Risks of Insertional Mutagenesis by DNA Transposons in Cancer Gene Therapy

Perry B. Hackett[a], Timothy K. Starr[a,b], & Laurence J.N. Cooper[c]

[a]Department of Genetics Cell Biology and Development, Center for Genome Engineering and Masonic Cancer Center, University of Minnesota, Minneapolis, MN 55455

[b]Department of Obstetrics, Gynecology & Women's Health, Center for Genome Engineering and Masonic Cancer Center University of Minnesota Medical School, Minneapolis, MN 55455

[c]Division of Pediatrics and Department of Immunology, University of Texas MD Anderson Cancer Center, Houston, TX

For correspondence: Perry Hackett: perryh@umn.edu
Others: TKS: star0044@umn.edu; LJNC: ljncooper@mdanderson.org

Running Head: Risks of Insertional Mutagenesis by Transposons

ABSTRACT

Recently the *Sleeping Beauty* (SB) transposon/transposase system has been applied in clinical trials to redirect T-cell specificity through the addition of a transgenic cassette that drives expression of a chimeric antigen receptor (CAR). We have focused on issues relating to insertional mutagenesis in the context of the plasticity of human genomes, the unexpected variability in human genomes elucidated by recent high-throughput, whole-genome sequencing projects of hundreds of individuals and cells therein, and the recently discovered high rates of remobilization of endogenous transposable elements. These findings are put into the context of the use of SB transposons in the treatment of human disease and suggest that integration by SB transposons is not likely to induce adverse events in the clinic.

KEY WORDS: aAPC, artificial antigen-presenting cell, CAR, chimeric antigen receptor; HSC, hematopoietic stem cell, ITR, inverted terminal repeat; LINE/SINE; long/short interspersed elements SB, *Sleeping Beauty*,

INTRODUCTION

Following a long probationary period, human gene therapy has entered the western commercial world with the approval of Glybera® to treat lipoprotein lipase deficiency using an adeno-associated virus gene vector that has a largely random pattern of integration into host chromosomes. For other diseases, gene therapy has turned the corner from eliciting a mixed reaction of hope associated with high risk to therapeutic benefits that could not otherwise be achieved. The potential of gene therapy to treat disease has always been obvious, but so are the risks of adverse outcomes. Accordingly, the principal objective now is to reduce risk as the applications are broadened to the point that gene therapy becomes commonplace in clinical practice. Although recent clinical trials have demonstrated the increasing promise of gene therapy, they have also illustrated the difficulties of assessing risks given the inherent uncertainty of trial outcomes [1].

This chapter is focused on a primary concern of gene therapy for treatment of cancer using genetically modified T cells, a procedure that was recognized as The Scientific Breakthrough of the Year 2013, and the risks associated with introducing new genes into human chromosomes. Chromosomal integration is necessary when outgrowth of the genetically engineered cells is a required condition for effective clinical treatment, as is the case with adoptive immunotherapy. Here we discuss the concerns of insertional mutagenesis, our recent increased understanding of the plasticity of the human genome, the newly discovered levels of background transposition by endogenous retro-transposable elements, and how that influences our understanding of risks associated with gene therapy.

Retrovirus-associated adverse events in gene therapy trials

Whenever any DNA sequence “randomly” integrates into a genome, there is a chance that the relationship of endogenous genes and their regulatory components will be affected. This is called *insertional mutagenesis*. There are many ways that endogenous genetic information can be corrupted by insertional mutagenesis of which the most commonly discussed are shown in **Figure 1**. Genetic alterations that are associated with adverse health consequences are 1) *Loss-of-function* (LOF) of a critical gene due to interruption by the insertion of an exogenous DNA sequence into an exon. 2) *Gain-of-function* (GOF) of an oncogene, wherein a gene associated with cellular expansion that should not be expressed in a differentiated cell is activated by integration of strong transcriptional regulatory elements that were designed to drive the incoming transgene. 3) LOF due to insertions into introns. Insertion of a vector into an intron generally is more likely than into an exon because introns are on average 6 to 10-fold longer than exons; although vectors that prefer to integrate proximal to

transcriptional start sites may have a preference for the first exon. As long as there are no cryptic splice-acceptor or splice-donor sites in the vector, normal splicing should occur, but the level of transcription may be reduced resulting in low expression of the endogenous gene. 4) Should there be *cryptic* (low level) splice-donor or splice-acceptor sites in the transgenic vector, then there is the possibility of synthesis of chimeric polypeptides that may have dominant-negative function (DNF) or dominant-gain of function (DGF) that could have adverse effects on cellular growth. Other potential mutagenic processes by transposons have been extensively described [2]. The important lesson of Figure 1 is that integration of transgenic expression cassettes inside a gene *does not* irrevocably lead to loss of normal gene expression. This point is discussed further in terms of recently discovered plasticity of regulation of gene expression in human cells.

FIGURE 1

The potential of unpredictable outcomes from insertional vectors was observed when recombinant retroviruses were employed to deliver genes semi-randomly into mouse genomes [3]. Two decades later, similar results were seen in humans undergoing gene therapy in which retroviral vectors were used to introduce a therapeutic IL2γc gene to treat X-linked severe combined immunodeficiency disease (X-SCID). Five of twenty boys developed leukemia three or more years after their treatments [4]. Two boys died, the other three were successfully treated for their leukemias as well as immunodeficiency. In each case, integration of a gamma-retroviral vector was found upstream of the endogenous LMO2 proto-oncogene (Fig. 1, example 2), which led to the hypothesis that in hematopoietic stem cells (HSCs) integrated retroviral vectors and the enhancer/promoter cassettes they carry could activate an oncogene [5]. In subsequent studies, therapeutic retroviruses containing the $gp91^{Phox}$ gene for treatment of chronic granulomatous disease (CGD) resulted in leukemia and one death [6]. In the CGD patients, retroviral integrations proximal to growth-related genes have been associated with enhanced proliferation [7]. Consequently, viral vectors that integrated into chromosomes were associated with leukemias that, in some cases, could be successfully treated. These early events in retroviral transduction of HSCs cemented the notion of the dangers of insertional mutagenesis in the minds of many researchers [8].

In contrast, retroviral vectors carrying either the adenosine deaminase (ADA) gene for treatment of ADA-linked immunodeficiency [9, 10] or WASp for Wiskott-Aldrich Syndrome [11] have not been associated with similar adverse events. Currently, by using altered retroviruses and appropriate doses, adverse events have been avoided [12-14]. Recently, successful use of lentiviral therapies have been demonstrated for metachromatic leukodystrophy, an inherited

lysosomal storage disease caused by arylsulfatase A deficiency [15], and CGD in which vector integration analyses revealed highly polyclonal and multilineage hematopoiesis that resulted from the gene-corrected HSCs. These lentiviral gene therapies have not induced selection of integrations near oncogenes nor aberrant clonal expansion [16]. Thus, although vector integration can be mutagenic, consequential adverse events are not a *fait accompli*. In some cases, there is limited, temporary clonal expansion of treated cells that actually increases the likelihood of successful gene therapy. Moreover, in contrast to early gene therapy of HSCs, viral-mediated transduction of T cells for adoptive immunotherapy has not been associated with adverse outcomes; comparable treatments of T-cells for adoptive immunotherapy has not resulted in undesirable clinical outcomes [12, 13, 17, 18].

There may be differences between cell types for integration sites as well as the availability of endogenous genes for activation. Analyses of populations of gamma retrovirus-modified HSCs show that transgene integration sites were proximal to genes involved in cell survival and growth control, e.g., 40% of insertions were found in <0.5% of the genome [19, 20]. These and other studies make it clear that only a small subset of all the integrations are responsible for most of the outgrowths of therapeutically treated cells. Retroviral parameters associated with integration effects and genotoxicity are still poorly understood [21-24].

The deliberate use of insertional vectors to uncover oncogenes by inducing leukemia and solid tumors has exacerbated the concerns of insertional mutagenesis following integration of the very same vectors when used for gene therapy. Retroviruses [25] as well as transposons [26] are used for this purpose, although predisposing mutations are often included in studies to raise the numbers of oncogenic events per mouse [27]. These studies have identified a number of mechanisms by which integrations of specialized retroviruses induce mutations that lead to cancer.

AIDS patients are at increased risk for some cancers due to their immune state [28], but lentiviral vectors have not been associated with adverse outcomes stemming from either activation, inactivation, and/or alteration of splicing of endogenous genes [29]. The absence of cancers/leukemias following treatments with lentiviruses might be due either to the ability of these vectors to infect non-dividing cells for therapeutic benefit [30] or their selection of a different set of genetic loci for integration [31].

THE *SLEEPING BEAUTY* TRANSPOSON/TRANSPOSASE SYSTEM FOR GENE THERAPY

Our research focuses on the use of the SB transposon system for gene therapy. We have used the SB system to introduce a CAR vector into a patient's T cells *ex vivo*. The T cells

are transplanted back into the patient in order to redirect the patient's immune system to attack the cancerous cells [32].

Non-viral gene therapy has distinct advantages over viral vectors for gene therapy [33, 34]. Transposons, specifically, have two major advantages over viruses as gene therapy vectors. First, clinical grade manufacture and quality control of purified transposon DNA is easier, more reliable and less expensive than employing clinical-grade virus. Second, unlike viral vectors that have an integration bias either into or proximal to transcriptional units, SB transposons have few known preferences for integration. Nevertheless, 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, the results from the 1000 Genomes Project and ENCODE Project [35] have demonstrated that the interactions of genetic elements in our chromosomes are far more variable and complex than previously thought. Second, it is evident that endogenous transposons are far more active in human cells than was surmised until very recently [36]. Two questions arise from these findings - Do these elements induce similar genetic consequences as therapeutic transposons and do cells have mechanisms to cope with insertional mutagenic agents? In light of these recent discoveries, we have three objectives in this review that pertain to the use of SB transposons to engineer human T cells for adoptive immunotherapy in the clinic. 1) Present the context of SB transposition in an environment of relatively high retro-transposition activity in human cells. 2) Evaluate SB transposon-mediated induction of cancer in mouse studies and their pertinence to patients treated with SB transposons. 3) Present approaches to ameliorate potential risks of transposon-mediated events in patients.

DNA Transposons

At least 45% of the human genome is composed of clearly identifiable transposable elements [37] and up to two-thirds of chromosomes may be derived from transposons that have been adapted to support cell function [38]. Transposons are divided into two classes (**Fig. 2**). 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 distinctive features of Class I and Class II transposons are important considerations in evaluating risks associated with any type of gene therapy that involves integration of genetic sequences into chromosomes.

FIGURE 2

DNA transposons are DNA sequences that are excised from the genome for insertion elsewhere in the same, or different chromosome, 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, with many transposons, when the donor DNA sequence is repaired, there is commonly a "footprint" that is left, which, in the case of SB, appears to vary according to cell type, but often is a 5-bp AC(A/T)GT insertion [39] flanked by the TA sites that flank all SB transposons. Approximately 300,000 Class II transposons comprise about 3% of the human genome; but all are inactive due to a lack of a functional transposase gene [37]. *Active* class II transposons are defined by inverted terminal repeats (ITRs) that flank a transposase gene, which does not have a promoter. Thus, remobilization is dependent on integration proximal to an active, endogenous promoter, which is infrequent due to evolutionary pressure to suppress insertional mutagenesis. However, DNA transposons under some circumstances can spread when the transposase gene is activated [40]. Transposons, like viruses, are natural mechanisms used in nature to introduce new genetic sequences into cells. However, the host-cell responses to viruses and transposons, both parasitic invaders, differ. The general strategy of most viruses is to *hit-and-run*, *i.e*., infect and replicate regardless of consequences to the cell or whole organism. As a result, cells have elaborate defenses against viruses. In contrast, transposons are just DNA and so they do not have the infective ability of viruses; they rarely insert into cellular genomes and so the defenses against their spread is generally, but not always, minimal. Transposons are *hit-and-hide* vectors that are maintained over evolutionary periods in all the offspring of the cell.

The *Sleeping Beauty* Transposon System

The potential of DNA transposons for use as vectors for transgene delivery was realized with the resurrection of an extinct DNA transposon in current salmonid genomes [41]. DNA transposons accommodate genetic cargos of up to, and on occasions, more than 10kbp [42]. There are not any active Class II transposons (transposases) in the human genome. The SB system (SB transposon plus SB transposase) gets its name from its resurrection from an evolutionary sleep of more than ten million-years [41]. Owing to approximately 500 million years of evolutionary separation, the salmonid-based SB transposase does not recognize remnants of class II transposons in the human genome. Since its first use as a vector for gene therapy in mice [43], SB transposase has been re-engineered for increased activity, e.g., SB11 and SB100X [42, 44]. Likewise, the inverted-terminal-repeat structure of original transposon, *T*, also has been re-engineered for greater activity to produce *T2* and other versions [33]. As noted later, maximal activity is not always desirable for gene therapy because multiple integrations increase the

probability of unwanted mutations and unwanted effects on the target cells. Currently the intermediate-strength SB11 [42] has been most extensively used in human trials.

FIGURE 3

For current clinical applications of the SB system, the SB11 expression cassette is supplied in *trans* (**Fig. 3**) rather than being incorporated into the same plasmid that carries the therapeutic transposon. There are two reasons for using two plasmids. First, small plasmids are more efficiently introduced into target cells for transposition [42]. Second, for clinical trials, once an SB transposase plasmid has been certified, it can be used with other plasmids that carry various therapeutic transposons. That said, a single plasmid with both the transposon and SB transposase gene (*cis*) could reduce cost as only one clinical-grade DNA plasmid would be needed, a consideration if transposons are used widely to treat cancer or other genetic diseases.

There are multiple advantages to using transposons as gene-delivery vectors. They can be straightforwardly delivered to single cells, the entire transposon is precisely integrated, and each integration reaction is independent so that concatemers are essentially never formed [45]. A further advantage of the SB system, but not all other transposons, is that it only requires a TA dinucleotide sequence for integration, with very few preferences for integration, unlike most integrating viruses [46] (**Fig. 4**). There are about 200 million TA sites in the human genome, which is appealing for obtaining a wide distribution of integrated vectors. Although about 10% of TA sites can be considered *preferential* due to an inherent flexibility of their flanking basepairs [47], the SB system has the least preference for integration in and around transcriptional units. As a consequence of its overall lack of preference for promoters and actively transcribed genes, unlike other transposons adapted for use in mammals, the SB transposon system has become the leading non-viral vector for gene therapy. The SB vector has been validated for *ex vivo* gene delivery to stem cells, including T-cells for the treatment of B-cell malignancies [48], and SB transposons have been delivered to liver for treatment of various diseases in mice [33].

FIGURE 4

When used in a clinical setting, the possibility of remobilization of a transposon leading to an adverse event, as a result of residual transposase activity, is a theoretical concern that is ameliorated by the understanding that SB-mediated transposition is based on the conserved nature of DDE transposases [49] that are designed to be extremely inefficient. Remobilization of a SB transposon would have two consequences. First, remobilization would leave a *footprint* (**Fig. 3**) that generally is an addition of 5bp, although sometimes deletions and larger insertions occur as a result of repair at the excision site [39]. Hence, remobilization out of a protein-encoding exon

would likely induce a frameshift mutation resulting in an abnormal polypeptide. Because exons comprise no more than 2% of the human genome and the rate of excision of a transposon in a cell is about 10^{-4}, we estimate the chance of remobilization of a transposon into an exon is about 10^{-6}, and the chances of integration into an exon in a haplo-insufficient oncogene to induce an adverse event would be more than 100-fold smaller. Of course, the chance of an adverse event from re-integration is the same as for the initial delivery wherein far more sites are hit.

Nevertheless, there are two ways to deal with residual SB activity. First, use a transiently active promoter, such as the CMV early promoter, to control SB transposase expression. This method results in SB activity being reduced about 10,000 fold over a couple of days [50]. Nevertheless, because the possibility of a few cells continuing to express transposase cannot be ruled out, a second alternative is to deliver the SB transposase as mRNA instead of a DNA plasmid [51]. Theoretically this is sound, but the instability of mRNA raises quality control issues that could hinder widespread use for gene therapy.

Experience suggests that more elaborate constructions may not be needed. Data from clinical quality-control criteria studies suggests that there is no more than 1 SB-mRNA/$2x10^{10}$ cellular mRNA molecules, which translates to a maximal number of about 1 SB-mRNA/10^8 T cells [45]. But, studies using SB transposons to identify oncogenes (following section) as well as transformation of hepatocytes [50] suggest that approximately 100 SB-mRNAs per cell are required to support transposition. Hence the rate of re-mobilization is below detectable limits. In any case a remobilization event does not imply adverse consequences to a greater extent than the initial delivery of the integrating vector.

After more than a decade of using SB transposons for gene therapy in mice, there have not been any reports of tumor formation or leukemia. However, as discussed in the following section, SB transposons are notable for their role in identifying putative oncogenes.

SB Transposon-Mediated Induction of Cancer in Mice

The SB transposon system has been used to discover a large number of candidate oncogenes and pathways that lead to cancer [26]. Based on early studies that showed MuLV can cause cancer by acting as an insertional mutagen [52], SB transposons were designed to mimic the abilities of retroviruses to cause both GOF and LOF following integration as well as tagging putative genes involved with induction of tumors. The transposons (originally T/Onc but a number of derivatives such as T2/Onc are in use; here we collectively refer to them as T/Onc transposons) contain splice acceptors (SA) followed by polyadenylation (pA) signals in both orientations (**Fig. 5**). 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) that contain strong, methylation-resistant, transcriptional motifs that will initiate expression from the site of integration [53]. Immediately downstream of the LTR is a splice donor (SD) for splicing of a transcript initiated from the LTR into a neighboring gene. Many variations on this theme have been developed [26]. As shown in Figure 5, integration of a T2/Onc transposon into a gene, exon or intron, can lead to its nearly complete disruption with termination of transcription regardless of orientation of insertion (right-to-left or left-to-right). Additionally, the strong MSCV promoter can cause overexpression of the entire gene, or a truncated portion of the gene This alters the expression of endogenous genes and can direct the synthesis of dominant-negative gain of function polypeptides that can direct cells down any of a number of pathways leading to cancer.

FIGURE 5

Figure 6 shows the general strategy for using SB transposons to not only disrupt genes, but also tag them so that the mutated genes can be identified. Two genetic constructs are required, while additional constructs can be used to limit SB activity to specific cell types or to predispose the mouse to cancer using a known cancer gene (Fig. 6). The first required element is a concatemer of T/Onc transposons that is introduced into mice using standard pro-nuclear injection techniques to generate transgenic mice. This technique results in multiple transposons inserted into a single location in the mouse genome. The multiple transposons, referred to as a concatemer, can range from 25 to 200 copies. The second required element is the SB transposase expression cassette. The expression cassette can be designed to be ubiquitously expressed, or conditionally expressed using a third element. Ubiquitous expression is commonly generated by "knocking-in" an SB cDNA to the ubiquitous Rosa26 promoter in mice. This method has been used to induce or accelerate tumors [54]. Conditional activity is achieved by including a *lox-stop-lox* signal between the Rosa26 promoter and the SB cDNA. The *lox-stop-lox* signal blocks expression of SB unless the cell also expresses Cre Recombinase, a bacteriophage enzyme capable of removing the *lox-stop-lox* signal. Many strains of mice expressing Cre Recombinase under the control of different tissue-specific promoters are available and can be used to direct the expression of SB to specific cell types. This method has been used to generate intestinal epithelial adenocarcinoma, hepatocellular carcinoma and many other cancer types [55, 56]. Finally, a known cancer pre-disposing mutation, such as mutant *p53* or *Apc* can be introgressed into the SB mice to identify mutations that cooperate with these pre-disposing mutations [57, 58]. Figure 6 shows a representative experiment using a predisposing mutation ($Onc^{+/-}$) and a conditional SB expression system using Cre Recombinase. In the outlined experiment, four genotypes are produced: the experimental group (active SB and

Onc$^{+/}$), the active SB only group (active SB, no Onc$^{+/-}$), the predisposed group (Onc$^{+/-}$, no Cre), and the control group (no Onc$^{+/-}$, no Cre). The expected outcome of this experimental design is a high rate of tumorigenesis in the experimental group, attenuated rates in the SB only and Onc$^{+/-}$ groups, and few or no tumors in the control group.

FIGURE 6

To identify the genomic location of the transposon insertions DNA is extracted from a tumor and linker-mediated PCR is used to amplify fragments of DNA where the transposon is adjacent to genomic DNA. These PCR amplicons are sequenced using next generation sequencing technology and the sequences are mapped to the mouse genome. In a typical experiment a single tumor will have several hundred transposon insertions that can be identified. By comparing the insertions in all the tumors generated in the experiment it is possible to identify common insertion sites (CISs) (Fig. 7). Clustering of insertions in certain regions, beyond what is expected by chance, indicates selection for these rare insertion events, among many neutral or negative insertion events, that can give the cell a selective growth or survival advantage. The genes present in these CISs are, therefore, candidate cancer genes.

FIGURE 7

Five overall observations are important for evaluation of SB transposons used for gene therapy. First, although there is efficient induction of T-cell malignancy [54] and myeloid leukemia [59] in mice that are not genetically predisposed to tumor-formation, 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. Second, the presence of either just a concatemer of T/Onc transposons or expression of the SB transposase by itself, does not result in tumorigenesis in mice, even when the transposase is expressed ubiquitously over the lifetime of the mice. These findings suggest that SB transposase does not induce chromosomal instability and/or mobilize endogenous mammalian transposons at a detectable rate. Third, an intermediate to strong promoter sequence to drive transposition is required for efficient tumor induction, consistent with transpositional studies in somatic tissues [50]. Fourth, a promoter/splice donor sequence appears to be required to direct efficient oncogenesis. Fifth, in aggregate, all of the various studies using SB transposons to identify oncogenes by inducing tumors and leukemias highlight the major influence that genetic background has on genotoxicity by integrating vectors.

In rough numbers, assuming a mouse organ has about 1 million cells, with each cell containing a total of 100 transposons, a single tumor per animal would mean that the incidence

of an adverse event is about $1/10^8$ transposons. This is a high number that does not reflect the actual incidence of a single transposon hit because in these animals there are hundreds of insertional mutations per tumor. This adverse event rate will be further reduced in gene therapy because therapeutic transposons will not contain the elements specifically designed to cause tumors (LTRs, splicing signals and transcription termination sequences) that are present in the mutagenesis experiments.

Thus, although the use of SB transposons to cause cancer may suggest danger, more careful consideration of the differences between the transposons used in mice to tag cancer genes and the SB transposons used in a clinical setting reveals that cancer is a highly unlikely outcome following therapeutic transposon delivery. To summarize, in the mouse experiments: 1) transposons are designed to maximally disrupt endogenous genetic pathways in multiple ways; 2) multiple copies of the disruptive transposon are mobilized in every targeted cell; 3) transposase enzyme is expressed continuously in all cells of a given tissue throughout the lifetime of the animal, which can lead to multiple remobilizations; and 4) efficient cancer induction is achieved in mice genetically predisposed to 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.

A central point of the above discussion is that for all the diagrams and models we have to explain mutagenesis that leads to cancer, it takes an extraordinarily high number of events that employ insertion of a collection of transcriptional and RNA processing motifs to elicit adverse events. Recent investigations into the regulation of gene expression, the unexpected variations in human genomes and the natural instability of genomes may explain the actual resistance to transformation events. These observations are important considerations in evaluating risks of insertional mutagenesis by SB transposons when used clinically.

PLASTICITY OF GENOMES AND GENE EXPRESSION IN HUMANS

Sequencing of human genomes from various normal tissue and many tumors and leukemias has revealed unexpected variations in chromosomal sequences including large alterations and complex rearrangements [60, 61]. Deep sequencing has found that there are approximately 22,000 insertions and deletions (indels), which are equivalent to insertions of transposons, in a given human genome with up to 50% being novel for any individual [62]. Moreover, the average human has about 250 to 300 loss-of-function mutations in his/her genome, 50 to 100 of which are in disease-related genes [63, 64]. Included in the variation are copy number variations, especially of

L1 and Alu retro-elements that are found in protein-encoding genes, including growth-promoting genes, in cells of somatic tissues [65-67]. Thus, the human genome is not only highly variable, but it can sustain genetic hits from transposons without apparent genetic consequences.

Retrotransposable elements

DNA transposons are relatively minor contributors to the transposon load in mammalian genomes and they all appear to be inactive [68] Class I transposons (**Fig. 2**) comprise about 42% of the genome and are divided into four types. The most important with respect to gene therapy are the 500,000 to 800,000 long interspersed elements (LINEs) that are up to 6-8 kb long. About 100 LINE-1 elements are active and about 10 are considered highly active (hot) because they encode all of the necessary products for integration [69]. The second subtype of Class I transposons are the approximate 1.5 million short interspersed elements (SINEs) that are about 100-300 bp and are transcribed by RNA polymerase III. SINEs are found everywhere in the genome. A small subset of one prominent SINE family, called *source elements,* also contributes to genomic instability [70]. The third category comprises the nearly half-million retrovirus-like, human *e*ndogenous retro*v*iral sequences (HERVs). HERVs have hallmarks of retroviruses that persistently invaded primate genomes over eons and since have lost their activities to mobilize. HERVs account for about 8% of the human genome but their integrations over millions of years have contributed substantially to primate-specific gene regulation [71]. Resurrection of a consensus HERV-K sequence indicated that these elements preferentially integrated into or proximal to transcriptional units [72]. The fourth category comprises a heterogeneous set of repetitive sequences that are several hundred bp in length and probably are transcribed by RNA polymerase II.

Of the nearly three million retrotransposable elements, only a small proportion of LINEs, those designated L1Hs, and SINEs, primarily *Alu* elements, are active and responsible for insertional mutagenic events that are associated with genetic disease [69, 73], including cancer [66, 74, 75]. For transpositional activity, the LINE or SINE first must be transcribed, generally from an endogenous promoter, and the RNA must be reverse-transcribed to form a double-stranded DNA that can integrate somewhere into the genome. L1 sequences have two open reading frames that support the integration step. The origin of the reverse transcriptase activity in cells is poorly understood. Regardless, the polypeptides encoded in the *hot* L1 elements can trans-mobilize both inactive LINE as well as SINE sequences [76] since both depend on reverse transcriptase and integration enzymes for their spread. With more than two million LINEs and SINEs in the genome, on average they will occur about every 2000 basepairs. However, they are not randomly distributed due to evolutionary selection that eliminates over-active

transposons and transposons in regions that lower the fitness of the cell, but this calculation indicates that many will be well within the influence of endogenous promoters with uncharacterized leakiness.

Decades of analyses of mutations that cause genetic diseases in humans have documented the ability of LINEs to inflict deleterious mutations in human genomes [77]. Nevertheless, owing to the huge number of retroelements in genomes and the paucity of mutagenic events, retroelements were thought to be largely inactive, mainly playing roles in genome evolution [78]. This notion was reconsidered when whole-genome sequencing became available [77, 79]. Mobilization of L1 transposons were found in many somatic neuronal cells [65, 80] and in cancers [81, 82]. The association of the loss of methyl-CpG-binding protein-2 activity in some cells with mobilization suggested that a failure in epigenetic silencing accounted for rare cases of L1 transposition [83, 84]. Investigations into the genome-wide DNA methylation status in human embryonic and adult tissues have identified unexpected tissue-specific and subfamily-specific hypomethylation signatures that can support transposition of retrotransposons [85]. Epigenetic silencing is relaxed in germline cells [86], hence, it is not surprising that epigenetic silencing fails to keep all retrotransposable elements quiet all of the time in all cells.

Complexity of transcription and gene expression in human genomes

The ENCODE project revealed that about 80% of the human genome was expressed in a regulated manner in one cell type or another and that 95% of the genome was subject to transcription. Expression is highly variable - transcription of different sequences spans six orders of magnitude for polyadenylated RNA, which is generally associated with pre-mRNAs with a large proportion of the transcripts being initiated from repetitive elements. The average protein-encoding gene has four splicing isoforms that are produced in unequal amounts [87]. The insertion of transgenic DNA into chromosomes can alter epigenetic marks that affect gene expression [84] and conversely retroviral vectors are subject to epigenetic silencing [88]. Moreover, enhancer elements can work over long distances via chromosomal looping with only about 7% of the looping interactions being to the nearest gene [89]. 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 [90]. Hence, physical proximity of an integrated vector is not a simple predictor for activation (or repression) of an endogenous gene. Transcriptional regulatory proteins selectively interact with each other and specific promoters [91], which may explain the pronounced lack of activation of endogenous genes that lead to cancer by enhancers in lentiviral vectors and the repetitive findings of LMO2 activations with some murine retroviral vectors.

These findings are important for risk assessment of integrating vectors used in gene therapy. 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. The bottom line is that transcriptional activation is far more complex than previously thought and the possible consequences of insertional mutagenesis are clearly more complex than shown in Fig. 1, which serves more as a model than a guide. In fact, experiments conducted decades ago analyzing transposition in nematodes showed that despite a strong preference for integration into transcriptional units most Tc1-transposon integrations did not have an obvious phenotypic effect because when they integrated into introns (Fig. 1, example 3) their sequences were removed during pre-mRNA processing thereby avoiding detrimental consequences [92].

Genetic consequences of natural transposition and therapeutic transposition

One way of addressing insertional mutagenesis by SB transposon vectors is to compare the nearly random insertion profile of transposons with natural variation and the germline mutation rate. For example, variations in transcriptional regulation in cells normally leads to expression of millions of transposable elements, including the ~10^{14} *hot* retrotransposable elements and an unknown number of active *Alu* elements. Combined with the background mutation rate in normal cells, the potential for deleterious mutations is potentially millions of fold greater than the number of mutations resulting from single integrations of SB transposons in several thousand cells. 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. Specifically, since only approximately 1.2% of the genome consists of protein-encoding exons [35], almost 99% of integrations are unlikely to affect protein sequences in any way other than their rates of expression, which can vary widely (**Fig. 1**, example 3). The remaining 1% 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, in many cases 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. Although retrotransposition and DNA transposition occur by different mechanisms (*copy-and-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 (**Fig. 1**). These controls over insertional damage by endogenous transposons will likely be relevant to transposons used for gene therapy.

In the clinical studies discussed next, a chimeric antigen receptor (CAR) is used to redirect the specificity of T cells to bind to a tumor-associated antigen (TAA). 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, potential mutations arising from insertions of SB transposons may be more likely to produce an undesirable effect than in cells from healthy individuals.

TRANSPOSON-MEDIATED GENE THERAPY IN THE CLINIC

Adoptive transfer of CAR^+ T cells with a CD-19 targeted CAR in early clinical trials resulted in the desired outcome, the loss of normal $CD19^+$ B cells and the destruction of malignant B cells in some patients [93, 94]. At present, most trials targeting CD19 have used viruses to deliver the CAR to T cells. There have not been any reports of insertional mutagenesis leading to either persistent clonal expansion or enrichment for integration sites near genes implicated in growth control and/or transformation [12]. These clinical data provide the foundation for the first human application at MD Anderson Cancer Center of the SB system to redirect T-cell specificity against B-cell leukemias and lymphomas. In these trials, a CAR (**Fig. 8A**) was constructed to recognize the CD19-lineage antigen, a common polypeptide found on the surface of normal and malignant B cells. In preparation for these trials, we accumulated data regarding the ability of the SB system to enforce expression of a second generation CAR that activates T cells via chimeric CD3-ζ and CD28 [48]. These two chimeric signaling motifs achieve a fully competent T-cell activation-event, which is defined, at a minimum, as CAR-mediated proliferation, cytokine production, as well as serial and specific lysis. The introduction of the CAR was achieved by co-electroporation (**Fig. 3**) of two supercoiled DNA plasmids, one containing the CAR-transposon and the second encoding SB11 [95]. T cells stably expressing the CAR could be expanded by co-culture with designer artificial antigen-presenting cells (aAPC) in the presence of the soluble recombinant cytokines IL-2 and IL-21 (**Fig. 8B**). The aAPC used for the initial clinical trials were derived from the human K-562 immortalized cells that were genetically modified using lentivirus to co-express CD19 as well as CD64 and the T-cell co-

stimulatory molecules CD86, CD137L, and a membrane-bound mutant form of human IL-15. By recursively adding γ-irradiated aAPC to the stably modified CAR$^+$ T cells, we achieved outgrowth of CAR$^+$ T cells. The electroporation method employs defined ratios of mononuclear cells and components of the SB transposon system that result in between 1 and 2 CAR insertions per T-cell genome without integration of SB11 transgene, based on the absence of a PCR signature for SB11. The combination of SB and aAPC platforms routinely results in the outgrowth of T cells with the majority expressing CAR within a few weeks after electroporation [96].

FIGURE 8

Risks associated with T cells modified to stably express CARs from transposons

Adverse events may be associated with 1) lympho-depletion of the recipient to improve engraftment of infused T cells, 2) massive cell death and the accompanying cytokine storm triggered by the CAR T-cell killing of the B-cells, and 3) rare adverse effects from integration of the CAR transgene.

. Recipients that receive myelosuppressive or myeloablative therapy 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 [97]; but exceptional events of this sort should be manageable by modifications in caring for medically-fragile patients. There have been expected and unexpected adverse events due to *on-target* toxicities as the CAR recognizes tumor-associated antigen on normal cells [98, 99]. For example, CAR$^+$ T cells will not distinguish between CD19 on normal and cancerous B cells. This results in damage and even loss of humoral immunity in patients that benefit from T-cell therapy for B-lineage leukemias and lymphomas [100-103]. Destruction of normal B cells can be managed by repeated administrations of intravenous immunoglobulin, which appears tolerable in most patients at risk from dying from progressive B-cell malignancies. In addition, the synchronous activation of CAR$^+$ T cells engaging large numbers of CD19$^+$ malignant B cells can result in supra-physiologic elevations of cytokines and attendant systemic side effects. In severe cases patients experience fever and degrading vital signs that can be complicated by tumor lysis syndrome associated with the desired destruction of large numbers of tumor cells. This type of adverse event can be managed by timely application of IL-6 receptor blockade and intravenous administration of corticosteroids.

Unforeseen adverse events can occur due to on-target, but off-tissue binding of the CAR to normal cells. For example, one patient that received a large number of ERBB2-specific CAR$^+$ T cells died due to off-tissue effects [99]. Expected 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. Whether unforeseen toxicities can be managed with a similar approach remains to be seen.

The third type of risk associated with infusing CAR^+ T cells stems from insertional mutagenesis. As noted earlier, unlike gene transfer into hematopoietic stem cells [8] there has not been a reported genotoxic event in genetically modified T cells. Indeed, hundreds of infusions have been delivered using mostly recombinant retrovirus to transduce T cells [12]. Extensive in-process testing and release testing is undertaken prior to infusion of the SB-modified T cells. The in-process evaluations include performing PCR to exclude integration of SB11 transgene, limiting the number of CAR integrants per T-cell genome to 1 or 2, and assessing the diversity of the T-cell receptor (TCR) to exclude the emergence of either a mono- or an oligo-clonal population of genetically modified T cells 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. Other release tests include measurements of 1) sterility (absence of bacteria, fungal, mycoplasma, endotoxin), 2) viability, 3) enumeration of CAR^+ T cells, 4) chain of custody (measurement of HLA class I to validate identity between CAR^+ T cells and the intended recipient), 5) identity (CD3 expression to validate the presence of T cells), 6) contamination to exclude B cells and aAPC, and 7) expression of the CAR. The SB-modified clinical-grade T cells are prepared in compliance with current good manufacturing practice for Phase I/II trials at MD Anderson Cancer Center.

Four phase I studies are underway to establish safety, feasibility and persistence of the infused T cells. Three of these trials infuse patient- and donor-derived CAR^+ T cells in a setting of minimal residual disease after autologous and allogeneic hematopoietic stem-cell transplantation, including after umbilical cord blood transplantation. The fourth trial infuses autologous CAR^+ T cells after lympho-depleting chemotherapy.. 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.

CONCLUSIONS

The fears of insertional mutagenesis during the course of adoptive cell therapy will likely persist – the memories of the X-SCID adverse events will not be forgotten. A recurrent idea has been to guide integrations into “safe havens” in the genome [104, 105]. Site-specific nucleases, including zinc-finger nucleases (ZFNs), TAL-Effector Nucleases (TALENs), the CRISPR/Cas9 system, are all feasible for efficiently supporting homologous recombination into specific

genomic sites [106]. The SB system can also be combined with genetic editing tools such as designer ZFNs and TALENs [107].

The barriers to widespread adoption of gene therapy are being lowered as we achieve greater understanding of gene expression regulation and the genomic environments into which vectors insert. The promise of genetically modifying the human genome to overcome cancer has never been greater. The SB system provides a clear alternative to viral vectors to achieve safe and efficient gene therapy. The low cost of manufacturing clinical-grade SB plasmids and availability of expertise to manufacture these expression vectors in a commercial setting contrasts with the expense, uncertainty, and lengthy timeline to manufacture recombinant clinical-grade virus. Lowering the cost and improving the reliability of the gene delivery method will be essential for widespread application of gene therapy. Several cohorts of patients have successfully received T cells genetically modified with the SB system. Facilitating the use of SB transposons is the availability of methods such as electroporation to introduce immunologically neutral naked DNA molecules into cells. Now it is possible to adopt the SB system for further human applications beyond T cells and CARs.

ACKNOWLEDGEMENTS

We thank Drs. Judy S. Moyes and Harjeet Singh for careful reading and editing of the manuscript 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 5R00CA151672-04 and P30 CA77598 (TKS).

REFERENCES

1. Deakin CT, Alexander IE, Hooker CA, Kerridge IH. Gene therapy researchers' assessments of risks and perceptions of risk acceptability in clinical trials. *Mol Ther* 2013; **21**: 806-15.
2. Huang CR, Burns KH, Boeke JD. Active transposition in genomes. *Ann Rev Genet* 2012; **46**: 651-75.
3. Wagner EF, Covarrubias L, Stewart TA, Mintz B. Prenatal lethalities in mice homozygous for human growth hormone gene sequences integrated in the germ line. *Cell* 1983; **35**: 647-55.
4. Hacein-Bey-Abina S, Hauer, J., Lim, A., Picard, C., Wang, G.P., Berry, C.C., Martinache, C., Rieux-Laucat, F., Latour, S., Belohradsky, B.H., Leiva, L., Sorensen, R., Debré, M., Casanova, J.L., Blanche, S., Durandy, A., Bushman, F.D., Fischer, A., Cavazzana-Calvo, M. Efficacy of gene therapy for X-linked severe combined immunodeficiency. *New Eng J Med* 2010; **363**: 355-64.
5. Gaspar HB, Cooray S, Gilmour KC, et al. Long-term persistence of a polyclonal T cell repertoire after gene therapy fo rX-linked severe combined immunodeficiency. *Sci Transl Med* 2011; **3**: 97ra79.

6. Kang EM, Choi, U., Theobald, N., Linton, G., Long Priel, D.A., Kuhns, D., Malech, H.L. 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**: 783-91.
7. Biasco L, Baricordi C, Aiuti A. Retroviral integrations in gene therapy trials. *Mol Ther* 2012; **20**: 709-16.
8. Baum C, von Kalle C, Staal FJ, et al. Chance or necessity? Insertional mutagenesis in gene therapy and its consequences. *Mol Ther* 2004; **9**: 5-13.
9. Aiuti A, Cassani B, Andolfi G, et al. Multilineage hematopoietic reconstitution without clonal selection in ADA-SCID patients treated with stem cell gene therapy. *J Clin Invest* 2007; **117**: 2233-40.
10. Cassani B, Montini E, Maruggi G, 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**: 3546-56.
11. Boztug K, Schmidt ,M., Schwarzer, A., Banerjee, P.P., Díez, I.A., Dewey, R.A., Böhm, M., Nowrouzi, A., Ball,C.R., Glimm, H., Naundorf,S., Kühlcke, K., Blasczyk, R., Kondratenko, I., Maródi, L., Orange, J.S., von Kalle, C., Klein, C. Stem-cell gene therapy for the Wiskott-Aldrich Syndrome. *New Eng J Med* 2010; **363**: 1918-27.
12. Scholler J, Brady TL, Binder-Scholl G, et al. Decade-long safety and function of retroviral-modified chimeric antigen receptor T cells. *Sci Transl Med* 2012; **4**: 132ra53.
13. Schwarzwaelder K, Howe, S.J., Schmidt, M., Brugman, M.H., Deichmann, A., Glimm, H., Schmidt, S., Prinz, C., Wissler, M., King, D.J., Zhang, F., Parsley, K.L., Gilmour, K.C., Sinclair, J., Bayford, J., Peraj, R., Pike-Overzet, K., Staal, F.J., de Ridder, D., Kinnon, C., Abel, U., Wagemaker, G., Gaspar, H.B., Thrasher, A.J., von Kalle, C. Gammaretrovirus-mediated corrrectionof SCID-X1 is associated with skewed vector integration site distribution in vivo. *J Clin Invest* 2007; **117**: 2241-9.
14. van der Loo JC, Swaney WP, Grassman E, et al. 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**: 872-6.
15. Biffi A, Montini E, Lorioli L, et al. Lentiviral Hematopoietic Stem Cell Gene Therapy Benefits Metachromatic Leukodystrophy. *Science* 2013; **341**: 1233158.
16. Aiuti A, Biasco L, Scaramuzza S, et al. Lentiviral hematopoietic stem cell gene therapy in patients with Wiskott-Aldrich syndrome. *Science* 2013; **341**: 1233151.
17. Newrzela S, Cornils K, Li Z, et al. Resistance of mature T cells to oncogene transformation. *Blood* 2008; **112**: 2278-86.
18. Cattoglio C, Maruggi G, Bartholomae 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.
19. Wang GP, Berry CC, Malani N, 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**: 4356-66.
20. Deichmann A, Brugman MH, Bartholomae CC, et al. Insertion sites in engrafted cells cluster within a limited repertoire of genomic areas after gammaretroviral vector gene therapy. *Mol Ther* 2011; **19**: 2031-9.
21. Baum C. Parachuting in the epigenome: the biology of gene vector insertion profiles in the context of clinical trials. *EMBO Mol Med* 2011; **3**: 75-7.
22. Corrigan-Curay J, Cohen-Haguenauer, O., O'Reilly, M., Ross, S.R., Fan, H., Rosenberg, N., Somia, N., King, N., Friedmann, T., Dunbar, C., Aiuti, A., Naldini, L., Baum, C., von Kalle, C., Kiem, H.P., Montini, E., Bushman, F,. Sorrentino, B.P., Carrondo, M., Malech, H., Gahrton, G., Shapiro, R., Wolff, L., Rosenthal, E., Jambou, R., Zaia, J., Kohn, D.B. Challenges in vector and trial design using retroviral vectors for long-term gene correction in hematopoietic stem cell gene therapy. *Mol Ther* 2012; **20**: 1084-94.

23. Montini E, Cesana D, Schmidt M, 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**: 964-075.
24. Kustikova O, Brugman M, Baum C. The genomic risk of somatic gene therapy. *Semin Cancer Biol* 2010; **20**: 269-78.
25. Uren AG, Kool J, Berns A, van Lohuizen M. Retroviral insertional mutagenesis: past, present and future. *Oncogene* 2005; **24**: 7656-72.
26. Copeland NG, Jenkins NA. Harnessing transposons for cancer gene discovery. *Nat Rev Genet* 2010; **10**: 696-706.
27. Bergerson RJ, Collier, L.S., Sarver, A.L., Been, R.A., Lugthart ,S., Diers, M.D., Zuber, J., Rappaport ,A.R., Nixon, M.J., Silverstein, K.A., Fan, D., Lamblin, A.F., Wolff, L., Kersey, J.H., Delwe, I. R., Lowe, S.W., O'Sullivan, M.G., Kogan, S.C., Adams, D.J., Largaespada, D.A. An insertional mutagenesis screen identifies genes that cooperate with Mll-AF9 in a murine leukemogenesis model. *Blood* 2012; **119**: 4512-23.
28. Oliveira-Cobucci RN, Saconato H, Lima, P.H., Rodrigues HM, et al. Comparative incidence of cancer in HIV-AIDS patients and transplant recipients. *Cancer Epidemiol* 2012; **36**: e69-73.
29. Cesana D, Squaldino 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**: 1667-76.
30. Brady T, Bushman FD. Nondividing cells: a safer bet for integrating vectors? *Mol Ther* 2011; **19**: 640-1.
31. Biffi A, Bartolomae CC, Cesana D, 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**: 5332-9.
32. Kebriaei P, Huls H, Jena B, 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**: 444-50.
33. 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.
34. Zhang Y, Satterlee A, Huang L. In vivo gene delivery by nonviral vectors: overcoming hurdles? *Mol Ther* 2012; **20**: 1298-304.
35. ConsortiumENCODE P, Bernstein BE, Birney E, et al. An integrated encyclopedia of DNA elements in the human genome. *Nature* 2012; **489**: 57-74.
36. Kaer K, Speek M. Retroelements in human disease. *Gene* 2013; **518**: 231-41.
37. Lander ES, al. e. Initial sequencing and analysis of the human genome. *Nature* 2001; **409**: 860-921.
38. 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.
39. 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**: 574-83.
40. Plasterk RHA, Izsvák Z, Ivics Z. Resident aliens: the Tc1/mariner superfamily of transposable elements. *Trends Genet* 1999; **15**: 326-32.
41. Ivics Z, Hackett PB, Plasterk RH, Izsvak Z. Molecular reconstruction of *Sleeping Beauty*, a Tc1-like transposon from fish, and its transposition in human cells. *Cell* 1997; **91**: 501-10.
42. Geurts AM, Yang Y, Clark KJ, et al. Gene transfer into genomes of human cells by the *Sleeping Beauty* transposon system. *Mol Ther* 2003; **8**: 108-17.
43. Yant SR, Meuse L, Chiu W, Ivics Z, Izsvak 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.

44. Mátés L, Chuah MK, Belay E, et al. Molecular evolution of a novel hyperactive *Sleeping Beauty* transposase enables robust stable gene transfer in vertebrates. *Nat Genet* 2009; **41**: 753-61.
45. Hackett PB, Largaespada DA, Switzer KC, Cooper LJ. Evaluating risks of insertional mutagenesis by DNA transposons in gene therapy. *Transl Res* 2013; **161**: 265-83.
46. Yant SR, Wu X, Huang Y, et al. Nonrandom insertion site preferences for the SB transposon in vitro and in vivo. *Mol Ther* 2004; **9**: S309.
47. Geurts AM, Hackett CS, Bell JB, et al. DNA structural patterns influence integration site preferences for mobile elements. *Nucl Acids Res* 2006; **34**: 2803-11.
48. Singh H, Huls H, Kebriaei P, Cooper LJN. A new approach to gene therapy using Sleeping Beauty to genetically modify clinical-grade T cells to target CD19. *Immunol Rev* 2014; **257**: 181-90.
49. Nesmelova IV, Hackett PB. DDE Transposases: structural similarity and diversity. *Adv Drug Del Rev* 2010; **62**: 1187-95.
50. 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**: 1792-802.
51. 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**: 625-30.
52. Mikkers H, Berns A. Retroviral insertional mutagenesis: tagging cancer pathways. *Adv Cancer Res* 2003; **88**: 53-99.
53. Cherry SR, Biniszkiewicz D, van Parijs L, Baltimore D, Jaenisch R. Retroviral expression in embryonic stem cells and hematopoietic stem cells. *Mol Cell Biol* 2000; **20**: 7419-26.
54. Dupuy AJ, Akagi K, Largaespada DA, Copeland NG, Jenkins NA. Mammalian mutagenesis using a highly mobile somatic *Sleeping Beauty* transposon system. *Nature* 2005; **436**: 221-6.
55. Starr TK, Allaei R, Silverstein KA, et al. A transposon-based genetic screen in mice identifies genes altered in colorectal cancer. *Science* 2009; **323**: 1747-50.
56. Keng VW, Villanueva A, Chiang DY, et al. A conditional transposon-based insertional mutagenesis screen for genes associated with mouse hepatocellular carcinoma. *Nature Biotech* 2009; **27**: 264-74.
57. Starr TK, Scott, P.M., Marsh, B.M., Zhao, L., Than, B.L., O'Sullivan, M.G., Sarver, A.L., Dupuy, A.J., Largaespada, D.A., Cormier, R.T. 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**: 5765-70.
58. Rahrmann EP, Watson AL, Keng VW, et al. Forward genetic screen for malignant peripheral nerve sheath tumor formation identifies new genes and pathways driving tumorigenesis. *Nature Genet* 2013; **45**: 756-66.
59. Vassiliou GS, Cooper JL, Rad R, et al. Mutant nucleophosmin and cooperating pathways drive leukemia initiation and progression in mice. *Nature Genet* 2011; **43**: 470-5.
60. Abyzov A, Mariani J, Palejev D, et al. Somatic copy number mosaicism in human skin revealed by induced pluripotent stem cells. *Nature* 2013; **492**: 438-42.
61. Handsaker RE, Korn JM, Nemesh J, McCarroll SA. Discovery and genotyping of genome structural polymorphism by sequencing on a population scale. *Nature Genet* 2011; **43**: 269-76.
62. Alkan C, Coe BP, Eichler EE. Genome structural variation discovery and genotyping. *Nature Rev Genet* 2011; **12**: 363-76.
63. Durbin RM, Abecasis GR, Altshuler DL, et al. A map of human genome variation from population-scale sequencing. *Nature* 2010; **467**: 1061-73.
64. MacArthur DG, Balasubramanian S, Frankish A, et al. A systematic survey of loss-of-function variants in human protein-coding genes. *Science* 2012; **335**: 823-8.
65. Baillie JK, Barnett, M.W., Upton, K.R., Gerhardt, D.J., Richmond, T.A., De Sapio, F., Brennan, P.M., Rizzu, P., Smith, S., Fell, M., Talbot, R.T., Gustincich, S., Freeman, T.C., Mattick, J.S., Hume, D.A.,

Heutink, P., Carninci, P., Jeddeloh, J.A., Faulkner, G.J. Somatic retrotransposition alters the genetic landscape of the human brain. *Nature* 2011; **479**: 534-7.
66. Lee E, Iskow R, Yang L, et al. Landscape of somatic retrotransposition in human cancers. *Science* 2012; **337**: 967-71.
67. 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**: 18018-23.
68. Pace II JK, Feschotte C. The evolutionary history of human DNA transposons: evidence for intense activity in the primate lineage. *Genome Res* 2007; **17**: 422-32.
69. Beck CR, Collier P, Macfarlane C, et al. LINE-1 retrotransposition activity in human genomes. *Cell* 2010; **141**: 1159-70.
70. Ade C, Roy-Engel AM, Deininger PL. Alu elements: an intrinsic source of human genome instability. *Curr Opin Virol* 2013; **3**: 639-45.
71. Faulkner GJ, Kimura Y, Daub CO, et al. The regulated retrotransposon transcriptome of mammalian cells. *Nature Genet* 2009; **40**: 563-71.
72. Brady T, Lee YN, Ronen K, et al. Integration target site selection by a resurrected human endogenous retrovirus. *Genes Dev* 2009; **23**: 633-42.
73. Brouha B, Schustak J, Badge RM, et al. Hot L1s account for the bulk of retrotransposition in the human population. *Proc Natl Acad Sci USA* 2003; **100**: 5280-8.
74. Solyom S, Ewing, A.D., Rahrmann, E.P., Doucet, T.T., Nelson, H.H., Burns, M.B., Harris, R.S., Sigmon, D.F., Casella, A., Erlanger, B., Wheelan, S., Upton, K.R., Shukla, R., Faulkner, G.J., Largaespada, D.A., Kazazian, H.H. Extensive somatic L1 retrotransposition in colorectal tumors. *Genome Res* 2012; **22**: 2328-38.
75. Shukla R, Upton KR, Muñoz-Lopez M, et al. Endogenous retrotransposition activates oncogenic pathways in hepatocellular carcinoma. *Cell* 2013; **153**: 101-11.
76. Wagstaff BJ, Hedges, D.J., Derbes, R.S., Campos-Sanchez, R., Chiaromonte, F., Makova, K.D., Roy-Engel, A.M. Rescuing Alu: Recovery of new inserts shows LINE-1 preserves Alu activity through A-tail expansion. *PLoS Genet* 2012; **8**: e1002842.
77. Ewing AD, Kazazian HHJ. Whole-genome resequencing allows detection of many rare LINE-1 insertion alleles in humans. *Genome Res* 2011; **21**: 985-90.
78. Cordaux R, Batzer MA. The impact of retrotransposons on human genome structure. *Nature Rev Genet* 2009; **10**: 691-703.
79. Burns KH, Boeke JD. Human transposon tectonics. *Cell* 2012; **149**: 740-52.
80. Coufal NG, Garcia-Perez, J.L., Peng, G.E., Yeo, G.W., Mu, Y., Lovci, M.T., Morell, M., O'Shea, K.S., Moran, J.V., Gage, F.H. L1 retrotransposition in human neural progenitor cells. *Nature* 2009; **460**: 1127-31.
81. Iskow RC, McCabe MT, Mills RE, et al. Natural mutagenesis of human genomes by endogenous retrotransposons. *Cell* 2010; **141**: 1253-61.
82. Rodić N, Burns KH. Long interspersed element-1 (LINE-1): passenger or driver in human neoplasms? *PLoS Genet* 2013; **9**: e1003402.
83. Garcia-Perez JL, Morell, M., Scheys, J.O., Kulpa, D.A., Morell, S., Carter, C.C., Hammer, G.D., Collins, K.L., O'Shea, K.S., Menendez, P., Moran, J.V. Epigenetic silencing of engineered L1 retrotransposition events in human embryonic carcinoma cells. *Nature* 2010; **466**: 769-73.
84. Doerfler W. Impact of foreign DNA integration on tumor biology and on evolution via epigenetic alterations. *Epigenomics* 2012; **4**: 41-9.
85. Xie M, Hong, C., Zhang, B., Lowdon, R.F., Xing, X., Li, D., Zhou, X., Lee, H.J., Maire, C.L., Ligon, K.L., Gascard, P., Sigaroudinia, M., Tlsty, T.D., Kadlecek, T., Weiss, A., O'Geen, H., Farnham, P.J., Madden, P.A., Mungall, A.J., Tam, A., Kamoh, B., Cho, S., Moore, R., Hirst, M., Marra, M.A., Costello, J.F., Wang, T. DNA hypomethylation within specific transposable element families associates with tissue-specific enhancer landscape. *Nature Genet* 2013; **45**: 836-41.

86. Bao J, Yan W. Male germline control of transposable elements. *Biol Reprod* 2012; **86**: 162, 1-14.
87. Djebali S, Davis CA, Merkel A, et al. Landscape of transcription in human cells. *Nature* 2012; **489**: 101-8.
88. Poleshko A, Palagin I, Zhang R, et al. Identification of cellular proteins that maintain retroviral epigenetic silencing: evidence for an antiviral response. *J Virol* 2008; **82**: 2313-23.
89. Sanyal A, Lajoie BR, Jain G, Dekker J. The long-range interaction landscape of gene promoters. *Nature* 2012; **489**: 109-13.
90. Zhang J, Markus J, J B, Paul T, Wolff L. Three murine leukemia virus integration regions within 100kb upstream of c-myb are proximal to the 5' regulatory region of the gene through DNA looping. *J Virol* 2012; **86**: 10524-32.
91. Zhang Y, Wong CH, Birnbaum RY, et al. Chromatin connectivity maps reveal dynamic promoter-enhancer long-range associations. *Nature* 2013; **504**: 306-10.
92. Rushforth AM, Anderson P. Splicing removes the Caenorhabditis elegans transposon Tc1 from most mutant pre-mRNAs. *Mol Cell Biol* 1996; **16**: 422-9.
93. Kershaw MH, Westwood JA, Darcy PK. Gene-engineered T cells for cancer therapy. *Nature Rev Cancer* 2013; **18**: 525-41.
94. Themeli M, Kloss CC, Ciriello G, et al. Generation of tumor-targeted human T lymphocytes from induced pluripotent stem cells for cancer therapy. *Nature Biotech* 2013; **31**: 928-33.
95. Singh H, Manuri PR, Olivares S, et al. Redirecting specificity of T-cell populations for CD19 using the *Sleeping Beauty* system. *Cancer Res* 2008; **68**: 2961-71.
96. Manuri PV, Wilson MH, Maiti SN, et al. piggyBac transposon/transposase system to generate CD19-specific T cells for the treatment of B-lineage malignancies. *Hum Gene Ther* 2010; **21**: 427-37.
97. 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**: 666-8.
98. Lamers CH, Sleijfer S, Vulto AG, 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**: e20-2.
99. 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**: 843-51.
100. Kochenderfer JN, Dudley ME, Feldman SA, 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**: 2709-20.
101. Porter DL, Levine BL, Kalos M, Bagg A, June CH. Chimeric antigen receptor-modified T cells in chronic lympholid leukemia. *New Eng J Med* 2011; **365**: 725-33.
102. Brentjens RJ, Rivière I, Park JH, 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**: 4817-28.
103. Kalos M, B.L. L, Porter DL, 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.
104. Sadelain M, Papapetrou EP, Bushman FD. Safe harbours for the integration of new DNA in the human genome. *Nature Rev Cancer* 2011; **12**: 51-8.
105. Papapetrou EP, Lee G, Malani N, et al. Genomic safe harbors permit high beta-globin transgene expression in thalassemia induced pluirpotent stem cells. *Nature Biotech* 2011; **29**: 73-8.
106. Gaj T, Gersbach CA, Barbas CFr. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol* 2013; **31**: 397-404.

107. Torikai H, Reik A, Liu PQ, 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**: 5697-705.

FIGURE LEGENDS

Fig. 1: Insertional mutagenesis by integration of therapeutic vectors. This figure uses a DNA transposon as a model vector, denoted by the inverted terminal repeats (blue triangles) but the same concepts apply to viral vectors. Transposon or viral integration can occur into four general types of chromatin. Integration into heterochromatin, essentially a black hole for expression cassettes, normally results in silencing of the vector. Integration into intergenic regions is desirable because a Therapeutic Gene (TG) will be expressed without affecting endogenous genes. Integration into or proximal to a transcriptional regulatory region that contains a protein-encoding gene (Gene X) can have several outcomes including gain-of-function (GOF) of the transgene, as well as either GOF or loss-of-function (LOF) of Gene X. Normal expression of Gene X is shown by the dashed green arrow representing RNA synthesis, light black lines representing splicing of exons (with alternative splicing of the first exon to either the second or third exon being indicated), and alternative translational products represented by the purple boxes. Four examples of insertional mutagenesis are shown: 1) integration of a vector into an exon, which can lead to formation of a hybrid polypeptide (purple/blue product) similar to normal proteins that come from alternative splicing; 2) integration of a vector with a strong promoter that both leads to expression of the transgene, but also may activate a silenced promoter (the red X) for Gene X thereby leading to unwanted expression of an endogenous gene; 3) integration into an intron of Gene X, which generally results in expression of the transgene with little effect on Gene X because the transgenic sequences are spliced out of the Gene X pre-mRNAs; and 4) integration of a vector with cryptic splicing sites that (red line) that can lead to hybrid proteins (purple/gold product) that will be loss of function with respect to Gene X. In both examples 1 and 4, as well as other possibilities not shown, the hybrid polypeptides could have dominant-negative functions (DNF) and/or dominant gain-of-function (DGF) that need not be “negative” *per se*. Many other types of genomic alterations, including induction or repression of epigenetic markings are possible.

Fig. 2: **Transposable elements in the human genome**. Transposable elements, which comprise nearly 50% of the human genome, are divided into two classes: Class I

retrotransposable elements that spread by a *copy-and-paste* mechanism whereby the transposon is transcribed by either RNA polymerase II (LINEs and LTR-like) or RNA polymerase III (SINEs) and the RNA transcript (dashed green line) is then reverse–transcribed for insertion as a double-stranded DNA elsewhere in the genome. The LTR-like transposons have long terminal repeat sequences (illustrated in the table to the right) that flank a set of genes that are related to retroviral genomes; the genes generally are defective. Only a very small proportion of Class I transposons are active and even these rarely transpose. Class II transposons, which include the synthetic SB system, are DNA sequences that can "hop/jump" from one site to the next via a *cut-and-paste* mechanism when an active transposase is available that recognizes their inverted terminal repeat (ITR) sequences. No active DNA transposase genes in the human genome have been identified and the SB transposase does not bind and transpose the inactive human Class II sequences. Blue entries in the tables indicate the numbers of *active* elements of each type; there are an unknown number of active SINEs, generally *Alu* elements.

Fig. 3: 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 (in this case a chimeric antigen receptor - CAR) and its promoter (green pentagon) is delivered through the plasma membrane of target cells; electroporation (yellow bolt) is the method of choice for delivery to T cells (**a**). After entry into the nucleus (dashed oval), the SB transposase is expressed (**b**) and SB transposase cuts the transposon out of the plasmid (**c**) and inserts it into a chromosome (**d**). The inverted terminal repeats (inverted set of double arrowheads) define the transposon. The second part of the SB system is SB transposase, which in this example is carried in a separate plasmid. The remaining plasmid, called an excision product, will have a footprint (red X) that is a sequence resulting from the repair of the cleavages on either side of the transposon. Excision products are a convenient way of evaluating the extent of transposition reactions. The SB transposon will almost always integrate into a TA-dinucleotide basepair. There are about 200 million in a diploid human genome, which allows nearly random insertion in the genome. Integration into a chromosome can confer sustained expression of the gene of interest through many rounds of cell amplification.

Fig. 4: Comparison of integration preferences of vectors used in gene therapy. The schematic illustrates the nearly random integration profile of SB transposons (blue arrows) that

can integrate at any of about 200 million TA sites in the human genome more-or-less regardless of genetic activity, compared to retroviruses that prefer transcriptional regulatory motifs and lentiviruses and adeno-associated (AAV) viruses that prefer integration into active transcriptional units.

Fig. 5: The T/Onc transposon. The T/Onc series of SB mutagenic vectors is illustrated by T2/Onc. T2/Onc contains elements designed to elicit either transcriptional activation from the *m*ouse *s*tem *c*ell *v*irus 5'-LTR (green MSCV pentagon) and splice donor (purple SD box) or inactivation (splice acceptors (gold SA boxes) and polyadenylation signals (red pA hexagons). The inverted terminal repeats are indicated by the blue arrows labeled ITR. The T2/Onc is effective in disrupting gene expression regardless of the orientation in which it integrates. Potential consequences of integration are shown in the bottom panel using the same conventions as in Fig. 1.

Fig. 6. **T2/Onc transposon, cancer-gene screens in mice**. The steps in mouse breeding for a tissue-specific cancer-gene screen are depicted. The steps include a mouse with a predisposing mutation ($Onc^{+/-}$). A double transgenic mouse (top left) is generated to contain a concatemer of T2/Onc transposons (20-250 in a single locus) and an SB transposase expression cassette that contains a blocking *lox-stop-lox* signal (gold star) between the promoter (green arrowhead) and the SB transpose gene (red box). The blocking *lox-stop-lox* signal permits tissue-specific expression of SB when the mice are bred to another mouse line that harbors a *Cre Recombinase* transgene driven by a tissue-specific promoter. A second double transgenic mouse (top right) is generated to contain the tissue-specific *Cre* transgene and another mutant tumor suppressor or oncogene. The two double-transgenic mice are mated resulting in offspring containing various combinations of these genes. Animals are genotyped and offspring with the four possible different genotypes (four middle boxes) are aged and monitored for tumor growth. Transposition in somatic cells causes random insertion mutations in mice that have an active transposase expression cassette and a T2/Onc concatemer (red and blue boxed mice on the left); owing to the presence of the $Onc^{+/-}$ predisposition (red box) there should be more oncogenic events than without it (blue box). Typical survival curves are shown at the bottom; generally there is no significant difference between control mice without any part of the SB system and mice that just have either a T2/Onc concatemer or SB transposase gene.

Fig. 7. **Analysis of T2/Onc-induced solid tumors and leukemias.** Cells from tumors (colored symbols in the top right mouse) caused by T/Onc transposition are collected, the genomes are

isolated and the transposon integration sites (colored arrows on the green lines representing the genome) are identified and mapped to chromosomal sites (lower left corner; star color relates to the transposons shown in the lower right corner). Genes that are repeatedly mutated in multiple, independent tumors are designated as *common insertion sites* or CIS. CIS can be analyzed to determine what genes and genetic pathways contribute to cancer.

Fig. 8. **Design and manufacture of CAR$^+$ T cells for non-viral gene therapy. A**) Structure of a prototypical chimeric antigen receptor (CAR) used at MD Anderson Cancer Center. A 2nd generation CAR is shown as a homodimer. Antigen recognition independent of human leukocyte antigen is achieved by scFv region derived from a CD19-specific mouse monoclonal antibody. The scFv is fused in frame to a modified hinge and Fc region derived from human IgG4. These extracellular structures are bound to the cell surface via chimeric CD28 and CD3-zeta, which activate T cells upon docking with CD19. **B)** Blood from either a recipient or a healthy donor is collected by venipuncture. Supercoiled DNA plasmids from *Sleeping Beauty* system coding for transposon (CAR) and transposase (SB11) are synchronously electroporated into mononuclear cells using a commercial electroporator. The electroporation conditions and transposon to transposase ratio are calibrated to deliver between one to two CAR integration events per T-cell genome. Intranuclear and intracellular events that follow are illustrated in Fig. 3. T cells that have stably integrated the CD19-specific CAR are retrieved by co-culture with irradiated artificial antigen presenting cells (aAPC) derived from K562 cells. The K562 cells function as aAPC as they are genetically modified to co-express CD19 and co-stimulatory molecules (CD86, CD137L, and a membrane-bound variant of IL-15). The CD19-specific CAR$^+$ T cells are numerically expanded before cryopreservation, release (and in-process) testing, and infusion into a recipient.

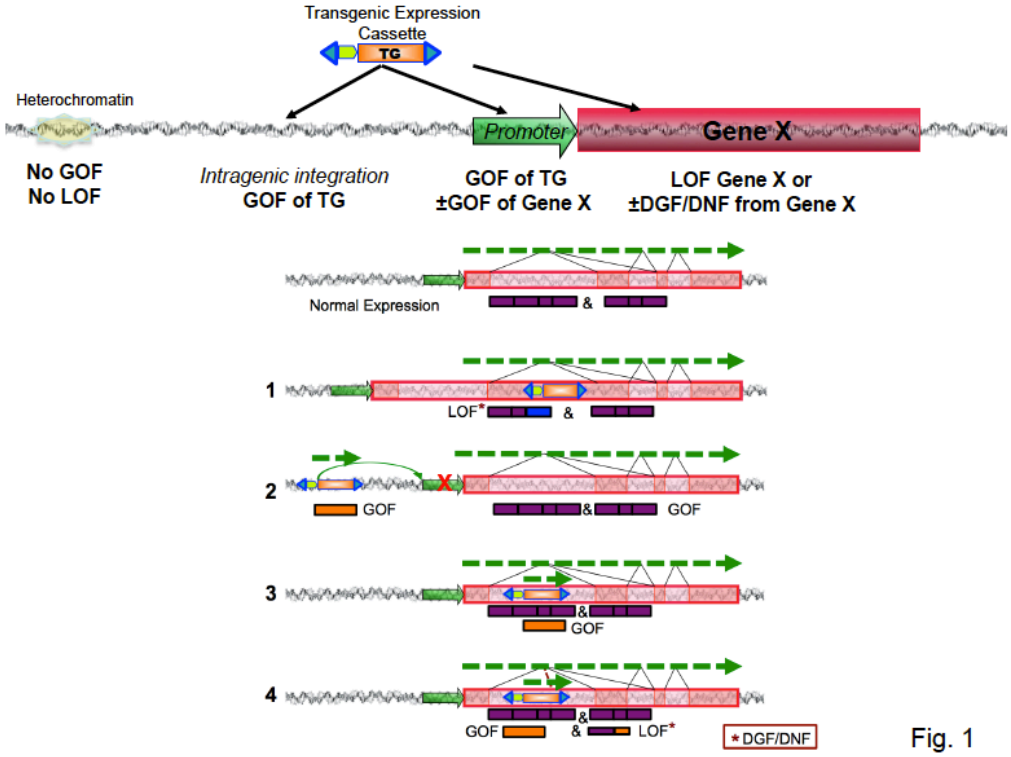
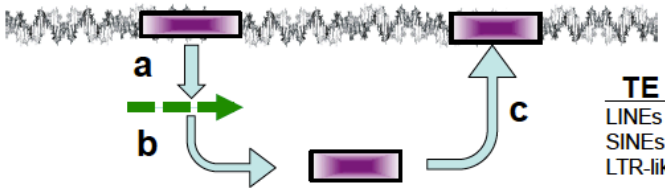


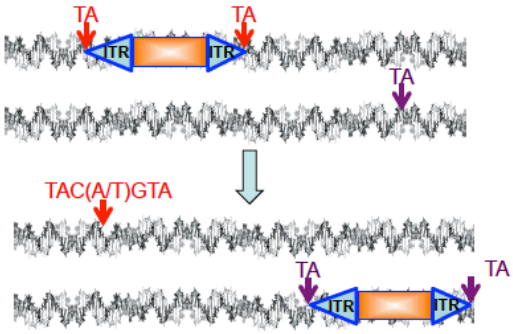
Fig. 1

Class I – Retro-transposable Elements (*copy-and- past*)



TE	number/genome
LINEs	850,000 (10-100)
SINEs	1,500,000 (source ?)
LTR-like	450,000 (0 ?)

Class II – DNA Transposable Elements (*cut-and-past*)



TE	number
DNA fossils	300,000 (0)

Fig. 2

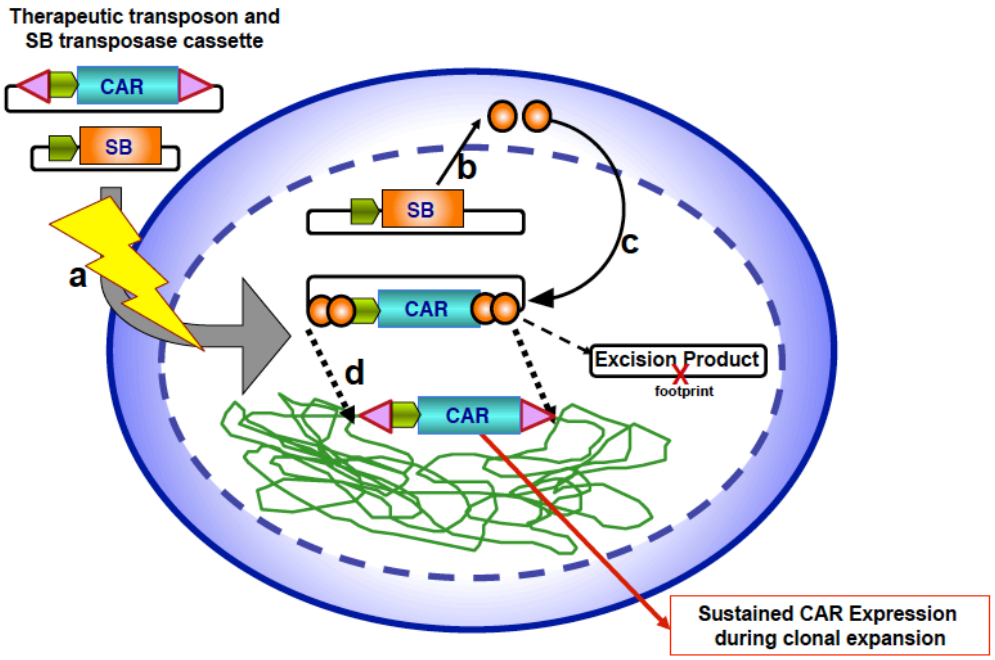

Therapeutic transposon and
SB transposase cassette
CAR
SB
a
b
SB
c
CAR
Excision Product
footprint
d
CAR
Sustained CAR Expression
during clonal expansion

Fig. 3

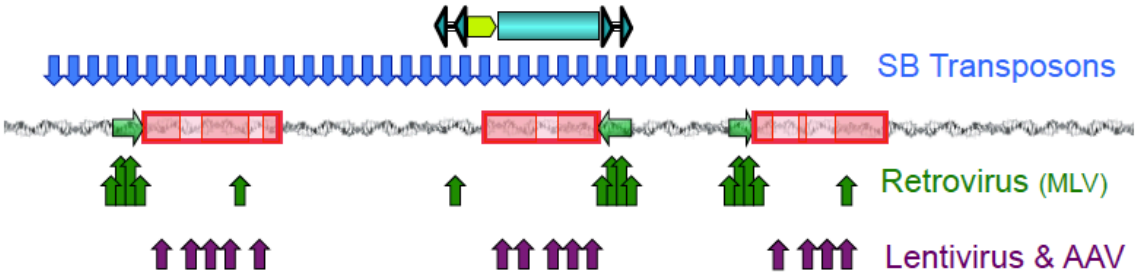

SB Transposons
Retrovirus (MLV)
Lentivirus & AAV

Fig. 4

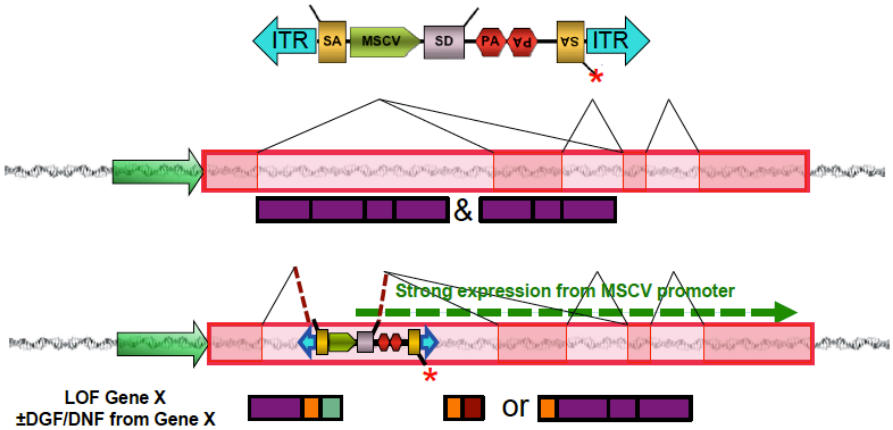

ITR
SA
MSCV
SD
PA
PA
SA
ITR
&
Strong expression from MSCV promoter
LOF Gene X
±DGF/DNF from Gene X
or

Fig. 5

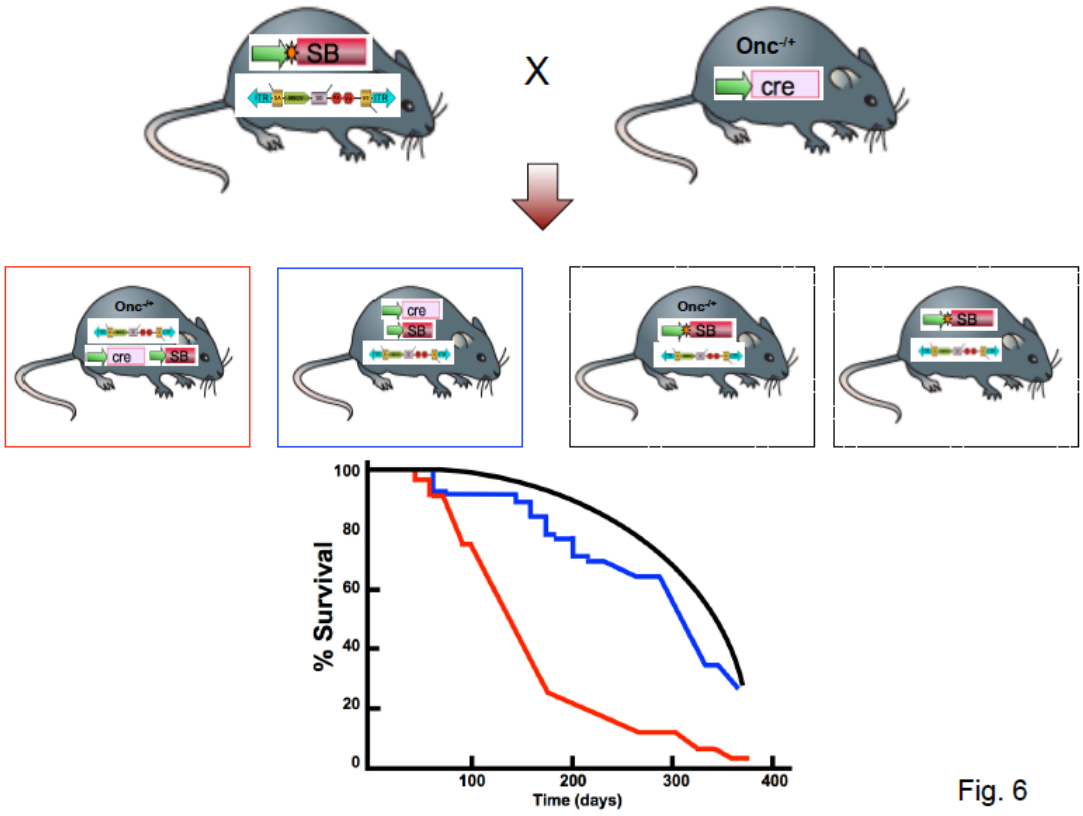

SB
X
Onc-/+
cre
Onc-/+
cre
SB
cre
SB
Onc-/+
SB
SB
% Survival
100
80
60
40
20
0
100
200
300
400
Time (days)

Fig. 6

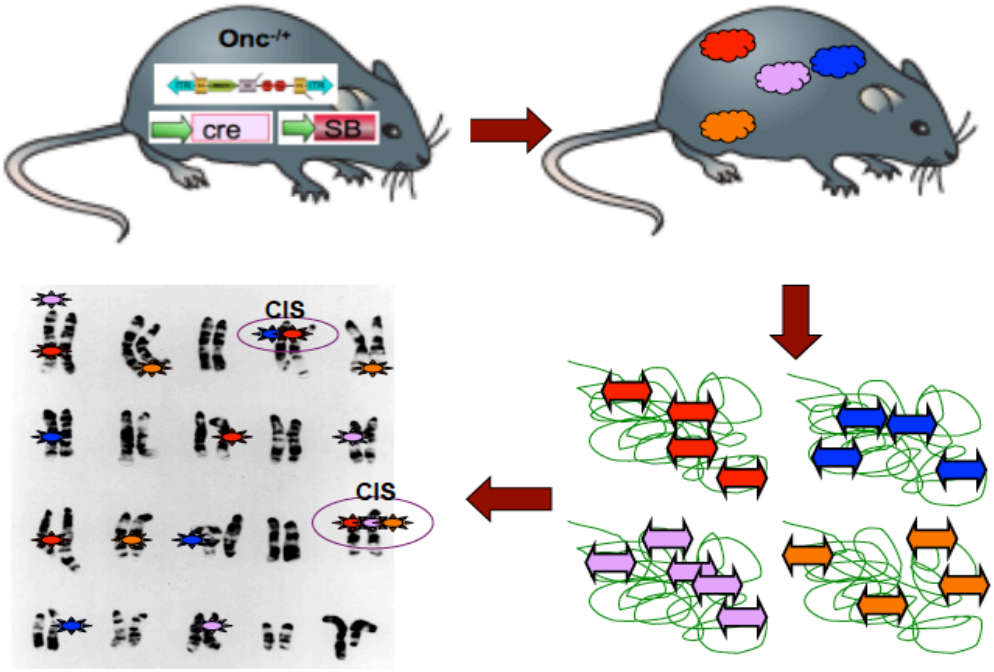

Onc-/+
cre
SB
CIS
CIS

Fig. 7

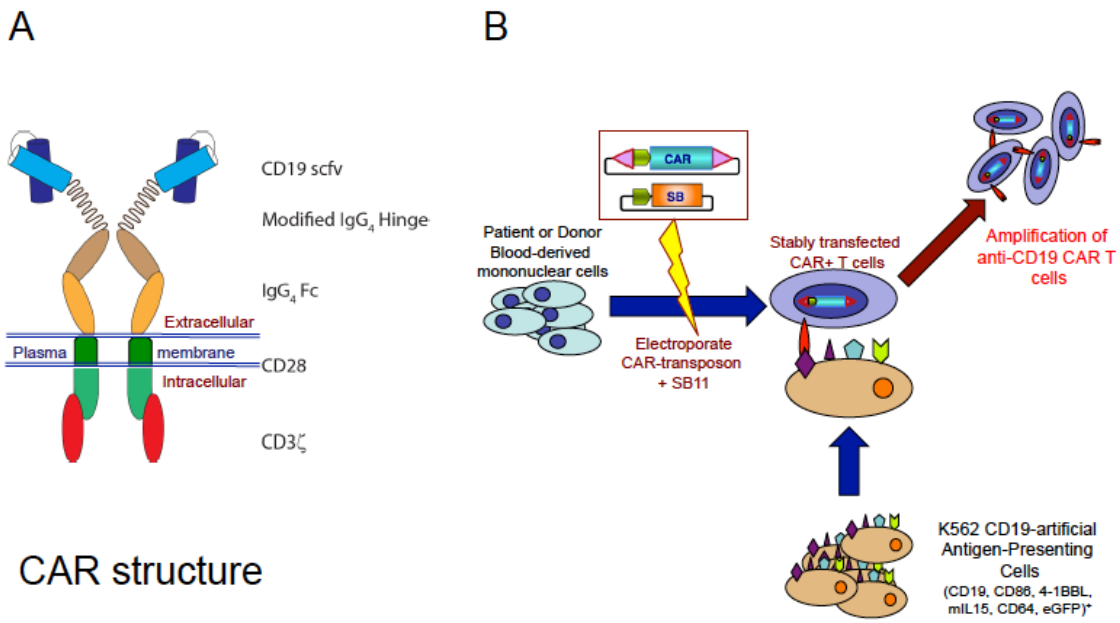

A
B
CD19 scfv
Modified IgG4 Hinge
IgG4 Fc
Extracellular
Plasma membrane
Intracellular
CD28
CD3ζ
CAR structure
CAR
SB
Patient or Donor Blood-derived mononuclear cells
Electroporate CAR-transposon + SB11
Stably transfected CAR+ T cells
Amplification of anti-CD19 CAR T cells
K562 CD19-artificial Antigen-Presenting Cells
(CD19, CD86, 4-1BBL, mIL15, CD64, eGFP)*

Fig. 8