Review

Transposons for cancer gene discovery: *Sleeping Beauty* and beyond
Lara S Collier and David A Largaespada

The Department of Genetics, Cell Biology and Development, The Cancer Center, The University of Minnesota Twin Cities, Church St SE, Minneapolis, Minnesota 55455, USA

Correspondence: Lara S Collier. E-mail: lsc5e@alumni.virginia.edu

Abstract

The use of *Sleeping Beauty* transposons as somatic mutagens to discover cancer genes in hematopoietic tumors and sarcomas has been documented. Here, we discuss the future of *Sleeping Beauty* for cancer genetic studies and the potential use of additional transposable elements for somatic mutagenesis.

Introduction

Cancer genomes are complex, and there are probably many yet undiscovered tumor suppressor genes (TSGs) and onco-genes. The multiple mechanisms by which cancer genes can suffer mutation, including amplification/deletion, point mutation, and epigenetic silencing, can complicate cancer gene discovery. Forward genetic approaches for cancer gene discovery are attractive because they allow unbiased, whole genome scans for cancer genes. In mice, treatment with chemical and radiologic mutagens can mutate cancer genes and promote tumor formation, but identification of the affected gene is difficult and is often achieved using a candidate gene approach. The use of insertional mutagens, such as the *Sleeping Beauty* (SB) transposon, is attractive because the inserted sequence serves as a molecular tag to facilitate identification of the affected gene. Two reports have been published, those by Collier, Carlson and co-workers [1] and Dupuy and colleagues [2], on the ability of somatically mobilized SB transposons to mutagenize, tag, and thereby lead to the identification of cancer genes. Both the scientific rationale for these experiments and the mechanisms by which SB mutagenizes and identifies cancer genes were recently reviewed [3-5]. In this discussion, we focus on what lessons can be learned from these two studies about how to improve the utility of transposon-based somatic mutagenesis for cancer gene discovery.

SB was the first nonviral insertional mutagen used for cancer gene identification. However, retroviruses have been used as powerful insertional somatic mutagens for cancer gene discovery in mice, as well as other laboratory animals, for many years [6-10]. Despite the important discoveries made, retroviruses have several limitations that the use of nonviral insertional mutagens, such as SB, are expected to overcome [5].

Retroviruses require cell infection, reverse transcription of the viral genome, and integration of the resulting provirus into the host genome to be mutagenic. Murine leukemia viruses (MuLVs) are frequently used as mutagens in mouse models of leukemia development. MuLVs and other mouse retroviruses are unable to infect nondividing cells and do so very inefficiently in poorly replicating cells [11]. This limits their utility for insertional mutagenesis in some tissues. Other tissues have physical barriers, such as the basement membrane or mucin layer, that prevent efficient infection with retroviruses [12]. In addition, MuLVs have profound insertion site bias, and therefore they do not mutagenize the entire genome equally because they have a strong preference for landing near the promoter region of actively transcribed genes [13,14]. Elements such as SB that do not exhibit such a strong insertion site preference [15] are likely to mutagenize the genome more completely. Indeed, lymphocytic leukemia associated genes uncovered using SB include many genes not previously identified by retroviruses, despite many years of this kind of work with MuLV [2]. Finally, retroviruses used in insertional mutagenesis screens must be capable of efficient infection and spread in the host animal, which

imposes tremendous limitations on the ways in which the retroviral cargo can be manipulated for specific mutagenesis projects. SB, and other cut-and-paste transposons, require only an inverted terminal repeat sequence for transposition and can therefore be engineered with diverse cargoes of mutagenic elements.

Lessons from *Sleeping Beauty*: transposon and transposase transgene design

The SB system consists of two parts: the transposon and the enzyme that mobilizes it, the transposase. The SB transposon used for somatic mutagenesis (T2/onc) contains splice acceptors in both orientations followed by polyadenylation signals, so that it can generate loss-of-function mutations in TSGs. T2/onc also contains sequences from the murine stem cell virus (MSCV) long terminal repeat (LTR) that contain enhancer/promoter elements, so that T2/onc can promote over-expression of proto-oncogenes that are near to where it lands. The version of T2/onc used by Dupuy and coworkers [2] contains a longer version of one splice acceptor, and it is therefore named T2/onc2 to denote this difference. Although T2/onc and T2/onc2 are similar, the transgenic lines generated from them harbor dramatically different numbers of transposon copies residing in a chromosomal concatomer. T2/onc lines contain approximately 25 copies of T2/onc [1], whereas the transgenic lines generated for T2/onc2 contain approximately 150 to 350 copies of transposons in their chromosomal concatomers [2]. To indicate this difference, we refer to T2/onc lines as 'low-copy lines' and T2/onc2 lines as 'high-copy lines'. Because tumor formation is hypothesized to require multiple hits in cancer genes in the same cell [16], it was hypothesized that having more transposons to mobilize would allow these hits to occur more rapidly.

Two different transposase transgenic lines have been used in somatic mutagenesis studies. Collier, Carlson and coworkers [1] used transgenic mice that express the SB10 version of the transposase under the control of the theoretically ubiquitous CAGGS promoter (CAGGS-SB10) [17]. CAGGS-SB10 mice were generated using standard pronuclear injection techniques for generating transgenic mice [18]. Dupuy and colleagues [2] generated mice (R26-SB11) in which the SB11 version of the transposase was knocked into the endogenous *Rosa26* locus using homologous recombination in embryonic stem cells. Similar targeting of transgenes to the *Rosa26* locus in mice has resulted in essentially ubiquitous expression during development and in adulthood [19,20]. Although mobilizing T2/onc from low-copy lines by CAGGS-SB10 could accelerate sarcoma formation in *Arf*[-/-] mice, it was not sufficient to result in tumor formation on an otherwise wild-type background [1]. Mobilizing T2/onc2 from high-copy lines by R26-SB11 resulted in high levels of embryonic lethality. Mice that survived to adulthood rapidly succumbed to tumor formation, primarily lymphocytic

leukemias, by 120 days [2]. The differences in tumor induction potential between these two studies [1,2] could be explained by either differences in transposase activity/ expression patterns, transposon copy number, or both.

Recently, a monoclonal antibody directed against the SB transposase useful for immunohistochemistry has become commercially available (R&D Systems, Minneapolis, MN, USA), allowing us to investigate transposase expression in transgenic mice on a cell-by-cell basis. Although our studies have not been exhaustive, we observed dramatic differences in expression between the two transgenics (Figure 1) (Rahrmann EP, Collier LS, Kuslak SL, Green LE, Largaespada DA, Marker PC, unpublished data) (Collier LS, Largaespada DA, unpublished data). In general, transposase expression in CAGGS-SB10 mice is rare, highly variegated, and mainly mesenchymal in nature (Figure 1b,c) (Rahrmann EP, Collier LS, Kuslak SL, Green LE, Largaespada DA, Marker PC, unpublished data). This variegated pattern of expression probably results from epigenetic silencing, which is known to occur with traditional transgenes that exist in multicopy arrays. Transposase is easily detected but somewhat variegated in the testis of CAGGS-SB10 mice (Rahrmann EP, Collier LS, Kuslak SL, Green LE, Largaespada DA, Marker PC, unpublished data). Testis expression is expected because of the ability of CAGGS-SB10 to promote transposon mobilization efficiently in the male germline [18,21]. At first glance, these findings appear to contradict the results of published excision assays, which detected transposase activity in all tissue types examined in CAGGS-SB10;T2/onc doubly transgenic mice [1]. However, the polymerase chain reaction based methods used to detect excision and therefore transposase activity in CAGGS-SB10;T2/onc mice are very sensitive and can detect few mobilization events. In addition, excision assays cannot accurately address the temporal and spatial pattern of transposase expression in transgenic mice.

In contrast to CAGGS-SB10 transgenics, R26-SB11 transgenics express high levels of transposase in virtually every cell type examined (Figure 1a) (Collier LS, Largaespada DA, unpublished data). However, extremely high levels of transposase expression may not always be ideal. In germline mutagenesis screens, we have generated seed mice for the same chromosomal concatomer with both R26-SB11 and CAGGS-SB10. Although the study of offspring generated from R26-SB11 seed mice was not exhaustive, we found that mobilization in the germline by CAGGS-SB10 resulted in more insertions per gamete than did R26-SB11 [21]. One potential explanation for lower germline transposition rates using R26-SB11 could be over-expression inhibition. Over-expression inhibition refers to the observation in cell culture models that, given a fixed number of SB transposons, increasing levels of SB transposase eventually lead to a decrease in transposition efficiency [22]. Although our observation by no means offers proof that this phenomenon
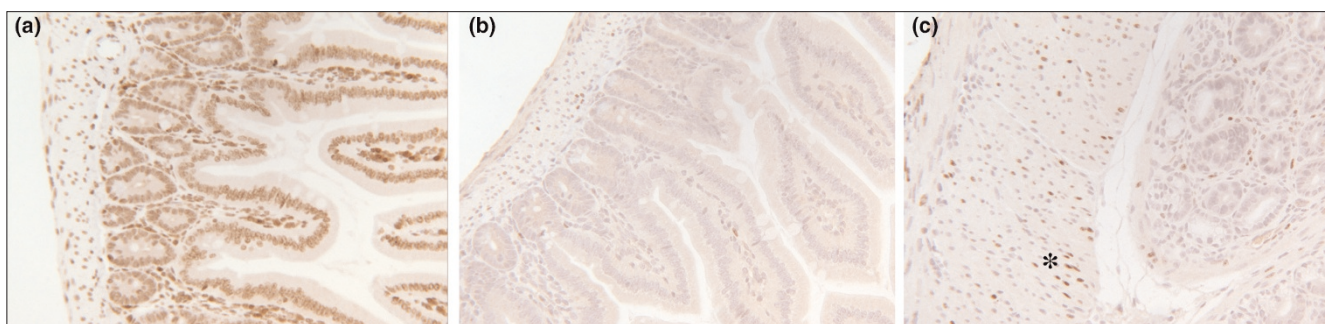
Figure 1
Immunohistochemistry staining for SB transposase. Shown is immunohistochemistry staining for *Sleeping Beauty* (SB) transposase in **(a)** R26-SB11 mice and **(b,c)** CAGGS-SB10 mice. Anti-transposase immunoreactivity is indicated by brown stain and nuclei are counter-stained blue. In panel (a), essentially all nuclei are positive for transposase in intestine from a R26-SB11 mouse. In panel (b), rare nuclei are positive for transposase in this section of intestine from a CAGGS-SB10 mouse. In panel (c), variegated mesenchymal transposase expression (asterisk) is detected in muscular tissue from the digestive tract of a CAGGS-SB10 mouse.

takes place *in vivo*, it is somewhat surprising - given the high, ubiquitous levels of transposase in R26-SB11 mice - that tumors or precancerous lesions have only been observed in a few tissues [2]. Alternatively, although SB mobilization has been detected in a wide variety of tissue and cell types *in vivo* and *in vitro* [1,23,24], *in vitro* work indicates that cell-specific factors may play a role in regulating transposition as transposition rates vary from cell line to cell line [24,25]. We must await the development of accurate methods to measure transposition in somatic tissues to examine the relationship between transposase expression level, transposition frequencies, and cancer development *in vivo*. Nevertheless, improved transposase transgenic lines, or the development of additional methods to deliver transposase expression constructs, may be useful for allowing cancer gene discovery in a wider range of tissue types [26-29].

To address whether transposon copy number influences tumor formation in somatic mutagenesis studies, we combined T2/onc low-copy lines with R26-SB11. Unlike what was reported for high-copy lines, we observed no evidence for embryonic lethality because the transgenes essentially follow Mendelian inheritance (Collier LS, Green LE, Davies M, Dupuy AJ, Copeland NG, Jenkins NA, Largaespada DA, unpublished data). T2/onc low-copy;R26-SB11 doubly trans-genic mice live longer than high-copy T2/onc2;R26-SB11 mice, but they also do eventually succumb to primarily lymphocytic leukemia (Collier LS, Green LE, Diers MD, Matise I, Largaespada DA, unpublished data). Therefore, although use of high-copy T2/onc2 lines are attractive because of the short tumor latency, the Mendelian inheritance observed with T2/onc low-copy lines and R26-SB11 allows the generation of larger cohorts of mice, and therefore more tumors to study, than would be practical with high-copy lines.

The mutagenic elements carried within a transposon may also influence its ability to promote tumor formation in various tissues. Although T2/onc was designed with elements to cause both loss-of-function and gain-of-function mutations, insertions that resulted in proto-oncogene over-expression predominated in both leukemias and sarcomas [1,2]. It is possible that use of transposons with only loss-of-function elements will greatly facilitate TSG identification, especially when coupled with a genetic background in which mitotic recombination is elevated. The results of an experimental approach like this were recently reported, in which retroviral-based mutagenesis was used to accelerate leukemia formation in *Bloom* mutant mice [30]. Viral integrations at certain common insertion sites (CISs) were shown to have undergone loss-of-heterozygosity, and thus they are potential new TSGs. However, the vast majority of viral integrations at all CISs in *Bloom* mutant leukemias were located 5' or 3' to genes, indicating that the strong promoter/enhancer activity of the retroviral LTR still imposes a bias toward proto-oncogene over-expression. Unlike retroviruses, it is possible to engineer SB vectors that contain loss-of-function only elements. In terms of oncogene identification, T2/onc is likely only to be useful in tissues in which the MSCV LTR is highly active. The use of alternative promoters in transposons for cancer gene discovery may allow more robust over-expression of proto-oncogenes in different cell types [3].

Caveats of cut-and-paste transposons for mutagenesis: local hops, genome rearrangements, and remobilization

In tumors initiated or promoted by SB mobilization, transposons have inserted into or near cancer genes. How-ever, each tumor contains many transposon integrations, because approximately 1,000 insertions were cloned from 28 sarcomas whereas 782 insertions were cloned from 16 lymphomas [1,2]. Although some of these insertions in tumors contribute to tumor formation, some are merely passenger or bystander insertions that happened to occur in

a cell that had also suffered insertions in cancer genes. In order to determine which insertions probably mark the chromosomal location of a cancer gene, CIS analysis is performed. CIS analysis looks for clustering of insertions in tumors above that which is expected by random chance. So-called Monte Carlo simulations, in which a randomly generated dataset of theoretical insertions is generated, are used to determine the amount of clustering that would be expected simply by random chance [31]. Unlike what is modeled by a Monte Carlo simulation, SB integration is not completely random [21,32-35]. On a macro scale, SB has a slight preference for integrating in and 5' to RefSeq genes [15]. A greater caveat with SB mobilized from chromosomal concatomers is the local hopping phenomenon. Local hopping refers to the observation that mobilized elements tend to reinsert near their original location. Local hopping increases the probability that, at loci linked to the concatomer, clustering will occur more frequently by random chance than at other locations in the genome. Therefore, although clustering of insertions near the concatomer is often seen, this is not always due to selection for insertions in or near cancer genes.

In the future, it may be possible to elucidate the influence of local hopping on CIS identification by performing Monte Carlo simulations that factor in local hopping rates. However, this type of analysis would be complicated by the observation that local hopping rates appear to vary from concatomer to concatomer and from transposase source to transposase source [1,2,21,32,34]. An alternative would be to determine the frequency of 'background' CISs linked to the donor locus by cloning large numbers of control SB insertions from tissues of transgenic mice that are not under selection for tumorigenesis. This was done using embryos from doubly transgenic high-copy T2onc2;R26-SB11 mice [2]. However, this analysis would have to be performed anew with every transposon concatomer, transposase combination.

In fact, the rate of SB local hopping in tumors appears to be much less than that observed in germline mutagenesis screens. In germline screens, the local hopping rate has been as high as approximately 58% of insertions occurring within 10 megabases on either side of the donor concatomer [34]. In somatic screens, the local hopping rate calculated from sarcomas generated using a low-copy donor locus on chromosome 1 and CAGGS-SB10 was 20% of insertions found within 20 megabases either side of the concatomer [1]. This is far less than reported germline rates of local hopping but far greater than the rate in lymphomas generated by mobilization using R26-SB11 [2]. Preliminarily, the rate of local hopping in lymphomas generated by R26-SB11 mobilization of transposons from a low-copy line is less than that observed in sarcomas (Collier LS, Adams DJ, Akagi K, Bradley A, Largaespada DA, unpublished data), indicating that both the donor locus and transposase activity influence local hopping rates.

There are several explanations for the decreased local hopping rate observed in R26-SB11 lymphomas as compared with CAGGS-SB10 sarcomas. The differences in rate could be explained by transposon remobilization. The R26-SB11 transposase may be so active in hematopoietic cells that transposons that do not provide a selective advantage to the cell may have many opportunities to remobilize during tumor development. Theoretically, the more opportunities a single transposon has to mobilize, the greater is the chance that it will eventually integrate in a location unlinked to the original concatomer. Another possibility emerges from observations recently made in our laboratory during a germline screen using SB [21]. In this study, deletions flanking the concatomer were frequently observed in mice with novel transposon insertions. Deletions flanking the concatomer were also visible using fluorescence *in situ* hybridization (FISH) in somatic cells from the spleens of transposon concatomer;R26-SB11 mice. Given this observation, a distinct possibility is that local hopping rates decrease when R26-SB11 is used as a source of transposase because the DNA immediately *cis* to the concatomer has been deleted and therefore no insertions can be cloned from that region.

These transposition-associated deletions linked to the concatomer can potentially complicate somatic screens using SB or other transposable elements. For example, these deletions could partially explain the high embryonic lethality rates observed with high-copy T2/onc2 lines and R26-SB11. One hypothesis is that high mobilization rates due to active transposase or high numbers of transposons could increase the frequency and size of these deletions. Should a large enough deletion occur during development, it could result in haploinsufficient lethality. These deletions could also promote tumor development if they happen to remove one copy of a haploinsufficient TSG. The use of array-based comparative genome hybridization can be used to determine whether deletions, local or genome wide, are consistently selected for in tumors induced by SB. In addition, it will be important to carefully characterize the genomic location of any additional concatomers generated for somatic screens. For example, concatomers located in a gene-poor region of the genome might be ideal for cancer gene screens. Although difficult with current technology, generating concatomers on artificial chromosomes would be an ideal way to address this issue and eliminate the concerns about local hopping and local genomic rearrangements and deletions. Because similar deletions have been observed with other endogenous cut-and-paste transposable elements [36], they may also occur in genetic screens in which such elements are mobilized from multi-copy, chromosomally resident concatomers.

Another potential complication to the use of SB for somatic mutagenesis is that, in the presence of continued transposase expression, transposon remobilization may occur. For this reason, the tumor extracted from the animal may not possess all of the original transposon integrations

that promoted tumorigenesis. However, there is likely to be strong selective pressure to maintain insertions in cancer genes. Remobilization of a transposon in a cancer gene that is providing a selective advantage could prove growth deleterious to a cell in which it occurs, probably eliminating that cell from the tumor clone. This would be especially true when the tumor cell relies on the MSCV LTR in T2/onc to over-express a proto-oncogene. However, the DNA repair process that accompanies SB excision generates a characteristic footprint, usually consisting of the addition of five nucleotides, at the former insertion site [35]. It remains possible that remobilization of a SB transposon that had landed in a TSG would leave behind this footprint, which - although invisible to linker-mediated polymerase chain reaction techniques - would result in a loss-of-function mutation. This may be true only for SB insertions into TSG exons because the introns are much bigger targets and the disruption of splicing caused by SB-induced footprints in introns is likely to be a rare event. Therefore, we suspect that most TSG mutations will remain marked by transposon insertion in tumors. Even if TSG mutation by transposon footprint occasionally occurs, it should be possible to identify insertions in the TSG with analysis of enough individual tumors. Although remobilization of SB transposons in the germline generally is rare [32,33,37], the remobilization rate in the soma has not been measured.

To date, the only published accounts of somatic screens using nonviral transposable elements for cancer gene discovery have involved the SB transposon system. This is probably because SB was the first vertebrate active cut-and-paste transposon characterized [23], and thus there has been more time to develop SB for such screens. In recent years, however, additional transposable elements that are capable of transposing in the mouse germline and soma have been described. Each element could have its own advantages and disadvantages in somatic screens.

The pros and cons of other transposable elements
Part 1: retroelements
LINE-1 elements (L1) and other retrotransposons transpose to new locations in the genome via an RNA intermediate. This RNA intermediate must exit the nucleus, be translated, and be returned to the nucleus with all the necessary enzymes to be reverse transcribed while being inserted at a new location in the genome [38]. As a result of this lifecycle, retroelements do not exhibit a preference for inserting near their donor locus, which would eliminate the local hopping complication discussed above for SB. Both mouse and human synthetic L1 elements have been reported to retrotranspose *in vitro* [39,40] and *in vivo* in mice transgenic for L1 expression cassettes [41-44]. Somatic retrotransposition has been detected in these transgenic mice [42-44]. The activities of

L1s have been improved by the identification of more active human elements and by optimizing the translational efficiency of murine L1 elements [40,42].

L1 elements do have potential advantages for use in somatic mutagenesis screens. Transposed L1 elements, unlike cut-and-paste elements, cannot remobilize. Therefore, any L1 transposition event would be fixed in all progeny of the initial cell that suffered the integration. An additional potential advantage for L1 as compared with cut-and-paste elements is the potential for L1 self-expansion [45]. It has been hypothesized that because a L1 transgene can continually generate L1 transcripts, a theoretically infinite expansion of L1 integrations in a somatic cell could occur. Because DNA transposons transpose in a cut-and-paste manner, the number of transposons in a somatic cell is essentially limited to the number of transposons in the chromosomal concatomer. However, there are ways in which the number of cut-and-paste transposons in a somatic cell could amplify, including transposition during S phase or duplication of the chromosome harboring the transposon concatomer [45]. However, it does appear that current technologies involving SB provide enough transposon copies for effective mutagenesis, because concatomers consisting of 25 copies contain sufficient numbers of transposons to promote tumorigenesis (Collier LS, Green LE, Davies M, Dupuy AJ, Copeland NG, Jenkins NA, Largaespada DA, unpublished data).

In terms of insertion site preference, analysis of a small number of insertions ($n = 48$) in mice transgenic for the active human L1 element revealed only a modest preference against inserting in RefSeq genes [42]. A survey of 170 insertions in mice transgenic for the murine L1 transgene identified no preference for or against insertion near or in RefSeq genes [44]. The cloning of additional insertion sites may be necessary to address fully whether L1 has any insertion site bias in transgenic models.

One potential obstacle to the use of L1 as a cancer gene discovery tool, compared with cut-and-paste elements, is the tendency of L1 to incompletely reverse transcribe the transposon RNA as it integrates, leading to truncation. *In vivo* work using a mouse transgenic that expresses a synthetic human L1 element indicates that very few novel L1 integrations are full length, with the vast majority being truncated at the 5' end [42]. An analysis of 25 *de novo* insertions from the synthetic murine L1 element revealed that all were truncated at the 5' end [44]. This could prove deleterious to cancer screens if such truncations removed mutagenic elements such as splice acceptors or over-expression elements. This issue could partially be alleviated by placing the mutagenic elements as far 3' as possible, but truncated elements are often quite small [42,44] and proper design of the element will only partially ameliorate the problem.

Part 2: cut-and-paste elements

Several cut-and-paste elements are known to be active in vertebrate cells. These include *Minos* [46-49], *Tol2* [25,50-52], *piggyBac* (PB) [25,53,54], *Frog Prince* [55], and a hyperactive version of *Himar1* [56-58]. Of these elements, only *Minos* [49,50] and PB [54] have been tested in transgenic mice. Although not tested for activity in transgenic mice, *Tol2* is known to transpose in mouse embryonic stem cells [51] and mouse liver cells when delivered using hydrodynamics-based techniques [52]. *Tol2* has several characteristics that could make it appealing for use in somatic mutagenesis studies. First, initial studies using *Tol2* indicated that, unlike SB, the transposon can carry larger cargo without decreased transposition efficiency [52]. This could also be an advantage when larger mutagenesis cassettes are needed. Second, unlike SB and PB (see below), no evidence of over-expression inhibition has been observed with *Tol2* [25,52]. Although it remains possible that even higher levels of transposase could result in decreased transposition efficiency, the current data are encouraging.

In vitro experiments have detected higher transposition rates for PB than SB [25,59]. *In vivo* studies have been more limited, but PB is known to be active in mouse one cell embryos and in the mouse germline [54]. Unlike SB, PB integrations studied from transposition events *in vivo* indicate a strong preference for inserting in genes, because 50 out of 104 integrations occurred in validated genes (an additional 20 integrations occurred in invalidated or predicted genes). A recent study *in vitro* in human cell lines also detected a strong preference for PB to insert into RefSeq genes (48.8% for PB versus 39.1% for SB versus 33.2% for random) [15,59]. If this trend holds true for somatic integrations *in vivo*, then this could be an advantage for PB in cancer genetic screens because it would increase the probability that any mobilization event could mutate a gene. Similar to local hopping, this preference could also complicate CIS analysis because a random PB insertion set would not be distributed evenly throughout the genome [14]. Initial studies indicate that, like *Tol2*, the PB transposon can carry larger cargo without decreased transposition efficiency [52,54]. However, like SB, PB is subject to over-expression inhibition *in vitro* [25]. It has also been reported that PB does not appear to local hop like many other cut-and-paste transposons. The story is probably incomplete, however, because this assumption was made based on only three integrations cloned from germline transposition events from a chromosomally resident concatomer [54]. It can be argued that although these three events mapped to three different chromosomes, the dataset is clearly too small to address this question completely.

To date, published reports on *Minos* have found it to be only weakly active *in vivo* in the mouse [48,49]. For example, sensitive excision assays could detect *Minos* mobilization from a donor concatomer in transposon/transposase doubly transgenic mice in which transposase was expressed from a transgene that drives expression in the thymus and spleen. FISH was used to detect transposon sequences in these somatic cells. In 19 of 3,114 metaphases, transposon probe signal was detected at chromosomal sites away from the donor concatomer [48]. However, FISH is not a very sensitive technique for detecting single copy elements. Given that, like SB, *Minos* is a member of the Tc1/*Mariner* family, it is possible that the events detected by FISH could represent chromosomal rearrangement events involving the concatomer, similar to those we have observed with SB [21]. In the *Minos* study [48], the transposon donor concatomer was located near the telomere of chromosome 14. A chromosome 14 telomeric probe was therefore used to determine that the donor chromosome was intact in cells where transposon signal was found on other chromosomes. However, it is possible that this probe did not lie sufficiently close to the donor locus to detect genomic inversions, translocations, or deletions involving the donor locus. Nevertheless, somatic transposition of *Minos* was not sufficiently high to be detectible by Southern analysis [48]. In the female mouse germline, a published report also indicates that *Minos* is active, but only 8.2% of offspring of transposon/transposase positive 'seed' females harbored novel transposon integrations [49].

For all other mouse-active DNA transposon systems, analysis is too preliminary to indicate the degree to which they will have caveats such as local hopping and deletions flanking the concatomer. However, potentially the greatest obstacle to use of all new transposable elements - both retroelements and cut-and-paste elements - in cancer gene discovery is the need to achieve mobilization rates in the soma that are sufficiently high to ensure that there is an opportunity to insertionally mutate a cancer gene. For example, in SB-induced lymphocytic tumors, the average number of novel SB insertions visible by Southern blot is about 30 [2]. In p19 *Arf*$^{-/-}$ sarcomas accelerated by SB mobilization, the average number is about five [1]. However, each tumor is likely to be highly polyclonal in terms of novel integrations because the actual number of insertions cloned per tumor was 48 in leukemias and 37 in sarcomas. To date, although limited, all published reports on the rate of new insertions per gamete for germline mobilizations of other transposable elements have been significantly lower than that achieved using SB [21,37,41,42,44,49,54]. The rate of SB mobilization does vary from concatomer to concatomer [37], and so the generation of additional donor concatomers for other transposons could allow greater activity in the germline. In addition, varying transposase (or RNA production for the case of L1) by the use of different promoters may improve these rates in the future. A final possibility is that no direct comparison can or should be made between somatic, germline, and *in vitro* mobilization rates, and that a comparison between elements will only be possible once somatic mobilization has been thoroughly tested.

Conclusion

Although we can speculate on the advantages and dis-advantages of each element, their different insertion site preferences may mean that each element will be useful for identifying a different set of cancer genes. For example, although both SB and retroviruses have been used as insertional mutagens for cancer gene discovery in lympho-mas, comparison of cancer genes tagged by both systems only partially overlap [2]. No matter which element is used, such somatic screens should be designed and executed with intense planning and forethought. Careful choice of trans-posable element type, the mutagenic elements carried by the transposon, the donor transposon concatomer location, promoters for transposase or RNA expression, and even predisposed genetic background will be instrumental in the success of these screens in the future.

Competing interests

DAL is a founding partner of Discovery Genomics, Inc. (DGI), a company that has previously licensed *Sleeping Beauty* technology. DGI focuses on gene therapy. LSC and DAL are inventors on a pending patent that covers the use of *Sleeping Beauty* for somatic mutagenesis for cancer gene discovery.

Acknowledgements

References

1. 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:** 272-276.
2. Dupuy AJ, Akagi K, Largaespada DA, Copeland NG, Jenkins NA: **Mammalian mutagenesis using a highly mobile somatic *Sleeping Beauty* transposon system.** *Nature* 2005, **436:**221-226.
3. Dupuy AJ, Jenkins NA, Copeland NG: *Sleeping Beauty***: a novel cancer gene discovery tool.** *Hum Mol Genet* 2006, **15 Spec No 1:**R75-R79.
4. Starr TK, Largaespada DA: **Cancer gene discovery using the *Sleeping Beauty* transposon.** *Cell Cycle* 2005, **4:**1744-1748.
5. Collier LS, Largaespada DA: **Hopping around the tumor genome: transposons for cancer gene discovery.** *Cancer Res* 2005, **65:**9607-9610.
6. Uren AG, Kool J, Berns A, van Lohuizen M: **Retroviral insertional mutagenesis: past, present and future.** *Oncogene* 2005, **24:** 7656-7672.
7. Payne LN: **Retrovirus-induced disease in poultry.** *Poult Sci* 1998, **77:**1204-1212.
8. Levy LS, Starkey CR, Prabhu S, Lobelle-Rich PA: **Cooperating events in lymphomagenesis mediated by feline leukemia virus.** *Leukemia* 1997, **11(Suppl 3):**239-241.
9. 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 nephroblas-toma.** *Oncogene* 2003, **22:**665-673.
10. Pajer P, Pecenka V, Kralova J, Karafiat V, Prukova D, Zemanova Z, Kodet R, Dvorak M: **Identification of potential human onco-genes by mapping the common viral integration sites in avian nephroblastoma.** *Cancer Res* 2006, **66:**78-86.
11. Yamashita M, Emerman M: **Retroviral infection of non-dividing cells: old and new perspectives.** *Virology* 2006, **344:**88-93.
12. Wang G, Williams G, Xia H, Hickey M, Shao J, Davidson BL, McCray PB: **Apical barriers to airway epithelial cell gene transfer with amphotropic retroviral vectors.** *Gene Ther* 2002, **9:**922-931.
13. Wu X, Li Y, Crise B, Burgess SM: **Transcription start regions in the human genome are favored targets for MLV integra-tion.** *Science* 2003, **300:**1749-1751.
14. Wu X, Luke BT, Burgess SM: **Redefining the common insertion site.** *Virology* 2006, **344:**292-295.
15. 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:**2085-2094.
16. Kinzler KW, Vogelstein B: **Lessons from hereditary colorectal cancer.** *Cell* 1996, **87:**159-170.
17. Okabe M, Ikawa M, Kominami K, Nakanishi T, Nishimune Y: **'Green mice' as a source of ubiquitous green cells.** *FEBS Lett* 1997, **407:**313-319.
18. Dupuy AJ, Fritz S, Largaespada DA: **Transposition and gene dis-ruption in the male germline of the mouse.** *Genesis* 2001, **30:** 82-88.
19. Soriano P: **Generalized lacZ expression with the ROSA26 Cre reporter strain.** *Nat Genet* 1999, **21:**70-71.
20. Mao X, Fujiwara Y, Chapdelaine A, Yang H, Orkin SH: **Activation of EGFP expression by Cre-mediated excision in a new ROSA26 reporter mouse strain.** *Blood* 2001, **97:**324-326.
21. Geurts AM, Collier LS, Geurts JL, Oseth LL, Bell ML, Mu D, Lucito R, Godbout SA, Green LE, Lowe SW, *et al.*: **Gene mutations and genomic rearrangements in the mouse as a result of trans-poson mobilization from chromosomal concatemers.** *PLoS Genet* 2006, **2:**e156.
22. Geurts AM, Yang Y, Clark KJ, Liu G, Cui Z, Dupuy AJ, Bell JB, Largaespada DA, Hackett PB: **Gene transfer into genomes of human cells by the sleeping beauty transposon system.** *Mol Ther* 2003, **8:**108-117.
23. Ivics Z, Hackett PB, Plasterk RH, Izsvak Z: **Molecular reconstruc-tion of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells.** *Cell* 1997, **91:**501-510.
24. Izsvak Z, Ivics Z, Plasterk RH: **Sleeping Beauty, a wide host-range transposon vector for genetic transformation in ver-tebrates.** *J Mol Biol* 2000, **302:**93-102.
25. Wu SC, Meir YJ, Coates CJ, Handler AM, Pelczar P, Moisyadi S, Kaminski JM: **piggyBac is a flexible and highly active transpo-son as compared to Sleeping Beauty, Tol2, and Mos1 in mammalian cells.** *Proc Natl Acad Sci USA* 2006, **103:**15008-15013.
26. Liu F, Song Y, Liu D: **Hydrodynamics-based transfection in animals by systemic administration of plasmid DNA.** *Gene Ther* 1999, **6:**1258-1266.
27. Zhang G, Budker V, Wolff JA: **High levels of foreign gene expression in hepatocytes after tail vein injections of naked plasmid DNA.** *Hum Gene Ther* 1999, **10:**1735-1737.
28. Yant SR, Ehrhardt A, Mikkelsen JG, Meuse L, Pham T, Kay MA: **Transposition from a gutless adeno-transposon vector stabi-lizes transgene expression in vivo.** *Nat Biotechnol* 2002, **20:**999-1005.
29. Wilber A, Frandsen JL, Geurts JL, Largaespada DA, Hackett PB, McIvor RS: **RNA as a source of transposase for sleeping beauty-mediated gene insertion and expression in somatic cells and tissues.** *Mol Ther* 2006, **13:**625-630.
30. Suzuki T, Minehata K, Akagi K, Jenkins NA, Copeland NG: **Tumor suppressor gene identification using retroviral insertional mutagenesis in Blm-deficient mice.** *EMBO J* 2006, **25:**3422-31.
31. Mikkers H, Allen J, Knipscheer P, Romeijn L, Hart A, Vink E, Berns A: **High-throughput retroviral tagging to identify components of specific signaling pathways in cancer.** *Nat Genet* 2002, **32:** 153-159.
32. Fischer SE, Wienholds E, Plasterk RH: **Regulated transposition of a fish transposon in the mouse germ line.** *Proc Natl Acad Sci USA* 2001, **98:**6759-6764.

33. Carlson CM, Dupuy AJ, Fritz S, Roberg-Perez KJ, Fletcher CF, Largaespada DA: **Transposon mutagenesis of the mouse germline.** *Genetics* 2003, **165**:243-256.
34. Horie K, Yusa K, Yae K, Odajima J, Fischer SE, Keng VW, Hayakawa T, Mizuno S, Kondoh G, Ijiri T, *et al.*: **Characterization of *Sleeping Beauty* transposition and its application to genetic screening in mice.** *Mol Cell Biol* 2003, **23**:9189-9207.
35. Luo G, Ivics Z, Izsvak Z, Bradley A: **Chromosomal transposition of a Tc1/mariner-like element in mouse embryonic stem cells.** *Proc Natl Acad Sci USA* 1998, **95**:10769-10773.
36. Gray YH: **It takes two transposons to tango: transposable-element-mediated chromosomal rearrangements.** *Trends Genet* 2000, **16**:461-468.
37. Geurts AM, Wilber A, Carlson CM, Lobitz PD, Clark KJ, Hackett PB, McIvor RS, Largaespada DA: **Conditional gene expression in the mouse using a *Sleeping Beauty* gene-trap transposon.** *BMC Biotechnol* 2006, **6**:30.
38. Han JS, Boeke JD: **LINE-1 retrotransposons: modulators of quantity and quality of mammalian gene expression?** *Bioessays* 2005, **27**:775-784.
39. Kubo S, Seleme Mdel C, Soifer HS, Perez JL, Moran JV, Kazazian HH, Jr., Kasahara N: **L1 retrotransposition in nondividing and primary human somatic cells.** *Proc Natl Acad Sci USA* 2006, **103**:8036-8041.
40. Han JS, Boeke JD: **A highly active synthetic mammalian retrotransposon.** *Nature* 2004, **429**:314-318.
41. Ostertag EM, DeBerardinis RJ, Goodier JL, Zhang Y, Yang N, Gerton GL, Kazazian HH Jr: **A mouse model of human L1 retrotransposition.** *Nat Genet* 2002, **32**:655-660.
42. Babushok DV, Ostertag EM, Courtney CE, Choi JM, Kazazian HH Jr: **L1 integration in a transgenic mouse model.** *Genome Res* 2006, **16**:240-250.
43. Muotri AR, Chu VT, Marchetto MC, Deng W, Moran JV, Gage FH: **Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition.** *Nature* 2005, **435**:903-910.
44. An W, Han JS, Wheelan SJ, Davis ES, Coombes CE, Ye P, Triplett C, Boeke JD: **Active retrotransposition by a synthetic L1 element in mice.** *Proc Natl Acad Sci USA* 2006, **103**:18662-18667.
45. Bestor TH: **Transposons reanimated in mice.** *Cell* 2005, **122**:322-325.
46. Franz G, Savakis C: **Minos, a new transposable element from *Drosophila hydei*, is a member of the Tc1-like family of transposons.** *Nucleic Acids Res* 1991, **19**:6646.
47. Klinakis AG, Zagoraiou L, Vassilatis DK, Savakis C: **Genome-wide insertional mutagenesis in human cells by the *Drosophila* mobile element *Minos*.** *EMBO Rep* 2000, **1**:416-21.
48. Zagoraiou L, Drabek D, Alexaki S, Guy JA, Klinakis AG, Langeveld A, Skavdis G, Mamalaki C, Grosveld F, Savakis C: **In vivo transposition of *Minos*, a *Drosophila* mobile element, in mammalian tissues.** *Proc Natl Acad Sci USA* 2001, **98**:11474-11478.
49. Drabek D, Zagoraiou L, deWit T, Langeveld A, Roumpaki C, Mamalaki C, Savakis C, Grosveld F: **Transposition of the *Drosophila hydei Minos* transposon in the mouse germ line.** *Genomics* 2003, **81**:108-11.
50. Koga A, Suzuki M, Inagaki H, Bessho Y, Hori H: **Transposable element in fish.** *Nature* 1996, **383**:30.
51. Kawakami K, Noda T: **Transposition of the Tol2 element, an Ac-like element from the Japanese medaka fish *Oryzias latipes*, in mouse embryonic stem cells.** *Genetics* 2004, **166**:895-899.
52. Balciunas D, Wangensteen KJ, Wilber A, Bell J, Geurts AM, Sivasubbu S, Wang X, Hackett PB, Largaespada DA, McIvor RS, *et al.*: **Harnessing a high cargo-capacity transposon for genetic applications in vertebrates.** *PLoS Genetics* 2006, **2**:e169.
53. Fraser MJ, Ciszczon T, Elick T, Bauser C: **Precise excision of TTAA-specific lepidopteran transposons piggyBac (IFP2) and tagalong (TFP3) from the baculovirus genome in cell lines from two species of Lepidoptera.** *Insect Mol Biol* 1996, **5**:141-151.
54. Ding S, Wu X, Li G, Han M, Zhuang Y, Xu T: **Efficient transposition of the piggyBac (PB) transposon in mammalian cells and mice.** *Cell* 2005, **122**:473-483.
55. Miskey C, Izsvak Z, Plasterk RH, Ivics Z: **The *Frog Prince*: a reconstructed transposon from *Rana pipiens* with high transpositional activity in vertebrate cells.** *Nucleic Acids Res* 2003, **31**:6873-6881.
56. Lampe DJ, Churchill ME, Robertson HM: **A purified mariner transposase is sufficient to mediate transposition in vitro.** *EMBO J* 1996, **15**:5470-5479.
57. Lampe DJ, Akerley BJ, Rubin EJ, Mekalanos JJ, Robertson HM: **Hyperactive transposase mutants of the Himar1 mariner transposon.** *Proc Natl Acad Sci USA* 1999, **96**:11428-11433.
58. Keravala A, Liu D, Lechman ER, Wolfe D, Nash JA, Lampe DJ, Robbins PD: **Hyperactive himar1 transposase mediates transposition in cell culture and enhances gene expression in vivo.** *Hum Gene Ther* 2006, **17**:1006-1016.
59. Wilson MH, Coates CJ, George AL Jr: ***PiggyBac* transposon-mediated gene transfer in human cells.** *Mol Ther* 2007, **15**:139-145.