DOI:10.1158/0008-5472.CAN-05-3085

Cancer Research

Hopping around the Tumor Genome: Transposons for Cancer Gene Discovery

Lara S. Collier and David A. Largaespada

Cancer Res 2005;65:9607-9610. Published online November 1, 2005.

DOI:10.1158/0008-5472.CAN-05-3085

Hopping around the Tumor Genome: Transposons for Cancer Gene Discovery

Lara S. Collier and David A. Largaespada

Department of Genetics, Cell Biology and Development, The Arnold and Mabel Beckman Center for Transposon Research, The Cancer Center, University of Minnesota-Twin Cities, Minneapolis, Minnesota

Abstract

Retroviruses are powerful insertional somatic mutagens that have been used for many landmark discoveries of cancer genes in model organisms. However, their use as a cancer gene discovery tool has been limited to only a few tissues, mainly the hematopoietic system and mammary gland. Recently, the *Sleeping Beauty* (SB) transposon system was shown to be useful for random somatic cell mutagenesis in mice, allowing the induction or acceleration of tumor formation both in the hematopoietic system and in sarcomas. In these tumors, SB transposons repeatedly "tagged" specific genes, both known and new cancer genes. These results indicate that the SB system has great potential both for generating specific mouse models of human cancer and for cancer gene discovery in a wide variety of tissues. (Cancer Res 2005; 65(21): 9607-10)

Cancer is hypothesized to result from the accumulation of multiple somatic mutations (1). Mouse models of cancer initiated by tumor suppressor gene inactivation or oncogene activation support this hypothesis (2, 3). Such mice are cancer-prone but generally only develop disease after a long latency, indicating that additional somatic "hits" must occur before an overt tumor forms. Chemical or radiological mutagens often decrease this latency, lending further support to the hypothesis (4, 5). Cloning the genes affected by these mutagens is a difficult proposition and usually occurs by a "candidate gene" approach. Somatic mutagens that allow easy identification of genes that promote tumor formation can facilitate genome-wide forward genetic screens for cancer genes. Retroviruses are one such somatic mutagen.

Many years ago, it became apparent that certain strains of mice develop leukemia and other strains of mice develop mammary tumors early in life due to chronic infection of ecotropic retroviruses (6, 7). In leukemias, the responsible slow-transforming retroviruses were called murine leukemia viruses (MuLV), and in mammary tumors, these were murine mammary tumor viruses. It was later elucidated that these retroviruses do not carry a virally encoded oncogene but promote tumor formation due to integration of proviruses into the host genome. Tumor initiation occurs when a provirus integrates by chance within a tumor suppressor gene and inactivates it or inserts near or within a proto-oncogene and activates its transcription, stabilizes its mRNA, or creates oncogenic truncated protein products. These effects are due to promoter/enhancer elements in the proviral long terminal repeats (LTR) and additional features of the provirus, such as polyadenylation and splicing signals. Eventually, the accumulation of cooperating mutations (which can themselves be proviral integrations) leads to the emergence of a malignant clone. Regions of the genome repeatedly mutagenized by proviral integration in multiple independent tumors are designated common integration sites (CIS). CIS result from the repeated selection for insertions in a chromosomal region because they affect the expression of nearby tumor suppressor genes or proto-oncogenes. In the mammary gland, retroviral CIS have indicated genes known to function in important cancer signal transduction pathways, including *Wnt* genes, *Fgf* genes, etc. (8). In the hematopoietic system, many genes known to play a role in human leukemiagenesis and lymphomagenesis have been identified at CIS: *Notch1, Flt3*, and *Lmo2*, to name a few (9). Despite this record of success, retroviruses do have limitations as somatic mutagens.

Retroviruses can mutate tumor suppressor genes in the tumors they cause, *Nf1* and *Trp53* being two examples, although the vast majority of CIS are thought to result from activation of nearby proto-oncogenes. One interpretation for this discrepancy is that proviral mutation of both copies of tumor suppressor genes in a single cell would be an extremely rare event. Nevertheless, this does occur at the *Nf1* locus in MuLV-based acute myelogenous leukemia in the BXH-2 mouse strain (10). Although retroviruses were originally thought to insert into the genome at random, this is now known to not be true (11). Therefore, it is possible that proviruses do not tend to mutate tumor suppressor genes because they prefer to insert near the 5′ end of genes (11). Indeed, this strong preference for proviral insertion near the promoter region of actively transcribed genes may severely limit the amount of the genome accessible to retroviral mutagenesis. Moreover, retroviruses may be ineffectual at disrupting gene expression on intronic insertion.

Retroviruses have also had very limited success as somatic mutagens for cancer gene identification outside of the hematopoietic system and mammary gland. In chicken nephroblastomas, *twist* was identified as a CIS by the avian retrovirus, MAV2 (12). Recently, a MuLV was used to identify several CIS near both known and potentially novel cancer genes in glioblastoma. This required the direct intracranial injection of a MuLV engineered to express *PDGFβ* chain (13). However, the engineering or discovery of a retroviral mutagen with widespread applicability to cancer gene discovery has been elusive. A highly flexible somatic mutagenesis system not limited by tissue type specificity would be an asset to cancer geneticists. Therefore, to expand on the success of retroviral mutagenesis and to make more tissues amenable to somatic mutagenesis, two separate but collaborating groups recently tested if a different type of mobile element, a DNA transposon, can mutate and "tag" cancer genes in the mouse (14, 15).

Although active DNA transposons are found in the genomes of invertebrates and plants, they are very rare in the genomes of vertebrates, especially mammals. Instead, vertebrate genomes

Requests for reprints: David A. Largaespada, Department of Genetics, Cell Biology and Development, University of Minnesota-Twin Cities, 6-160 Jackson Hall, 321 Church Street Southeast, Minneapolis, MN 55455. Phone: 612-626-4979; E-mail: larga002@umn.edu.

doi:10.1158/0008-5472.CAN-05-3085

DOI:10.1158/0008-5472.CAN-05-3085

contain inactive DNA transposon remnants that are incapable of transposition due to accumulated mutations. DNA transposons have proven invaluable for genetic studies of *Drosophila, Caenorhabditis elegans*, unicellular prokaryotes, bacteria, and plants. In the mid-1990s, a vertebrate-active DNA transposon was engineered by first identifying long-dormant DNA transposons of the Tc1/mariner family in the genomes of salmonid fish. Directed mutagenesis was used to restore the activity of the transposon, which was named *Sleeping Beauty* (SB; ref. 16). SB has since been shown to be active in zebrafish, human cells in culture, mouse embryonic stem cells, the mouse germ line, and the mouse soma when delivered exogenously on plasmid DNA (17).

SB is a two-component system consisting of a transposase, the enzyme responsible for mobilization, and the transposon, the actual mobilized piece of DNA. The minimal transposon consists of a left and a right inverted repeat/direct repeat (IRDR) elements, each of ~230 bp, flanking a cargo sequence. The cargo of the transposon can be any sequence of choice; however, transposition efficiency decreases with increased cargo size (18). For cancer gene identification experiments (14, 15), the cargos were mutagenic elements designed to mimic retroviral insertional mutagenesis. Specifically, transposons were engineered with splice acceptors/polyadenylation sequences in both orientations to disrupt the expression of genes in which they land. In addition, the transposons contained sequences from the 5′ LTR of the murine stem cell virus (MSCV LTR) to serve as promoter/enhancer elements to drive expression of nearby genes. The MSCV LTR was followed by a splice donor, so a transcript initiated in the LTR can splice into downstream exons of endogenous genes (Figs. 1 and 2). Both transposons, called T2/Onc (14) and T2/Onc2 (15), were constructed in essentially the same manner, the one exception being that T2/Onc2 contains a larger fragment for one of the two splice acceptors. Transgenic lines were established that harbor chromosomal concatomers of T2/Onc or T2/Onc2. For T2/Onc, the two transgenic lines used contain ~25 transposons, whereas T2/Onc2 transgenic lines selected for in-depth characterization contained 148, 214, or 358 transposons. Both experiments used transgenic mice that express transposase under the control of theoretically ubiquitous promoters. The experiment employing T2/Onc used the "original" SB10 transposase expressed under the control of the ubiquitous CAGGS promoter (CAGGS-SB10; ref. 19). Experiments using T2/Onc2 used the SB11 version of the transposase knocked into the *Rosa26* locus (Rosa26SB11), which is expected to result in widespread transposase expression.

Crossing of transposase and transposon transgenic mice yields experimental mice in which the transposon is mobilizing in the

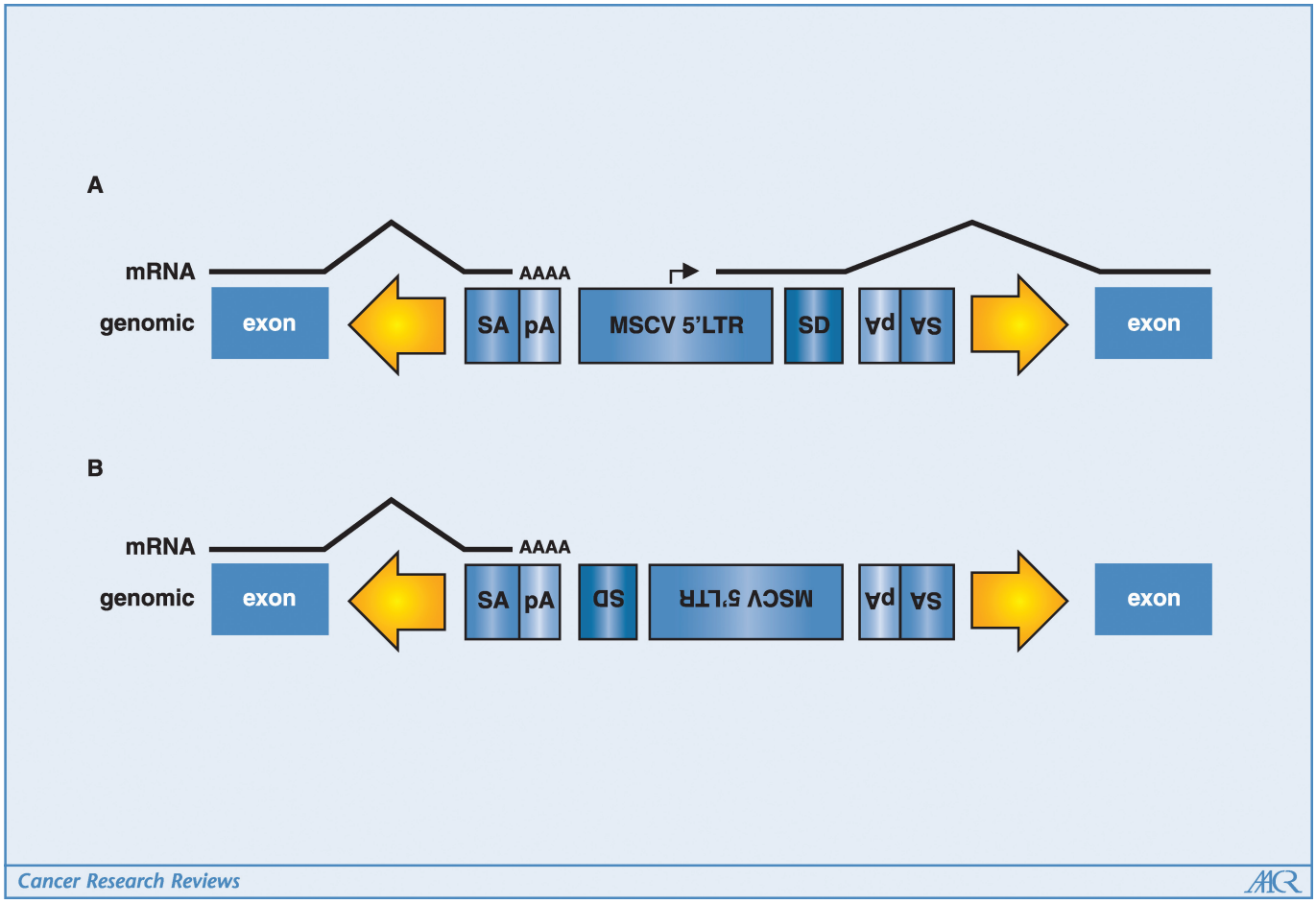

Figure 1. T2/Onc and T2/Onc2 are designed with both loss-of-function [splice acceptors/polyadenylation signals (*SA/pA*)] and gain-of-function [MSCV LTR splice donor (*MSCV 5′ LTR-SD*)] elements so that it can mutate both tumor suppressor genes and oncogenes. (*A*) When a transposon lands in a gene in the "forward" orientation, it can accept splicing from downstream endogenous exons and produce premature transcript termination due to polyadenylation signals present in the transposon. In this orientation, a transcript that initiates in the MSCV LTR can splice from the transposon splice donor into downstream endogenous exons. This can result in protein overexpression if the downstream exons contain an in-frame ATG for translational initiation. (*B*) In the "reverse" orientation, the transposon causes premature transcriptional termination.

DOI:10.1158/0008-5472.CAN-05-3085

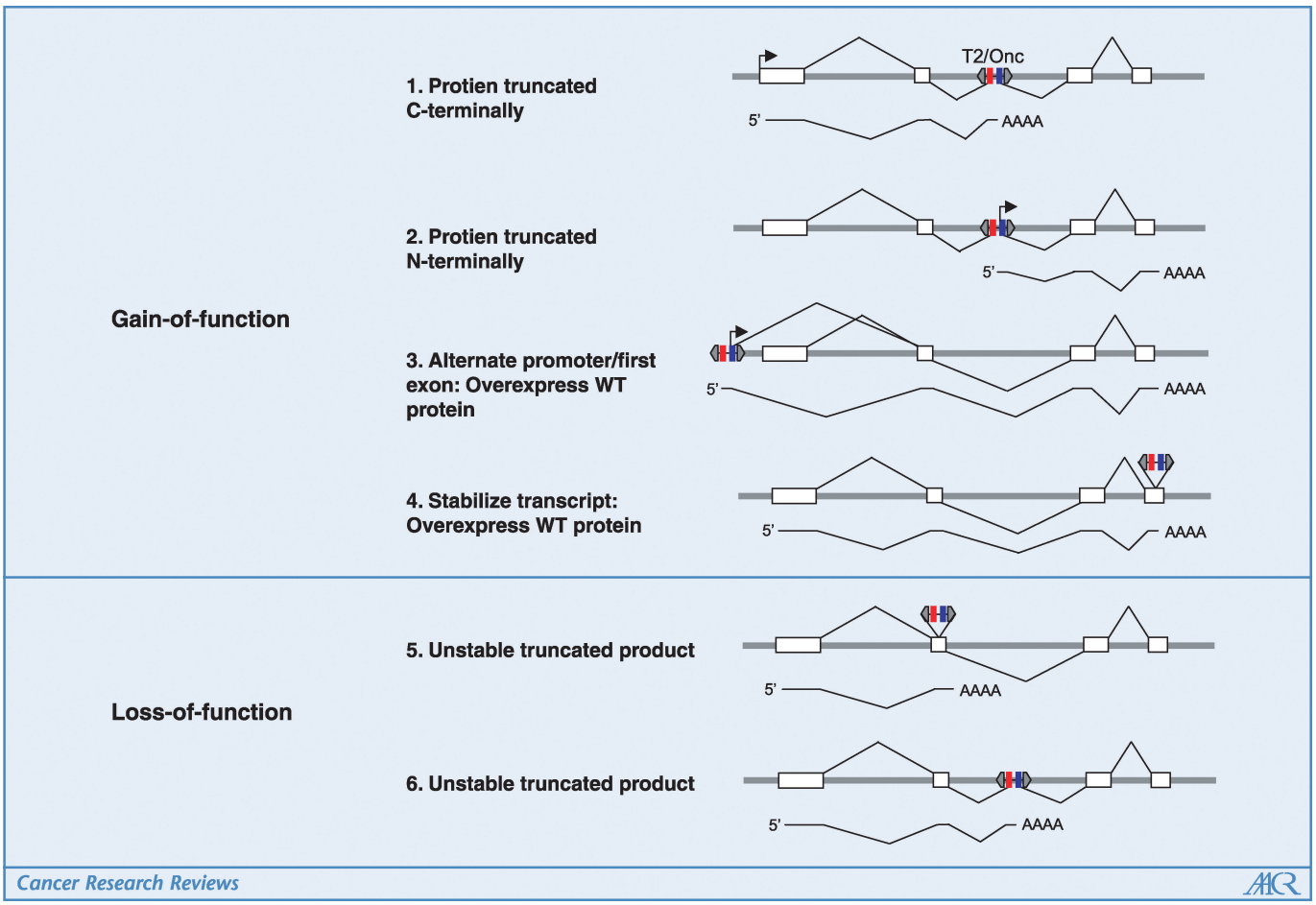

Figure 2. Potential mechanisms for gain-of-function mutations in proto-oncogenes and loss-of-function mutations in tumor suppressor genes that could be caused by T2/Onc insertion. Position and/or orientation bias are expected for mechanisms 1 to 4, pointing toward activation of proto-oncogene function. Mechanisms 5 and 6, resulting in loss-of-function mutations in tumor suppressors, would be expected to be orientation independent and perhaps show less bias for the region of the gene targeted for insertion.

soma due to the presence of both transposase and transposons. Singly transgenic littermates serve as controls. Crossing T2/Onc2 transgenics to Rosa26SB11 transgenics resulted in 5.6% to 12.5% doubly transgenic offspring depending on the T2/Onc2 concatomer used. The observed sub-Mendelian ratio of genotypes was due to high rates of embryonic lethality in double transgenics. All T2/Onc2;Rosa26SB11 mice that survived to birth were moribund by 114 days of age. Twenty-three of 24 mice developed hematopoietic malignancies (mainly T-cell lymphoma) and 2 mice developed medulloblastomas. In addition to frank neoplasia, 4 mice also had hyperplasia of the intestine or pituitary gland (15). These results contrast to experiments using the lower-copy T2/Onc concatomers and CAGGS-SB10 transposase in which doubly transgenic mice had life spans comparable with controls. Although somatic mobilization of T2/Onc on an otherwise wild-type genetic background was not sufficient to promote tumor formation, it was able to accelerate tumor formation in mice deficient for the p19 *Arf* (*Arf*) tumor suppressor. The majority of T2/Onc;CAGGS-SB10;$Arf^{-/-}$ mice developed sarcomas, similar to what had been observed previously in $Arf^{-/-}$ mice (4, 14).

In tumors in both experiments, novel subclonal and clonal transposon integrations could be observed by Southern analysis, indicating that certain transposon integrations were being maintained because they conferred a growth advantage to the tumor clone. Cloning of T2/Onc integrations in 28 sarcomas and T2/Onc2 integrations in 15 hematopoietic malignancies and 1 medulloblastoma revealed the presence of CIS, similar to those found in retroviral mutagenesis screens. In sarcomas, integrations in the ninth intron of *Braf* were very common, occurring in at least 80% of tumors. These integrations resulted in the production of high levels of a transcript that initiated in the MSCV LTR of T2/Onc and spliced from the T2/Onc splice donor into exon 10 of *Braf*. This transcript produces an amino-terminally truncated version of *Braf* that is capable of morphologic transformation of NIH 3T3 cells. Additional CIS were identified near or within both novel and known cancer genes, such as *Ptpr2* (14). In the hematopoietic malignancies, several CIS were identified near or in genes known to play a role in human tumor development (e.g., *Erg* and *Ets1*) and several were previously identified CIS in retroviral mutagenesis screens, such as *Runx2* and *Rasgrp1*. CIS were also identified near genes not implicated previously as cancer genes (15). In addition, both experiments showed strong genetic interactions between cancer genes. For example, most $Arf^{-/-}$ sarcomas possessed T2/onc integrations in *Braf*, indicating that *Arf* loss and *Braf* activation can cooperate strongly in tumorigenesis (14). This cooperating pair of genes was also seen in a study of human melanoma in which 15 of 41 samples had both loss of *Arf* expression and activation of *Braf* by point mutation (20). In T-cell tumors induced by T2/Onc2, three of six tumors that contained activating *Notch1* integrations also contained *Rasgrp1* activating integrations, showing an interaction between *Notch* and *Ras* pathways in tumorigenesis (15).

DOI:10.1158/0008-5472.CAN-05-3085

These studies suggest that SB may have certain advantages over retroviral mutagenesis. First, unlike retroviruses, SB transposons are known to have only a slight inherit preference for inserting into genes. Furthermore, the ability of SB to insert into a gene is unaffected by transcriptional status; therefore, SB is likely to mutagenize a more representative sample of the genome (21). Comparison of retroviral mutagenesis and SB mutagenesis in hematopoietic malignancies reveals that SB can target genes not accessible to proviral integration as only 7 of 25 CIS identified by SB were CIS identified previously in retroviral screens (15). Conversely, there are many CIS identified by retroviruses that were not identified by SB (9). The relatively small number of SB-induced tumor samples analyzed may not have allowed saturation mutagenesis by SB in the hematopoietic system. Indeed, analysis of only a limited number of new tumors by our laboratory has identified additional CIS,[1] indicating that studying a large number of additional tumors will identify more CIS.

Although these two studies show the power of SB for somatic mutagenesis, they also leave unanswered several questions about how to make the SB system amenable to cancer gene discovery in a wider variety of tissues. First, why is T2/Onc2 mobilization from higher-copy concatomers by Rosa26SB11 sufficient for tumor formation, whereas T2/Onc mobilization by CAGGS-SB10 only accelerates tumor formation in already predisposed animals? The subtle difference between T2/Onc and T2/Onc2 seems unlikely to account for the discrepancy. More likely, either differences in T2/Onc and T2/Onc2 concatomer copy number or in activity between CAGGS-SB10 and RosaSB11 are responsible for the different results. In cell culture–based transposition assays, SB11 is more active than SB10 when "first-generation" IRDRs (in so-called pT-based vectors) are used as substrate. However, SB11 and SB10 are equally efficient at mobilization of the "second-generation" IRDRs contained in the pT2 vector used for constructing both pT2/Onc and pT2/Onc2 (18). Therefore, it is likely that differences between the two experiments are not due to inherent differences between SB10 and SB11 but in dissimilarities between strength and spatiotemporal activity of transcription regulated by CAGGS and the *Rosa26* locus.

Second, could different combinations of transposon concatomers, with different copy numbers and differently regulated transposase transgenes, promote tumorigenesis in different tissues? Our preliminary experiments show that T2/Onc;RosaSB11 mice are born at Mendelian ratios, indicating that transposon copy number influences the rate of embryonic lethality. Although the doubly transgenic mice survive longer than T2/Onc2;RosaSB11 mice, they do eventually succumb to mainly hematopoietic disease, indicating no striking increase in the tumor spectrum available for analysis.[1] It remains to be seen if mobilization of T2/Onc2 from high-copy concatomers by CAGGS-SB10 is sufficient to promote tumor formation and, if so, in what tissues. In the future, tissue-specific regulation of transposase activity by conditional expression from the *Rosa26* locus may improve the utility of the system. However, the expression of transposase from *Rosa26* may not be as strong in every tissue as it is in the hematopoietic system. For example, our preliminary germ line mutagenesis experiments indicate that, in a direct comparison (i.e., mobilizing transposons from the same concatomer), RosaSB11 is less efficient in the germ line than is CAGGS-SB10.[2] Therefore, the generation of transgenic mice expressing transposase under control of tissue-specific promoters may further the utility of the system. Additional transposon designs, such as different promoters or enhancers in the place of the MSCV LTR, may facilitate the mutagenesis of proto-oncogenes in specific types of tumors or simply increase the breadth of cancer genes discovered. Conversely, a transposon consisting of only splice acceptors may increase the yield of loss-of-function mutations in tumor suppressor genes.

In summary, somatic mutagenesis using SB is a novel tool for cancer gene discovery in mice in tissues not amenable previously to high-throughput forward genetic screens. We believe that future improvements to both the transposase and the transposon will further improve the utility of the system. In addition to cancer gene discovery, we expect that the SB system will also be useful in generating genetically diverse, ever-evolving mouse tumor models for drug discovery and validation that more faithfully mimic the genetic complexity of human cancer.

Acknowledgments

Received 9/2/2005; accepted 9/21/2005.

[1] L.S. Collier et al., unpublished observations.

[2] A.M. Geurts, et al. unpublished observations.

References

1. Kinzler KW, Vogelstein B. Lessons from hereditary colorectal cancer. Cell 1996;87:159–70.
2. Jacks T. Tumor suppressor gene mutations in mice. Annu Rev Genet 1996;30:603–36.
3. Berns A, Breuer M, Verbeek S, van Lohuizen M. Transgenic mice as a means to study synergism between oncogenes. Int J Cancer Suppl 1989;4:22–5.
4. Kamijo T, Bodner S, van de Kamp E, Randle DH, Sherr CJ. Tumor spectrum in ARF-deficient mice. Cancer Res 1999;59:2217–22.
5. Fero ML, Randel E, Gurley KE, Roberts JM, Kemp CJ. The murine gene p27Kip1 is haplo-insufficient for tumour suppression. Nature 1998;396:177–80.
6. van Lohuizen M, Berns A. Tumorigenesis by slow-transforming retroviruses—an update. Biochim Biophys Acta 1990;1032:213–35.
7. Gross L. Viral etiology of cancer and leukemia: a look into the past, present and future—G.H.A. Clowes Memorial Lecture. Cancer Res 1978;38:485–93.
8. Callahan R, Smith GH. MMTV-induced mammary tumorigenesis: gene discovery, progression to malignancy and cellular pathways. Oncogene 2000;19:992–1001.
9. Akagi K, Suzuki T, Stephens RM, Jenkins NA, Copeland NG. RTCGD: retroviral tagged cancer gene database. Nucleic Acids Res 2004;32:D523–7.
10. Largaespada DA, Shaughnessy JD, Jr., Jenkins NA, Copeland NG. Retroviral integration at the Evi-2 locus in BXH-2 myeloid leukemia cell lines disrupts Nf1 expression without changes in steady-state Ras-GTP levels. J Virol 1995;69:5095–102.
11. Wu X, Li Y, Crise B, Burgess SM. Transcription start regions in the human genome are favored targets for MLV integration. Science 2003;300:1749–51.
12. Pajer P, Pecenka V, Karafiat V, Kralova J, Horejsi Z, Dvorak M. The twist gene is a common target of retroviral integration and transcriptional deregulation in experimental nephroblastoma. Oncogene 2003;22:665–73.
13. Johansson FK, Brodd J, Eklof C, et al. Identification of candidate cancer-causing genes in mouse brain tumors by retroviral tagging. Proc Natl Acad Sci U S A 2004;101:11334–7.
14. Collier LS, Carlson CM, Ravimohan S, Dupuy AJ, Largaespada DA. Cancer gene discovery in solid tumours using transposon-based somatic mutagenesis in the mouse. Nature 2005;436:272–6.
15. Dupuy AJ, Akagi K, Largaespada DA, Copeland NG, Jenkins NA. Mammalian mutagenesis using a highly mobile somatic Sleeping Beauty transposon system. Nature 2005;436:221–6.
16. Ivics Z, Hackett PB, Plasterk RH, Izsvak Z. Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. Cell 1997;91:501–10.
17. Ivics Z, Kaufman CD, Zayed H, Miskey C, Walisko O, Izsvak Z. The Sleeping Beauty transposable element: evolution, regulation and genetic applications. Curr Issues Mol Biol 2004;6:43–55.
18. Geurts AM, Yang Y, Clark KJ, et al. Gene transfer into genomes of human cells by the sleeping beauty transposon system. Mol Ther 2003;8:108–17.
19. Dupuy AJ, Fritz S, Largaespada DA. Transposition and gene disruption in the male germline of the mouse. Genesis 2001;30:82–8.
20. Daniotti M, Oggionni M, Ranzani T, et al. BRAF alterations are associated with complex mutational profiles in malignant melanoma. Oncogene 2004;23:5968–77.
21. Yant SR, Wu X, Huang Y, Garrison B, Burgess SM, Kay MA. High-resolution genome-wide mapping of transposon integration in mammals. Mol Cell Biol 2005;25:2085–94.