver, J., and C. Kozak. 1986. Common proviral integration region on mouse chromosome 7 in lymphomas and myelogenous leukemias induced by Friend murine leukemia virus. *J. Virol.* 57:526-533.
 50. Silver, J., J. B. Whitney 3d, C. Kozak, G. Hollis, and I. Kirsch. 1985. *erbB* is linked to the alpha-globin locus on mouse chromosome 11. *Mol. Cell. Biol.* 5:1784-1786.
 51. Sola, B., S. Fichelson, D. Bordereaux, P. E. Tamborin, and S. Gisselbrecht. 1986. *fim-1* and *fim-2*: two new integration regions of Friend murine leukemia virus in myeloblastic leukemias. *J. Virol.* 60:718-725.
 52. Steffen, D. 1984. Proviruses are adjacent to *c-myc* in some murine leukemia virus-induced lymphomas. *Proc. Natl. Acad. Sci. USA* 81:2097-2101.
 53. Steffen, D., and R. A. Weinberg. 1978. The integrated genome of murine leukemia virus. *Cell* 15:1003-1010.
 54. Swan, D., M. Oskarsson, D. Keithley, F. H. Ruddle, P. Deustachio, and G. F. Vande Woude. 1982. Chromosomal localization of the Moloney sarcoma virus mouse cellular (*c-mos*) sequence. *J. Virol.* 44:752-754.
 55. Tschlis, P. N., M. A. Lohse, C. Szpirer, J. Szpirer, and G. Levan. 1985. Cellular DNA regions involved in the induction of rat thymic lymphomas (*MLvi-1*, *MLvi-2*, *MLvi-3*, and *c-myc*) represent independent loci as determined by their chromosomal map locations in the rat. *J. Virol.* 56:938-942.
 56. Tschlis, P. N., P. G. Strauss, and L. F. Hu. 1983. A common region for proviral DNA integration in MoMuLV-induced rat thymic lymphomas. *Nature (London)* 302:445-449.
 57. Tschlis, P. N., P. G. Strauss, and C. A. Kozak. 1984. Cellular DNA region involved in induction of thymic lymphomas (*MLvi-2*) maps to mouse chromosome 15. *Mol. Cell. Biol.* 4:997-1000.
 58. Tschlis, P. N., P. G. Strauss, and M. H. Lohse. 1985. Concerted DNA rearrangements in Moloney murine leukemia virus-induced thymomas: a potential synergistic relationship in oncogenesis. *J. Virol.* 56:258-267.
 59. Vennstrom, B., L. Fanshier, C. Moscovici, and J. M. Bishop. 1980. Molecular cloning of the avian erythroblastosis virus genome, and recovery of oncogenic virus by transfection of chicken cells. *J. Virol.* 36:575-585.
 60. 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.
 61. Voronova, A. F., and B. M. Sefton. 1986. Expression of a new tyrosine protein kinase is stimulated by retrovirus promoter insertion. *Nature (London)* 319:682-685.
 62. Wadsworth, S. C., K. Madhaven, and D. Bilodeau-Wentworth. 1985. Maternal inheritance of transcripts from three *Drosophila* *src*-related genes. *Nucleic Acids Res.* 13:2153-2170.
 63. Watson, D. K., M. J. McWilliams-Smith, C. Kozak, R. Reeves, J. Gearhart, M. F. Nunn, W. Nash, J. R. Fowle 3d, P. Duesberg, T. S. Papas, and S. J. O'Brien. 1986. Conserved chromosomal positions of dual domains of the *ets* protooncogene in cats, mice, and humans. *Proc. Natl. Acad. Sci. USA* 83:1792-1796.
 64. Westaway, D., J. Papkoff, C. Moscovici, and H. E. Varmus. 1986. Identification of a provirally activated *c-Ha-ras* oncogene in an avian nephroblastoma via a novel procedure: cDNA cloning of a chimaeric viral-host transcript. *EMBO J.* 5:301-309.
 65. Yamamoto, T., T. Nishida, N. Miyajima, S. Kawai, T. Ooi, and K. Toyoshima. 1983. The *erbB* gene of avian erythroblastosis virus is a member of the *src* gene family. *Cell* 35:71-78.
 66. Zabel, B. U., R. E. Fournier, P. A. Lalley, S. L. Naylor, and A. Y. Sakaguchi. 1984. Cellular homologs of the avian erythroblastosis virus *erbA* and *erbB* genes are syntenic in mouse but asyntenic in man. *Proc. Natl. Acad. Sci. USA* 81:4874-4878.