Published in final edited form as:

Transl Res. 2013 April ; 161(4): 265–283. doi:10.1016/j.trsl.2012.12.005.

Evaluating Risks of Insertional Mutagenesis by DNA Transposons in Gene Therapy

Perry B. Hackett^a, David A. Largaespada^a, Kirsten C. Switzer^b, and Laurence J.N. Cooper^b

^aDepartment of Genetics Cell Biology and Development, Center for Genome Engineering and Masonic Cancer Center, University of Minnesota, Minneapolis, MN 55455

^bDivision of Pediatrics and Department of Immunology, University of Texas MD Anderson Cancer Center, Houston, TX

Abstract

Investigational therapy can be successfully undertaken using viral- and non-viral-mediated *ex vivo* gene transfer. Indeed, recent clinical trials have established the potential for genetically modified T cells to improve and restore health. Recently the *Sleeping Beauty* (SB) transposon/transposase system has been applied in clinical trials to stably insert a chimeric antigen receptor (CAR) to redirect T-cell specificity. We discuss the context in which the SB system can be harnessed for gene therapy and describe the human application of SB-modified CAR⁺ T cells. We have focused on theoretical issues relating to insertional mutagenesis in the context of human genomes that are naturally subjected to remobilization of transposons and the experimental evidence over the last decade of employing SB transposons for defining genes that induce cancer. These findings are put into the context of the use of SB transposons in the treatment of human disease.

INTRODUCTION

Gene therapy – in many the phrase elicits a reaction of increased therapeutic possibilities and yet high risk to patients. The high potential is obvious. Natural and chimeric genes can produce products to restore homeostasis and provide clinical benefit. The question now is reducing the risk and broadening the appeal so that gene therapy can enter into the mainstream of clinical practice.

Retroviral-associated adverse events in gene therapy trials

The high risk of virus-mediated transduction using recombinant retroviruses, first recognized in mice in 1983 [1], became clear in pilot clinical trials when five of 20 patients treated for X-linked severe combined immunodeficiency disease (X-SCID) unfortunately developed leukemia three or more years after administration of the therapeutic retroviral vector [2–4]. Whereas three of the five were successfully treated for both immunodeficiency and leukemia, two have died. A linkage between the retrovirus carrying the therapeutic IL2 γ c gene and the leukemias was inferred based on a common region of integration

© 2012 Mosby, Inc. All rights reserved

For correspondence: Perry Hackett: perryh@umn.edu.

Others: DAL: larga002@umn.edu; KCS: KCSwitzer@mdanderson.org; LJNC: ljncoper@mdanderson.org

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

All authors have read the journal's policy on conflicts of interest.

upstream of a resident LMO2 oncogene in hematopoietic stem cells (HSCs) [5–11]. Similarly, gamma-retroviruses, encoding the gp91^{Phox} gene for treatment of chronic granulomatous disease (CGD) [12, 13], have led to leukemia and one death [14], demonstrating that the vectors were not synergistic with just one disease. These events in HSCs cemented the notion of the dangers of insertional mutagenesis and biological selection for the retrovirus-induced outgrowths of treated cells [15–18]. In contrast, retroviral vectors carrying either the adenosine deaminase (ADA) gene for treatment of ADA-linked immunodeficiency [19–24] or WASp for Wiskott-Aldrich Syndrome [25] have not been associated with similar adverse events. In the CGD patients, retroviral integrations regional to growth-related genes have been associated with enhanced proliferation [26–29]. Currently, by using altered retroviruses and appropriate doses, adverse events have been avoided [30–32]

Thus, the linkage of vector integration and adverse mutagenesis is not unequivocal. In some cases clonal expansion of treated cells was temporary and devoid of adverse effects. Indeed, limited expansion of transgenic cells had the potential of increasing the likelihood of successful gene therapy. Moreover, in contrast to early treatments of HSC, viral-mediated transduction of T cells for adoptive immunotherapy has not resulted in adverse outcomes [33]. Thus, cell-type does make a difference. Whereas HSC with elevated expression of growth-related genes can result in T-cell leukemia and/or lymphoma, comparable treatments of T-cells for adoptive immunotherapy has not resulted in undesirable clinical outcomes [34–36].

There may be differences between cell types for integration sites as well as the availability of endogenous genes for activation. Analyses of gamma retrovirus-modified hematopoietic stem cells show that integration sites were clustered often near genes involved in growth control and cell survival, e.g., in one study 40% of insertions were clustered in only 0.36% of the genome [37, 38]. Retroviral parameters associated with integration effects and genotoxicity are under intense scrutiny [39–47]. These and other studies make it clear that only a small subset of all the integrations near proto-oncogenes actually cause transformation [48–50].

The deliberate use of insertional vectors to uncover oncogenes by inducing leukemia and solid tumors has exacerbated the concerns of insertional mutagenesis by the very same vectors when used for gene therapy. Retroviruses [51–53] as well as transposons [54–56] have long been used for this purpose, although to obtain a meaningful number of effects, cooperating mutations are required either by pre-selection of mice with identified gene knockout genotypes or by delivering multiple genetic hits to single genomes [57–60]. Notably, as discussed below, screens with both types of vectors have revealed the involvement of multiple genetic alterations to induce cancer. One consequence of these studies has been the elucidation of the multiplicity of mechanisms by which insertions can induce mutations that lead to adverse events, as summarized in Fig. 1.

Thus, although AIDS patients are at increased risk for some cancers due to their immune state [61], it has been a surprise to many that lentiviral vectors based on the HIV-1 LTRs have not been found in preclinical and clinical studies to have associated adverse events stemming from either activation, inactivation, and/or alteration of splicing of endogenous genes [62, 63]. Although the absence of cancers and leukemias following integrations of lentiviruses (including those with AIDS) might be due to the ability of the lentiviral vectors to infect non-dividing cells for therapeutic benefit [64] or their selection of a different set of genetic loci for integration [38, 39, 65–67] even though the integration of both is directed into outward-facing major grooves on nucleosomal DNA [68].

Non-viral gene therapy and insertional mutagenesis

Non-viral gene therapy is an alternative to using viral vectors that presents an opportunity to avoid some of the issues such as preferential integration sites associated with most viruses [69, 70]. Our research has revolved around the use of the *Sleeping Beauty* (SB) transposon/transposase for gene therapy [71] with a recent focus on employing the SB system to introduce a chimeric antigen receptor (CAR) to redirect the specificity of human T cells. Transposons have two major advantages over viruses as gene therapy vectors. First, clinical grade manufacture and quality control, for use in many patients is easier, more reliable and less expensive than employing clinical-grade virus. Second, unlike viral cargos that often are integrated either into or proximal to genes that can incur mutagenic risks, SB transposons have few known preferences for integration sites [72]. Nevertheless, as discussed later, insertional mutagenesis is always a concern.

Over the past two years there have been significant findings in areas that pertain to genotoxicity and its impact on gene therapy. First, it is evident that endogenous transposons are far more active in human cells than was surmised until very recently. Two questions arise from these findings: 1) Do these elements induce similar genetic consequences as therapeutic transposons? 2) Do cells have mechanisms to cope with insertional mutagenic agents? Second, the results from the 1000 Genomes Project [73] and ENCODE Project [74] have demonstrated that the interactions of genetic elements in our chromosomes are far more variable and complex than customarily thought. These findings are reviewed in the sections below.

In light of these recent discoveries, we have three objectives in this review that pertain to the use of SB transposons in the clinic. 1) Evaluate the relative impact of SB transposition in consideration of the apparent relatively high activity of endogenous genetic elements. 2) Evaluate the data from transposon-mediated induction of cancer in mouse studies in terms of risks to patients undergoing SB-based gene therapy. 3) Discuss approaches that are required for dealing with the identified risks of transposon-mediated adverse events in the clinic.

TRANSPOSONS AND NATURAL GENETIC VARIATION IN HUMAN CELLS

Approximately 45% of the human genome is composed of transposable elements [75, 76] and up to two-thirds [77, 78] of chromosomes may be derived from transposons that have been adapted to support cell function. Transposons are divided into two classes. Class I transposons are retro-elements that spread by a copy-and-paste mechanism whereby the transposon is transcribed and the RNA transcript reverse-transcribed for insertion elsewhere in the genome. Class II transposons, which include the SB system, are DNA sequences that can “hop/jump” from one site to the next via a cut-and-paste mechanism. The two classes of transposons have distinctive features that are important considerations in gene therapy.

Retrotransposable elements

Class I transposons comprise about 42% of the genome and are categorized into four subtypes. The first are long interspersed elements (LINEs) that are up to 6–8 kb in length; there are about 500–800,000 copies of these elements, of which about 100 are active and about 10 may be highly active (“hot”) as a result of encoding the necessary products for integration of their RNA polymerase II transcription product [79]. The second subtype of Class I transposons comprises about 1.5 million short interspersed elements (SINEs) of about 100–300 bp that are transcribed by RNA polymerase III and litter the entire genome. The third category comprises a relatively undefined set of sequences that are several hundred bp in length and probably transcribed by RNA polymerase II. A fourth category comprises the nearly half-million retrovirus-like, human endogenous retroviral sequences (HERVs) that appear to be remnants of retroviruses that invaded the genome over eons and since

degraded; these account for about 8% of the human genome. Resurrection of a consensus HERV-K sequence indicated that these elements preferentially integrated into or proximal to transcriptional units [80]. Of a total set of nearly 3 million retrotransposable elements, only a subclass of LINEs, designated L1s, are active and responsible for insertional mutagenic events that result in genetic disease [79, 81], including cancer [82, 83].

The activity of LINEs in cells and their potential for causing deleterious mutations in human genomes has been known for decades [84–86]. Nevertheless, the huge number of retro-elements, which have not been linked to disease-causing mutations, led to the inference that they were largely inactive and played roles mainly in genome evolution [87–89]. This notion was reconsidered when whole genome sequencing became available [90, 91]. Extensive L1 insertions were found in somatic neuronal cells [92, 93]. L1 mobilization appeared to occur in the absence of methyl-CpG-binding protein-2 (MeCP2), a protein involved in global DNA methylation and human neurodevelopmental diseases [94]. Other studies found abundant new retroelement integrations, some of which were associated with cancer [95, 96]. A failure in epigenetic silencing may account for the rare cases of L1s introducing somatic genetic lesions [97, 98], which are relaxed in germline cells [99]. L1 activity can trans-mobilize SINE sequences as well [100]. Moreover, in the mouse genome endogenous retroviral elements can also spread [101], suggesting that there might be even more hopping elements in the human genome.

Assuming about 10^{13} cells per human and just the 10 hot L1 retrotransposons, there would be around 10^{14} potential mutagenic sequences in the average human; more than 10,000 per basepair of genomic sequence. Clearly the controls over these elements are extraordinarily tight such that few actually remobilize. Nevertheless, given the immensity of this background of elements, there is a strong hint that the human genome has adapted to accommodate insertion of transposon sequences. The identity of these controls is at least in part epigenetic and their induction remains to be clarified. Although retro-transposition and DNA transposition occur by different mechanisms (copy-paste compared to cut-and-paste), the phenotypic results will be similar when integration interrupts a genomic sequence that either encodes proteins or regulates their expression. As discussed in detail below, the SB transposase has little preference for integration into protein-encoding genes whereas retro-transposable elements appear to have a preference for integration into protein-coding genes. These controls over insertional damage by endogenous transposons will likely be relevant to transposons used for gene therapy.

DNA Transposons

Class II transposons are DNA sequences that are excised from the genome for insertion elsewhere in the same, or different DNA molecule, by a cut-and-paste mechanism (Fig. 2). The transposon sequence that is excised is precise in that the termini of the transposed sequence are exact. However, when the donor DNA sequence is repaired, there is often a “footprint” that is left, which in the case of SB, appears to vary according to cell type, but often is a 5bp TAC(A/T)G insertion [102]. The approximate 300,000 Class II transposons comprise about 3% of the human genome. Active class II transposons are defined by inverted terminal repeats (ITRs) that flank a transposase gene that commonly does not have a promoter and thus is dependent on integration proximal to an endogenous promoter. This feature keeps the transposon from remobilizing in most cases, but does allow spread of the transposon under some circumstances when the transposase gene is activated [103]. In general, transposons and viruses are both natural pathways for introduction of new genetic material into cells. However, the general evolutionary strategy of many viruses is to infect and make many viral particles, regardless of consequences to the cell (organism), for further infection of other hosts. In contrast, transposons generally only occasionally insert into cellular genomes, but are carried over evolutionary periods in all the offspring of the cell.

The Sleeping Beauty Transposon/Transposase System for Gene Therapy

We, and others, explored the potential of DNA transposons for use as vectors for transgene delivery [104] based on their capacity to accommodate genetic cargos of up to, and on occasions, more than 10kbp [69, 105–107] (Fig. 2). No active Class II transposon has been uncovered in the human genome, although many sequencing projects mask repetitive elements for convenience of analysis. Indeed, because active class II transposons are exceedingly rare but useful for gene transfer, the SB system (transposon plus cognate transposase) was resurrected from an approximate 14 million-year sleep from salmonid genomes [104]. As a result of its origin in fish (last common ancestor for fish and humans lived about 500 million years ago), the SB transposase does not recognize class II transposons in the human genome. Since its first use as gene therapy vector [105], the transposase has been re-engineered for increased activity, from SB10 [104] through SB11 [106] and other intermediates [108] to the extremely hyperactive SB100× [109]. Likewise, the inverted terminal repeat structure of original transposon, *T*, also has been re-engineered for greater activity to produce *T2* and *T3* [106] as well as other versions [107].

For current clinical applications of the SB system, the active transposase gene is supplied in *trans* rather than being incorporated into the plasmid carrying the therapeutic transposon. As a result, the plasmid is smaller than it would be with the SB transposase expression cassette, which leads to more efficient delivery and transposition [69, 106]. However, a SB transposon can be on the same plasmid (*cis*) rather than on a different plasmid (*trans*) as the transposase expression cassette. This would reduce cost, as only one clinical-grade DNA plasmid would be needed, but this would be at the risk of decreased efficiency of gene transfer. This has “real world” implications since patients enrolled on clinical trials infusing genetically modified T cells have varying abilities to donate peripheral blood affecting the quality and quantity of T cells available for transposition. Thus, the ability to titrate the amount and ratio of two DNA plasmids coding for transposon and transposase facilitates the generation of patient-specific T cells.

The advent of the SB system demonstrated advantages of transposons as vectors – they are easy to use if they can be delivered into cells, they result in precise integration of a defined genetic sequence that is flanked exactly by the inverted repeats of the transposon, and each integration event is separate [110], such that concatemers do not arise from transposition (note, concatemers of transposons are often introduced purposely into genomes by illegitimate recombination for insertional mutagenic studies described later in the review). A further advantage of the SB system, but not all other transposons, is that it requires only a TA sequence for integration, with very few preferences for integration unlike most integrating viruses [111]. There are about 200 million TA sites in the human genome, which is appealing for obtaining a wide distribution of integrated vectors. However, not all TA sites are apparently equal. SB transposons, unlike other transposons, appear to prefer “flexible” TA sites [112, 113] that can be screened in any genetic locus [72]. Nevertheless, of all of the integrating vectors currently under study, the SB system appears to have the least preference for integration either into or proximal to transcriptional units [114]. As a consequence, the SB transposon system has been developed as a leading non-viral vector for gene therapy (currently in clinical trials) that has the advantages of using naked DNA and chromosomal integration of the therapeutic expression cassette. The SB vector has been validated for *ex vivo* gene delivery to stem cells, including T-cells for the treatment of lymphoma, and SB transposons have been delivered to liver for treatment of various systemic diseases in mice, including hemophilia [105, 115, 116] and mucopolysaccharidoses types I and VII [117].

The possibility of remobilization of a transposon from residual transposase activity is a theoretical concern (Fig. 3). In most studies with a promoter that has a limited duration of

expression such as the CMV early promoter, expression of SB transposase from an episomal plasmid is transient, *e.g.*, in the liver transcription of the SB transposase gene is reduced about 10,000 fold over the first few days following uptake by hepatocytes [118]. Human cells have developed self-defense mechanisms to prevent the continued transposases as derived protein degradation products can interfere with their enzymatic activity [119]. In addition, the transposase, since it is derived from fish, would be immunogenic a subject to immune mediated clearance. Nevertheless, the possibility of a few cells continuing to express transposase cannot be ruled out. When a SB transposon is re-mobilized, it will leave a “footprint”, an insertion or deletion that is variable in length, but often an addition of 5bp [102]. In an exon that encodes a polypeptide, this would result in a frameshift leading to an abnormal protein. Because 1) SB transposons do not have a pronounced preference for transcriptional units, 2) exons comprise no more than 2% of the human genome, and 3) the rate of excision of a transposon in a cell is less than 10^{-4} based on gene deliveries to liver [118] and in transgenic mice (Section on Transposon-Mediated Induction of Cancer, below), the chance of an adverse event is estimated to be less than 10^{-6} . It may be far lower - too rare to detect. Since few genes are haplo-insufficient, the probability of an adverse event occurring from mutation of a single allele will be significantly lower.

One reoccurring question is whether SB transposons might “skip along a genome” (serially inserting and excising) like a rock skipping across the surface of a pond before sinking into a final resting site. This is thought to be highly unlikely for two reasons. The first is evolutionary [120]; such a process would damage genomes by leaving footprints and thus be counterproductive to minimal disturbance of the host genome. The second is based on the conserved nature of DDE transposases, including SB [121], in which the flexible target sites for integration [72, 112] undergo bending to render strand-transfer irreversible by facilitating the snapping back of the integrated DNA following the concerted single-strand invasion steps [122]. That is, upon integration, the two ends the transposon comprising the synaptimal complex would snap, like a mousetrap that has been sprung. In addition to explaining how transposases have evolved to avoid skipping, this mechanism explains why transposition reactions are so infrequent even when all the necessary components are in excess.

Alternative measures to prohibit sustained activity include supplying mRNA transcribed *in vitro* as the source of transposase [123], although the instability of mRNA and reliable delivery of sufficient quantities may introduce quality control issues if used for gene therapy. An alternative is to construct transposon-expression cassettes in which the promoter for the transposase gene is inside the transposon [124]; with this configuration, expression of the transposase gene would cease although there is the drawback the outward directed transposase-driving promoter then would be positioned to potentially activate transcription of nearby genes.

We have estimated potential problems of remobilization of SB transposons from data accrued from multiple studies conducted with mice. The multitude of remobilization experiments in mice using specialized SB transposons is discussed in detail below. In addition to the experimental approaches in mice, we can use evaluations of clinical quality-control criteria to estimate the possible extent of transposon hopping in transduced cells. Currently, SB mRNA can be detected by qRT-PCR in 100ng but not in 20ng of T-cell DNA. 1 mRNA/20ng of T-cell DNA corresponds to about 1 SB mRNA/ 2×10^{10} cellular mRNA molecules. Since the average cell expresses about 10,000 mRNA species [125], the detectable limit is equivalent to one SB mRNA in 2×10^6 T cells. In fibroblasts, transposition efficiencies are about 1% (unpublished compilations), which would suggest a maximal rate of 1 remobilization event in about 10^8 T cells if a single mRNA is adequate to provide the necessary level of transposase. Moreover, remobilization *per se* does not imply adverse

consequences. The range in new sites into which a remobilized transposon carrying a therapeutic expression cassette integrates is the same as the range of original sites into which the transposon originally integrated. Because less than 2% of the genome comprises exons [74] and since integration of transposons causes little effect on endogenous genes [126], the chances of remobilization into an exon would be about 1 event in 10^{10} cells. Assuming that 10% of genes could contribute to a transformed phenotype, the chances of remobilization leading to an adverse event would be about 10^{-11} . This is a very low number that is consistent with the mouse studies described later in this review. The number is much lower than the background level of retro-transposition, described next. In the case of the current clinical trials, discussed below, up to 10^{10} SB-modified T cells may be infused, which would suggest the odds of a remobilization event are up to 10%. However, the footprint would be inconsequential because the repaired integration site would have been mutated by the original insertion that would direct an alternative polypeptide that was partially encoded by the transposon's inverted terminal repeat. The new integration event would have the same probability of leading to an adverse consequence as any of the far more frequent original integration events.

Recent insights into chromosomal organization and gene expression in humans

Sequencing of genomes obtained from cancer cells has revealed the first direct information on the variation of mutation rates in chromosomes of somatic cells and the importance of chromatin organization and conformation [127, 128]. The ENCODE project [74], which employed 147 different cell types found that far from being composed largely of “junk” DNA, about 80% of the human genome was involved in some sort of regulated event and that 95% of the genome was within 8 kbp of a protein-binding site. Moreover, the landscape of transcription in human cells is diverse with about 40% of the genome being transcribed at levels that spanned six orders of magnitude for polyadenylated RNA and with the average gene having four splicing isoforms although there is always a dominant species that comprises 30% or more of the splicing alternatives for a given transcriptional unit [129]. Strikingly, a large proportion of the transcripts in the human genome appear to be initiated from repetitive elements, mainly retrotransposons. It is now clear that transcriptional activation is far more complex than thought heretofore [130]. Enhancer elements can work over long distances via chromosomal looping with only about 7% of the looping interactions being to the nearest gene [131, 132]. Clearly, physical proximity is not a simple predictor for gene activation by incoming transgenic expression cassettes. The selectivity of transcriptional regulatory proteins to interact with selective promoters may explain the pronounced lack of activation of endogenous genes that lead to cancer by enhancers in lentiviral vectors. These findings support earlier findings that murine leukemia virus integration up to 100 kbp upstream of the *c-myb* locus could activate a linearly distal locus through chromosomal looping loci [133].

The insertion of transgenic DNA into chromosomes can alter epigenetic marks that affect gene expression [98], and integrated retroviral cassettes are subject to epigenetic silencing through identified cellular proteins [134]. The advent of high throughput sequencing revealed a number of surprises. In sum, the interactions between regions of the genome that affect gene expression are complex and not based entirely on physical (linear) proximity. Cells have a bewildering level of variety in gene expression at both the transcriptional and RNA processing levels. Moreover, the genome appears to have the ability to assimilate transposable elements and even make use of them. All of these features strongly suggest that the genome is unexpectedly flexible in terms of events that would be expected to destabilize its panoply of functions. Recent findings from the 1000 Genomes Project and other whole genome sequencing projects support this assertion.

Genetic variation in the human population

Deep sequencing has found that insertions and deletions (indels), which are equivalent to insertions of transposons, are about 10-fold less frequent in the human genome than single nucleotide polymorphisms (22,000 vs 1,800 per genome compared to reference) with up to 50% of the indels being novel in any given individual [135, 136]. More pertinent to the consequences of insertional mutagenesis, these studies have revealed that the average human genome has approximately 250 to 300 loss-of-function mutations, with 50 to 100 in genes related to human disease [137, 138] and about 20 completely inactivated genes [139] as classified by the Human Gene Mutation Database [<http://www.hgmd.org>]. Studies of genome samples from individuals and families have recently shown that copy number variations, especially of L1 and Alu retro-elements, were found in somatic tissues from the same individual, with preferred integration into protein-coding genes, many of which regulate growth [82, 93, 140–142] but nearly none associated with disease. Thus, the human genome is not only highly variable [143, 144], but it can sustain genetic hits from transposons without apparent genetic consequences. This conclusion is consistent with early studies of transposition in nematodes where it was found that despite a strong preference for integration into transcriptional units [145] many, if not most, integrations did not have a phenotypic effect. Detailed studies revealed that the transposons could be removed during pre-mRNA processing [126].

Genetic consequences of natural transposition and therapeutic transposition

One way of addressing insertional mutagenesis by SB transposon vectors is to compare nearly random insertion with natural variation and the germline mutation rate. The background mutational rate from the approximate 1800 indels/individual and the approximate 10^{14} *hot* retrotransposable elements in a human may not be apparent due to epigenetic silencing or activation of immune responses that are able to detect aberrant cells. Thus the number of potentially mutagenic elements is millions, if not billions, of fold greater than the number originating from gene therapy using SB transposons. Superficially, with the histories of gamma-retrovirus- and lentivirus-mediated gene therapies in mind, it would appear that the consequences of insertions of SB transposons carrying a therapeutic cassette with enhancers that are *not* designed to interact promiscuously with all promoters should be relatively small. The probabilities of adverse consequences can be further analyzed. As only approximately 1.2% of the genome encodes proteins [74], about 98% of integrations are unlikely to affect protein sequences in any way other than their rates of expression, which can vary widely. The remaining 2% of integrations that might occur into exons must be viewed in terms of other conclusions from deep sequencing of human genomes. The average gene is multiply spliced giving isoforms that, for the most part, have undefined specific roles. Thus, only one of the isoforms of multiply spliced genes would be affected by an exonic integration, which may be the reason that background integration of retrotransposable elements does not cause as many adverse events as might be expected. As with retroviruses that can be epigenetically silenced [146], SB transposons, epigenetic silencing and RNA interference are probable causes of either complete suppression or reduction in transgene expression [147–150].

Although some transposase-like enzymes used for gene transfer, e.g., the phiC31 recombinase, can cause chromosomal translocations [151], the SB transposase is not associated with such activities when used for integration of a therapeutic expression cassette [152]. We note that the situation is different when an SB transposon is hopped out of a concatemer of transposons for integration elsewhere. As discussed in the following section, in this case chromosomal rearrangements have been observed close to the locus in which the concatemer resides. This situation will not be applicable in gene therapy settings where the chromosomes will not be pre-established to have multiple transposons in a single locus.

TRANSPOSON-MEDIATED INDUCTION OF CANCER/LEUKEMIA IN MICE

a. Induction of cell transformation by SB transposons in mice

The SB and *piggyBac* transposon systems have been used to induce cancer in transgenic mice via an insertional mutagenesis mechanism [153, 154]. This may indicate, upon initial consideration, that transposons are potentially dangerous from a genotoxicity perspective. However, more careful consideration of the differences between the experimental conditions generated in mice (described in the following sections) and those actually encountered by patients in a clinical setting reveals that cancer is an unlikely outcome as a side effect of therapeutic transposon delivery. First of all, in the mouse experiments the transposon vectors are *especially designed to induce alterations* to endogenous genes. Secondly, the mouse experiments are done in a way that results in the mobilization of multiple copies of the transposon in every cell of a given tissue. Thirdly, in the mouse experiments the transposase enzyme is *expressed continuously* in all cells of a given tissue throughout the lifetime of the animal. Lastly, in most of the experiments described so far in mice, efficient cancer induction was only achieved if the mouse was *genetically predisposed to cancer*, often by including a tumor suppressor gene mutation in the background. These points suggested to us, even in consideration that humans have far more cells and longer life-spans than mice, that unless the above conditions are met, cancer would be an unlikely sequelae of one time transposon vector mobilization and transposition into human chromosomes. Indeed, our published and unpublished observations do demonstrate that cancer is not induced at any significant level unless a transposon is used that carries a strong promoter/splice donor, is present in all cells in the body in multiple copies, and mobilized continuously for the lifetime of the animal. In addition, we have found that for most tissues, efficient cancer development requires a genetic background that confers predisposition to induction of cancer. Thus, given the limitations of using mice as a model system (fewer cells per organ and considerably shorter lives) in the setting of human gene therapy, we do not expect cancer as a likely side effect of use of transposon vectors.

i. General scheme for transposition-mediated mutagenesis in mice—We were led in our efforts to develop a flexible system for somatic insertional mutagenesis using transposons by a large body of literature, and our own work, on the induction of cancer in mice using murine leukemia viruses (MuLV). MuLV can cause cancer by acting as an insertional mutagen, either inserting near and activating proto-oncogenes or inserting within and inactivating tumor suppressor genes [155]. The features of the integrated provirus that can cause these effects on endogenous genes, are the enhancer and promoter sequences within the long terminal repeats (LTR), the splice donor and acceptor within the body of the virus, or the polyadenylation site within the long terminal repeat [156–158]. The tumor DNAs can be used to isolate new candidate cancer genes, by using the integrated provirus as a molecular tag. Sites of the genome that are recurrently mutated by proviral insertion (i.e., in multiple, independent tumors) are called common integration sites (CIS) and have been shown by experience to harbor cancer genes.

We hypothesized that an SB transposon, designed to mimic the ability of a retroviral element to cause both gene loss- and gain-of-functions, could be used to “tag” cancer genes in solid tumors. The transposon, T2/Onc contains splice acceptors (SA) followed by polyadenylation (pA) signals in both orientations (Fig. 4). Upon insertion into introns, these elements are designed to intercept upstream splice donors and elicit premature transcript truncation. Between the two SA's are sequences from the 5'LTR of the murine stem cell virus (MSCV), which contain strong promoter and enhancer elements that are methylation-resistant and active in stem cells [159–161]. Immediately downstream of the LTR is a splice donor (SD) for splicing of a transcript initiated from the LTR into a neighboring gene. In

subsequent work, other promoter sequences replaced the MSCV LTR sequences and similarly effective transposons for myeloid leukemia [162] or carcinoma induction were created [163]. The original T2/Onc transposon and later derivatives are thus specialized to identify both tumor suppressors and oncogenes.

Transgenic lines harboring multi-copy (usually ~25–200 copies) chromosomal concatemers of T2/Onc, or similar transposons, are used in these studies. Transposase can be supplied in a ubiquitous or tissue-specific manner. In both cases cohorts of mice are aged for tumor development. Once tumors are harvested, the transposon insertions are PCR amplified and sequenced. Usually hundreds or thousands of insertions per tumor genomic DNA are identified. Regions of the genome, and associated genes, that are mutated by transposon insertion in multiple, independent tumors are sought using bioinformatic and statistical analyses. Clustering of insertions in certain regions, beyond what is expected by chance, indicates selection for the rare insertion events, among many neutral or negatively insertion events, that can give the cell a selective growth or survival advantage. The accumulation of novel insertions through out their genomes causes the eventual appearance of tumors in mice. The general use of the SB system for identification of cancer genes is illustrated in Fig. 5. The insertion sites are analyzed by looking for T2/Onc insertions at reproducibly mutated genes. In this way, we can develop a cancer genetic fingerprint for various tumor types.

When T2/Onc was combined with SB transposase in both wild-type and cancer-predisposed mice, there was either induction or acceleration of tumors and/or leukemia [164, 165]. In both cases, the SB-accelerated or initiated tumors were characterized by recurrent somatic, tumor-specific insertions that occurred at dozens of known and novel cancer-genes. The hyperactive SB11 transposase [106], expressed from the nearly ubiquitous Rosa26 promoter in transgenic mice, has been used for most of these studies. When two different versions of T2/Onc were combined with the Rosa26-SB11 transgene in mice, the result was almost uniform development of T-cell malignancy [165, 166], which in some mice co-occurred with glioma brain tumors [166].

Tissue-specific mutagenesis with SB has been used to develop informative models of various forms of human solid tumors. Tissue-specificity was achieved by using Cre-recombinase that could be activated in a tissue-specific manner to excise a loxP-flanked stop (LSL) cassette that separated the SB11 cDNA from the Rosa26 promoter. Cancers and/or pre-neoplasias could be induced in mice in a variety of tissues – including the liver, gastrointestinal tract, brain (glioma and medulloblastoma), and the prostate [167–169]. For instance, analyses of over 16,000 transposon insertions identified 77 candidate colorectal cancer (CRC) genes, 60 of which are mutated and/or dysregulated in human CRC and thus are most likely to drive tumorigenesis. Moreover, analysis of the co-occurrence of transposon insertion sites within GI tumors, using a model that assumes a Poisson distribution of insertions at TA dinucleotides, identified *common co-occurring insertions* (CCIs) that were detected more frequently than expected by chance, e.g., Apc and Nsd1 - two known tumor suppressors [55]. A screen for pancreatic cancer genes on a *Kras*^{G12D} background has identified many candidate cancer genes and validated a novel epigenetically silenced tumor suppressor gene, USP9X [170].

ii. Control lines of mice and induction of tumors/leukemia—The experience of performing many of these studies *in mice* has shown that unless many conditions are met, tumor induction by transposon mobilization is not efficient. First of all, expression of the SB transposase by itself, neither has *ever* revealed any abnormalities in mice, nor an increase in cancer incidence, despite being expressed ubiquitously for the lifetime of the mice. This suggests that SB transposase does not cause significant chromosomal instability or mobilize

endogenous mammalian *Tc1/mariner*-family transposons at a rate high enough to cause cancer. Similarly, mice carrying just the transposon vectors are normal with no increase in cancer when aged, compared to non-transgenic controls.

Our work also suggests that a strong promoter sequence within the transposon is required for efficient tumor induction. When the Rosa26-SB11 transgene was employed to cause body-wide, lifelong mobilization of a transposon vector called T2/GT3, which lacks a promoter but contains a splice acceptor and polyadenylation signal, there was no significant increase in cancer incidence compared to control mice. In stark contrast, as mentioned above, lifelong mobilization of T2/Onc-like transposons does cause an increase in cancer incidence attributable to insertional mutagenesis. This suggests that T2/GT3 lacks a critical element for efficient cancer induction, probably a promoter/splice donor sequence that allows efficient oncogene activation to occur. However, SB-catalyzed T2/GT3 mobilization in p53^{+/-} mutant mice caused an apparent decrease in tumor latency, suggesting that a transposon without a promoter can accelerate tumor development in a cancer-prone genetic background (DAL, unpublished observations). Such studies highlight the influence that genetic background could have on genotoxicity by integrating vectors.

iii. The importance of genetic background for oncogene discovery—As mentioned above, the efficient induction of T-cell malignancy [165] and myeloid leukemia [162] did not require tumor-predisposed genetic backgrounds. Moreover, a T2/Onc3, with a CAG promoter in place of the MSCV LTR efficiently induces a variety of solid tumor types at low frequencies without a predisposing background [163]. Tissue-specific mobilization of T2/Onc is able to induce gastrointestinal track epithelial adenomas and adenocarcinomas in Villin-Cre cells in transgenic mice [167] and liver-specific mobilization of T2/Onc using albumin-Cre and Rosa26-LSL-SB11 mice can cause full-blown hepatocellular carcinoma in otherwise wild-type mice [169]. However, in most other tissue-specific mutagenesis experiments, mice do not develop tumors efficiently unless also carrying tumor-susceptibility mutations such as *p53*^{R270H} or *Kras*^{G12D} [170]. Moreover, the expression cassette in the transposons contained only transcriptional control elements and RNA processing motifs that were designed to cause maximal chaos when inserted into or proximal to chromosomal genes. Such would not be the case when SB transposons are used for gene therapy.

These studies have a bearing on the estimated risk for SB transposon vectors for causing cancer when used for human gene therapy. In the case of human gene therapy, the copy number of integrating vectors would be far fewer, ideally one or two, rather than 25–200 per cell, as in mouse mutagenesis experiments. The vectors would not be designed for disruption of endogenous genes, by loss- or gain-of-function. Importantly, the period of transposon mobilization would be transient. Thus, the experience of mouse somatic mutagenesis experiments suggests that the development of malignancies as a side effect of therapeutic gene delivery to human cells using SB should be an unlikely event. The vector design, copy number and duration of SB transposase expression are all fundamentally different than what occurs during the mouse somatic mutagenesis screens. Cancer induction in mice using SB mutagenesis was revealed to be an inefficient process requiring great efforts to maximize the number of transposon insertion mutations per cell and have many cells at risk for cancer development. Moreover, most screens have been done in strongly cancer-predisposed genetic backgrounds, a situation not likely to be often encountered in clinical settings.

As discussed next, the SB system has been recently deployed for the genetic transformation of autologous and allogeneic T cells for investigational treatment of B-cell malignancies. In these clinical studies, a chimeric antigen receptor (CAR) is used to redirect the specificity of T cells to bind to the tumor-associated antigen (TAA) CD19. It is likely that about 5% of

cancer patients treated with CAR-modified T cells carry some cancer-predisposing genetic alterations since it is thought that such a percentage of cases may have an inherited genetic component. Thus, it may be that potential mutations arising from insertions of SB transposons may be more likely to produce a secondary leukemia when used in T cells from patients with cancer

TRANSPOSON-MEDIATED GENE THERAPY IN THE CLINIC

Concerns

Transposition was applied to practice in 2012 using the SB system to enforce expression of a second generation [171] CD19-specific CAR in T cells. The CAR redirects T-cell specificity for CD19 independent of major histocompatibility complex as antigen-recognition is mediated through the scFv domain of a CD19-specific monoclonal antibody and T-cell activation is initiated through chimeric endodomain. Adoptive transfer of CD19-specific T-cells in early clinical trials has shown therapeutic efficacy based on expected loss of normal CD19⁺ B cells and significant destruction of large numbers of malignant B cells [172–176]. Patients with large tumor burdens that are rendered lymphopenic through prior conditioning chemotherapy can experience the numeric expansion of infused CAR⁺ T-cells. If synchronous activation of the infused T cells occurs, then there can be supra-physiologic elevation of inflammatory cytokines and systemic toxicities. These adverse events may be temporarily delayed from the time of infusion as the numbers of *in vivo* CAR⁺ T cells swells. However, the signs and symptoms can be managed through the administration of systemic corticosteroids as well as biologic agents that interrupt the inflammatory cascade. Infusion of CD19-specific CAR⁺ T cells is currently warranted in patients at high risk from dying due to progressive B-cell malignancies. At present, most trials targeting CD19 infuse T cells that express CAR as a result of viral transduction. Indeed, hundreds of transductions and infusions of genetically modified T cells have been undertaken without apparent evidence of insertional mutagenesis [177]. Furthermore, targeting CD19 in patients with aggressive leukemias and lymphomas is justified as the “on-target” side effects due to lysis of normal B cells and associated hypogammaglobulinemia is tolerated given the potential benefit of controlling or curing the underlying B-cell malignancy. Thus, the use of T cells as a cellular substrate and the expression of a CD19-specific CAR provide an appealing platform to assess the safety and feasibility of using the SB system for gene therapy.

The first-in-human trials using the SB system are currently underway at MD Anderson Cancer Center (INDs 14193, 14577, 14739, 15180). Three of these pilot studies infuse patient- and donor-derived CD19-specific CAR⁺ T cells after standard-of-care autologous and allogeneic hematopoietic stem-cell transplantation, including umbilical cord blood transplantation. A fourth trial infuses autologous CD19-specific CAR⁺ T cells after lymphodepleting chemotherapy. These trials have passed institutional (CRC, IRB, IBC) and federal (NIH-OBA and FDA) regulatory committees. The standard operating procedures (SOPs) and associated worksheets are in place to manufacture and release the genetically modified T cells in compliance with current good manufacturing practice (cGMP) for Phase I/II trials. In preparation for these trials, data was accumulated regarding the ability of the SB system to enforce expression of a CAR. The CAR itself was modified to not only redirect T-cell specificity to CD19, but to employ an endodomain with two chimeric signaling motifs to achieve a fully-competent T-cell activation event defined as CAR-mediated proliferation, cytokine production, and specific lysis. The introduction of the CAR was achieved by electro-transfer (using a commercially-available Nucleofector device from Lonza) of supercoiled DNA coding for SB transposon, and transposition was achieved by co-electro-transfer of a supercoiled DNA plasmid coding for the hyperactive SB transposase, SB11 [178]. T cells expressing stable integrants of CAR could be readily achieved through co-culture on designer artificial antigen-presenting cells (aAPC) in the

presence of the soluble recombinant cytokines interleukin IL-2 and IL-21. The aAPC, derived from K562, were genetically modified to co-express CD19 as well as CD64 and the T-cell co-stimulatory molecules CD86, CD137L, and a membrane-bound mutein of IL-15. The recursive addition of these γ -irradiated aAPC to the electroporated T cells results in the reproducible, rapid, and massive numeric expansion of CAR⁺ T cells. The electroporation uses defined ratios of T cells, SB transposon and SB transposase resulting in between 1 and 2 copy numbers of CAR per T-cell genome [179]. The propagation also uses defined ratios of T cells and aAPC to ensure efficient outgrowth of T cells with at least 80% expressing CAR emerging within a few weeks of electroporation. Long-term co-culture (28 days) with aAPC doesn't affect T-cell activity or function because the T cells can effectively lyse target tumor cells and produce cytokine in an antigen-dependent manner while maintaining a naïve/memory phenotype. The CAR⁺ T-cells can further effectively control B-cell tumor in an immunocompromised NSG-mouse model thus providing evidence for their persistence. Moreover, the T-cell telomere lengths are not shortened after co-culture [180].

Meeting the concerns

Toxicities correlated with infusion of genetically modified CAR⁺ T cells can be divided into adverse events associated with (i) lymphodepletion of the recipient to improve engraftment, (ii) triggering the CAR protein and (iii) those associated with integration of the CAR transgene. Patients that receive myelosuppressive or myeloablative chemotherapy to induce lymphopenia are at risk from opportunistic infections and toxicity from the administered drugs. Indeed, there has been one death attributed to administering cyclophosphamide in an elderly patient with bulky CLL that received CD19-specific CAR⁺ T cells [172]. This adverse event appears to be an exception as lymphodepleting chemotherapy is associated with expected complications and can be managed by clinical teams' experiences in caring for medically fragile patients. There are toxicities arising from infusing CAR⁺ T cells. There have been unexpected adverse events due to “on target” toxicities as the CAR recognizes tumor-associated antigen (TAA) on normal cells. For example, reversible liver and biliary tract damage occurred in patients receiving CAR⁺ T cells rendered specific for carbonic anhydrase IX (CAIX) [181]. However, one patient has died due to HER2/neu-specific T cells recognizing this TAA, which was inadvertently expressed on normal lung parenchyma [182]. Such unforeseen toxicities may be ameliorated using an inter-patient T-cell dose-escalation schema and pausing between cohorts of patients before dose-escalation to evaluate for emergence of delayed adverse events. Some expected on-target toxicities have occurred and these may not lead to modification of the T-cell dose. This occurs when T cells target the B-lineage TAA, CD19. Recipients can lose their normal B-cell repertoire as the infused CAR⁺ T cells do not distinguish between normal and malignant B cells [173–176]. Loss of humoral immunity appears tolerable in patients at risk from dying from progressive B-cell malignancies.

However, these patients need monitoring of serum immunoglobulin levels and if hypogammaglobulinemia develops then they benefit from intravenous immunoglobulin as they can succumb to viral infections. The coordinated activation of CAR by large amounts of TAA can lead to infusion-related toxicities. These can be delayed in time relative to the actual intravenous administration of T cells and are manifested as a “cytokine storm”. In severe cases patients experience fever and degrading vital signs. These events can also be complicated by “tumor lysis syndrome” associated with (welcomed) destruction of large amounts of tumor cells. These adverse events are typically managed by administration of systemic corticosteroids. If there are continued symptoms then etanercept and tocilizumab may be given to control high levels of cytokines. Indeed, serial measurement of IFN- γ and IL-6 in peripheral blood may serve as a biomarker for those patients at risk of developing a cytokine storm. The third type of risk associated with infusing CAR⁺ T cells stems from

insertional mutagenesis. Unlike gene transfer into hematopoietic stem cells [17, 183, 184] there has not been a genotoxic event in genetically modified T cells. Indeed, hundreds of infusions have been delivered using mostly recombinant retrovirus to transduce T cells [177]. Non-viral gene transfer has been safely undertaken based on the electro-transfer of DNA into clinical-grade T cells [185, 186]. We have improved the integration efficiency of DNA plasmids using the SB system to integrate CD19-specific CAR into T cells for human application. Extensive in-process testing and release testing is undertaken prior to infusion of the SB-modified T cells [178, 180]. The in-process evaluations include undertaking PCR to exclude integration of SB11 transgene, measurement of the number of CAR integrants per T-cell genome, and assessing the diversity of the T-cell receptor (TCR) to exclude the emergence of a clonal or oligoclonal population of genetically modified T cells that have superior proliferative advantage in culture and that might herald unwanted uncontrolled growth *in vivo*. To guard against this possibility we undertake a culturing assay to measure the potential for autonomous proliferation. This assay is in compliance with CLIA and is part of the release testing and thus recorded on the T-cell certificate of analysis. Other release tests include measurements of (i) sterility (bacteria, fungal, mycoplasma, endotoxin), (ii) viability, (iii) enumeration, (iv) chain of custody (measurement of HLA class I), (v) identity (CD3 expression), (vi) contamination (expression of CD19 and CD32 to exclude B cells and aAPC) and (vi) expression of transgene. These SB-modified clinical-grade T cells are prepared in compliance with current good manufacturing practice for Phase I/II trials at MDACC from peripheral and umbilical cord bloods.

Phase I studies are underway to establish safety. Safety is a primary concern to be balanced with the medical condition of the patient for these early-phase clinical trials employing a new approach to gene therapy. A detailed background on the safety data have been published [187].

CONCLUSIONS

The barriers to widespread adoption and impact of gene therapy are chiefly in translating promising pre-clinical data to the human laboratory. Never before has there been so much promise to the therapeutic potential of genetically modifying the human genome. This golden era is tarnished by the contracting financial support that erodes the ability of investigators to evaluate the human application of gene therapy, which has been successfully practiced on the bench and in animals. The SB system provides a clear line of sight to test new gene therapy. The low cost of manufacturing clinical-grade SB plasmids and availability of expertise to manufacture these expression vectors within the non-profit (*e.g.*, Production Assistance for Cellular Therapies under the auspices of the National Heart, Lung, and Blood Institute) or for-profit sectors contrasts with the expense, uncertainty, and lengthy timeline to manufacture recombinant clinical-grade virus. After passing institutional and federal regulatory reviews, the first patients have now successfully received T cells genetically modified with the SB system [188, 189]. This new approach to gene therapy employs a commercially available electroporation device and readily available APCs. This lays the groundwork for other groups to adopt the SB system for human application of cells other than T cells and transgenes other than CARs.

Future Directions - Expanding the potential of transposon-mediated gene therapy

Until a few years ago, many labs pursued the goal of identifying “safe havens” for integration of therapeutic vectors [190, 191]. Zinc finger nucleases (ZFNs) were first used for introducing double-strand breaks at specific sites in the genome to facilitate site-specific recombination of an expression cassette [192–197]. The first attempts to combine the targeted integration ability of zinc fingers with SB transposons consisted of adding a zinc-finger DNA-binding domain to SB transposase [198, 199], which resulted in diminished SB

activity and marginal targeting. More recent approaches have employed either a targeted zinc-finger nuclease-mediated double-strand DNA break [200], the adeno-associated virus Rep protein [201] or targeting ribosomal RNA genes [202, 203]. Both strategies report better outcomes. The advent of TAL-Effector Nucleases (TALENs) for directing double-strand DNA breaks with increased resolution and ease, and perhaps accompanied with lower rates of off-targeting, should facilitate insertions of therapeutic cassettes into selected sites in the genome [204, 205].

The electro-transfer of DNA plasmids coding for the SB system has been used to stably integrate a CAR into clinical-grade T-cells. The transposition event can be harnessed to introduce multiple transgenes by synchronous electro-transfer of SB transposons from multiple DNA plasmids. Using the DNA plasmids as “building blocks” allows investigators to generate populations of cells by combining sets of DNA plasmids expressing one or more transgenes while employing the same methodology and importantly the same standard operating procedures in a GMP facility. The SB system can also be combined with genetic editing tools such as designer ZFNs and TALENs [206].

Although we have calculated extremely low odds that integrated SB transposons will cause an adverse effect, either by initial integration or by potential remobilization that leaves a frame-shift footprint in a protein-encoding gene, the chance is not zero. In consideration of this possibility, the addition of a suicide gene, *iCasp9*, has been engineered into a *piggyBac* transposon. In this system, administration of a small chemical inducer, AP20187, will direct dimerization of mutant Caspase-9 polypeptides leading to cell death [207]. This approach is available for other transposons.

Clearly, the optimal method of gene therapy is to remove a deleterious mutation by genome editing/replacement; which requires a clinically meaningful level of correction in a large number of cells. The first steps along these lines have been taken for restoration of homeostasis in a mouse model of hemophilia [208]. The expansion to humans for this and other diseases is anticipated.

Acknowledgments

We thank Dr. Elena Aronovich for careful reading and editing of our manuscript, Mr. Jason Bell for help with the figures, and our colleagues in the Center for Genome Engineering for many insightful discussions. We acknowledge the financial support of NIH grants 1R01DK082516 and P01-HD32652 (PH), CA16672, CA124782, CA120956, CA141303, CA163587, and CA148600 (LJNC), and R01CA113636, R01CA134759 (DAL). PBH and DAL have equity in Discovery Genomics, Inc., a small biotech company that receives funding from the NIH to explore the use of transposons for gene therapy.

REFERENCES

1. Wagner EF, Covarrubias L, Stewart TA, Mintz B. Prenatal lethality in mice homozygous for human growth hormone gene sequences integrated in the germ line. *Cell*. 1983; 35:647–655. [PubMed: 6580963]
2. Hacein-Bey-Abina S, et al. Sustained correction of X-linked severe combined immunodeficiency by ex vivo gene therapy. *New Eng J Med*. 2002; 346:1185–1193. [PubMed: 11961146]
3. Hacein-Bey-Abina S, et al. A serious adverse event after successful gene therapy for X-linked severe combined immunodeficiency. *New Engl J Med*. 2003; 348:255–256. [PubMed: 12529469]
4. Hacein-Bey-Abina S, Hauer J, Lim A, Picard C, Wang GP, Berry CC, Martinache C, Rieux-Laucat F, Latour S, Belohradsky BH, Leiva L, Sorensen R, Debré M, Casanova JL, Blanche S, Durandy A, Bushman FD, Fischer A, Cavazzana-Calvo M. Efficacy of gene therapy for X-linked severe combined immunodeficiency. *New Eng J Med*. 2010; 363:355–364. [PubMed: 20660403]
5. Hacein-Bey-Abina S, et al. LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science*. 2003; 302:415–419. [PubMed: 14564000]

6. McCormack MP, Forster A, Drynan L, Pannell R, Rabbitts TH. The LMO2 T-cell oncogene is activated via chromosomal translocations or retroviral insertion during gene therapy but has no mandatory role in normal T-cell development. *Mol Cell Biol*. 2003; 23:9003–9013. [PubMed: 14645513]
7. Dave' UP, Jenkins NA, Copeland NG. Gene therapy insertional mutagenesis insights. *Science*. 2004; 303:33. [PubMed: 14704408]
8. McCormack MP, Rabbitts TH. Activation of the T-cell oncogene LMO2 after gene therapy for X-linked severe combined immunodeficiency. *New Eng J Med*. 2004; 350:913–922. [PubMed: 14985489]
9. Nam CH, Rabbitts TH. The role of LMO2 in development and in T cell leukemia after chromosomal translocation or retroviral insertion. *Mol Ther*. 2006; 13:15–25. [PubMed: 16260184]
10. Woods NB, Bottero V, Schmidt M, von Kalle C, Verma IM. Gene therapy: therapeutic gene causing lymphoma. *Nature*. 2006; 440:1123. [PubMed: 16641981]
11. Gaspar HB, et al. Long-term persistence of a polyclonal T cell repertoire after gene therapy for X-linked severe combined immunodeficiency. *Sci Transl Med*. 2011; 3:97ra79.
12. Ott MG, et al. Correction of X-linked chronic granulomatous disease by gene therapy, augmented by insertional activation of MDS1-EV11, PRDM16 or SETBP1. *Nat Med*. 2006; 5:401–409. [PubMed: 16582916]
13. Kang EM, Choi U, Theobald N, Linton G, Long Priel DA, Kuhns D, Malech HL. Retrovirus gene therapy for X-linked chronic granulomatous disease can achieve stable long-term correction of oxidase activity in peripheral blood neutrophils. *Blood*. 2010; 115:783–791. [PubMed: 19965657]
14. Stein S, Ott MG, Schultze-Strasser S, Jauch A, Burwinkel B, Kinner A, Schmidt M, Krämer A, Schwäble J, Glimm H, Koehl U, Preiss C, Ball C, Martin H, Göhring G, Schwarzwaelder K, Hofmann WK, Karakaya K, Tchatchou S, Yang R, Reinecke P, Kühlcke K, Schlegelberger B, Thrasher AJ, Hoelzer D, Seger R, von Kalle C, Grez M. Genomic instability and myelodysplasia with monosomy 7 consequent to EVI1 activation after gene therapy for chronic granulomatous disease. *Nature Med*. 2010; 16:198–204. [PubMed: 20098431]
15. Baum C, et al. Side effects of retroviral gene transfer into hematopoietic stem cells. *Blood*. 2003; 101:2099–2114. [PubMed: 12511419]
16. Baum C, Fehse B. Mutagenesis by retroviral transgene insertion: risk assessment and potential alternatives. *Curr Opin Mol Ther*. 2003; 5:458–462. [PubMed: 14601513]
17. Baum C, et al. Chance or necessity? Insertional mutagenesis in gene therapy and its consequences. *Mol Ther*. 2004; 9:5–13. [PubMed: 14741772]
18. Kustikova O, et al. Clonal dominance of hematopoietic stem cells triggered by retroviral gene marking. *Science*. 2005; 308:1171–1174. [PubMed: 15905401]
19. Muul L, et al. Persistence and expression of the adenosine deaminase gene for 12 years and immune reaction to gene transfer components: long-term results of the first clinical gene therapy trial. *Blood*. 2003; 101:2563–2569. [PubMed: 12456496]
20. Schmidt M, et al. Clonality analysis after retroviral-mediated gene transfer to CD34+ cells from the cord blood of ADA-deficient SCID neonates. *Nature Med*. 2003; 9:463–468. [PubMed: 12640448]
21. Aiuti A, Ficara F, Cattaneo F, Bordignon C, Roncarolo MG. Gene therapy for adenosine deaminase deficiency. *Curr Opin Allergy Clin Immunol*. 2004; 3:461–466. [PubMed: 14612670]
22. Gaspar HB, et al. Successful reconstitution of immunity in ADA-SCID by stem cell gene therapy following cessation of PEG-ADA and use of mild preconditioning. *Mol Ther*. 2006; 14:505–513. [PubMed: 16905365]
23. Aiuti A, et al. Multilineage hematopoietic reconstitution without clonal selection in ADA-SCID patients treated with stem cell gene therapy. *J Clin Invest*. 2007; 117:2233–2240. [PubMed: 17671653]
24. Cassani B, et al. Integration of retroviral vectors induces minor changes in the transcriptional activity of T cells from ADA-SCID patients treated with gene therapy. *Blood*. 2009; 114:3546–3556. [PubMed: 19652199]
25. Boztug K, Schmidt M, Schwarzer A, Banerjee PP, Díez IA, Dewey RA, Böhm M, Nowrouzi A, Ball CR, Glimm H, Naundorf S, Kühlcke K, Blasczyk R, Kondratenko I, Maródi L, Orange JS,

- von Kalle C, Klein C. Stem-cell gene therapy for the Wiskott-Aldrich Syndrome. *New Eng J Med*. 2010; 363:1918–1927. [PubMed: 21067383]
26. Buonamici S, Chakraborty S, Senyuk V, Nucifora G. The role of EVI1 in normal and leukemic cells. *Blood Cells Mol Dis*. 2003; 31:206–212. [PubMed: 12972028]
 27. Buonamici S, et al. EVI1 induces myelodysplastic syndrome in mice. *J Clin Invest*. 2004; 114:713–719. [PubMed: 15343390]
 28. Li Z, et al. Murine leukemia induced by retroviral gene marking. *Science*. 2002; 296:497. [PubMed: 11964471]
 29. Biasco L, Baricordi C, Aiuti A. Retroviral integrations in gene therapy trials. *Mol Ther*. 2012; 20:709–716. [PubMed: 22252453]
 30. Gabriel R, Schmidt M, von Kalle C. Integration of retroviral vectors. *Curr Opin Immunol*. 2012; 24:592–597. [PubMed: 22981243]
 31. Scholler J, et al. Decade-long safety and function of retroviral-modified chimeric antigen receptor T cells. *Sci Transl Med*. 2012; 4:132ra153.
 32. van der Loo JC SW, Grassman E, Terwilliger A, Higashimoto T, Schambach A, Hacein-Bey-Abina S, Nordling DL, Cavazzana-Calvo M, Thrasher AJ, Williams DA, Reeves L, Malik P. Critical variables affecting clinical-grade production of the self-inactivating gamma-retroviral vector for the treatment of X-linked severe combined immunodeficiency. *Gene Ther*. 2012; 19:872–876. [PubMed: 22551777]
 33. June CH, Blazar BR, Riley JL. Engineering lymphocyte subsets: tools, trials and tribulations. *Nature Rev Immunol*. 2009; 9:704–716. [PubMed: 19859065]
 34. Schwarzwaelder K, Howe SJ, Schmidt M, Brugman MH, Deichmann A, Glimm H, Schmidt S, Prinz C, Wissler M, King DJ, Zhang F, Parsley KL, Gilmour KC, Sinclair J, Bayford J, Peraj R, Pike-Overzet K, Staal FJ, de Ridder D, Kinnon C, Abel U, Wagemaker G, Gaspar HB, Thrasher AJ, von Kalle C. Gammaretrovirus-mediated correction of SCID-X1 is associated with skewed vector integration site distribution in vivo. *J Clin Invest*. 2007; 117:2241–2249. [PubMed: 17671654]
 35. Newrzela S, et al. Resistance of mature T cells to oncogene transformation. *Blood*. 2008; 112:2278–2286. [PubMed: 18566328]
 36. Cattoglio C, et al. High-definition mapping of retroviral integration sites defines the fate of allogeneic T cells after donor lymphocyte infusion. *PLoS One*. 2010; 5:e15688. [PubMed: 21203516]
 37. Wang GP, et al. Dynamics of gene-modified progenitor cells analyzed by tracking retroviral integration sites in a human SCID-X1 gene therapy trial. *Blood*. 2010; 115:4356–4366. [PubMed: 20228274]
 38. Deichmann A, et al. Insertion sites in engrafted cells cluster within a limited repertoire of genomic areas after gammaretroviral vector gene therapy. *Mol Ther*. 2011; 19:2031–2039. [PubMed: 21862999]
 39. Montini E, et al. The genotoxic potential of retroviral vectors is strongly modulated by vector design and integration site selection in a mouse model of HSC gene therapy. *J Clin Invest*. 2009; 119:964–975. [PubMed: 19307726]
 40. Kustikova O, Brugman M, Baum C. The genomic risk of somatic gene therapy. *Semin Cancer Biol*. 2010; 20:269–278. [PubMed: 20600920]
 41. Almarza E, et al. Correction of SCID-X1 using an enhancerless Vav promoter. *Hum Gene Ther*. 2011; 22:263–270. [PubMed: 20887212]
 42. Baum C. Parachuting in the epigenome: the biology of gene vector insertion profiles in the context of clinical trials. *EMBO Mol Med*. 2011; 3:75–77. [PubMed: 21254403]
 43. Wu C, Dunbar CE. Stem cell gene therapy: the risks of insertional mutagenesis and approaches to minimize genotoxicity. *Front Med*. 2011; 5:356–371. [PubMed: 22198747]
 44. Trobridge GD. Genotoxicity of retroviral hematopoietic stem cell gene therapy. *Expert Opin Biol Ther*. 2011; 11:581–593. [PubMed: 21375467]
 45. Huston MW, Brugman MH, Horsman S, Stubbs AP, van der Spek PJ, Wagemaker G. Comprehensive investigation of the parameters of viral integration site analysis and their effects on the gene annotations produced. *Hum Gene Ther*. 2012 in press.

46. Corrigan-Curay J, Cohen-Haguenauer O, O'Reilly M, Ross SR, Fan H, Rosenberg N, Somia N, King N, Friedmann T, Dunbar C, Aiuti A, Naldini L, Baum C, von Kalle C, Kiem HP, Montini E, Bushman F, Sorrentino BP, Carrondo M, Malech H, Gahrton G, Shapiro R, Wolff L, Rosenthal E, Jambou R, Zaia J, Kohn DB. Challenges in vector and trial design using retroviral vectors for long-term gene correction in hematopoietic stem cell gene therapy. *Mol Ther.* 2012; 20:1084–1094. [PubMed: 22652996]
47. Aiuti A, Bacchetta R, Seger R, Villa A, Cavazzana-Calvo M. Gene therapy for primary immunodeficiencies: Part 2. *Curr Opin Immunol.* 2012; 24:1–7. [PubMed: 22277983]
48. Scobie L, et al. A novel model of SCID-X1 reconstitution reveals predisposition to retrovirus-induced lymphoma but no evidence of gammaC gene oncogenicity. *Mol Ther.* 2009; 17:1031–1038. [PubMed: 19337236]
49. Deichmann A, et al. Vector integration is nonrandom and clustered and influences the fate of lymphopoiesis in SCID-X1 gene therapy. *J Clin Invest.* 2007; 117:2225–2232. [PubMed: 17671652]
50. Gabriel R, et al. Comprehensive genomic access to vector integration in clinical gene therapy. *Nature Med.* 2009; 15:1431–1436. [PubMed: 19966782]
51. Kolb AF, Coates CJ, Kaminski JM, Summers JB, Miller AD, D.J. S. Site-directed genome modification: nucleic acid and protein modules for targeted integration and gene correction. *Trends Biotech.* 2005; 23:399–406.
52. de Jong J, et al. Computational identification of insertional mutagenesis targets for cancer gene discovery. *Nucl Acids Res.* 2011; 39:e105. [PubMed: 21652642]
53. March HN, et al. Insertional mutagenesis identifies multiple networks of cooperating genes driving intestinal tumorigenesis. *Nature Genet.* 2011; 43:1202–1209. [PubMed: 22057237]
54. Dupuy AJ. Transposon-based screens for cancer gene discovery in mouse models. *Semin Cancer Biol.* 2010; 20:261–268. [PubMed: 20478384]
55. Bergemann TL, et al. New methods for finding common insertion sites and co-occurring common insertion sites in transposon- and virus-based genetic screens. *Nucl Acids Res.* 2012; 40:3822–3833. [PubMed: 22241771]
56. Howell VM. Sleeping Beauty - a mouse model for all cancers? *Cancer Lett.* 2012; 317:1–8. [PubMed: 22079740]
57. Mann KM, Ward JM, Yew CC, Kovochich A, Dawson DW, Black MA, Brett BT, Sheetz T, E, Dupuy AJ, Australian Pancreatic Cancer Genome Initiative. Chang DK, Biankin AV, Waddell N, Kassahn KS, Grimmond SM, Rust AG, Adams DJ, Jenkins NA, Copeland NG. Sleeping Beauty mutagenesis reveals cooperating mutations and pathways in pancreatic adenocarcinoma. *Proc Natl Acad Sci USA.* 2012; 109:5934–5941. [PubMed: 22421440]
58. van der Weyden L, et al. Loss of rassf1a synergizes with deregulated runx2 signaling in tumorigenesis. *Cancer Res.* 2012; 72:3817–3827. [PubMed: 22710434]
59. Bergerson RJ, Collier LS, Sarver AL, Been RA, Lugthart S, Diers MD, Zuber J, Rappaport AR, Nixon MJ, Silverstein KA, Fan D, Lamblin AF, Wolff L, Kersey JH, Delwe l. R. Lowe SW, O'Sullivan MG, Kogan SC, Adams DJ, Largaespada DA. An insertional mutagenesis screen identifies genes that cooperate with Mll-AF9 in a murine leukemogenesis model. *Blood.* 2012; 119:4512–4523. [PubMed: 22427200]
60. van der Weyden L, Arends MJ, Rust AG, Poulogiannis G, McIntyre RE, Adams DJ. Increased tumorigenesis associated with loss of the tumor suppressor gene Cadm. *Mol Cell.* 2012; 11 in press.
61. Oliveira-Cobucci RN, et al. Comparative incidence of cancer in HIV-AIDS patients and transplant recipients. *Cancer Epidemiol.* 2012; 36:e69–73. [PubMed: 22236649]
62. Wang GP, et al. Analysis of lentiviral vector integration in HIV+ study subjects receiving autologous infusions of gene modified CD4+ T cells. *Mol Ther.* 2009; 17:844–850. [PubMed: 19259065]
63. Cesana D, Squaidino J, Rudilosso L, Merella S, Naldini L, Montini E. Whole transcriptome characterization of aberrant splicing events induced by lentivirus vector integrations. *J Clin Invest.* 2012; 122:1667–1676. [PubMed: 22523064]

64. Brady T, Bushman FD. Nondividing cells: a safer bet for integrating vectors? *Mol Ther.* 2011; 19:640–641. [PubMed: 21455211]
65. Cattoglio C, Facchini G, Sartori D, Antonelli A, Miccio A, Cassani B, Schmidt M, von Kalle C, Howe S, Thrasher AJ, Aiuti A, Ferrari G, Recchia A, Mavilio F. Hot spots of retroviral integration in human CD34+ hematopoietic cells. *Blood.* 2007; 110:1770–1778. [PubMed: 17507662]
66. Nowrouzi A, Glimm H, von Kalle C, Schmidt M. Retroviral vectors: post entry events and genomic alterations. *Viruses.* 2011; 3:429–455. [PubMed: 21994741]
67. Biffi A, et al. Lentiviral vector common integration sites in preclinical models and a clinical trial reflect a benign integration bias and not oncogenic selection. *Blood.* 2012; 117:5332–5339. [PubMed: 21403130]
68. Roth SL, Malani N, Bushman FD. Gammaretroviral integration into nucleosomal target DNA in vivo. *J Virol.* 2011; 85:7393–7401. [PubMed: 21561906]
69. Hackett PB. Integrating DNA vectors for gene therapy. *Mol Ther.* 2007; 15:10–12. [PubMed: 17164769]
70. Zhang Y, Satterlee A, Huang L. In vivo gene delivery by nonviral vectors: overcoming hurdles? *Mol Ther.* 2012; 20:1298–1304. [PubMed: 22525514]
71. Hackett PB, Largaespada DA, Cooper LJN. A transposon and transposase system for human application. *Mol Ther.* 2010; 18:674–683. [PubMed: 20104209]
72. Hackett CS, Geurts AM, Hackett PB. Predicting preferential DNA vector insertion sites: Implications for functional genomics and gene therapy. *Genome Biol.* 2007; 8(Suppl. 1):S12. [PubMed: 18047689]
73. Consortium1000 GP. A map of human genome variation from population-scale sequencing. *Nature.* 2010; 467:1061–1073. [PubMed: 20981092]
74. ConsortiumENCODE P. et al. An integrated encyclopedia of DNA elements in the human genome. *Nature.* 2012; 489:57–74. [PubMed: 22955616]
75. Lander ES, et al. Initial sequencing and analysis of the human genome. *Nature.* 2001; 409:860–921. [PubMed: 11237011]
76. Venter JC, et al. The sequence of the human genome. *Science.* 2001; 291:1304–1351. [PubMed: 11181995]
77. Cordaux R. The human genome in the LINE of fire. *Proc Natl Acad Sci USA.* 2008; 105:19033–19034. [PubMed: 19057007]
78. de Koning AP, Gu W, Castoe TA, Batzer MA, Pollock DD. Repetitive elements may comprise over two-thirds of the human genome. *PLoS Genet.* 2011; 7:e1002384. [PubMed: 22144907]
79. Beck CR, et al. LINE-1 retrotransposition activity in human genomes. *Cell.* 2010; 141:1159–1170. [PubMed: 20602998]
80. Brady T, et al. Integration target site selection by a resurrected human endogenous retrovirus. *Genes Dev.* 2009; 23:633–642. [PubMed: 19270161]
81. Brouha B, et al. Hot L1s account for the bulk of retrotransposition in the human population. *Proc Natl Acad Sci USA.* 2003; 100:5280–5288. [PubMed: 12682288]
82. Lee E, et al. Landscape of somatic retrotransposition in human cancers. *Science.* 2012; 337:967–971. [PubMed: 22745252]
83. Solyom S, Ewing AD, Rahrmann EP, Doucet TT, Nelson HH, Burns MB, Harris RS, Sigmon DF, Casella A, Erlanger B, Wheelan S, Upton KR, Shukla R, Faulkner GJ, Largaespada DA, Kazazian HH. Extensive somatic L1 retrotransposition in colorectal tumors. *Genome Res.* 2012 in press.
84. Dombroski BA, Mathias SL, Nanthakumar E, Scott AF, Kazazian HHJ. Isolation of an active human transposable element. *Science.* 1991; 254:1805–1808. [PubMed: 1662412]
85. Moran JV, Holmes SE, Naas TP, DeBerardinis RJ, Boeke JD, Kazazian HHJ. High frequency retrotransposition in cultured mammalian cells. *Cell.* 1996; 87:917–927. [PubMed: 8945518]
86. Kazazian HH, Goodier JL. LINE drive: retrotransposition and genome instability. *Cell.* 2002; 110:277–280. [PubMed: 12176313]
87. Deininger PL, Moran JV, Batzer MA, Kazazian HHJ. Mobile elements and mammalian genome evolution. *Curr Opin Genet Dev.* 2003; 13:651–658. [PubMed: 14638329]

88. Kazazian HH. Mobile elements: drivers of genome evolution. *Science*. 2004; 303:1626–1632. [PubMed: 15016989]
89. Cordaux R, Batzer MA. The impact of retrotransposons on human genome structure. *Nature Rev Genet*. 2009; 10:691–703. [PubMed: 19763152]
90. Ewing AD, Kazazian HHJ. Whole-genome resequencing allows detection of many rare LINE-1 insertion alleles in humans. *Genome Res*. 2011; 21:985–990. [PubMed: 20980553]
91. Burns KH, Boeke JD. Human transposon tectonics. *Cell*. 2012; 149:740–752. [PubMed: 22579280]
92. Coufal NG, Garcia-Perez JL, Peng GE, Yeo GW, Mu Y, Lovci MT, Morell M, O'Shea KS, Moran JV, Gage FH. L1 retrotransposition in human neural progenitor cells. *Nature*. 2009; 460:1127–1131. [PubMed: 19657334]
93. Baillie JK, Barnett MW, Upton KR, Gerhardt DJ, Richmond TA, De Sapio F, Brennan PM, Rizzu P, Smith S, Fell M, Talbot RT, Gustincich S, Freeman TC, Mattick JS, Hume DA, Heutink P, Carninci P, Jeddelloh JA, Faulkner GJ. Somatic retrotransposition alters the genetic landscape of the human brain. *Nature*. 2011; 479:534–537. [PubMed: 22037309]
94. Muotri AR, et al. L1 retrotransposition in neurons is modulated by MeCP2. *Nature*. 2010; 443:443–446. [PubMed: 21085180]
95. Huang CR, Schneider AM, Lu Y, Niranjana T, Shen P, Robinson MA, Steranka JP, Valle D, Civin CI, Wang T, Wheelan SJ, Ji H, Boeke JD, Burns KH. Mobile interspersed repeats are major structural variants in the human genome. *Cell*. 2010; 141:1171–1182. [PubMed: 20602999]
96. Iskow RC, et al. Natural mutagenesis of human genomes by endogenous retrotransposons. *Cell*. 2010; 141:1253–1261. [PubMed: 20603005]
97. Garcia-Perez JL, Morell M, Scheys JO, Kulpa DA, Morell S, Carter CC, Hammer GD, Collins KL, O'Shea KS, Menendez P, Moran JV. Epigenetic silencing of engineered L1 retrotransposition events in human embryonic carcinoma cells. *Nature*. 2010; 466:769–773. [PubMed: 20686575]
98. Doerfler W. Impact of foreign DNA integration on tumor biology and on evolution via epigenetic alterations. *Epigenomics*. 2012; 4:41–49. [PubMed: 22332657]
99. Bao J, Yan W. Male germline control of transposable elements. *Biol Reprod*. 2012; 86:162, 161–114. [PubMed: 22357546]
100. Wagstaff BJ, Hedges DJ, Derbes RS, Campos-Sanchez R, Chiaromonte F, Makova KD, Roy-Engel AM. Rescuing Alu: Recovery of new inserts shows LINE-1 preserves Alu activity through A-tail expansion. *PLoS Genet*. 2012; 8:e1002842. [PubMed: 22912586]
101. Zhang Y, Maksakova IA, Gagnier L, van de Lagemaat LN, Mager DL. Genome-wide assessments reveal extremely high levels of polymorphism of two active families of mouse endogenous retroviral elements. *PLoS Genet*. 2008; 4:e1000007. [PubMed: 18454193]
102. Liu G, Aronovich EL, Cui Z, Whitley CB, Hackett PB. Excision of *Sleeping Beauty* transposons: parameters and applications to gene therapy. *J Gene Med*. 2004; 6:574–583. [PubMed: 15133768]
103. Plasterk RHA, Izsvák Z, Ivics Z. Resident aliens: the Tc1/mariner superfamily of transposable elements. *Trends Genet*. 1999; 15:326–332. [PubMed: 10431195]
104. Ivics Z, Hackett PB, Plasterk RH, Izsvák Z. Molecular reconstruction of *Sleeping Beauty*, a Tc1-like transposon from fish, and its transposition in human cells. *Cell*. 1997; 91:501–510. [PubMed: 9390559]
105. Yant SR, Meuse L, Chiu W, Ivics Z, Izsvák Z, Kay MA. Somatic integration and long-term transgene expression in normal and haemophilic mice using a DNA transposon system. *Nat Genet*. 2000; 25:35–41. [PubMed: 10802653]
106. Geurts AM, et al. Gene transfer into genomes of human cells by the *Sleeping Beauty* transposon system. *Mol Ther*. 2003; 8:108–117. [PubMed: 12842434]
107. Zayed H, Izsvák Z, Walisko O, Ivics Z. Development of hyperactive Sleeping Beauty transposon vectors by mutational analysis. *Mol Ther*. 2004; 9:292–304. [PubMed: 14759813]
108. Yant SR, Park J, Huang Y, Mikkelsen JG, Kay MA. Mutational analysis of the N-terminal DNA-binding domain of *Sleeping Beauty* transposase: critical residues for DNA binding and hyperactivity in mammalian cells. *Mol Cell Biol*. 2004; 24:9239–9247. [PubMed: 15456893]

109. Mátés L, et al. Molecular evolution of a novel hyperactive *Sleeping Beauty* transposase enables robust stable gene transfer in vertebrates. *Nat Genet.* 2009; 41:753–761. [PubMed: 19412179]
110. 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. [PubMed: 14632258]
111. Yant SR, et al. Nonrandom insertion site preferences for the SB transposon in vitro and in vivo. *Mol Ther.* 2004; 9:S309.
112. Liu G, et al. Target-site preference for *Sleeping Beauty* transposons. *J Mol Biol.* 2005; 346:161–173. [PubMed: 15663935]
113. Geurts AM, et al. DNA structural patterns influence integration site preferences for mobile elements. *Nucl Acids Res.* 2006; 34:2803–2811. [PubMed: 16717285]
114. Berry C, Hannenhalli S, Leipzig J, Bushman FD. Selection of target sites for mobile DNA integration in the human genome. *PLoS Comp Biol.* 2006; 2:e157.
115. Ohlfest JR, et al. Phenotypic correction and long-term expression of factor VIII in hemophilic mice by immunotolerization and nonviral gene transfer using the *Sleeping Beauty* transposon system. *Blood.* 2005; 105:2691–2698. [PubMed: 15576475]
116. Liu L, Mah C, Fletcher BS. Sustained FVIII expression and phenotypic correction of hemophilia A in neonatal mice. *Mol Ther.* 2006; 13:1006–1015. [PubMed: 16464640]
117. Aronovich EL, McIvor RS, Hackett PB. The *Sleeping Beauty* transposon system: a non-viral vector for gene therapy. *Hum Mol Genet.* 2011; 20(R1):14–20.
118. Bell JB, Aronovich EL, Schreifels JM, Beadnell TC, Hackett PB. Duration of expression of *Sleeping Beauty* transposase in mouse liver following hydrodynamic DNA delivery. *Mol Ther.* 2010; 18:1792–1802.
119. Aravin AA, Hannon GJ, Brennecke J. The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. *Science.* 2007; 318:761–764. [PubMed: 17975059]
120. Fedoroff NV. Transposable elements, epigenetics, and genome evolution. *Science.* 2012; 338:758–767. [PubMed: 23145453]
121. Nesmelova IV, Hackett PB. DDE Transposases: structural similarity and diversity. *Adv Drug Del Rev.* 2010; 62:1187–1195.
122. Montañó SP, Pigli YZ, Rice PA. The Mu transpososome structure sheds light on DDE recombinase evolution. *Nature.* 2012; 491:413–417. [PubMed: 23135398]
123. Wilber AC, Frandsen JL, Geurts AM, Largaespada DA, P.B. H, 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. [PubMed: 16368272]
124. Urschitz J, et al. Helper-independent piggyBac plasmids for gene delivery approaches: strategies for avoiding potential genotoxic effects. *Proc Natl Acad Sci USA.* 2010; 107:8117–8122. [PubMed: 20404201]
125. Hastie ND, Bishop JO. The expression of three abundance classes of messenger RNA in mouse tissues. *Cell.* 1976; 9:761–774. [PubMed: 1017013]
126. Rushforth AM, Anderson P. Splicing removes the *Caenorhabditis elegans* transposon Tc1 from most mutant pre-mRNAs. *Mol Cell Biol.* 1996; 16:422–429. [PubMed: 8524324]
127. Hodgkinson A, Chen Y, Eyre-Walker A. The large-scale distribution of somatic mutations in cancer genomes. *Hum Mutat.* 2012; 33:136–143. [PubMed: 21953857]
128. Schuster-Böckler B, Lehner B. Chromatin organization is a major influence on regional mutation rates in human cancer cells. *Nature.* 2012; 488:504–507. [PubMed: 22820252]
129. Djebali S, et al. Landscape of transcription in human cells. *Nature.* 2012; 489:101–108. [PubMed: 22955620]
130. Cheng C, Alexander R, Min R, Leng J, Yip KY, Rozowsky J, Yan KK, Dong X, Djebali S, Ruan Y, Davis CA, Carninci P, Lassman T, Gingeras TR, Guigó R, Birney E, Weng Z, Snyder M, Gerstein M. Understanding transcriptional regulation by integrative analysis of transcription factor binding data. *Genome Res.* 2012; 22:1658–1667. [PubMed: 22955978]
131. Sanyal A, Lajoie BR, Jain G, Dekker J. The long-range interaction landscape of gene promoters. *Nature.* 2012; 489:109–113. [PubMed: 22955621]

132. Thurman RE, et al. The accessible chromatin landscape of the human genome. *Nature*. 2012; 489:75–82. [PubMed: 22955617]
133. Zhang J, Markus J, J B, Paul T, Wolff L. Three murine leukemia virus integration regions within 100kb upstream of c-myc are proximal to the 5' regulatory region of the gene through DNA looping. *J Virol*. 2012 in press.
134. Poleshko A, et al. Identification of cellular proteins that maintain retroviral epigenetic silencing: evidence for an antiviral response. *J Virol*. 2008; 82:2313–2323. [PubMed: 18094192]
135. Alkan C, Coe BP, Eichler EE. Genome structural variation discovery and genotyping. *Nature Rev Genet*. 2011; 12:363–376. [PubMed: 21358748]
136. Marth GT, et al. The functional spectrum of low-frequency coding variation. *Genome Biol*. 2011; 12:R84. [PubMed: 21917140]
137. Durbin RM, et al. A map of human genome variation from population-scale sequencing. *Nature*. 2010; 467:1061–1073. [PubMed: 20981092]
138. Pelak K, et al. The characterization of twenty sequenced human genomes. *PLoS Genet*. 2010; 6:e1001111. [PubMed: 20838461]
139. MacArthur DG, et al. A systematic survey of loss-of-function variants in human protein-coding genes. *Science*. 2012; 335:823–828. [PubMed: 22344438]
140. Wada T, Candotti F. Somatic mosaicism in primary immune deficiencies. *Curr Opin Allergy Clin Immunol*. 2008;8.
141. Abyzov A, Urban AE, Snyder M, Gerstein M. CNVnator: an approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. *Genome Res*. 2011; 21:974–984. [PubMed: 21324876]
142. O'Huallachain M, Karczewski KJ, Weissman SM, Urban AE, Snyder MP. Extensive genetic variation in somatic human tissues. *Proc Natl Acad Sci USA*. 2012; 109:18018–18023. [PubMed: 23043118]
143. Kidd JM, et al. A human genome structural variation sequencing resource reveals insights into mutational mechanisms. *Cell*. 2010; 143:837–847. [PubMed: 21111241]
144. Boyle AP, et al. Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res*. 2012; 22:1790–1797. [PubMed: 22955989]
145. Eide D, Anderson P. Insertion and excision of *Caenorhabditis elegans* transposable element Tc1. *Mol Cell Biol*. 1988; 8:737–746. [PubMed: 2832734]
146. Ellis J. Silencing and variegation of gammaretrovirus and lentivirus vectors. *Hum Gene Ther*. 2005; 16:1241–1246. [PubMed: 16259557]
147. Blumenstiel JP. Evolutionary dynamics of transposable elements in a small RNA world. *Trends Genet*. 2011; 27:23–31. [PubMed: 21074888]
148. Siomi MC, Sato K, Pezic D, Aravin AA. PIWI-interacting small RNAs: the vanguard of genome defence. *Nature Rev Mol Cell Biol*. 2011; 12:246–258. [PubMed: 21427766]
149. Garrison BS, Yant SR, Mikkelsen JG, Kay MA. Post-integrative gene silencing within the *Sleeping Beauty* transposition system. *Mol Cell Biol*. 2007; 27:8824–8833. [PubMed: 17938204]
150. Rauschhuber C, Ehrhardt A. RNA interference is responsible for reduction of transgene expression after Sleeping Beauty transposase mediated somatic integration. *PLoS One*. 2012 in press.
151. Chalberg TW, et al. Integration specificity of phage phiC31 integrase in the human genome. *J Mol Biol*. 2006; 357:28–48. [PubMed: 16414067]
152. Ehrhardt A, Engler JA, Xu H, Kay MA. Molecular analysis of chromosomal rearrangements in mammalian cells after phiC31-mediated integration. *Human Gene Ther*. 2006; 17:1077–1094. [PubMed: 17069535]
153. Largaespada DA. Transposon mutagenesis in mice. *Methods Mol Biol*. 2009; 530:379–390. [PubMed: 19266336]
154. Copeland NG, Jenkins NA. Harnessing transposons for cancer gene discovery. *Nat Rev Genet*. 2010; 10:696–706.
155. Mikkers H, Berns A. Retroviral insertional mutagenesis: tagging cancer pathways. *Adv Cancer Res*. 2003; 88:53–99. [PubMed: 12665053]

156. Jonkers J, Berns A. Retroviral insertional mutagenesis as a strategy to identify cancer genes. *Biochim Biophys Acta*. 1996; 1287:29–57. [PubMed: 8639705]
157. Lohuizen, Mv; Berns, A. Tumorigenesis by slow transforming retroviruses--an update. *Biochim Biophys Acta*. 1990; 1032(2–3):213–235. [PubMed: 2261495]
158. Largaespada DA. Genetic heterogeneity in acute myeloid leukemia: maximizing information flow from MuLV mutagenesis studies. *Leukemia*. 2000; 14:1174–1184. [PubMed: 10914539]
159. Lu M, Zhang N, Maruyama M, Hawley RG, Ho AD. Retrovirus-mediated gene expression in hematopoietic cells correlates inversely with growth factor stimulation. *Hum Gene Ther*. 1996; 7:2263–2271. [PubMed: 8953317]
160. Hawley RG, Lieu FH, Fong AZ, Hawley TS. Versatile retroviral vectors for potential use in gene therapy. *Gene Ther*. 1994; 1:136–138. [PubMed: 7584069]
161. Cherry SR, Binischewicz D, van Parijs L, Baltimore D, Jaenisch R. Retroviral expression in embryonic stem cells and hematopoietic stem cells. *Mol Cell Biol*. 2000; 20:7419–7426. [PubMed: 11003639]
162. Vassiliou GS, et al. Mutant nucleophosmin and cooperating pathways drive leukemia initiation and progression in mice. *Nature Genet*. 2011; 43:470–475. [PubMed: 21441929]
163. Dupuy AJ, et al. A modified Sleeping Beauty transposon system that can be used to model a wide variety of human cancers in mice. *Cancer Res*. 2009; 69:8150–8156. [PubMed: 19808965]
164. Collier LS, Carlson CM, Ravimohan S, Dupuy AJ, Largaespada DA. Cancer gene discovery in solid tumors using transposon-based somatic mutagenesis in the mouse. *Nature*. 2005; 436:272–276. [PubMed: 16015333]
165. 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. [PubMed: 16015321]
166. Collier LS, et al. Whole-body Sleeping Beauty mutagenesis can cause penetrant leukemia/lymphoma and rare high-grade glioma without associated embryonic lethality. *Cancer Res*. 2009; 69:8429–8437. [PubMed: 19843846]
167. Starr TK, et al. A transposon-based genetic screen in mice identifies genes altered in colorectal cancer. *Science*. 2009; 323:1747–1750. [PubMed: 19251594]
168. Starr TK, Scott PM, Marsh BM, Zhao L, Than BL, O'Sullivan MG, Sarver AL, Dupuy AJ, Largaespada DA, Cormier RT. A Sleeping Beauty transposon-mediated screen identifies murine susceptibility genes for adenomatous polyposis coli (Apc)-dependent intestinal tumorigenesis. *Proc Natl Acad Sci USA*. 2011; 108:5765–5770. [PubMed: 21436051]
169. Keng VW, et al. A conditional transposon-based insertional mutagenesis screen for genes associated with mouse hepatocellular carcinoma. *Nature Biotech*. 2009; 27:264–274.
170. Pérez-Mancera PA, Rust AG, van der Weyden L, Kristiansen G, Li A, Sarver AL, Silverstein KA, Grützmann R, Aust D, Rümmele P, Knösel T, Herd C, Stemple DL, Kettleborough R, Brosnan JA, Li A, Morgan R, Knight S, Yu J, Stegeman S, Collier LS, ten Hoeve JJ, de Ridder J, Klein AP, Goggins M, Hruban RH, Chang DK, Biankin AV, Grimmond SM, Australian Pancreatic Cancer Genome Initiative, Wessels LF, Wood SA, Iacobuzio-Donahue CA, Pilarsky C, Largaespada DA, Adams DJ, Tuveson DA. The deubiquitinase USP9X suppresses pancreatic ductal adenocarcinoma. *Nature*. 2012; 486:266–270. [PubMed: 22699621]
171. Dotti JB, Cooper LNJ. Redirecting T-cell specificity by introducing a tumor-specific chimeric antigen receptor. *Blood*. 2010; 116:1035–1044. [PubMed: 20439624]
172. Brentjens R, Yeh R, Bernal Y, Riviere I, Sadelain M. Treatment of chronic lymphocytic leukemia with genetically targeted autologous T cells: Case report of an unforeseen adverse event in a Phase I clinical trial. *Mol Ther*. 2010; 18:666–668. [PubMed: 20357779]
173. Brentjens RJ, et al. Safety and persistence of adoptively transferred autologous CD19-targeted T cells in patients with relapsed or chemotherapy refractory B-cell leukemias. *Blood*. 2011; 118:4817–4828. [PubMed: 21849486]
174. Kalos M, et al. T cells with chimeric antigen receptors have potent antitumor effects and can establish memory in patients with advanced leukemia. *Sci Transl Med*. 2011; 3:95ra73.

175. Kochenderfer JN, et al. B-cell depletion and remissions of malignancy along with cytokine-associated toxicity in a clinical trial of anti-CD19 chimeric-antigen-receptor-transduced T cells. *Blood*. 2012; 119:2709–2720. [PubMed: 22160384]
176. Porter DL, Levine BL, Kalos M, Bagg A, June CH. Chimeric antigen receptor-modified T cells in chronic lymphoid leukemia. *New Eng J Med*. 2011; 365:725–733. [PubMed: 21830940]
177. Bonini C, et al. Safety of retroviral gene marking with a truncated NGF receptor. *Nature Med*. 2003; 9:367–369. [PubMed: 12669036]
178. Singh H, et al. Redirecting specificity of T-cell populations for CD19 using the *Sleeping Beauty* system. *Cancer Res*. 2008; 68:2961–2971. [PubMed: 18413766]
179. Manuri PV, et al. piggyBac transposon/transposase system to generate CD19-specific T cells for the treatment of B-lineage malignancies. *Hum Gene Ther*. 2010; 21:427–437. [PubMed: 19905893]
180. Singh H, et al. Reprogramming CD19-specific T cells with IL-21 signaling can improve adoptive immunotherapy of B-lineage malignancies. *Cancer Res*. 2011; 71:3516–3527. [PubMed: 21558388]
181. Lamers CH, et al. Treatment of metastatic renal cell carcinoma with autologous T-lymphocytes genetically retargeted against carbonic anhydrase IX: first clinical experience. *J Clin Oncol*. 2006; 24:e20–22. [PubMed: 16648493]
182. Morgan RA, Yang JC, Kitano M, Dudley ME, Laurencot CM, S/A. R. Case report of a serious adverse event following the administration of T cells transduced with a chimeric antigen receptor recognizing ERBB2. *Mol Ther*. 2010; 18:843–851. [PubMed: 20179677]
183. Montini E, et al. Hematopoietic stem cell gene transfer in a tumor-prone mouse model uncovers low genotoxicity of lentiviral vector integration. *Nature Biotech*. 2006; 24:687–696.
184. Modlich U, et al. Leukemias following retroviral transfer of multidrug resistance 1 (MDR1) are driven by combinatorial insertional mutagenesis. *Blood*. 2005; 105:4235–4246. [PubMed: 15713797]
185. Jensen MC, et al. Human T lymphocyte genetic modification with naked DNA. *Mol Ther*. 2000; 1:49–55. [PubMed: 10933911]
186. Cooper LJ, et al. T lymphocytes for autologous cellular therapy for lymphoma. *Cytotherapy*. 2006; 8:105–117. [PubMed: 16698684]
187. Maiti SN, et al. *Sleeping Beauty* system to redirect T-cell specificity for human applications. *J Immunol*. in press.
188. Jena B, Dotti G, Cooper LJ. Redirecting T-cell specificity by introducing a tumor-specific chimeric antigen receptor. *Blood*. 2010; 116:1035–1044. [PubMed: 20439624]
189. Kebriaei P, et al. Infusing CD19-directed T cells to augment disease control in patients undergoing autologous hematopoietic stem-cell transplantation for advanced B-lymphoid malignancies. *Hum Gene Ther*. 2012; 23:444–450. [PubMed: 22107246]
190. Sadelain M, Papapetrou EP, Bushman FD. Safe harbours for the integration of new DNA in the human genome. *Nature Rev Cancer*. 2011; 12:51–58. [PubMed: 22129804]
191. Papapetrou EP, et al. Genomic safe harbors permit high beta-globin transgene expression in thalassemia induced pluripotent stem cells. *Nature Biotech*. 2011; 29:73–78.
192. Porteus MH, Baltimore D. Chimeric nucleases stimulate gene targeting in human cells. *Science*. 2003; 300:763. [PubMed: 12730593]
193. Urnov FD, et al. Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature*. 2005; 435:646–651. [PubMed: 15806097]
194. Urnov FD, Rebar EJ, Holmes MC, Zhang HS, Gregory PD. Genome editing with engineered zinc finger nucleases. *Nature Rev Genet*. 2010; 11:636–646. [PubMed: 20717154]
195. Lombardo A, et al. Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Nature Biotech*. 2007; 25:1298–1306.
196. Perez EE, et al. Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nature Biotech*. 2008; 26:808–816.

197. Holt N, Wang J, Kim K, Friedman G, Wang X, Taupin V, Crooks GM, Kohn DB, Gregory PD, Holmes MC, Cannon PM. Human hematopoietic stem progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 in vivo. *Nature Biotech.* 2012; 28:839–847.
198. Ivics Z, Katzer A, Stuve EE, Fiedler D, Knespel S, Izsvak Z. Targeted *Sleeping Beauty* transposition in human cells. *Mol Ther.* 2007; 15:1137–1144. [PubMed: 17426709]
199. Yant SR, Huang Y, Akache B, Kay MA. Site-directed transposon integration in human cells. *Nucl Acids Res.* 2007; 35:e50. [PubMed: 17344320]
200. Voigt K, et al. Retargeting Sleeping Beauty transposon insertions by engineered zinc finger DNA-binding domains. *Mol Ther.* 2012; 20:1852–1862. [PubMed: 22776959]
201. Ammar I, et al. Retargeting transposon insertions by the adeno-associated virus Rep protein. *Nucl Acids Res.* 2012; 40:6693–6712. [PubMed: 22523082]
202. Wang Z, Lisowski L, Finegold MJ, Nakai H, Kay MA, Grompe M. AAV vectors containing rDNA homology display increased chromosomal integration and transgene persistence. *Mol Ther.* 2012; 20:1902–1911. [PubMed: 22990673]
203. Lisowski L, et al. Ribosomal DNA integrating rAAV-rDNA vectors allow for stable transgene expression. *Mol Ther.* 2012; 20:1912–1923. [PubMed: 22990671]
204. Bogdanove AJ, Voytas DF. TAL effectors: customizable proteins for DNA targeting. *Science.* 2011; 333:1843–1844. [PubMed: 21960622]
205. Carlson DF, Fahrenkrug SC, Hackett PB. Targeting DNA with fingers and TALENs. *Mol Ther Nuc Acids.* 2012; 1:e3.
206. Torikai H, et al. A foundation for universal T-cell based immunotherapy: T cells engineered to express a CD19-specific chimeric-antigen-receptor and eliminate expression of endogenous TCR. *Blood.* 2012; 119:5697–5705. [PubMed: 22535661]
207. Nakazawa Y, et al. Optimization of the PiggyBac transposon system for the sustained genetic modification of human T lymphocytes. *J Immunother.* 2009; 32:826–836. [PubMed: 19752751]
208. Li H, et al. In vivo genome editing restores haemostasis in a mouse model of haemophilia. *Nature.* 2011; 475:217–221. [PubMed: 21706032]

SB Transposons integrate almost randomly into TA sites
($\approx 2 \times 10^8$ / mammalian genome)

Fig. 1. Genetic consequences of integration of therapeutic vectors into genomes

In this instance a transposon vector is illustrated, but the considerations apply to any vector system. A therapeutic transposon can integrate into four general categories of chromatin. Heterochromatin will suppress expression of the transgene – no gain-of-function (GOF) or loss-of-function (LOF) will occur. The most desirable integration events will be into intergenic regions where the Therapeutic Gene (TG) will be expressed. Integration into or proximal to a transcriptional regulatory region can have several outcomes including GOF of the transgenic cassette, as well as either enhancement or loss of expression of the neighboring gene (Gene X). As reviewed in the text, in some cases the transcriptional regulatory elements of the transgene can activate quiescent chromosomal genes. Integration of the transposon into a transcriptional unit may allow expression of the transgene but block expression of the host gene leading to a phenotypic loss of function due to blockage of the gene or alterations in splicing. Integration within some genes can also lead to a dominant gain-of-function (DGF) and/or production of a dominant-negative form (DNF) of the original Gene X.

Fig. 2. DNA transposition

DNA transposition, as exemplified by the *Sleeping Beauty* (SB) transposon system, is a cut-and-paste reaction in which a transposon containing an expression cassette with a therapeutic gene (TG) and its promoter (pentagon) is delivered to target cells wherein the transposon is cut out of the plasmid and inserted into a chromosome. The inverted terminal repeats (inverted set of double arrowheads) define a transposon. The second part of the SB system is SB transposase, which in this example is carried in a separate expression cassette that is on the same plasmid, but not in the transposon. The SB transposon will only integrate into TA-dinucleotide basepairs (about 200 million in mammalian genomes). 1) The plasmid is delivered into a cell by any of several means and proceeds through the nuclear membrane (dotted oval) by a poorly understood process. 2) The SB gene is expressed. 3) The transposase molecules enter the nucleus and bind to the transposon. 4) Four transposases work in concert to cut the transposon out of the plasmid and paste it (dotted lines) into an AT sequence in chromatin (tangled line). A plasmid excision product is left behind in this reaction (the transposon site is marked by an X). Integration into a chromosome can confer sustained expression of the gene of interest that is contained within the transposon.

Fig. 3. Potential consequences of remobilization of SB transposons

The schematic illustrates the initial transposon integration site (orange) and subsequent hopping to other TA sites in chromatin, one of which might induce an adverse transforming event.

Fig. 4. The T2/onc transposon vector

This SB vector contains elements designed to elicit either transcriptional activation (Mouse Stem Cell Virus 5'-LTR and splice donor [SD]) or inactivation (splice acceptors [SA] and polyadenylation signals [pA]). The inverted terminal repeats are indicated by the arrows labeled ITR

Figure 5.

Using SB for cancer-gene screens in mice. Step 1. Breed SB transposon and transposase transgenes together. In some cases, arrangements for tissue-specific expression of the transposase will be made or specific cancer-predisposed backgrounds could be used. Step 2. Transposition in somatic cells causes random insertion mutations. A correctly designed transposon vector can cause gain or loss-of-function mutations. Step 3. Mice are aged for tumor development. Step 4. Tumors develop as a result of transposon-induced mutations. Step 5. Transposon insertions are cloned from tumor genomic DNA. Step 6. Clones are sequenced. Step 7. Insertion sites are mapped and annotated with respect to nearest genes. Those genes repeatedly mutated in multiple, independent tumors are designated as common insertion sites or CIS. Step 8. CIS can be analyzed to determine what genes and genetic pathways contribute to cancer. *Cancer genetic fingerprints* are obtained that can be characterized by networks of interacting cancer-gene mutations. These genes can be queried in relevant human cancers.