



Identification of candidate target genes for EVI-1, a zinc finger oncoprotein, using a novel selection strategy

Jeong H Kim, Pei Hui, Dongxian Yue, Joyce Aycock, Cynthia Leclerc, Anna R Bjoring and Archibald S Perkins

Yale University, Department of Pathology, PO Box 208023, New Haven, Connecticut 06520-8023, USA

We have sought to identify and isolate target genes for the zinc finger protein, EVI-1, which has been implicated in the genesis of myelogenous leukemia both in mouse and human. We have approached this with a two-step selection: we first selected for genomic fragments of mouse DNA that bind to the protein with high affinity; second, we employed cDNA hybrid selection to identify gene sequences contained within these fragments. We show that we have constructed a sublibrary of genomic fragments that contains a significant fraction of the EVI-1-binding sites in the mouse genome. Our data has allowed us to estimate that there are approximately 4300 binding sites per haploid genome in the mouse. We further demonstrate that by using cDNA hybrid selection, it is relatively straightforward to isolate cDNAs that correspond to genes embedded in the EVI-1-binding sublibrary. Several of these are novel, but are represented in databases of anonymous human or mouse cDNAs (expressed sequence tags). One selected gene is *Itpr2*, encoding the inositol trisphosphate type two receptor, which is transcriptionally regulated during myelopoiesis. Finally, using a chimeric EVI-1-VP16-fusion protein under the control of a tetracycline-regulated system, we have shown that this chimeric activator can directly regulate *Itpr2*.

Keywords: myeloid leukemia; transcription factor; target gene

Introduction

Evi-1 (for ecotropic viral integration site 1) is a zinc finger oncogene that is activated by retroviral insertion in murine myelogenous leukemia (Morishita *et al.*, 1988; Mucenski *et al.*, 1988). *EVI-1* is also involved in human leukemia, as shown by the presence of inversions, deletions, and translocations of chromosome 3q25–26, where the gene is located (Fichelson *et al.*, 1992; Morishita *et al.*, 1992; Levy *et al.*, 1994; Mitani *et al.*, 1994; Nucifora *et al.*, 1994; Suzukawa *et al.*, 1994; Peeters *et al.*, 1997). These alterations are associated with inappropriate expression of the gene, and in some instances (e.g., t(3;21) and t(3;12); (Mitani *et al.*, 1994; Nucifora *et al.*, 1994; Peeters *et al.*, 1997)), changes in the structure of the encoded protein. The mechanism by which *Evi-1* causes leukemia is not

entirely clear. However, certain experiments suggest that *Evi-1* contributes to leukemia by causing an abnormal response to cytokines that induce terminal differentiation of hematopoietic cells (Morishita *et al.*, 1992; Kreider *et al.*, 1993). *Evi-1* appears to also play an important role in mammalian development: it is expressed in a temporally and spatially restricted manner in mouse development (Perkins *et al.*, 1991), and homozygous null mutant mice die at day 10 of gestation (Hoyt *et al.*, 1997).

The murine *Evi-1* encodes several isoforms, the most well-studied of which is a 1042 amino acid protein with ten zinc finger motifs separated into two domains: fingers 1–7 in an N-terminal domain, and fingers 8–10 in a more C-terminal region (Morishita *et al.*, 1988). These zinc finger motifs mediate sequence-specific binding to DNA, and binding sites for the first (GACAAGATAA (Perkins *et al.*, 1991) or GAC/TAAGATAAGATAA (Delwel *et al.*, 1993)) and second (GAAGATGAG (Funabiki *et al.*, 1994)) sets of zinc fingers have been identified. The binding of fingers 1–7 to the GACAAGATAA motif is highly specific: first, methylation interference studies revealed multiple points of contact between the protein and the binding site, including bases 1–8 on the top strand and scattered bases on the bottom strand, with strongest interactions at the two G residues on the top and the one on the bottom strand. Second, single base pair changes at positions 1, 2, 3, 5, 6 or 8 of the consensus site reduce the affinity of binding to less than 5% relative to the 'wildtype' site, in an *in vitro* binding assay, while changes at position ten had a less drastic effect (the other positions were not assessed) (Perkins and Kim, 1996). Third, changes at positions 1 (to T), 3 (to A) or 6 (to T) abolished high affinity binding *in vivo*, as assessed by cotransfection studies using reporter constructs bearing mutant or wildtype binding sites. It is likely that T substitutions at position 3 allow high affinity binding, since oligonucleotides containing the sequence GATAAGATA were isolated by site selection (Delwel *et al.*, 1993).

The highly specific nature of this binding, and the capacity of the protein to bind to two distinct motifs, suggested to us the possibility that in the mouse genome, there may be a defined number of sites with which EVI-1 interacts, and that these may represent the functional targets of EVI-1 action.

Several studies support the notion that EVI-1 acts as a transcriptional regulatory protein. In cotransfection studies in fibroblasts using synthetic reporters bearing the first zinc finger binding motif, EVI-1 can repress transcription of the reporter when it is activated by another protein, such as GATA1, which can bind specifically to the first EVI-1 binding motif (Kreider *et*

al., 1993; Perkins and Kim, 1996). By itself, it appears that EVI-1 neither activates nor represses transcription on reporters containing only a single type of EVI-1 binding site. When the reporter contains both the first and the second binding motifs, EVI-1 can act as an activator (Morishita *et al.*, 1995). By what appears to be an indirect mechanism, EVI-1 can also raise AP1 levels in cells (Tanaka *et al.*, 1994). Recently, a domain in EVI-1 has been identified that is required for repression of activated transcription (Bartholomew *et al.*, 1997). This region, which is located between the two zinc finger domains (amino acids 514–724) and is very rich in proline residues, is also required for transformation of Rat1 fibroblasts (Bartholomew *et al.*, 1997).

While it is almost certain that EVI-1 is a regulator of transcription, the target genes for EVI-1 have not been reported. In an effort to better understand the biological role of *Evi-1* in leukemogenesis as well as in normal cells, we wish to identify genes that are regulated by EVI-1. Given that EVI-1 appears to act as a transcriptional repressor, we considered it unlikely to attain our goal with a differential cloning strategy using the native protein, such as one employed to identify *CIP1* (*WAF1*) (El-Deiry *et al.*, 1993) or the *PIG* genes (Polyak *et al.*, 1997) as targets of p53. In this report, we describe a novel approach to identify the target genes that EVI-1 regulates. One of the potential target genes we have identified is a known gene *Itp2*, which encodes a key component of the inositol trisphosphate signaling pathway.

Results

Testing of the binding site selection protocol

Our strategy to identify EVI-1 target genes involved two steps: first, the creation of a library of genomic fragments that bind the protein with high affinity; second, use of the cDNA hybrid selection to isolate cDNAs that correspond to genes within the genomic fragments (Figure 1). We expected that if the fragments of genomic DNA were large enough, they would span between an EVI-1 binding site and an exon of the target gene located in *cis*. To achieve the first step, we performed cycles of DNA-protein binding and selection as described (Sompayrac and Danna 1990) using EVI-1 (produced in either *E. coli* or in insect cells with the baculovirus expression system) and a plasmid-based mouse genomic library. This library was created with size-selected (2–7 kb) *Sau3A*I-digested mouse DNA. Protein-DNA complexes were isolated by filtration through nitrocellulose filters, and amplified by transforming *E. coli*. We decided to use this approach rather than PCR-based approaches (Kinzler and Vogelstein, 1989), since PCR tends to select for small (<600 bp) fragments.

To test whether this binding site selection approach would work for EVI-1, we performed a reconstruction experiment by performing cycles of selection on a mixture of two plasmids, p5EVBS-CAT (*lacZ*⁻, containing five GACAAGATAAGATAA motifs) and pBluescript (*lacZ*⁺). For these experiments, we employed an *E. coli*-produced protein that contained only zinc fingers 1–7 (MBP-EVI-1(1–254)) (Perkins

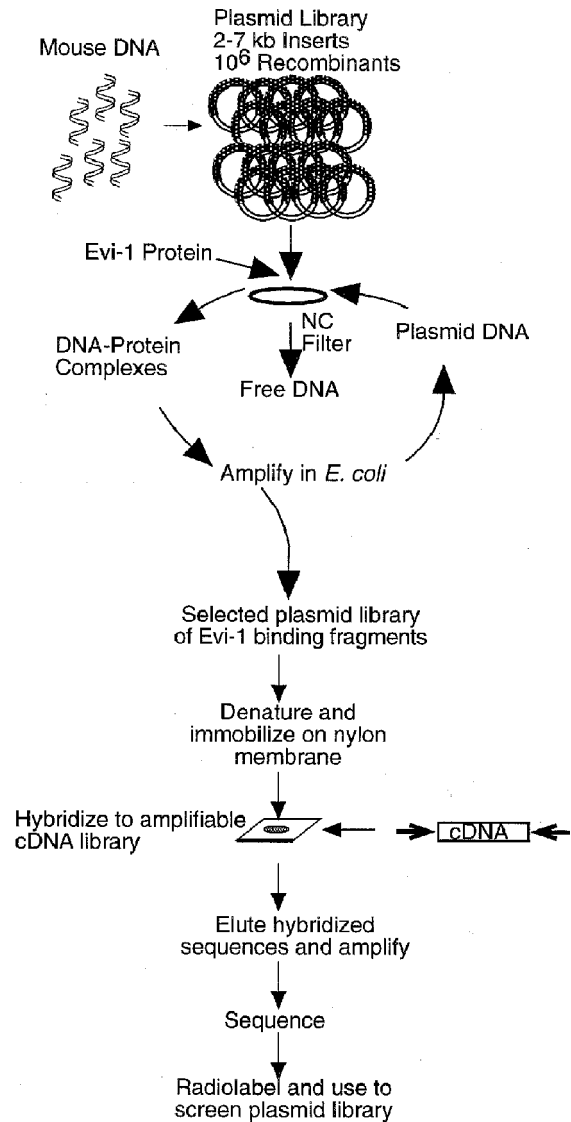


Figure 1 Scheme for isolating target genes for EVI-1, as described in text

et al., 1991). After each cycle of selection, the types of plasmids retained on the filter were assessed on *lacZ* indicator plates. With two rounds of selection, the ratio of p5EVBS-CAT to pBluescript went from 1:100 to 98:2 (Figure 2a), indicating an efficient selection for plasmids containing EVI-1 binding sites. To show that this same approach would be able to isolate binding fragments from a library of mouse genomic fragments, we performed a similar reconstruction experiment by selecting plasmids for EVI-1 binding from a pool composed of 95 parts pBluescript and five parts genomic library plasmids that had been enriched with two rounds of EVI-1 binding selection. The ratio of pBluescript to library plasmids was assessed on *lacZ* indicator plates. Each successive round of selection yielded significant enrichment of the *lacZ*⁻, library-derived plasmids, indicating effective selection against pBluescript and selection for genomic fragments with EVI-1 binding sites (Figure 2b).

To show that the selected plasmids contained binding sites, we assessed their affinity for EVI-1 by competitive gel shift analysis (Figure 2c). Using a

radiolabeled oligonucleotide containing the first EVI-1 binding motif (GACAAGATAAGATAA) as a probe, we competed for binding to MBP-EVI-1(1–254) with selected plasmid DNAs. While plasmid vector alone did not compete, the selected plasmid DNAs did compete (representative shown in Figure 2c), indicating that the latter had EVI-1 binding sites. The presence of EVI-1 binding sites on this particular clone (P25) was confirmed by DNA sequencing (Table 1, line 13).

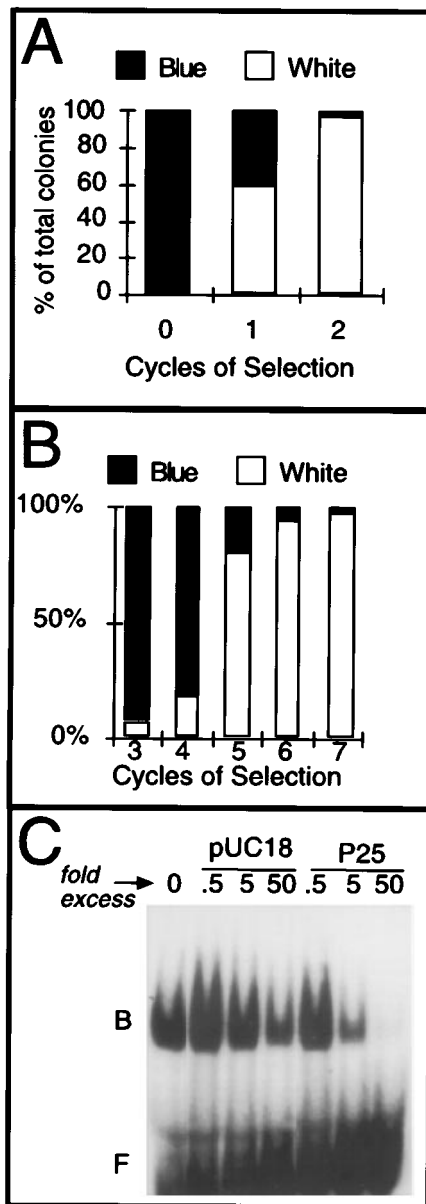


Figure 2 Assessment of binding site selection scheme. (a) Reconstruction experiment to test selection technique. Binding site selection was performed with a mixture of two plasmids: 1% *lacZ*⁻ plasmid containing five copies of the first EVI-1 binding site and 99% *lacZ*⁺ (Bluescript) devoid of binding sites. The proportion of *lacZ*⁻ and *lacZ*⁺ plasmids was assessed by transformation into *E. coli* and growth on indicator plates. The elimination of plasmids lacking binding sites is indicated by the loss of *lacZ*⁺ plasmids with each round of selection for binding sites, performed as described in Materials and methods section. (b) Similar reconstruction to that shown in a, this time using a starting mixture 95% *lacZ*⁺ (Bluescript), and 5% *lacZ*⁻ plasmids from genomic library of mouse DNA that had been through two rounds of selection for EVI-1 binding. Bar graph displays the percent of *lacZ*⁻ and *lacZ*⁺ plasmids with each round of binding site selection. (c) Competition gel shift to assess relative affinity of one of the selected plasmids, using MBP-EVI-1 (1–254) protein and a radiolabeled oligonucleotide containing the first EVI-1 binding site (GACAAGATAAGATAA) as a substrate. Unlabeled competitors include plasmid vector (pUC18) and the selected plasmid (P25)

Creation of a sublibrary of genomic fragments that bind EVI-1

These experiments indicated the feasibility of using this selection approach to isolate EVI-1-binding fragments from a library of mouse genomic DNA. We scaled this selection up to create a complete library of EVI-1-binding fragments. We performed nine independent selections using His₆-tagged, full-length EVI-1 (EVI-1-His) produced in insect cells using the baculovirus expression system. Each selection consisted of 6 to 9 cycles of protein-DNA binding, selection, and amplification. During the course of this selection, the efficacy of the selection was assessed by doping the selected plasmids with pUC18, and then monitoring the elimination of the *lacZ*⁺ colonies with successive rounds of selection. The fold enrichment (see Materials and methods) ranged from 2–30 (typically around four) for each cycle.

We were interested in obtaining the DNA sequence for the EVI-1 binding sites within the selected plasmids. Since most of the inserts of genomic DNA within the plasmid library were greater than 2 kb, it seemed inefficient to sequence the entire insert to find the binding site. Therefore, we shotgun-subcloned *Sau3A1* fragments from individual representative plasmids into pUC18, prepared purified plasmid DNA from the population of amp^r *E. coli* transformants, and selected for EVI-1 binding plasmids from the collection of *Sau3A1* subclones by filter binding as described above. This allowed identification of a single *Sau3A1* subfragment derived from most all of the selected genomic plasmid clones that exhibited EVI-1 binding by filter binding assay, and was likely to contain an EVI-1 binding site. Indeed, from approximately 80% of the parental genomic plasmids we were able to enrich selectively for a single *Sau3A1* fragment, strongly suggesting the presence of an EVI-1 binding site within the subclones.

Forty-seven of the retained *Sau3A1* fragments were partially sequenced; 20 (42%) were found to harbor EVI-1 binding sites (Table 1), indicating that the selection for EVI-1 binding fragments was successful. We suspect that a higher percentage (about 80%) harbored EVI-1 binding sites, but since the fragments were not sequenced completely, not all of the binding sites were identified. Interestingly, we found exclusively the binding motif for the first set of zinc fingers (GACAAGATA; Table 1), and did not find the GAAGATGAG that is recognized by zinc fingers 8–10 (Funabiki et al., 1994). Table 2 shows only 100 bp of sequence surrounding the binding site; the complete sequence for each clone has been deposited with GenBank.

Table 1 Sequence of 18 different genomic DNAs containing EVI-1 binding sites

		GenBank#
1	K127	AF031110
2	K142	AF031111
3	KJ14	AF031112
4	KJ23	AF031113
5	KJ31	AF031114
6	KJ24	AF031115
7	KJ25	AF031116
8	KJ26	AF031117
9	KJ27	AF031118
10	KJ33	AF031119
11	KJ42	AF031120
12	KJ5C (<i>Itp2</i>)	AF031127
13	P25	AF031121
14	sub3	AF031122
15	sub12	AF031123
16	CG6.3	AF031126
17	CG5	AF031124
18	CG7	AF031125

Sequence analysis of selected *Sau3A1* subclones derived from selected parental plasmids. Displayed are the regions of the sequence containing the EVI-1 binding site, exhibited in bold. To the right is given the GenBank Accession number

Table 2 Number of occurrences of individual binding site clones in the nine independently initiated selections

Number of occurrences	Number of fragments
1	231
2	20
3	8
4	2
5	0
6	0
7	0
8	1
Total	1596

Analysis of the sequence data obtained from the 47 *Sau3A1* fragments for similarity to sequences in GenBank and EMBL revealed only limited homology; those with highest significance had *P* values = 10⁻⁷, indicating a lack of identity to previously characterized sequences.

We utilized the TESS DNA sequence analysis program (URL: <http://agave.humgen.upenn.edu/tess/index.html>) to identify the presence of binding motifs for transcription factors other than EVI-1 in those *SauA1* fragments with EVI-1 binding sites. While this revealed potential binding motifs for other factors, there was no consistent or repeating relationship between EVI-1 motifs and other motifs. In one fragment (KJ25, Table 1, line 7), there was a direct repeat of 68 bp that contained an EVI-1 binding site and an NF-Zz-binding site (Tsang *et al.*, 1990) (data not shown). We have also analysed each sequence for splice acceptor, splice donor, and potential starts of transcription using neural networking software available from Lawrence Berkeley Laboratories (<http://www-hgc.lbl.gov/projects/splice.html> and <http://www-hgc.lbl.gov/projects/promoter.html>), which revealed potential splicing sites and starts of transcription in several of the clones; however, significant open reading frames were not associated with these features, and thus their significance remains unclear.

Estimating the total number of EVI-1 binding sites per haploid genome

By comparing the *Sau3A1* restriction fragment fingerprints of 262 randomly picked plasmids from the nine different pools, it was evident that certain plasmids had been selected more than once in the nine

selections. This was confirmed by Southern blotting experiments, in which single *Sau3A1* fragments derived from individual selected plasmids were radiolabeled and hybridized to panels of selected plasmid clones (data not shown). From this analysis, we were able to tabulate the number of clones that were present in more than one plasmid pool obtained from independently initiated selections, and in how many pools they were present (Table 2). This has the form of a Poisson distribution (Figure 3), and has allowed us to calculate an estimate for the total number of EVI-1-binding sites within the genome, by employing statistical analysis (Wang and Brown, 1991; Mohler, 1977; see Materials and methods). We estimate there are approximately 1600 EVI-1 binding fragments in our library. Since the original library was estimated to be 37% complete (Sompayrac, personal communication), we can estimate that there are approximately 4300 binding sites total per haploid genome.

Selection of cDNAs

The limited sequence determination and analysis that we performed on the genomic fragments selected for EVI-1 binding failed to reveal any homology with known genes, indicating that this was not an efficient approach for identifying genes located near the EVI-1 binding sites. We therefore used cDNA hybrid selection (Parimoo *et al.*, 1993) to identify cDNAs that corresponded to exons contained within the genomic fragments selected for EVI-1 binding. Plasmids DNAs from the nine independent EVI-1-binding selections were pooled and immobilized onto nylon membrane, and used to select cDNAs from a short-insert, random-primed mouse cDNA library that was prepared from high complexity brain mRNA. The selected cDNAs were cloned into a lambda cloning vector (GT10).

With the goal of determining the complexity of the library of selected cDNAs, we arrayed the insert DNAs from 91 individual cDNA clones on a dot blot matrix and hybridized this with individual radiolabeled cDNAs. Of the 91 cDNAs screened, one third contained repetitive sequences, mostly B1 repeats; three contained ribosomal sequences. Of the ones that were free of repeats, five cDNAs were present more than once (2–4 times), comprising a total of 14 (15%) of the 91 cDNAs screened. Thus, the cDNA library following hybrid selection was of low complexity, suggesting that the selection was successful.

One would predict that if the cDNA hybrid selection were successful, the cDNA clones would hybridize to sequences in the library of genomic fragments selected for EVI-1 binding, and would hybridize only rarely to clones in the unselected genomic library. To assess this, we radiolabeled repeat-free selected cDNAs and hybridized these against filter lifts of either bacterial colonies containing selected EVI-1-binding plasmids, or those containing unselected library plasmids. Most cDNAs hybridized exclusively to filter lifts of the

selected library rather than the unselected one (data not shown), indicating that the genomic fragments that corresponded to these cDNAs were enriched during the EVI-1-binding selection.

Identification of an exon for Itp2 within a genomic fragment selected for EVI-1 binding

Our analysis of the library of selected cDNAs suggested that we had successfully derived a relatively low complexity library that was enriched for cDNAs that hybridized to the EVI-1-binding genomic fragments, and thus may represent EVI-1-regulated genes. We thus determined the sequence of 22 nonrepetitive cDNAs. The majority of these (17) showed no significant homology to sequences in the databases, and represented novel genes (Table 3). Three (JA36, JA66, and JA69) had significant homology to expressed sequence tags (ESTs; Table 3). More interestingly, three selected cDNAs showed significant homology to a known gene: JA15, JA18 and JA19 were nearly identical to rat *Itp2* (Sudhof et al., 1991), which encodes the inositol trisphosphate receptor, type 2. Since three *Itp2* cDNA clones were identified within the 91 clones analysed, it seemed unlikely that this was an artifactual result.

We confirmed that the cDNA JA18 was appropriately selected by identifying a corresponding genomic clone by colony hybridization screening of the EVI-1 binding fragment library that was used as bait in the cDNA hybrid selection. This clone, CG18 (for corresponding genomic for cDNA JA18) was found to have a 2.3 kb insert that, upon further analysis, was found to contain a 111 bp exon of *Itp2* and 296 bp 3' of this, an EVI-1 binding site (Figure 4). This exon encoded an open reading frame with nearly 100% identity to amino acids 2141–2150 of the rat *Itp2* gene (Sudhof et al., 1991). By comparing the

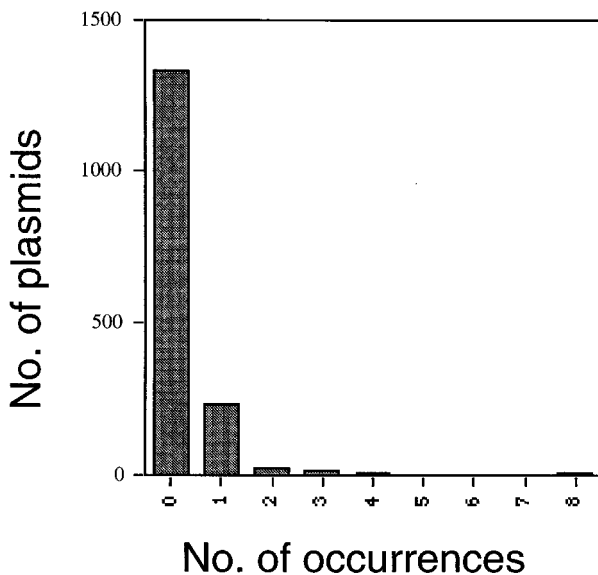


Figure 3 Distribution of EVI-1 binding plasmids relative to the number of occurrences in nine parallel selections. The number of occurrences is depicted on the horizontal axis, and the number of different plasmids is given on the vertical axis. The number of plasmids occurring zero times was calculated as described in the Materials and methods

Table 3 Sequence analysis of cDNAs isolated by hybrid selection using the EVI-1-binding genomic fragments as bait

	cDNA	Length	Homology	P value	GenBank accession number
1	JA008	274	none		AF031090
2	JA010	402	msy-3 gene LB9 mRNA	5×10^{-24} 1×10^{-29}	AF031091
3	JA016	260	none		AF031092
4	JA018	325	<i>Itp2</i>	5.7e-65	
5	JA024	241	none		AF031093
6	JA033	236	none		
7	JA036	337	Human EST, GenBank #R52089	3.4e-64	AF031094
8	JA045	261	none		AF031095
9	JA047	344	none		AF031096
10	JA053	170	none		AF031097
11	JA055	309	none		AF031098
12	JA057	124	none		AF031099
13	JA064	216	E. multilocularis U1 SN RNA	7×10^{-7}	AF031100
14	JA066	362	Human EST, GenBank #AA009478	5.2e-37	AF031101
15	JA067	321	none		AF031102
16	JA069	308	Mouse brain EST, GenBank #W51211	2.8e-25	AF031103
17	JA071	268	none		AF031104
18	JA072	310	none		AF031105
19	JA074	281	phospholipase A2 receptor	3.7e-06	AF031106
20	JA076	135	none		AF031107
21	JA084	239	none		AF031108
22	AB005	295	none		AF031109

Any homologies with known genes or expressed sequence tags (ESTs) and the associated P value are given, as were determined by BLAST searches of the DNA sequence databases. In the rightmost column is the GenBank Accession number for the selected cDNA sequence

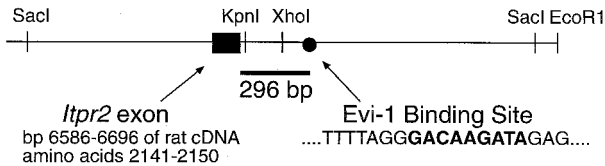


Figure 4 Map of the EVI-1 binding site in *Itp2*. The relative locations of the *Itp2* exon encoding amino acids 2141–2150 of the inositol trisphosphate receptor type 2, and the EVI-1 binding site are indicated. The sequence at the site, given in the Figure, is the same as that for clone KJ5 (Table 1, line 12)

sequence of CG18 to that obtained from EVI-1-binding genomic fragments, it was evident that fragment KJ5 represented an independently isolated, overlapping fragment of *Itp2* (Table 1, line 12). Furthermore, by *Sau3A1* fingerprint analysis, we had found that related clones had been isolated in eight of the nine independently initiated binding site selections (data not shown).

To directly demonstrate that EVI-1 binds to the GACAAGATA motif within CG18, we performed electromobility shift analysis using purified EVI-1 prepared in *E. coli*. EVI-1 interacted specifically with a radiolabeled oligonucleotide of 30 base pairs containing the GACAAGATA motif and adjacent bases (Figure 5): the binding could be competed with a different oligonucleotide containing the GACAAGATA motif, but not effectively by a competitor bearing a G₁→T mutation (Figure 5). These data show that the EVI-1 binds directly with the GACAAGATA motif within this intron of *Itp2*.

A chimeric EVI-1-derived transcriptional activator can transactivate the endogenous Itp2 gene

We were interested in determining if *Itp2* is a physiological target gene for EVI-1 within cells. It seemed likely that the EVI-1 binding site within *Itp2* was located many kilobases from the start of transcription since the exon adjacent to it in CG18 represents bp 6586–6696 of the rat cDNA (Sudhof *et al.*, 1991) (Figure 4). Binding sites for transcriptional regulatory factors are typically localized relatively close to the start of transcription. Thus, we considered it singular that the binding site for EVI-1 was situated adjacent to an exon that was likely at some distance from the start of transcription.

Nonetheless, we sought to determine if EVI-1 regulated the transcription of *Itp2*. To that end, we initially performed Northern blot analysis of RNA samples obtained from a variety of cell lines that are believed to contain EVI-1. When these blots were hybridized with a probe for the murine *Itp2*, there was no dramatic difference in level of *Itp2* transcript relative to control cell lines (data not shown). However, it has been shown that EVI-1 can act as a repressor (Kreider *et al.*, 1993; Perkins and Kim, 1996; Bartholomew *et al.*, 1997) or an activator (Morishita *et al.*, 1995) on artificial reporter constructs, depending on the promoter region, we could not predict what response to expect.

To gain support for the possibility that *Itp2* is a target gene of EVI-1 in mammalian cells, we tried to

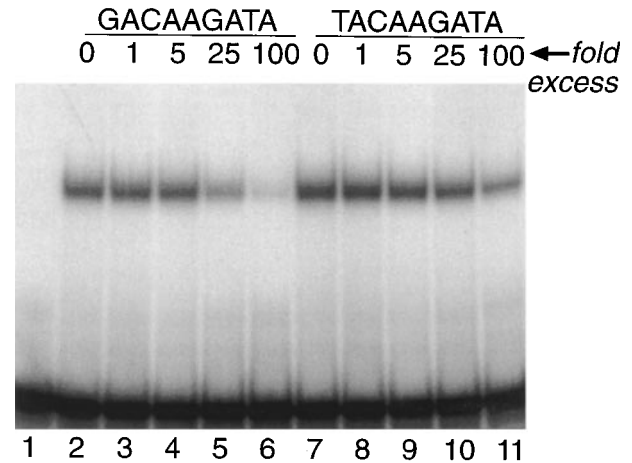


Figure 5 Electromobility shift assay using radiolabeled oligonucleotide probe with sequence derived from the EVI-1 binding site in *Itp2*. Lane 1: no added protein; lanes 2–11: added EVI-1 amino-terminal protein comprising zinc fingers 1–7. Unlabeled competitor oligonucleotide of the sequence indicated and in fold excess relative to probe as noted, was added in lanes 3–6 and 8–11

activate the endogenous *Itp2* gene with an EVI-1-VP16 fusion protein. We have described the construction and testing of EVI-1-VP16, a chimeric EVI-1-derived activator comprised of the zinc fingers 1–7 of EVI-1 fused to the transcriptional activation domain of VP16 (Perkins and Kim, 1996). We placed this construct under transcriptional control of *tet* operons (creating plasmid pHP-3) so that its expression could be controlled by tetracycline (Tc) via TetR-VP16 (Gossen and Bujard, 1992). We transferred pHP-3 into S2-6 cells, a mouse fibroblast cell line that expresses TetR-VP16 under control of *tet* operons so that its expression was positively autoregulated upon removal of Tc and highly inducible (Shockett *et al.*, 1995). We tested 24 different clonally derived cell lines for appropriate regulation of EVI-1-VP16 by Tc by transfecting in the 5EVBS-CAT reporter plasmid with five EVI-1 binding motifs upstream of minimal promoter from the HSV thymidine kinase gene, and assaying for activation of CAT activity in the presence (uninduced) or absence (induced) of Tc. In one line, 6D, EVI-1-VP16 activity was undetectable in the presence of Tc, but was induced several hundred fold upon Tc removal (Figure 6).

To see if EVI-1-VP16 could regulate the endogenous *Itp2*, in mouse fibroblasts, we harvested RNAs from 6D cells at various timepoints following removal of Tc, and performed Northern blot analysis, hybridizing for *Evi-1* or for *Itp2*. As a control, RNAs were also harvested from the parental S2-6 cell line that contained TetR-VP16 but not EVI-1-VP16. We found that EVI-1-VP16 was rapidly and maximally inducible within 8 h of Tc removal (Figure 7, compare lanes 1 and 2). When this same Northern blot was hybridized with a probe for *Itp2*, we saw no transcript in 6D cells at zero hours of induction, nor in the parental S2-6 line, with or without induction (Figure 7, lanes 1, 9 and 10). However, *Itp2* RNA was present following removal of Tc, as early as 8 h of induction (Figure 7, lane 2).

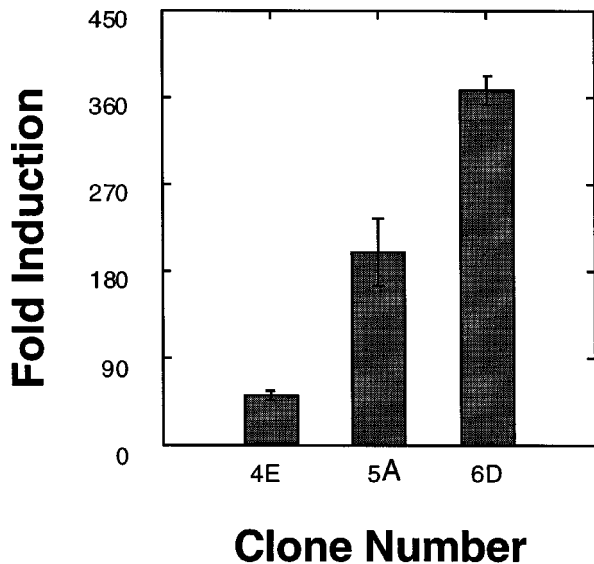


Figure 6 Fold induction of CAT reporter activity in three clonal S2-6-derived cell lines containing pHP-3, the Tc-regulated EVI-1-VP16 expression plasmid. Induced or uninduced cells were transfected with the 5EVBS-CAT construct in triplicate and assayed for CAT activity as described in Material and methods. The fold induction is the ratio of mean CAT activity induced cells to that in uninduced cell.

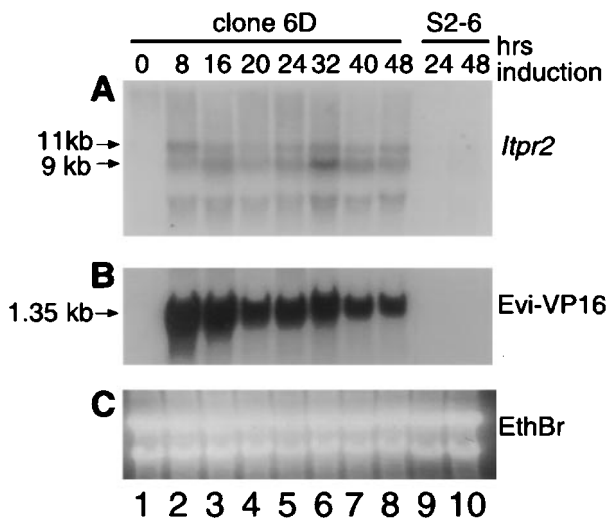


Figure 7 Induction of *Itp2* expression in clone 6D cells upon removal of Tc. Shown is Northern blot analysis (a and b) and ethidium bromide staining (c) of RNA samples from clone 6D and S2-6 cells, either uninduced (0 h induction) or induced for different time periods as indicated. (a) Shows hybridization with a murine *Itp2* probe; (b) hybridization with an *Evi-1* probe. Sizes of observed transcripts are indicated to the left

The induction of *Itp2* that we observed was not a reflection of widespread changes in gene transcription due to the expression of Evi-1-VP16: hybridization of Northern blots with RNA from induced and uninduced 6D cells with several probes, including novel cDNAs under consideration as Evi-1 targets, showed equal levels of expression in the two cell lines (data not shown). Equal loading of RNA in the lanes of the gel is indicated by the ethidium-stained 18S and 28S RNAs (Figure 7c).

Discussion

The identification of downstream target genes for transcriptional regulatory proteins is an important but infrequently approached question. One of the most direct and successful methods is to molecularly clone transcripts that are rapidly induced following the expression of active factor (El-Deiry *et al.*, 1993; Braun *et al.*, 1995; Polyak *et al.*, 1997). However, this approach depends on the factor being a transcriptional activator. When the factor is a repressor, differential cloning approaches become more problematic, since this requires a cell line in which the target genes are actively transcribed to begin with, and then are repressed upon expression of the repressor. This approach also requires that the factor be able to dramatically alter levels of transcription of the target genes. However, regulation of transcriptional is often dependent on combinations of factors that act in concert, and change in the level of one factor alone may not induce an effect on target gene transcription that is of sufficient magnitude to allow the transcripts of these loci to be differentially cloned. In this paper, we describe an approach to target gene cloning that does not rely on the activation properties of the factor, nor on the accessibility of the gene to the transcription factor. This selection is based on a physical selection for fragments of genomic DNA that contain binding sites, followed by identification of those that contain exons. We apply this technique to the isolation of candidate target genes for the zinc finger protein EVI-1, which has been shown to play an important role in myeloid leukemogenesis.

Creation of a binding fragment library

We used purified EVI-1 to select EVI-1-binding DNAs from a plasmid-based mouse genomic library. It is clear that the selection scheme that we have employed was successful in isolating high affinity EVI-1-binding fragments, based on: (1) efficient elimination of nonbinding pUC18 plasmid from the populations (Figure 2b); (2) competitive EVI-1 binding assays performed with selected plasmids (Figure 2c); (3) repeated isolation of the same plasmids in independently initiated selections; (4) successful further selection of *Sau3A1* subfragments derived from individual plasmids within the EVI-1-binding pool; (5) DNA sequence analysis of these *Sau3A1* fragments, which revealed the presence of the core EVI-1 binding motif, GACAAGATA (Table 1); and (6) *in vivo* induction of one selected gene *Itp2* by a chimeric EVI-1-derived activator.

While the per cent false positives in the selected library is not zero, we contend that it is 20% or less. This estimate is based on data from the selection of *Sau3A1* subclones for EVI-1 binding, in which selected plasmids are digested with *Sau3A1*, subcloned into pUC18 and the pool of subclones is then subjected to selection for EVI-1 binding. This analysis was performed on 47 plasmids. For over 80% of these plasmids we found that either single *Sau3A1* fragment was selected with cycles of filter binding with EVI-1, providing strong evidence of preferential binding to EVI-1, and/or a perfect EVI-1 consensus site was identified upon sequencing. This was true for DNA

fragments from plasmids that were identified repeatedly in the independent selections and also for DNAs from plasmids that were isolated in only one selection, which argues that we have obtained a population of plasmids with reasonably similar affinities for EVI-1, rather than a rarefied high affinity class with a lot of low affinity plasmids mixed in at lower representation. Thus, while not a pure population of EVI-1 binding fragments, the sublibrary we have created is remarkably enriched for such regions of DNA.

Enumeration of binding sites for EVI-1 per haploid genome

The EVI-1 core recognition site in DNA has been identified through site selection and further analysis as GACAAGATA (Perkins *et al.*, 1991; Delwel *et al.*, 1993; Perkins and Kim, 1996), which ought to occur by chance every 260 000 bp or 11 000 times within the 3×10^9 bp of mouse genomic DNA. Based on the number of recombinants in the initial plating (400 000) and the average insert size (3.5 kb), the starting library of mouse DNA has an estimated complexity of approximately 1.4×10^9 bp. Thus there is a probability of 0.37 that any sequence is present in the library. Within this library, the number of occurrences of a random ninemer is thus approximately $0.37 \times 11\ 000$ or 4070 sites.

We have performed nine parallel selections for plasmids bearing EVI-1-binding genomic fragments and analysed 576 randomly picked plasmids from the resultant selected pools by *Sau3A1* fingerprinting. This allowed us to identify 262 different plasmids, many of which were repeatedly present within each of the nine selections, and 31 that were present in two or more of the nine pools (see Table 2 and Figure 3). We have used these data to estimate the total number of EVI-1-binding fragments in the original library to be 1600. It is likely that our selected sublibrary of binding plasmids, comprising the plasmids from the nine independent selections, is a nearly complete representation of the EVI-1 binding plasmids in the original starting library. If one corrects for the 37% representation of the original library, there are 4300 EVI-1 binding sites per haploid mouse genome, which represents 40% of the total number of occurrences one would expect if they were present by chance (11 000 sites).

Short of an exhaustive analysis of all of the plasmids in the selected library, we can only estimate the number of false positives is 20% or lower, as discussed above. It is by far more likely that the false positives will occur in the group of plasmids that only occurred in one of the nine selections, and the net result of removing these from the Poisson distribution is to lower the estimate of binding sites. Thus, it is likely that our estimate of 4300 binding sites per haploid genome is the high end of the possible range.

A likely explanation for the discrepancy between the number of sites thought to occur by chance and the number of sites that we have calculated to exist, is that there has been a selection against random occurrence of the site during evolution, due to the biological activity of the EVI-1 binding motif as a site of action of transcription factors. It is also possible that EVI-1 binds with higher affinity to a subset of GACAAGA-

TAA sites that are located within a particular sequence context, and this explains the smaller number of sites that we arrived at. While analysis of the twenty binding sites that we identified (Table 1) fails to indicate any obvious contextual features that are similar between the sites and that might determine binding affinity, it is clear that a small number of sites have a much higher affinity for EVI-1: eight different plasmids were selected three times, two were selected four times, and one plasmid was selected in eight of the nine selections (Table 2). If one assumes that our calculation for the mean of the Poisson distribution is correct, then the expected numbers of plasmids for those numbers of occurrences is 1–2, zero, and 0, respectively. Indeed, the probability of isolating one plasmid in eight of nine selections is less than one in 10^7 , unless this plasmid has a higher affinity. This particular plasmid is from the *Itpr2* locus, and it is not at all clear why it exhibits such a remarkably high affinity for EVI-1. Nonetheless, this argues strongly that *Itpr2* is a true functional target for the protein.

Number of genes represented in EVI-1-binding sublibrary

We have used the selected library of EVI-1 binding plasmids to isolate corresponding cDNAs by performing cDNA hybrid selection. Given the calculated presence of 1600 different clones in our sublibrary of genomic fragments that bind EVI-1, and an average insert size of 3.5 kb, there is likely over 5600 kb of complexity in the library. Exons represent 2% of genomic DNA and have a length of 200 bp and thus occur on average every ten kilobases. Thus in this sublibrary, there may be as many as 560 exons. There may be more, if one considers that selection for binding sites may not be independent of selection for exons. Within the initial 91 clones, we have identified 22 different cDNAs, and we estimate that 60–70% of these are present in the selected library due to hybridization to exons within the genomic DNA; the remainder were retained through the cDNA hybrid selection nonspecifically. These data are consistent with those of Weissman and coworkers (Parimoo *et al.*, 1993), who found that 80% of their selected cDNA clones hybridized back to the 'bait' genomic DNA that was used to select them. Exhaustive retrieval of cDNAs from this pool will likely yield cDNA fragments that represent nearly all of the exon sequences located within our EVI-1-binding genomic fragments. Of course, to be present within the collection of selected cDNAs the corresponding mRNA must be present in brain RNA. While brain represents a good source of high complexity mRNA, not all messages are present in this tissue. Thus, to completely exhaust the capture exon-hybridizing cDNAs, the cDNA hybrid selection will need to be repeated with a cDNA library of a different source, such as embryo.

Use of Tc-regulated EVI-1-VP16

While the finding of cDNAs for genes located in *cis* to EVI-1 binding sites certainly narrows the search for EVI-1 target genes, there is no assurance that all of these represent true targets; an unquantifiable fraction of the EVI-1 binding sites are very likely to be near genes, yet nonfunctional. To approach this, we devised

a cell culture system in which we can rapidly induce the expression of EVI-1-VP16, a chimeric activator with the binding specificity of zinc fingers 1–7 and the potent transcriptional activating properties of VP16, under the control of Tc-regulated promoter (Gossen and Bujard, 1992). We used this system to prepare Northern blots of RNAs from induced and uninduced cells, which were hybridized with radiolabeled DNA probes for potential target genes. In addition, we have used this system directly to isolate potential target EVI-1 target genes, through the different cloning of mRNAs that are rapidly induced by EVI-1-VP16 (Hui and Perkins, in preparation).

Itp2 as a candidate target gene for EVI-1

Using this Tc-regulated system, we were able to show that EVI-1-VP16, a potent EVI-1-derived transcriptional activator, can regulate endogenous *Itp2* transcription. While this analysis does not assure that native EVI-1 can regulate the candidate target gene, it provides support for that possibility since it suggests that the EVI-1-VP16 chimera has access to and can bind *in cis* to the gene, and that it can influence the basal transcriptional machinery from that site. Furthermore, the genomic fragment of *Itp2* exhibited a very high affinity for EVI-1, being selected in eight of the nine independent selections for EVI-1 binding, and exhibiting specific binding for EVI-1 in a electromobility shift assay. These data provide compelling suggestive evidence that native EVI-1 also binds to this site in *Itp2 in vivo* and thereby directly influences *Itp2* transcription. This is a remarkable finding since it is highly likely that the binding site for EVI-1 that we have identified is located at some distance from the start of transcription. While being at such a distance is not the most common arrangement for the binding site of a transcriptional regulator, there are numerous examples, such as the locus control region within the globin locus, which is located 50 kb from some of the genes it regulates (Dillon and Grosfeld, 1993). That we were able to identify this site attests to the power of our approach, since it is very unlikely that this binding site could have been identified by commonly employed approaches such as promoter analysis, which is typically confined to sequences close to the start of transcription.

Is *Itp2* a physiologic target for native EVI-1? To answer this will require additional experiments that may entail the use of a particular cell type in which *Itp2* transcription is critically dependent on EVI-1. Nonetheless, *Itp2* provides an attractive candidate target gene for EVI-1. The inositol trisphosphate receptor (IP3R) resides in both plasma and endoplasmic reticulum membranes, where it mediates the release of calcium ion upon binding of IP3 or IP4, which, in turn, are generated by hydrolysis of membrane substrates by phospholipase C (PLC). PLC activation can occur through either G-linked receptors, or through tyrosine kinase receptors. *Itp2* may be regulated during normal myelopoiesis, since it is inducible in several myeloid cell differentiation models (Sugiyama *et al.*, 1994). The increase in its encoded protein, the inositol trisphosphate receptor, type 2 (IP3R, type 2), is thought to be important in the functional maturation of granulocytes. For instance, IP3R is thought to play a critical role in signaling from

the F-met-leu-phe receptor, which is important in mature granulocytes for the detection of bacterial proteins. In mature granulocytes, the resultant rise in intracellular calcium activates granulocyte degranulation, and secretion of bacteriolytic enzymes (Bradford *et al.*, 1992), leukemic cells are typically defective in these calcium-mediated responses. *Itp2* transcription is known to be activated by both retinoic acid, and by phorbol esters in cultured myeloid cells (Sugiyama *et al.*, 1994), and may also be activated during normal myelopoiesis, although this has not been reported. It is possible that overexpression of EVI-1 in leukemic cells negatively effects the inositol-calcium signaling pathway, perhaps through repressing transcription of *Itp2*, although this remains to be shown.

Absence of iterated GATA motif in selected Evi-1-binding fragments

A selection for EVI-1-binding genomic fragments has been reported, in which the amino-terminal zinc finger domain of EVI-1, as a GST fusion, was bound to nitrocellulose and this was used to isolate EVI-1-binding *Sau3A1* fragments of mouse DNA (Matsugi *et al.*, 1995). Sequence analysis of these fragments failed to reveal the complete GACAAGATA motif, but within several of the selected fragments, did identify repeats of the sequence GATA, to which GST/Evi-1 was able to bind in a DNaseI footprint analysis (Matsugi *et al.*, 1995). None of the genomic fragments that we isolated contained such repeats of the GATA sequence, an outcome that is likely due to differences in the approach that we employed. It is difficult to know if there are any sequences in common between the Evi-1-binding fragments that we identified and those reported by Matsugi *et al.* (1995) since the latter group restricted the sequence data to that comprising the GATA motifs. They were able to show that one of the genes that they identified (termed clone 150) was expressed at lower levels in leukemic cells expressing *Evi-1*. Whether this was due directly to Evi-1 action was not shown (Matsugi *et al.*, 1995).

Materials and methods

Purification of EVI-1

Truncated EVI-1 protein [MBP-EVI-1(1–254)], comprising amino acids 1–254 was purified from *E. coli* as a fusion with the maltose binding protein as described (Perkins *et al.*, 1991). Full length EVI-1, tagged at the carboxyl terminus with six histidines (EVI-1H₆), was isolated from Sf9 insect cells by nickel column chromatography by first creating a recombinant EVI-1 H₆ baculovirus expression plasmid as follows. The codons encoding the six histidine residues were added to the 3' end of the gene by polymerase chain reaction (PCR), using the following primers: (sense primer: 5'-CACAGGCATATGCTATGATG-3'; and antisense primer: 5'-ATTAGGTACCTCATCAGTGATGGTGATGGTGATGTACATGGCTTATGGACTGGATGGC-3'). This amplified a 170 bp piece of DNA extending from bp 3467 to bp 3603 of *Evi-1* (Morishita *et al.*, 1988), which was ligated into pCRII. From an appropriate recombinant, a 170 bp *NdeI*–*KpnI* fragment was isolated. This was ligated into *BamHI* and *KpnI*-digested pVL941 (a baculovirus expression vector (Luckow and Summers, 1988)) together with a *BamHI*–

NdeI fragment of pKS-Evi-1. pKS-Evi-1 was generated by amplifying the *Evi-1* gene from p58.2-1 (Morishita *et al.*, 1988) by PCR (a 50 μ l reaction containing 200 μ M dNTPs, 1 μ M of each primer, 1 ng template DNA, 50 mM KCl, 10 mM Tris, pH 8.3, 1.5 mM MgCl₂, 0.01% gelatin and 2.5 U Amplitaq (Perkin-Elmer)) using sense primer 5'-CGGAATTCGATGAAGAGTGAAGAGGACCCG-3' and antisense primer 5'-GCATCGATCCACTCTGGTCAACCTTGACAA-3', digesting the 3.6 kb product with *EcoRI* and *ClaI*, and ligating into similarly cut pBluescript KS⁺ (Stratagene). Using this strategy, the putative initiator methionine at bp 478 of EVI-1 (Morishita *et al.*, 1988) was situated 62 bp from the 5' end of the RNA transcript of the pVL941 vector. This recombinant (called pVL-Evi-1-His) was analysed by restriction enzyme mapping and partial DNA sequencing. pVL-Evi-1-His was cotransfected with Baculogold DNA (Pharmingen) into Sf9 cells as described (Summers and Smith, 1987). Viral supernatant from transfected cells was cloned by limiting dilution infection of Sf9 cells in 96 well microtiter plates as described (Summers and Smith, 1987). Positive wells were identified by dot blotting of Sf9 cells (Kafatos *et al.*, 1979), and hybridizing with a probe for *Evi-1*. One of these was subjected to another round of cloning, and was then expanded to create a high titer virus stock. This was used to infect 8×10^8 Sf9 cells, which were cultured in TNM-FH medium (Summers and Smith, 1987) at a density of 10^6 cells per ml in a 1 litre stir flask at 27°C for 72 h. The infected cells were washed with ice cold phosphate-buffered saline (PBS), and lysed in hypotonic lysis buffer (10 mM NaPO₄, pH 7, 10 mM HEPES, pH 7.9). Nuclei were isolated by centrifugation at 4°C, 900 g, 10 min. The pellet was resuspended in 50 ml of 5 M guanidine, 50 mM NaPO₄, pH 8.0 (GP8 buffer). This solution was stirred at 4°C for 15 min, clarified by centrifugation at 30 000 g, and passed over a nickel affinity matrix (NTA, Qiagen) at 4°C. Following elution of nonbinding proteins with five column volumes of GP8 buffer, proteins were eluted with a pH 8 to 3.5 gradient in five column volumes. Peak fractions were dialyzed against 20% glycerol, 20 mM NaPO₄, pH 7.0, 400 mM KCl 0.1% NP40, 5 mM MgCl₂, 50 μ M ZnCl₂, 0.2 mM phenylmethyl sulfonyl fluoride. The protein was stored in 0.2 ml aliquots at -80°C.

Electromobility shift assays

EVI-1 binding to a double stranded radiolabeled oligonucleotide containing the EVI-1 binding site (annealed primers 5'-GGATCTCCGTGACAAGATAAGGATTCCCTG-3' and 5'-GGCAGGGAATCCTTATCTTGTACAGGAGAT-3'; labeled with ³²P- γ -ATP and polynucleotide Kinase (NEB) or with ³²P- α -dCTP and Klenow) was assayed on low-ionic strength nondenaturing PAGE as described (Chodosh *et al.*, 1986). DNA binding reactions (20 μ l; 25 mM HEPES, pH 7.5, 50 mM KCl, 4 mM MgCl₂, 1 mM DTT, 20% glycerol, 250 μ g/ml bovine serum albumin and 250 μ g/ml poly dIdC (Pharmacia)) contained 5500 c.p.m. of ³²P-labeled oligonucleotide (10 fmol) and 250 ng of MBP-Evi-1(1-254). For the electromobility shift assay in Figure 5, the radiolabeled oligonucleotide had sequence derived from *Itpr2*: 5'-GTATTTTGTGACAA-GATAGTAGAGAATCT-3'.

Creation of binding site sublibrary

We employed the technique of Sompayrac and Danna (1990) with several modifications. A C3H mouse genomic DNA library based in a modified pSP64 vector (3.5 kb average insert), was a generous gift from L Sompayrac. Plasmid DNA comprising the entire library was purified by alkaline lysis and CsCl gradient centrifugation (Sambrook *et al.*, 1989) 2.75 μ g of library DNA was mixed with 4 μ g

of purified EVI-1 protein in a 500 μ l reaction containing 25 mM HEPES (pH 7.5), 50 mM KCl, 4 mM MgCl₂, 250 μ g/ml BSA, 48 μ g/ml poly dIdC (Pharmacia), 1 mM dithiothreitol and 10% glycerol. After a 30 min incubation at 30°C, DNA-protein complexes were isolated by filtration through nitrocellulose (McEntee *et al.*, 1980) at a flow rate of 1 ml/min followed by two 0.5 ml washes with wash buffer (25 mM HEPES (pH 7.5), 50 mM KCl, 4 mM MgCl₂ and 1 mM dithiothreitol). DNA was eluted from the membrane by incubating for 2 h at 42°C with 0.9 ml of 0.2% SDS, 20 mM Tris (pH 7.8) and 0.3 M sodium acetate, precipitated with three volumes of ethanol, and then used to transform *E. coli* by electroporation. Plasmid was isolated from ampicillin-resistant colonies, purified by Qiagen kit, and then subjected to another round of selection. To determine the fold enrichment for EVI-1-binding plasmids, the C3H library plasmids (*lacZ*⁻) were doped with pUC18 (*lacZ*⁺). The fold enrichment was calculated by (number of *lacZ*⁻)/number of *lacZ*⁺ obtained on indicator plates for filtration reactions with EVI-1, divided by the same ratio obtained on plates with reactions containing the same input plasmids but no protein.

Calculation of number of plasmids selected

The selection for EVI-1 binding plasmids was performed nine independent times, starting each time with the entire complexity of the plasmid-based C3H genomic library. Five hundred and seventy-six minipreps were prepared of plasmids that were selected in these nine isolation procedures. Miniprep DNAs were digested with *Sau3A1*, fractionated on ethidium bromide-stained agarose gels and photographed. This provided a unique pattern of fragment sizes or *Sau3A1* 'fingerprint' for each clone, allowing it to be identified and distinguished from other clones. This analysis revealed that there were 262 different plasmids isolated from the nine selections, some of which were present in more than one of the nine selection experiments. The repeated isolation of particular plasmids in two or more independently initiated experiments indicated a finite number of binding sites in the starting library. Tabulation of the number of different plasmids relative to the number of times they were isolated (Table 2) revealed the relationship had the form of a Poisson distribution (Figure 3), which can be described by the equation $N \cdot P(k)$ where N is the total number of EVI-1 binding plasmids in the library and $P(k)$ is equal to $m^k e^{-m} / k!$, where m is the mean of the Poisson distribution, and k is the number of occurrences for which one wishes to determine the probability (Mohler, 1977; Wang and Brown, 1991). The total number of EVI-1 binding plasmids (N) equals the number that were identified in the miniprep analysis (262 total) plus the number that occurred zero times. The latter number is equal to $N \cdot P(0)$, and can be derived using the data in Table 2 and the equation given as follows:

$$\text{To calculate } m : \frac{\text{no of fragments with 2 occurrences}}{\text{no of fragments with 1 occurrence}} = \frac{Nm^2 e^{-m}}{2Nme^{-m}} = \frac{m}{2}$$

or, for our data, $m = 2 \left(\frac{20}{231} \right) = 0.173$.

Using m , one can calculate that the number that occurred zero times, i.e., those that were not identified:

$$Ne^{-m} = \frac{\text{no of fragments with one occurrence}}{m} = \frac{231}{0.173} = 1334 \text{ fragments}$$

The total number of EVI-1-binding fragments in the starting library can be estimated as $1334 + 262 = 1596$.

cDNA hybrid selection

The selection of cDNAs was performed as described (Parimoo *et al.*, 1993) with the following modifications. A short insert mouse brain cDNA library was prepared in cloning vector lambda GT10 from poly A-selected RNA from mouse brain using standard techniques (Sambrook *et al.*, 1989), except that in the reverse transcription step, random hexamers were employed to prime the synthesis of cDNA. Two hundred nanograms of the target sequences for the cDNA hybrid selection (purified plasmid DNAs from the binding site selections) were immobilized on Hybond (Amersham). Quenching reagents were *E. coli* ribosomal sequences (pKK3535), human ribosomal sequences (plasmids pR7.3 and pR5.8), mouse C₀t1 DNA (Gibco-BRL), and poly dI.dC (Pharmacia). Hybridization, washing, PCR amplification, cloning, and sequencing of selected cDNAs were performed as described (Parimoo *et al.*, 1993). The corresponding genomic fragments for certain cDNAs were identified by colony hybridization. Bacteria containing plasmids with inserts of genomic DNA that were selected for EVI-1 binding were gridded on 150 mm LB-amp agar plates. Nitrocellulose lifts of these hybridized with radiolabeled cDNA fragments that were derived from individual cDNA lambda clones by PCR as described (Parimoo *et al.*, 1993) using standard conditions (Sambrook *et al.*, 1989). The presence of an exon corresponding to the cDNA was confirmed by shotgun subcloning of *Sau3A1* fragments and colony hybridization, using the appropriate cDNA fragment as a probe. To identify the EVI-1 binding site within the corresponding genomic clone, the plasmid DNA was digested with *Sau3A1*, shotgun ligated into *Bam*HI-cut pUC18, and the resultant library of fragments was selected for EVI-1-binding fragments by several rounds of nitrocellulose filter binding, as described above. Following 3–4 rounds of selection and amplification, individual plasmid clones were analysed by restriction enzyme and gel electrophoresis. The predominant species typically harbored an EVI-1-binding site, as identified through DNA sequence determination and analysis. DNA sequencing was performed by the Keck Biotechnology Center at Yale University, using an automated ABI Sequencer.

Construction and testing of the 6D cell line containing tetracycline-regulated EVI-1-VP16

Plasmid construction A tetracycline (Tc)-responsive expression construct driving *EVI-1-VP16* (Perkins and Kim, 1996) expression was generated by inserting the *EVI-1-VP16* moiety of *pEVI-1-VP16* (Perkins and Kim, 1996) as a *Hind*III–*Xba*I fragment into the pUHD10-3 vector ((Gossen and Bujard, 1992), a Tc-responsive expression vector, kindly provided by H Bujard). This construct, designated as pHP-3, places *EVI-1-VP16* under transcriptional control of the *Tet* operons. An EVI-1-responsive reporter construct (p5EVBS1-CAT) was generated as follows. Complementary oligonucleotides (5'-GATCTGCAAGATAAGATAAG-3' and 5'-GATCCTTATCTTATCTTGTC-3') were concatemerized as described (Kadonaga and Tjian, 1986): 22 μ g of each oligonucleotide were annealed in 75 μ l containing 66 mM TrisHCl, pH 7.6, 13.3 mM MgCl₂, 6 mM dithiothreitol (DTT), 1.3 mM spermidine and 1.3 mM EDTA, by heating to 88°C and slow cooling to room temperature. The reaction was brought to 100 μ l and adjusted to 3 mM ATP. 5 μ Ci [³²P]- γ -ATP and 50 units of polynucleotide kinase (New England Biolabs (NEB)) were added, and the reaction was incubated at 37°C for 2 h. The DNA was purified by

phenol:chloroform extraction, and then precipitated as described (Kadonaga and Tjian, 1986). The DNA was resuspended in to 50 μ l and concatemerized with T4 DNA ligase (2 μ l, NEB; 12 h incubation at 4°C. The DNA was again purified and precipitated as described above, resuspended to 200 μ l, and digested at 37°C with 114 U each of *Bam*HI and *Bg*III (NEB) in the manufacturer's buffer. The multimers were purified on an 8% nondenaturing PAGE buffered with 1 \times TBE (89 mM Tris base/89 mM boric acid/2 mM EDTA). Following visualization by autoradiography, bands were excised, the DNA was eluted, and ligated into *Bam*HI-digested, dephosphorylated pTK-CAT (Miksicek *et al.*, 1986). Clones were sequenced with M13 forward primer, which revealed the presence of five tandem copies of the binding site oligonucleotide monomer, all in head-to-tail orientation, in one isolate, designated 5EVBS-CAT.

Transfection and selection of cell lines NIH3T3 cells expressing pTet-tAK (designated as S2-6 cells) were generously provided by DG Schatz (Shockett *et al.*, 1995). S2-6 Cells were maintained in the histidine-deficient Dulbecco's Modified Eagle's (DME, Irvine Scientific Inc.), containing l-histidinol (0.5 μ M) and Tc (0.5 μ g/ml), at 37°C/5% CO₂. Stable S2-6 cells containing pHP-3 were generated by cotransfection of pHP-3 with pBABE-puro (Morgenstern and Land, 1990) followed by puromycin selection (2 μ g/ml). Single colonies were expanded in medium containing puromycin and Tc (0.5 μ g/ml) to obtain individual clonal cell lines. For induction of EVI-1-VP16 expression, cells were washed two times with PBS and then cultured in deficient DME medium without Tc for various times. Transient transfection with CAT reporter construct (p5EVBS-CAT) was performed with individual cell lines after 48 h of induction of EVI-1-VP16 (achieved with removal of Tc), followed by CAT assay of the cell lysates, which were performed as described (Sambrook *et al.*, 1989). A *lacZ* expression construct, driven by a CMV promoter (pCMV β ; Clontech), was included in the transfection as an internal control for transfection efficiency. Cell transfections were performed with a total of 15 μ g of DNA/10 cm plate by the calcium phosphate precipitate method (Graham and vanderEb 1973) with 2 min of 'shock' with 15% glycerol at 4 h after transfection (Sambrook *et al.*, 1989). Cells were assayed 24–48 h later. All assays were done in triplicate. CAT and β -galactosidase assays were performed as described (Sambrook *et al.*, 1989).

RNA analysis RNA was isolated from cells as described (Chomczynski and Sacchi, 1987) using the Trizol reagent (Gibco/BRL). Northern blots were performed with 30 μ g of total cellular RNAs which were fractionated on formaldehyde agarose gels, and blotted to nylon membranes (Hybond, Amersham) as described (Sambrook *et al.*, 1989). The blots were prehybridized and hybridized in 7% sodium dodecyl sulfate, 0.5 M NaPO₄, pH 7.5, 1 mM EDTA (Church and Gilbert, 1984) at 65°C. DNA hybridization probes were radiolabeled with ³²P- α -dCTP by random-primed synthesis (Feinberg and Vogelstein, 1983). The probe for *Evi-1-VP16* was the 0.75kb *Eco*RI fragment of *Evi-1-VP16*, which comprised sequences encoding zinc fingers 1–7 of *Evi-1*. The probe for murine *Itpr2* was a partial cDNA kindly provided by H deSmedt.

Acknowledgements

We thank Tracy Zhu for creating the VL-Evi-1-His construct; Youlan Li for deriving the AcEvi-1-His virus from this construct; Nagy Ayad for the initial analysis of the Evi-1-His protein; Lauren Sompayrac for providing the C3H plasmid library; SR Patanjali and Sherman Weissman for providing plasmids pR7.3, p5.8, and pKK3535;

H Bujard for providing the pUHD10-3 plasmid; David Schatz and Penny Shockett for providing the S2-6 cell line; Eric Fearon for providing the pNL-VP16 plasmid;

Humbert deSmedt for providing the probe for *Itpr2*; Mark Biggin for a critical reading of the manuscript; and Heping Zhang for reviewing the statistical analysis.

References

- Bartholomew C, Kilbey A, Clark A and Walker M. (1997). *Oncogene*, **14**, 569–577.
- Bradford P, Wang X, Jin Y and Hui P. (1992). *J. Biol. Chem.*, **267**, 20959–20964.
- Braun D, Reiden R, Lessnick S, May W and Denny C. (1995). *Mol. Cell Biol.*, **15**, 4623–4630.
- Chodosh LA, Carthew RW and Sharp PA. (1986). *Mol. Cell Biol.*, **16**, 4723–4733.
- Chomczynski P and Sacchi N. (1987). *Anal. Biochem.*, **162**, 156–159.
- Church G and Gilbert W. (1984). *Proc. Natl. Acad. Sci. USA*, **81**, 1991–1995.
- Delwel R, Funabiki T, Kreider B, Morishita K and Ihle J. (1993). *Mol. Cell Biol.*, **13**, 4291–4300.
- Dillon N and Grosveld F. (1993). *Trends Genet.*, **9**, 134–137.
- El-Deiry W, Tokino T, Velculescu V, Levy D, Parsons R, Trent J, Lin D, Mercer W, Kinzler K and Vogelstein B. (1993) *Cell*, **75**, 817–825.
- Feinberg A and Vogelstein B. (1983). *Analyt. Biochem.*, **132**, 6–13.
- Fichelson S, Dreyfus F, Berger R, Melle J, Bastard C, Miclea J and Gisselbrecht S. (1992). *Leukemia*, **6**, 93–99.
- Funabiki T, Kreider BL and Ihle JN. (1994). *Oncogene*, **9**, 1575–1581.
- Gossen M and Bujard H. (1992). *Proc. Natl. Acad. Sci. USA*, **89**, 5547–5551.
- Graham F and vanderEb A. (1973). *Virology*, **52**, 456–467.
- Hoyt PR, Bartholemew C, Davis AJ, Yutzey K, Gamer LW, Potter SS, Ihle JN and Mucenski ML. (1997). *Mechanisms of Development*, **65**, 55–70.
- Kadonaga JT and Tjian R. (1986). *Proc. Natl. Acad. Sci. USA*, **83**, 5889–5893.
- Kafatos F, Jones C and Efstratiadis A. (1979). *Nucl. Acids Res.*, **7**, 1541–1552.
- Kinzler KW and Vogelstein B. (1989). *Nucleic Acids Res.*, **17**, 3645–3653.
- Kreider B, Orkin S and Ihle J. (1993). *Proc. Natl. Acad. Sci. USA*, **90**, 6454–6458.
- Levy ER, Parganas E, Morishita K, Fichelson S, James L, Oscier D, Gisselbrecht S, Ihle JN and Buckle VJ. (1994). *Blood*, **83**, 1348–1354.
- Luckow VA and Summers MD. (1988). *BIO/TECH.*, **6**, 47–55.
- Matsugi T, Kreider B, Delwel R, Cleveland J, Askew D and Ihle J. (1995). *Oncogene*, **11**, 191–198.
- McEntee K, Weinstock G and Lehman I. (1980). *PNAS.*, **77**, 857–861.
- Miksicek R, Heber A, Schmid W, Danesch U, Posseckert G, Beato M and Schutz G. (1986). *Cells*, **46**, 283–290.
- Mitani K, Ogawa S, Tanaka T, Miyoshi H, Kurokawa M, Mano H, Yazaki Y, Ohki M and Hirai H. (1994). *EMBO J.*, **13**, 504–510.
- Mohler J. (1977). *Genetics*, **85**, 259–272.
- Morgenstern J and Land H. (1990). *Nucl. Acids Res.*, **18**, 3587–3596.
- Morishita K, Parganas E, Matsugi T and Ihle JN. (1992). *Mol. Cell Biol.*, **12**, 183–189.
- Morishita K, Parganas E, Willman CL, Whittaker MH, Drabkin H, Oval J, Taetle R, Valentine MB and Ihle JN. (1992). *Proc. Natl. Acad. Sci. USA*, **89**, 3937–3941.
- Morishita K, Parker DS, Mucenski ML, Jenkins NA, Copeland NG and Ihle JN. (1988). *Cell*, **54**, 831–840.
- Morishita K, Suzukawa K, Taki T, Ihle JN, and Yokota J. (1995). *Oncogene*, **10**, 1961–1967.
- Mucenski ML, Taylor BA, Ihle JN, Hartley JW, Morse III HC, Jenkins NA and Copeland NG. (1988). *Mol. Cell Biol.*, **8**, 301–308.
- Nucifora G, Begy CR, Kobayashi H, Roulston D, Claxton D, Pedersen-Bjergaard J, Parganas E, Ihle JN and Rowley JD. (1994). *Proc. Natl. Acad. Sci. USA*, **91**, 4004–4008.
- Parimoo S, Patanjali S and Weissman S. (1993). *Meth. Mol. Genet.*, **1**, 23–50.
- Peeters P, Wlodarska I, Baens M, Criel A, Selleslag D, Hagemeyer A, Van den Berghe H and Marynen P. (1997). *Cancer Res.*, **57**, 564–569.
- Perkins AS and Kim J. (1996). *J. Biol. Chem.*, **271**, 1104–1110.
- Perkins AS, Fishel R, Jenkins NA and Copeland NG. (1991). *Mol. Cell Biol.*, **11**, 2665–2674.
- Perkins AS, Mercer JA, Jenkins NA and Copeland NG. (1991). *Development*, **111**, 479–487.
- Polyak K, Xia Y, Zweier J, Kinzler K and Vogelstein B. (1997). *Nature*, **389**, 300–305.
- Sambrook J, Fritsch E and Maniatis T. (1989). *Molecular Cloning, A Laboratory Manual*. Cold Spring Harbor, NY, Cold Spring Harbor Laboratory.
- Shockett P, Diflippantonio M, Hellman and Schatz D. (1995). *Proc. Natl. Acad. Sci. USA*, **92**, 6522–6526.
- Sompayrac L and Danna K. (1990). *Proc. Natl. Acad. Sci. USA*, **87**, 3274–3278.
- Sudhof T, Newton C, Archer B, Ushkaryov Y, and Mignery G. (1991). *EMBO J.*, **10**, 3199–3206.
- Sugiyama T, Yamamoto-Hino M, Miyawaki A, Furuichi T, Mikoshiba K and Hasegawa M. (1994). *FEBS Lett.*, **349**, 191–196.
- Summers MD and Smith GE. (1987). *Texas Agri. Exper. Station Bull.* 1555.
- Suzukawa K, Parganas E, Gajjar A, Abe T, Takahashi S, Tani K, Asano S, Asou H, Kamada N, Yokota J, Morishita K and Ihle JN. (1994). *Blood*, **84**, 2681–2688.
- Tanaka T, Nishida J, Mitani K, Ogawa S, Yazaki Y and Hirai H. (1994). *J. Biol. Chem.*, **269**, 24020–24026.
- Tsang SY, Nakanishi M and Peterlin B. (1990). *Mol. Cell Biol.*, **10**, 711–719.
- Wang Z and Brown D. (1991). *Proc. Natl. Acad. Sci. USA*, **88**, 11505–11509.

