LETTERS

Cancer gene discovery in solid tumours using transposon-based somatic mutagenesis in the mouse

Lara S. Collier[1]*, Corey M. Carlson[1]*, Shruthi Ravimohan[1]†, Adam J. Dupuy[1]† & David A. Largaespada[1]

Retroviruses, acting as somatic cell insertional mutagens, have been widely used to identify cancer genes in the haematopoietic system and mammary gland[1,2]. An insertional mutagen for use in other mouse somatic cells would facilitate the identification of genes involved in tumour formation in a wider variety of tissues. Here we report the ability of the *Sleeping Beauty* transposon to act as a somatic insertional mutagen to identify genes involved in solid tumour formation. A *Sleeping Beauty* transposon, engineered to elicit loss-of-function or gain-of-function mutations, transposed in all somatic tissues tested and accelerated tumour formation in mice predisposed to cancer. Cloning transposon insertion sites from these tumours revealed the presence of common integration sites, at known and candidate cancer genes, similar to those observed in retroviral mutagenesis screens. *Sleeping Beauty* is a new tool for unbiased, forward genetic screens for cancer genes *in vivo*.

DNA 'cut-and-paste' transposons have recently been developed for use in vertebrates. Dormant *Tc1/mariner*-family transposable elements in the genomes of salmonid fish have been identified[3]. Directed mutagenesis was used to correct mutations that had silenced the activity of the transposases. A resurrected transposable element, called *Sleeping Beauty* (SB), was shown to be active *in vivo*[3]. SB functions as a germline insertional mutagen[4–6], but the use of SB in the mouse soma has previously been limited to gene therapy studies[7,8]. Here we show that chromosomally resident SB transposons can contribute to cancer development when mobilized in the soma of SB10 transposase transgenic mice.

A SB transposon, called T2/Onc, was engineered to induce both loss-of-function and gain-of-function mutations (Fig. 1a). T2/Onc contains splice acceptors followed by a polyadenylation signal in both orientations to intercept upstream splice donors after intronic insertion to generate loss-of-function mutations. Between the two splice acceptors are sequences from the 5′ long terminal repeat (LTR) of the murine stem cell virus (MSCV) that contain strong promoter and enhancer elements that have been shown to be active in stem cells[9–11]. Immediately downstream of the LTR is a splice donor for splicing of a transcript initiated from the LTR into downstream exons of endogenous genes. Two lines (nos 68 and 76) of T2/Onc transgenic mice were used for analysis (see Methods).

The ability of T2/Onc to mobilize in the soma was tested by breeding T2/Onc transgenic animals to transgenic mice expressing the SB transposase regulated by the ubiquitous CAGGS promoter[12] (CAGGS-SB10)[4]. Excision of T2/Onc from the concatemer was detected in every somatic tissue tested from T2/Onc;CAGGS-SB10 doubly transgenic animals but not in tissue from singly transgenic controls (Fig. 1b), whereas Southern blotting of normal tissue revealed few or no clonal, somatically acquired, T2/Onc insertions (Supplementary Fig. 1a). New subclonal T2/Onc insertions ($n = 12$) could be cloned from doubly transgenic somatic genomic DNA (data not shown). These data showed that SB transposition occurs readily in somatic cells.

Mice doubly transgenic for both T2/Onc and CAGGS-SB10 were aged for more than 1 year ($n = 26$) but did not show evidence of cancer susceptibility different from background (data not shown). We proposed that somatic T2/Onc mobilization by CAGGS-SB10 alone is insufficient to promote rapid, highly penetrant tumour formation in wild-type animals but might accelerate tumorigenesis in animals that are predisposed to cancer. Both T2/Onc concatemers and CAGGS-SB10 were crossed to $Arf^{-/-}$ mice, animals deficient for the p53 pathway regulator and tumour suppressor p19*Arf* (ref. 13). Mice carrying either or both of the T2/Onc and CAGGS-SB10 transgenes were generated on the $Arf^{-/-}$ background. The total number (n) and genotype of each group was as follows: $Arf^{-/-}$;T2/Onc ($n = 54$), $Arf^{-/-}$;CAGGS-SB10 ($n = 48$)

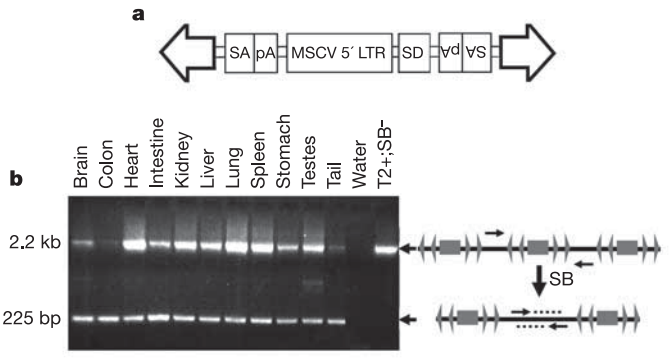

Figure 1 | Vector design and somatic transposition. a, The T2/Onc transposon contains elements to elicit transcriptional activation (MSCV 5′ LTR and SD) and inactivation (SA and polyadenylation signals (pA)). **b**, A PCR excision assay shows somatic transposon excision within mice doubly transgenic for transposon and transposase. Primers detect a roughly 2.2-kb product (the size of the T2/Onc transposon) if a transposon has not mobilized from within the concatemer. If transposition and excision repair occur anywhere within the concatemer, a 225-base-pair PCR product is generated.

[1]The Department of Genetics, Cell Biology and Development, The Arnold and Mabel Beckman Center for Transposon Research, The Cancer Center, The University of Minnesota Twin Cities, Minneapolis, Minnesota 55455, USA. †Present addresses: The Johns Hopkins School of Medicine, East Baltimore Campus, Ross 1064 – IM, 600 North Wolfe Street, Baltimore, Maryland 21287, USA (S.R.); Mouse Cancer Genetics Program, National Cancer Institute, Room 234, Building 539, 1050 Boyles Street, Frederick, Maryland 21702, USA (A.J.D.).
*These authors contributed equally to this work.

and *Arf*$^{-/-}$;T2/Onc;CAGGS-SB10 ($n = 64$). A statistically significant decrease in time to morbidity in *Arf*$^{-/-}$ mice doubly transgenic for T2/Onc and CAGGS-SB10 compared with singly transgenic *Arf*$^{-/-}$ control animals ($P < 0.001$, by log-rank Mantel–Cox test) was observed (Fig. 2a). The tumour spectrum of *Arf*$^{-/-}$;T2/Onc;CAGGS-SB10 mice was similar to that previously reported in *Arf*$^{-/-}$ mice on the C57BL/6 genetic background[14]. Of 52 *Arf*$^{-/-}$;T2/Onc;CAGGS-SB10 animals analysed, 36 had soft tissue sarcomas or osteosarcomas (Fig. 2b, c, and data not shown). Lymphomas, malignant meningiomas, myeloid leukaemias and a pulmonary adenocarcinoma were also observed (data not shown). Comparing the two control groups *Arf*$^{-/-}$;CAGGS-SB10 and *Arf*$^{-/-}$;T2/Onc revealed no difference in time to morbidity ($P = 0.19$, log-rank Mantel–Cox test).

A Southern blot analysis of sarcoma genomic DNA from *Arf*$^{-/-}$;T2/Onc;CAGGS-SB10 mice detected the presence of multiple, clonal, T2/Onc transposon insertions (an average of five), whereas genomic DNA isolated from normal tissues from various locations of the same mice showed either no subclonal T2/Onc insertions or only one or two (Supplementary Fig. 1b, c). Genomic sequences immediately flanking somatically acquired transposon integration events in sarcomas from *Arf*$^{-/-}$ mice were amplified by linker-mediated polymerase chain reaction (PCR)[15]. A total of 1,053 distinct tumour-associated transposon integration events were cloned and sequenced from 28 tumours. In addition to cloning genomic integration events, we obtained the sequences immediately flanking the transgene concatemer in line 76. One end of this concatemer donor locus was cloned and mapped to chromosome 1 at 164,879,699 base pairs (data not shown). This was confirmed by fluorescence *in situ* hybridization (FISH, data not shown).

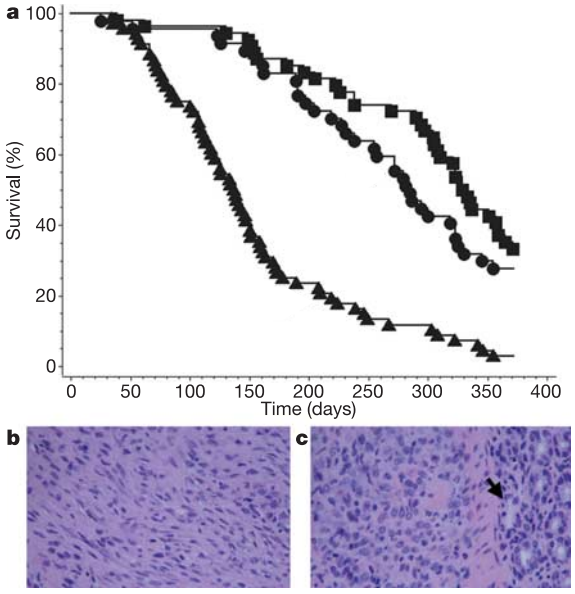

Figure 2 | *Arf*$^{-/-}$;T2/Onc;CAGGS-SB10 mice have a shorter tumour latency than singly transgenic controls. **a**, The Kaplan–Meier survival curve compares time to morbidity for *Arf*$^{-/-}$;T2/Onc;CAGGS-SB10 mice (triangles), *Arf*$^{-/-}$;CAGGS-SB10 mice (circles) and *Arf*$^{-/-}$;T2/Onc mice (squares). There is a statistically significant decrease in time to morbidity in *Arf*$^{-/-}$ mice doubly transgenic for T2/Onc and CAGGS-SB10 compared with singly transgenic *Arf*$^{-/-}$ controls ($P < 0.001$, log-rank Mantel–Cox test). This difference is evident when both distinct lines of T2/Onc transgenes are grouped or when they are assessed separately ($P < 0.001$ for both, data not shown). **b**, **c**, Examples of sarcomas from *Arf*$^{-/-}$;T2/Onc;CAGGS-SB10 mice: spindle cell tumour (undifferentiated sarcoma) found growing on the hindlimb (**b**); soft-tissue sarcoma infiltrating stomach glands (**c**; arrow).

In SB germline mutagenesis screens, 50–80% of transposons tend to reinsert within about 6 megabases (Mb) on either side of the donor concatemer[16–18]. For somatic transposition in sarcomas, this 'local hopping' interval seems to be broadened, because only 23% of somatic integrations cloned from tumours from mice in line 76 occurred within the 40 Mb surrounding the concatemer (data not shown). The cloning of a large number of insertion sites also permitted mapping of the concatemer donor to chromosome 15 in line 68, confirmed by FISH (data not shown), and revealed a similar percentage of local hopping.

A genomic region mutated by the integration of T2/Onc in multiple different tumours, a common integration site (CIS), indicates possible selection for that event during tumorigenesis. On the basis of published Monte Carlo simulations, a CIS was defined as two integrations from two independent tumours in 13 kb, three or more integrations from three independent tumours in 269 kb, or two or more integration events from two independent tumours within the

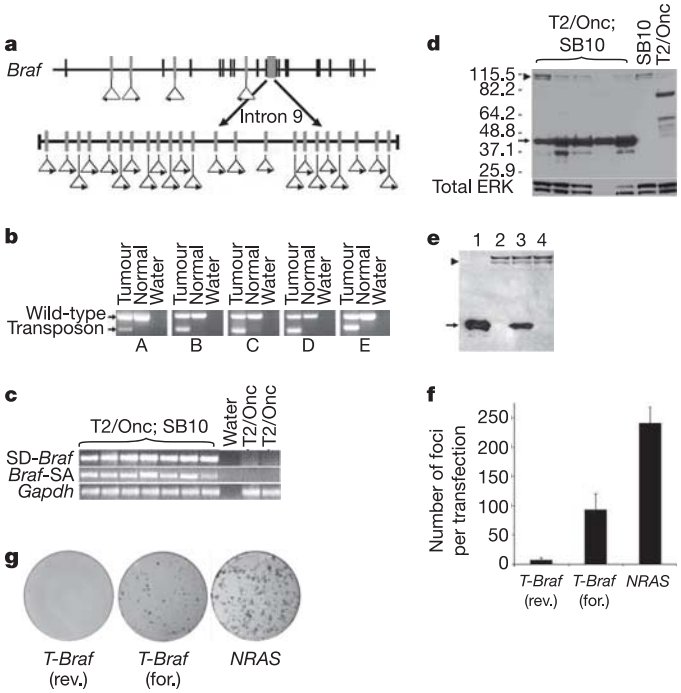

Figure 3 | Activation of *Braf* by T2/Onc insertion. **a**, Position and orientation of *Braf* T2/Onc insertions (grey). *Braf* exons are indicated by vertical black lines. The ninth intron is expanded to show detail of insertions. **b**, Three-primer PCR for ninth-intron *Braf* T2/Onc insertions (A–E) shows tumour specificity of insertions. **c**, RT–PCR reveals the presence of fusion transcripts, present in several T2/Onc;CAGGS-SB10 tumours harbouring ninth-intron insertions. The T2/Onc splice donor splices into the tenth exon of *Braf* (SD-*Braf*) and the *Braf* exon nine splice donor splices into the T2/Onc splice acceptor (*Braf*-SA). **d**, Western blot analysis detects a truncated C-terminal kinase domain of the BRAF protein (arrow) in sarcomas that harbour ninth-intron *Braf* integrations. Full-length BRAF protein is also detected (arrowhead). Total ERK was used as a loading control. (The size of molecular weight markers in kDa is shown on the left.) **e**, The SD-*Braf* transcript was amplified from tumours, cloned into an expression vector in the reverse (rev.) and forward (for.) orientations. Western analysis of 293T cells detects a truncated C-terminal kinase domain of the BRAF protein (arrow) and full-length BRAF protein (arrowhead): lane 1, tumour; lane 2, GFP-transfected; lane 3, forward; lane 4, reverse. **f**, **g**, Expression of truncated BRAF results in the formation of foci in NIH 3T3 cells. *NRAS* (G12V) is an acutely transforming oncogene for comparison. Error bars in **f** indicate standard deviation. **f**, Expression of truncated BRAF (T-*Braf* for) results in focus formation in NIH 3T3 cells. Transfection with the T-*Braf* rev control vector results in only baseline levels of focus formation. **g**, Representative plates showing focus formation.

same annotated gene[19,20]. By these definitions, 54 CISs were identified by T2/Onc in *Arf*$^{-/-}$ sarcomas (Table 1 and Supplementary Table 1). The local hopping phenomenon does increase the possibility of identifying CISs by random chance when they are linked to the concatemer donor locus. On the basis of published Monte Carlo simulations when insertions are distributed randomly and about 1,000 independent insertions are studied, 10 of our 54 CISs are predicted to occur simply by random chance[19]. As SB transposon integration favours sites linked to the donor locus, traditional Monte Carlo simulation (which assumes a completely random distribution of insertions throughout the genome) cannot accurately predict the number of false CISs occurring at loci linked to the transposon donor locus. The actual false-positive rate is therefore probably higher than this, because many CISs are linked to donor loci on chromosome 1 (for line 76) and chromosome 15 (for line 68) (Table 1 and Supplementary Table 1). However, it is likely that some linked CISs are not identified merely by random chance. For example, several T2/Onc integrations in *Arf*$^{-/-}$ tumours occurred in *Rabgap1l*, which is linked to the donor locus on chromosome 1 in line 76. *Rabgap1l* is a CIS identified in an accompanying paper[21], providing additional evidence that *Rabgap1l* has a function in tumorigenesis. Despite local hopping, it seems that the entire genome is accessible to SB somatic mutagenesis, because T2/Onc integration events were cloned from all mouse chromosomes (data not shown) and CISs were also identified on chromosomes 2, 4, 6, 12, 13 and X. In addition, T2/Onc integrations were found near several CISs previously identified in leukaemias or lymphomas in retroviral mutagenesis screens[22,23] (Supplementary Table 2).

The gene most frequently disrupted by T2/Onc was *Braf* (Table 1 and Supplementary Table 1). Integrations in or near *Braf* were cloned from 22 of 28 sarcomas and were found in tumours from mice transgenic for both T2/Onc concatemers, 68 and 76. Each sarcoma

Table 1 | Common integration sites in *Arf*$^{-/-}$;T2/Onc;CAGGS-SB10 transgenic sarcomas

CIS name	Mouse chromosome	Approximate location (Mb)	Number of integrations	Number of independent tumours
Bai3	1	25.8	3	2
Dst	1	34.3	2	2
ENSMUST00000042986.3	1	56	2	2
Spag16	1	70	3	3
ENSMUSG00000042581	1	129	2	2
Daf1	1	130	3	2
NG-1-143	1	143	3	3
Uchl5	1	143.7	4	4
Rgs	1	144	10	6
B830045N13Rik	1	146.8	6	6
NG-1-147	1	147.2	4	4
NG-1-147b	1	147.7	3	3
Laminin	1	153	4	4
Creg	1	158	8	3
C80879	1	159	4	3
Rabgap1l	1	160	17	8
Tnfsf	1	161	4	4
Fmo	1	162	3	3
Prrx1	1	163	2	2
Dpt	1	164	4	4
ENSMUSG00000038473	1	170	2	2
Ptprt	2	161	2	2
Ptch2	4	115	2	2
NG-6-23	6	22	2	2
Cadps2	6	23.4	5	4
Braf	6	39	37	22
E330009J07Rik	6	40.3	4	4
ENSMUST00000071875.1	6	43	3	3
Cntnap2	6	46	3	3
Baiap1	6	94	2	2
ENSMUSESTT00000078632	12	83	3	3
Adarb2	13	8.2	4	3
ENSMUSG00000039828	15	7.8	2	2
4933421G18Rik	15	8.1	2	2
ENSMUST00000082227.1	15	16	4	3
MGC92959	15	21.4	2	2
15-NG-22	15	22	3	3
ENSMUSG00000043556	15	26	3	3
ENSMUST00000075169.1	15	29	4	4
Catnd2	15	30.5	3	2
Sema5a	15	32	3	3
Coh	15	35.7	4	4
Rims2	15	39.3	2	2
2610028F08Rik	15	43.3	10	6
Trhr	15	43.9	25	11
Csmd3	15	47.8	14	8
ENSMUSESTG00000033246	15	48.8	3	3
Rad21	15	52	6	4
LOC277923	15	55	2	2
BC026439	15	57.4	4	4
NG-15-69	15	69	6	5
ENSMUSESTG00000029680	15	70	4	3
krt2	15	102	3	3
ENSMUST00000074972	X	137.7	2	2

See Supplementary Table 1 for additional details on integrations.

with *Braf* insertions had at least one within a TA dinucleotide in the ninth intron (Fig. 3a). All *Braf* ninth-intron insertions analysed seemed to be tumour-specific and absent from normal tissue from the same mouse (Fig. 3b); this is also true of other T2/Onc gene insertions studied (Supplementary Fig. 2).

All ninth-intron *Braf* integrations were directional with the MSCV LTR and splice donor oriented toward the tenth exon (Fig. 3a). This 'sense' orientation predicts that transcripts initiated from the MSCV LTR would splice into the tenth *Braf* exon, and this was confirmed by RT–PCR in seven sarcomas (Fig. 3c). This transcript could result in the expression of a truncated protein, translationally initiated in exon 10, containing the kinase domain of BRAF. An antibody against the carboxy-terminal fragment of BRAF detected a protein of the expected size (about 40 kDa) specifically in five sarcoma lysates that harbour intron 9 *Braf*-gene T2/Onc insertions (Fig. 3d). Moreover, an amino-terminal specific BRAF antiserum did not detect a truncated BRAF peptide (data not shown) despite the fact that a truncated *Braf* mRNA, generated by splicing from the *Braf* exon 9 splice donor into the splice acceptor upstream of the MSCV LTR sequences in T2/Onc, was detected (Fig. 3c).

Braf is a known oncogene, which has been shown to contain activating point mutations in 9% of human sarcoma cell lines and 0.5–5% of primary human sarcomas[24,25]. This provides a proof of principle that SB somatic mutagenesis identifies genes associated with, and clinically relevant to, specific human cancers. The truncated BRAF protein expressed in sarcomas with T2/Onc integrations in the ninth intron of *Braf* contains only the kinase domain, lacks N-terminal negative regulatory elements of the protein and is capable of morphological transformation of NIH 3T3 cells (Fig. 3e–g). Previous work has shown the oncogenic potential of a truncated kinase domain of the closely related *Craf* (ref. 26). On the basis of these results, it seems that *Braf* is capable of collaborating with *Arf* loss to elicit sarcoma development.

We have shown that *Sleeping Beauty* can be used for somatic-cell insertional mutagenesis in the mouse for the identification of cancer genes in solid tumours. In the future, T2/Onc mobilization combined with tissue-specific loss of a tumour suppressor might permit the identification of tumour-predisposing genes for any tissue type. In an accompanying paper[21] it is shown that the improved SB transposase, SB11 (ref. 27), expressed from the *Rosa26* locus can cause the efficient somatic mobilization of a T2/Onc-like transposon and induce tumour formation in the absence of a cancer-predisposed genetic background. The differences in the ability of T2/Onc mobilization by CAGGS-SB10 and *Rosa26*-SB11 to initiate and promote tumour formation might be due to differences in activity of SB10 and SB11, differences in levels of protein expression or differences in spatial or temporal expression of the transposase. The use of a conditionally expressed SB transposase may improve future studies by allowing control of its spatial and temporal expression. In addition, new transposon vector designs may further enhance the utility of the system.

METHODS

Vector construction. The T2/Onc vector contains the MSCV 5′ LTR from the MSCVneo vector (Clontech). The splice donor is from exon 1 of the mouse *Foxf2* gene. One splice acceptor is derived from exon 2 of the mouse *engrailed-2* gene and the other from the carp β-actin gene. Each is followed by the bi-directional SV40 poly(A).

Mice. Transgenic lines of T2/Onc were generated on the FVB/N genetic background. Southern blot analysis was performed on genomic DNA from a tail biopsy, and two lines (nos 68 and 76) with high copy numbers (about 25 copies) and a lack of transgene methylation were chosen for further analysis. The copy number of T2/Onc elements within the concatemer was estimated by comparison of T2/Onc signal intensity from transgenic genomic DNA to known amounts of T2/Onc plasmid DNA by Southern analysis. Methylation status was investigated by Southern analysis after digestion with a methylation-sensitive restriction enzyme. It was proposed that methylation of the transposon transgene might silence the activity of the MSCV LTR after transposition to new sites in the genome. FISH and spectral karyotyping analysis were used to map concatemer 68 to chromosome 15 and concatemer 76 to chromosome 1. To generate the $Arf^{-/-}$ cohort, $Arf^{+/-}$;CAGGS-SB10 and $Arf^{+/-}$;T2/Onc mice were first generated by crossing $Arf^{-/-}$ mice with CAGGS-SB10 or T2/Onc mice, respectively. $Arf^{+/-}$;CAGGS-SB10 mice were intercrossed to generate $Arf^{-/-}$;CAGGS-SB10 mice. $Arf^{+/-}$;T2/Onc mice were also intercrossed to generate $Arf^{-/-}$;T2/Onc mice. $Arf^{-/-}$;CAGGS-SB10 mice and $Arf^{-/-}$;T2/Onc mice were crossed to generate $Arf^{-/-}$;CAGGS-SB10;T2/Onc, $Arf^{-/-}$;T2/Onc and $Arf^{-/-}$;CAGGS-SB10 mice.

PCR primers. Sequences of PCR primers are given in Supplementary Information.

Histopathology. Tissues were fixed overnight in 10% formalin at 4 °C, stored in 70% ethanol, embedded in paraffin, sectioned and stained with haematoxylin and eosin.

Linker-mediated PCR and 'shot-gun cloning'. Linkers used to clone insertions were described previously[15]. Genomic DNA was digested with *Nla*III and *Xho*I (for cloning from the right side of T2/Onc) or *Bfa*I and *Bam*HI (for cloning from the left side of T2/Onc) and ligated to the linker. Primary PCR was used to amplify sequences flanking the IR/DR. Primary PCR products were diluted 1:50 and used in a secondary PCR. Secondary PCR products were ligated to pGEM-T Easy (Promega) and electroporated into DH10B Electromax-competent cells (Invitrogen). Library plating, colony picking and sequencing with the SP6 primer in 96-well format was performed by Agencourt Biosciences. Automated database searches of T2/Onc integration sites were performed as described previously by K. Akagi[22]. The closest gene to each integration (within 100 kb on either side of the integration) was determined with a combination of the UCSC (http://www.genome.ucsc.edu) and Ensembl (http://www.ensembl.org) mouse whole-genome annotations (NCBI m33 build).

***Braf* three-primer PCR.** Genomic DNA (500 ng) was used in a PCR with three separate primers. One transposon-specific primer was used for all three-primer PCRs together with specific primers to detect the wild-type locus of each cloned insertion event.

***Braf* RT–PCR.** Total RNA was isolated from tumour tissues by TRIzol (Invitrogen), contaminating DNA was removed by treatment with DNase (Invitrogen) and RT–PCR was performed with 500 ng of RNA with the RobusT I RT–PCR kit (Finnzymes, MJ Research). A splice-donor (SD)-specific primer and a *Braf* tenth-exon-specific primer were used. A *Braf* exon-seven-specific primer and a Carp β-actin splice-acceptor (SA)-specific primer were used. The resulting products were sequenced to verify the fidelity of each splicing event.

BRAF western blot analysis. Protein lysates were prepared by homogenizing tissue in IPWB lysis buffer (50 mM Tris-HCl pH 7.4, 14.6 mg ml^{-1} NaCl, 2 mM EDTA, 2.1 mg ml^{-1} NaF, 1% Nonidet P40, 1 mM $NaVO_4$ and 1 mM Na_2PO_4, containing protease inhibitors (Roche)) or by following the manufacturer's protocols for protein isolation from the organic phase of TRIzol. Samples (about 30 μg) were subjected to electrophoresis on a 4–12% Bis–Tris gel and transferred to nitrocellulose (BioRad) with the NuPAGE system (Invitrogen). Blots were probed with a primary antibody specific to the carboxy terminus of BRAF (Santa Cruz Biotechnology). An antibody specific for Erk-1 (Santa Cruz Biotechnology) was used as a loading control.

Cloning of truncated *Braf* and NIH 3T3 transformation assay. The cDNA of the truncated C-terminal fusion transcript of *Braf* generated in tumours (T2/Onc SD exons 10–19 of *Braf*) was amplified with the RobusT I RT–PCR kit. The amplified product was subcloned into pCR 2.1-TOPO (Invitrogen), excised with *Eco*RI and ligated in the forward and reverse orientations into an *Eco*RI site of a CAGGS vector. These plasmids and a CAGGS plasmid expressing the activated human *NRAS* oncogene (G12V) (C.M.C. and D.A.L., unpublished observations) were each transfected in duplicate into NIH 3T3 cells with the SuperFect Transfection Reagent (Qiagen). Cells were cultured in DMEM medium with 10% FBS, 2 mM L-glutamine, 0.1 mM non-essential amino acids, 55 μM 2-mercaptoethanol and 10 μg ml^{-1} gentamycin, split into two 100-mm plates 2 days after transfection, cultured for 10 days and stained with methylene blue.

Received 4 February; accepted 27 April 2005.

1. Neil, J. C. & Cameron, E. R. Retroviral insertion sites and cancer: fountain of all knowledge? *Cancer Cell* **2**, 253–255 (2002).
2. Jonkers, J. & Berns, A. Retroviral insertional mutagenesis as a strategy to identify cancer genes. *Biochim. Biophys. Acta* **1287**, 29–57 (1996).
3. Ivics, Z., Hackett, P. B., Plasterk, R. H. & Izsvak, Z. Molecular reconstruction of *Sleeping Beauty*, a Tc1-like transposon from fish, and its transposition in human cells. *Cell* **91**, 501–510 (1997).
4. Dupuy, A. J., Fritz, S. & Largaespada, D. A. Transposition and gene disruption in the male germline of the mouse. *Genesis* **30**, 82–88 (2001).
5. Horie, K. *et al.* Efficient chromosomal transposition of a *Tc1/mariner*-like

transposon *Sleeping Beauty* in mice. *Proc. Natl Acad. Sci. USA* **98,** 9191–9196 (2001).
6. Fischer, S. E., Wienholds, E. & Plasterk, R. H. Regulated transposition of a fish transposon in the mouse germ line. *Proc. Natl Acad. Sci. USA* **98,** 6759–6764 (2001).
7. Yant, S. R. *et al.* Somatic integration and long-term transgene expression in normal and haemophilic mice using a DNA transposon system. *Nature Genet.* **25,** 35–41 (2000).
8. Montini, E. *et al. In vivo* correction of murine tyrosinemia type I by DNA-mediated transposition. *Mol. Ther.* **6,** 759–769 (2002).
9. Abdallah, B. *et al.* A powerful nonviral vector for *in vivo* gene transfer into the adult mammalian brain: polyethylenimine. *Hum. Gene Ther.* **7,** 1947–1954 (1996).
10. Hawley, R. G., Lieu, F. H., Fong, A. Z. & Hawley, T. S. Versatile retroviral vectors for potential use in gene therapy. *Gene Ther.* **1,** 136–138 (1994).
11. Cherry, S. R., Biniszkiewicz, D., van Parijs, L., Baltimore, D. & Jaenisch, R. Retroviral expression in embryonic stem cells and hematopoietic stem cells. *Mol. Cell. Biol.* **20,** 7419–7426 (2000).
12. Okabe, M., Ikawa, M., Kominami, K., Nakanishi, T. & Nishimune, Y. 'Green mice' as a source of ubiquitous green cells. *FEBS Lett.* **407,** 313–319 (1997).
13. Kamijo, T. *et al.* Tumor suppression at the mouse *INK4a* locus mediated by the alternative reading frame product p19ARF. *Cell* **91,** 649–659 (1997).
14. Kamijo, T., Bodner, S., van de Kamp, E., Randle, D. H. & Sherr, C. J. Tumor spectrum in *ARF*-deficient mice. *Cancer Res.* **59,** 2217–2222 (1999).
15. Wu, X., Li, Y., Crise, B. & Burgess, S. M. Transcription start regions in the human genome are favored targets for MLV integration. *Science* **300,** 1749–1751 (2003).
16. Vigdal, T. J., Kaufman, C. D., Izsvak, Z., Voytas, D. F. & Ivics, Z. Common physical properties of DNA affecting target site selection of sleeping beauty and other Tc1/mariner transposable elements. *J. Mol. Biol.* **323,** 441–452 (2002).
17. Carlson, C. M. *et al.* Transposon mutagenesis of the mouse germline. *Genetics* **165,** 243–256 (2003).
18. Horie, K. *et al.* Characterization of *Sleeping Beauty* transposition and its application to genetic screening in mice. *Mol. Cell. Biol.* **23,** 9189–9207 (2003).
19. Mikkers, H. *et al.* High-throughput retroviral tagging to identify components of specific signalling pathways in cancer. *Nature Genet.* **32,** 153–159 (2002).
20. Johansson, F. K. *et al.* Identification of candidate cancer-causing genes in mouse brain tumors by retroviral tagging. *Proc. Natl Acad. Sci. USA* **101,** 11334–11337 (2004).
21. Dupuy, A. J., Akagi, K., Largaespada, D. A., Copeland, N. G. & Jenkins, N. A. Mammalian mutagenesis using a highly mobile somatic *Sleeping Beauty* transposon system. *Nature* doi:10.1038/nature03691 (this issue).
22. Akagi, K., Suzuki, T., Stephens, R. M., Jenkins, N. A. & Copeland, N. G. RTCGD: retroviral tagged cancer gene database. *Nucleic Acids Res.* **32,** D523–D527 (2004).
23. Lund, A. H. *et al.* Genome-wide retroviral insertional tagging of genes involved in cancer in *Cdkn2a*-deficient mice. *Nature Genet.* **32,** 160–165 (2002).
24. Davies, H. *et al.* Mutations of the *BRAF* gene in human cancer. *Nature* **417,** 949–954 (2002).
25. Seidel, C. *et al.* Alterations of cancer-related genes in soft tissue sarcomas: hypermethylation of *RASSF1A* is frequently detected in leiomyosarcoma and associated with poor prognosis in sarcoma. *Int. J. Cancer* **114,** 442–447 (2005).
26. Heidecker, G. *et al.* Mutational activation of *c-raf-1* and definition of the minimal transforming sequence. *Mol. Cell. Biol.* **10,** 2503–2512 (1990).
27. Geurts, A. M. *et al.* Gene transfer into genomes of human cells by the *Sleeping Beauty* transposon system. *Mol. Ther.* **8,** 108–117 (2003).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank M. Roussel for providing $Arf^{-/-}$ mice; P. Woll, S. Rosenthal, R. Wang, P. Lobitz, J. Zinggeler and M. Davies for technical assistance; P. Hackett, S. McIvor, S. Ekker, P. Marker and members of the Largaespada laboratory for critical reading of the manuscript; N. Kirchhof for performing the pathology; T. L. Bergemann for providing assistance with statistical analysis; and K. Akagi for performing automated sequence analysis and BLAT searches. T2/Onc transgenic mice were generated by the University of Minnesota Mouse Genetics Laboratory. This work was supported by the Arnold and Mabel Beckman Foundation, the National Institute of Drug Abuse and the National Cancer Institute.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare competing financial interests: details accompany the paper on www.nature.com/nature. Correspondence and requests for materials should be addressed to D.A.L. (larga002@umn.edu).