Received: 15 February 2021 | Revised: 21 March 2021 | Accepted: 23 March 2021

DOI: 10.1111/cas.14901

REVIEW ARTICLE

Cancer Science WILEY

Identification of cancer driver genes using *Sleeping Beauty* transposon mutagenesis

Haruna Takeda[1] | Nancy A. Jenkins[1] | Neal G. Copeland[2]

[1]Laboratory of Molecular Genetics, National Cancer Center Research Institute, Tokyo, Japan

[2]Genetics Department, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

Correspondence
Haruna Takeda, Laboratory of Molecular Genetics, National Cancer Center Research Institute, 5-1-1, Tsukiji, Chuo-ku, Tokyo, 104-0045, Japan.
Email: hartaked@ncc.go.jp

Funding information
Grant-in-Aid for Scientific Research, Grant/Award Number: 20H03522; MSD Foundation

Abstract

Cancer genome sequencing studies have identified driver genes for a variety of different cancers and helped to understand the genetic landscape of human cancer. It is still challenging, however, to identify cancer driver genes with confidence simply from genetic data alone. In vivo forward genetic screens using *Sleeping Beauty* (*SB*) transposon mutagenesis provides another powerful genetic tool for identifying candidate cancer driver genes in wild-type and sensitized mouse tumors. By comparing cancer driver genes identified in human and mouse tumors, cancer driver genes can be identified with additional confidence based upon comparative oncogenomics. This review describes how *SB* mutagenesis works in mice and focuses on studies that have identified cancer driver genes in the mouse gastrointestinal tract.

KEYWORDS

cancer driver genes, colorectal cancer, mouse model, screening, *Sleeping Beauty* transposon mutagenesis

1 | INTRODUCTION

Whole-genome sequencing, exome sequencing, transcriptomic analyses and copy number analyses[1,2] have helped to elucidate the genomic landscape of human cancer and piece together the order of genetic events that drive cancer development. However, it is still challenging to identify genes that promote cancer progression, metastasis, and therapy resistance from genomic data alone. To truly understand how cancer develops, new approaches are needed. One of these approaches makes use of comparative oncogenomics; in vivo forward genetic screens in mice using *Sleeping Beauty* (SB) transposon mutagenesis has been used to identify candidate cancer driver genes (CCGs).[3,4] These CCGs were then compared with the CCGs identified by genomic analyses in human cancer. Genes that are commonly mutated in mouse and human are likely to be cancer driver genes as their functions in cancer development are conserved between species. In addition, *SB* mutagenesis can identify human cancer genes that were missed by genomic analyses because, for example, they are deregulated in human cancer by mutations in intergenic regulatory domains, by epigenetic mechanisms or by transcriptional deregulation (Figure 1).

For decades, in vivo insertional mutagenesis in the mouse has relied on retroviral mutagenesis, which proved to be a potent genetic tool for identifying CCGs in the hematopoietic system and mammary glands[3,5,6]; tissues in which retroviruses could be easily transduced. However, due to the nature of retroviral integration, there has been a strong bias for identifying oncogenes, and retroviral long terminal repeats often have shown strong promoter activity, which could result in the wrong annotation of CCGs.

More recently, in vivo transposon insertional mutagenesis has been used to identify CCGs for a variety of cancer types. This approach is similar in concept to the technique used in organisms such as *Drosophila melanogaster*[7]; however, applying this technique to mice has been hampered by the lack of transposons that actively transpose in mammalian cells. In 2005, 2 papers published in *Nature* showed that *Sleeping Beauty* (*SB*), a transposon originally isolated

from salmonoid fish, could transpose in mouse somatic tissues at high enough frequencies to induce cancer, and therefore could be used as a genetic tool for identifying CCGs.[8,9] Since then, *SB* transposon mutagenesis screens have been performed in more than 30 different organ systems, including pancreas,[10,11] skin,[12] liver,[13-15] breast,[16-18] brain[19,20] and the gastrointestinal tract,[21-26] and hundreds, if not thousands, of CCGs identified.

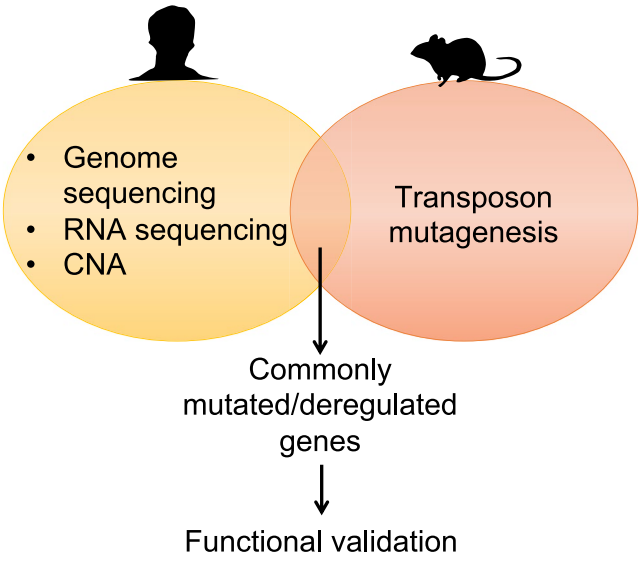

FIGURE 1 The approach to identify novel CRC driver genes by comparative oncogenomics. CNA, copy number alterations

In addition to *SB*, *PiggyBac* (*PB*) transposons, originally isolated from moths, have also been used for insertional mutagenesis screens in mice.[27] These 2 transposons have different integration site preferences. *SB* preferentially integrates into TA sites, whereas *PB* integrates into TTAA sites. *SB* integration events are little affected by the chromatin state, whereas *PB* integrations are affected by chromatin structure.[28] Also, *SB* transposons leave a 5-bp footprint behind following transposition, which may induce mutations if this occurs in exons or splice sites, whereas *PB* transposons leave no footprints behind following transposition.[3] Although each of these transposon systems has advantages and disadvantages, this review will focus on studies performed using *SB* mutagenesis, primarily in the gastrointestinal (GI) tract.

2 | ENGINEERING THE *SB* TRANSPOSON FOR MUTAGENESIS SCREENS IN MICE

There are basically 2 types of transposons that show different modes of transposition. One is typified by the LINE1 retrotransposon, in which transposition is mediated through an RNA intermediate that is reverse transcribed into a DNA copy and then re-inserted into the genome in a "copy"-and-paste manner (Figure 2A). The other

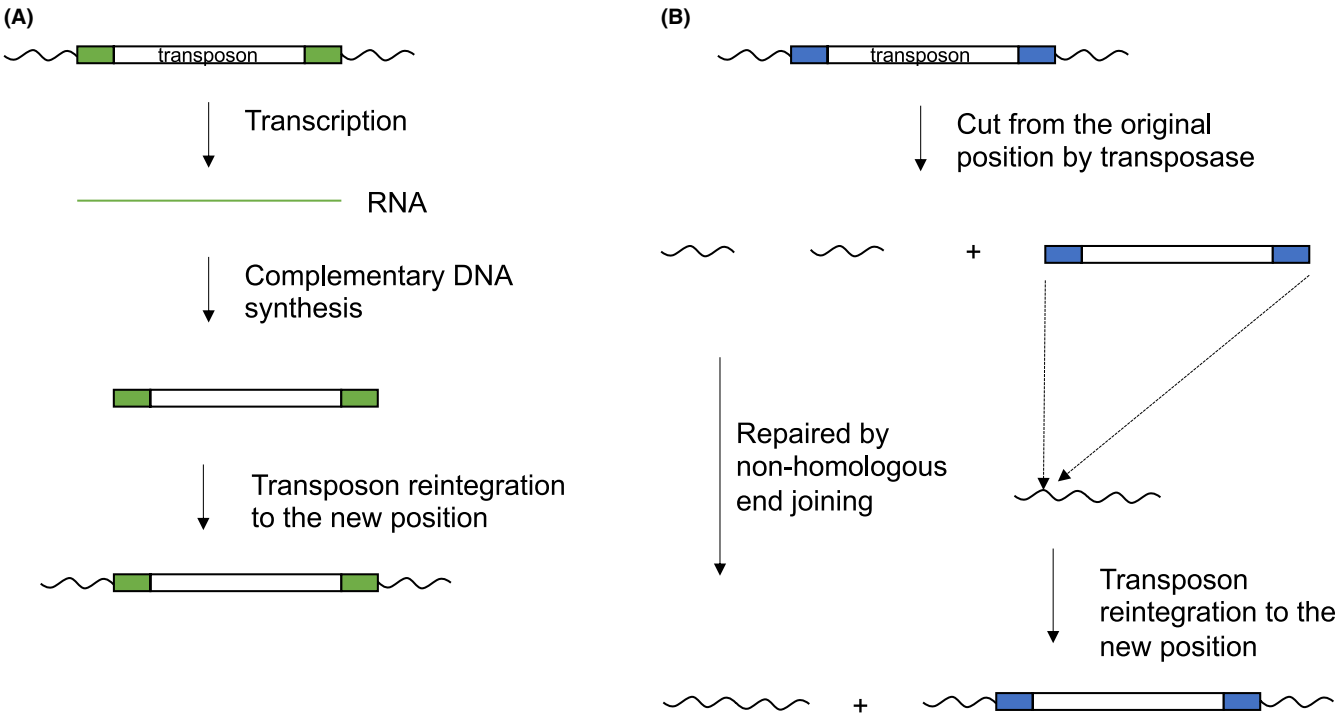

FIGURE 2 There are 2 types of transposons. A, Retrotransposons are transcribed into RNA, which then serves as the template for complementary DNA synthesis. A newly synthesized retrotransposon cDNA is then integrated into a new position in the genome. B, DNA transposons are excised from the genome by transposase, which makes double-strand breaks at each end of the transposon. The original insertion site is then repaired by non-homologous end joining leaving behind a 5 base-pair footprint. The excised transposon is subsequently re-inserted into a new TA site in the genome. The TA site is duplicated during insertion and now flanks each end of the transposon. Green or blue boxes indicate transposon-specific repeat sequences that are recognized by the transposase

is the DNA transposon, which is excised from the genome through double-strand breaks at each end of the transposon and then re-inserted into a new position in the genome in a "cut"-and-paste manner[7,29] (Figure 2B). The *SB* transposon belongs to the latter type of transposon. At the time of its isolation from fish, *SB* was inactive due to the accumulation of a series of mutations that occurred over millions of years of evolution, which prevented its transposition and the unintended consequences of deleterious mutations in the fish genome. To reactivate the *SB* transposon, a synthetic transposon was generated through a series of amino acid substitutions by reverse genetic re-engineering. The resultant transposon, called SB10, was subsequently shown to be able to transpose in mouse and human somatic cells by a cut-and-paste mechanism at a high frequency.[30]

To develop an in vivo *SB* transposon mutagenesis system for cancer studies, the 2 components of the transposition system, the DNA transposon and the transposase that catalyzes its transposition, were engineered separately. The DNA transposon carries inverted repeats (IRs) at both ends. These are described as inverted repeat, direct repeat, left end (IRDRL) and inverted repeat, direct repeat, right end (IRDRR) in Figure 2A. Between the IRs, one can put any sequences of interest, such as cDNAs for genes of interest, or genetic elements to deregulate cancer genes following integration. The transposon used in cancer screens is engineered to be a dual-function oncogenic transposon, which carries several genetic elements between the IRs. These include promoter sequences, a splice donor site (SD), 2 splice acceptor sites (SA) and a bidirectional polyA site (Figure 3A). When the transposon is inserted upstream or in the 5' end of an oncogene, in the same transcriptional orientation, it can drive overexpression of the oncogene through promoter insertion, where the promoter and downstream SA site fuses to the oncogene to drive its overexpression (Figure 3B). Conversely, when the transposon inserts into a tumor suppressor gene, in either orientation, the transposon can act as a gene trap by splicing into the SA site and prematurely terminating expression of the tumor suppressor gene transcript using the polyA site carried by the transposon (Figure 3C). Therefore, the *SB* transposon can activate the expression of oncogenes and also inactivate the expression of tumor suppressor genes. This is one of the great advantages of transposon mutagenesis, as other genetic screening systems involving retroviruses, CRISPR-Cas9, siRNA, or shRNA are biased for identification of oncogenes or tumor suppressor genes, but not both.

Since *SB* is a DNA transposon, which transposes through a cut-and-paste manner, many copies of the transposon are needed to generate enough cooperating mutations in tumor cells to induce a tumor. Therefore, several lines of transposon transgenic mice have been created, some carrying hundreds of copies of

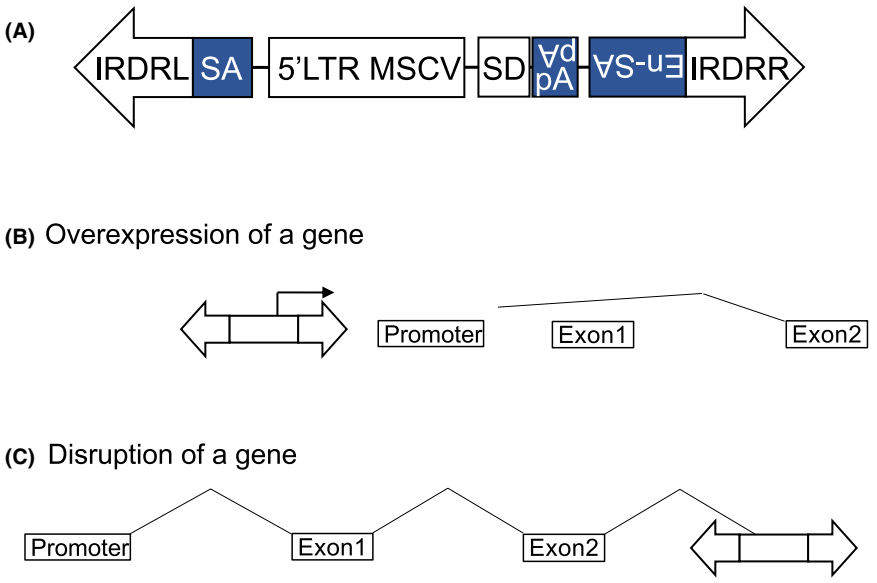

FIGURE 3 The structure of the T2/Onc2 transposon used in *SB* screening. A, The T2/Onc2 transposon has unique inverted repeat sequences (IRDRL and IRDRR) at each end that are recognized by the *SB* transposase. Between the IRDRL and IRDRR sequences, additional sequences have been inserted, including the 5′ long terminal repeat (LTR) of the murine stem cell virus (MSCV) that contains strong promoter and enhancer elements, a splice donor (SD) site derived from exon 1 of the mouse *Foxf2* gene, and bidirectional polyA (pA) sites that serve a transcriptional termination sites in cases in which the transposon inserts into the coding region of a tumor suppressor gene. There are 2 splice acceptor sites in the transposon located in inverse orientations, 1 from the carp β-actin gene (SA) and the other from the mouse *Engrailed2* gene (En-SA). B, When T2/Onc2 inserts into the promoter region or intron 1 of an oncogene, in the same transcriptional orientation, it can induce overexpression of the oncogene by the MSCV LTR promoter and downstream SD site. Oncogenes are recognized in *SB* screens because the transposons in different tumors are located in the 5′ end of the gene in the same transcriptional orientation. C, When T2/Onc2 inserts into the coding region of a tumor suppressor gene, in either orientation, the tumor suppressor transcript can be prematurely truncated by splicing into a SA site in the transposon followed by transcriptional termination at the pA site. Candidate tumor suppressor genes are identified in *SB* screens as genes in which the transposons in different tumors are located across the coding region, in both forward and reverse transcriptional orientations. Figures modified from the review[3]

TABLE 1 Genetically engineered mice used for *SB* screens in the intestinal tract

Study	Cre line	Transposon	Transposase	Sensitizing alleles
Starr et al, 2009[21]	Villin-Cre Tg[36]	T2/Onc2 Tg[9]	Rosa26-lsl-SB11 KI[21]	-
Starr et al, 2011[22]	Villin-Cre Tg[36]	T2/Onc2 Tg[9]	Rosa26-lsl-SB11 KI[21]	Apc Min/+[40]
March et al, 2011[23]	Ah-Cre Tg[58]	T2/Onc Tg[9]	Rosa26Lox66SBLox71 KI[23]	Apc flox/+,[59] Apc Min/+[40]
Takeda et al, 2015[24]	Villin-CreER^{T2} Tg[36]	T2/Onc2 Tg[8]	Rosa26-lsl-SB11 KI[21]	Apc Min/+,[40] Smad4KO/+,[44] Kras-lsl-G12D/+,[60] p53-lsl-R172H/+[46]
Morris et al, 2016[61]	Villin-Cre Tg[36]	T2/Onc2 Tg[8]	Rosa26-lsl-SB11 KI[21]	Tgfbr2 flox/flox[62]

Abbreviations: KI, knock-in; Tg, transgenic.

the transposon, all linked together at a single site in the genome (termed the transposon concatamer). Interestingly, when different promoter sequences were included in the transposon, different tumor spectrums were observed. The T2/Onc2 *SB* transposon, which carries the 5′ long terminal repeat (LTR) of murine stem cell virus (MSCV) as the promoter, induced primarily hematopoietic tumors,[8] whereas the T2/Onc3 *SB* transposon, which carries the ubiquitous CAG promoter, induced primarily solid tumors.[31] Therefore, it is necessary to use the right *SB* transposon transgenic lines for different cancer models. *SB* transposons integrate at TA sites. There are approximately 350 million TA sites distributed across the mouse genome. *SB* is therefore thought to be essentially a random insertional mutagen.

The other component of the *SB* mutagenesis system is the transposase. The *SB* transposase, which was originally carried in the transposon, is a protein that recognizes the IRs at both ends of the transposon and mediates its transposition. A couple of lines of mice expressing SB10 or a genetically enhanced SB11 transposase protein, have been generated.[8,9] One of the most frequently used lines is the conditional SB11 knock-in mouse in which the cDNA for SB11 together with upstream lox-STOP-lox sequences were knocked-in to the ubiquitously expressed and dispensable mouse *Rosa26* locus.[21] Activation of *SB* transposase, and therefore the *SB* mutagenic system, can therefore be spatiotemporally controlled in these mice using tamoxifen-inducible tissue-specific Cre recombinases.

3 | IDENTIFICATION OF CANDIDATE COLORECTAL CANCER (CRC) DRIVER GENES

Colorectal cancer is the second leading cause of cancer deaths worldwide. Cancer genome sequencing has identified numerous genes mutated in CRC.[32-35] So far, 5 studies have reported *SB* screens for the GI tract (Table 1).[21-25] To activate *SB* transposition specifically in the GI tract, most studies used mouse lines carrying Cre recombinase driven by the *villin* promoter, which showed an efficient transcriptional activity in all cell types of small intestinal epithelial cells, including stem cells, but showed limited activity in the colon.[36] One study used Ah-Cre transgenic mice in which Cre could be activated by the administration of β-naphthoflavone in several tissues, including the intestine.[23]

The first landmark paper used T2/Onc2 transposon transgenic mice and conditional SB11 knock-in mice to induce mutagenesis in intestinal epithelial cells. The triple compound mice carrying T2/Onc2, SB11, and Villin-Cre survived longer than 1.5 y and developed intestinal tumors with long latency. Analysis of 135 tumors identified 77 CCGs genes. CCGs are identified in these screens as genes that are insertionally mutated by *SB* at a higher frequency than predicted by random chance.[24] *Apc* was the most commonly mutated CCG identified in the screen. This was not surprising as *Apc* (*Adenomatous polyposis coli*) is the most important gatekeeper gene in the small and large intestines, as shown by genetic studies.[37,38] Mutations in *APC* are also found in nearly 80% of sporadic human CRC[33] and it is known that patients with familial adenomatous polyposis (FAP), a genetic syndrome caused by germline mutations in *APC*, are predisposed for colorectal cancers.[39] These results showed that *SB*-induced tumors faithfully recapitulated human CRC, and that *SB* screens are a powerful tool for the identification of CRC driver genes.

To identify genes promoting malignant tumor progression, 1 mutant allele of *Apc* (*ApcMin/+*[37,40] to model FAP or a conditional *Apc* knockout to model sporadic CRC) was also incorporated into the screen, resulting in the acceleration of tumor development.[22,23] Large-scale analysis of 446 tumors from these mice identified 867 CCGs in which *Apc* was the most highly mutated gene. Presumably, *SB* integrated into the 1 wild-type *Apc* allele present in these tumors to induce bi-allelic inactivation of *Apc* and accelerate tumor development. In addition, 183 out of 867 CCGs were candidate Wnt target genes. This result was very telling as *Apc* is a tumor suppressor gene that regulates Wnt signaling. These data showed that bi-allelic inactivation of *Apc* was induced in tumor cells and, in this context, genes enhancing Wnt signaling were preferentially selected to promote tumor progression.[23] However, introduction of an *Apc* mutant allele did not accelerate tumor histopathology, probably because these mice died early due to increased tumor multiplicity.[21,23]

4 | IDENTIFICATION OF CCGS IN DIFFERENT SENSITIZED MOUSE BACKGROUNDS

Vogelstein's genetic model shows that CRC progression is achieved by the multi-step acquisition of mutations, which is initiated by two-hit loss-of-function mutations in *APC*, followed by an activating

mutation in *KRAS*, loss-of-function mutations in *SMAD4*, which activates TGFβ signaling, and inactivating mutations in *TP53*.[39,41] Mouse mutant alleles corresponding to these mutations found in human CRC have been generated. The *KrasLSL-G12D/+* allele carries an activating point mutation that results in amino acid substitution of Gly to Asp at codon 12 in *Kras*. This mutation is one of the common mutations observed in human CRC.[42] *KrasLSL-G12D/+*: *Villin-Cre* mice, which expressed KrasG12D in the intestine, showed hyperplasia throughout the colonic epithelium, as illustrated by an extreme lengthening of the crypts and the development of large, prominent goblet cells. However, these mice did not develop any colonic tumors.[43] *Smad4* heterozygous knockout mice developed only a few benign tumors in the duodenum,[44] whereas compound mutant mice carrying mutations in *Apc* and *Smad4* developed numerous adenocarcinomas in the mouse intestine.[45] Although *TP53* mutations are observed in nearly half of late-stage human CRC, *Trp53LSL-R172H* knock-in mice seldom developed intestinal tumors.[46] These results showed that a single sensitizing mutation could not induce efficient intestinal tumor development and that cooperating mutations are required for tumor development. These 4 different sensitizing mutations (*ApcMin/+*, *KrasLSL-G12D/+*, *Smad4KO/+*, and *Trp53R172H/+*) were introduced in *SB* mutagenesis screens in the mouse intestine,[24] to provide a comprehensive understanding of genes and evolutionary forces promoting CRC progression.

Quadruple mutant mice carrying T2/Onc2, SB11, *Villin-CreERT2*, and 1 of 4 sensitizing mutations, were then generated by 2 rounds of mouse crosses (Figure 4), and the mice aged until they showed signs of anemia. Mice undergoing *SB* mutagenesis in the intestine showed shortened survivals and developed increased numbers of tumors compared with their non-*SB* controls, indicating that insertional mutations induced by *SB* cooperated specifically with each sensitizing mutation. Analysis of 65 KrasG12D:SB tumors, 100 Smad4KO:SB tumors and 55 p53R172H:SB tumors, identified 338, 449, and 500 CCGs, respectively. These CCGs were combined and added to the datasets obtained from previous large-scale *SB* mutagenesis screens on the *Apc* mutant background, which provided 1333 unique mouse CCGs. Comparison of these mouse CCGs to the CCGs identified in human CRC identified CCGs that were commonly mutated in human and mouse tumors, including known human CRC driver genes such as *Apc*, *Smad4*, *Pten*, *Trp53*, *Fbxw7*, and *Tgfbr2* as well as genes whose functions in human CRC are unknown. Interestingly, nearly half of all tumors were adenocarcinomas. Subsequent classification of tumors, depending on the malignancy grade, and analysis of corelated gene mutations, identified candidate CCGs that are likely to be involved in malignant tumor progression. These data showed that *SB* mutagenesis can also identify genes that function at different stages of tumor development.[24]

In addition, comparison of CCGs with the dataset of genes that were epigenetically silenced in CRC enriched 10 genes. These genes are likely to function as cancer drivers. Furthermore, SB mutagenesis screens identified several common integration sites in which CCGs were not annotated. These genomic loci may function as cancer-specific regulatory elements for gene expression, therefore it will be interesting to compare these with the loci identified from Assay for Transposase-Accessible Chromatin (ATAC)-seq, which identified the active chromatin regions, to enrich functional intergenic regions in the development of cancer.

5 | UNDERSTANDING CANCER GENOME EVOLUTION IN THE INTESTINE

Interestingly, the most frequently mutated gene in ApcMin:SB, KrasG12D:SB and p53R172H:SB tumors was *Apc*. This finding is consistent with APC's role as a gate-keeping gene in human CRC. Surprisingly, *Apc* was not the most commonly mutated gene in

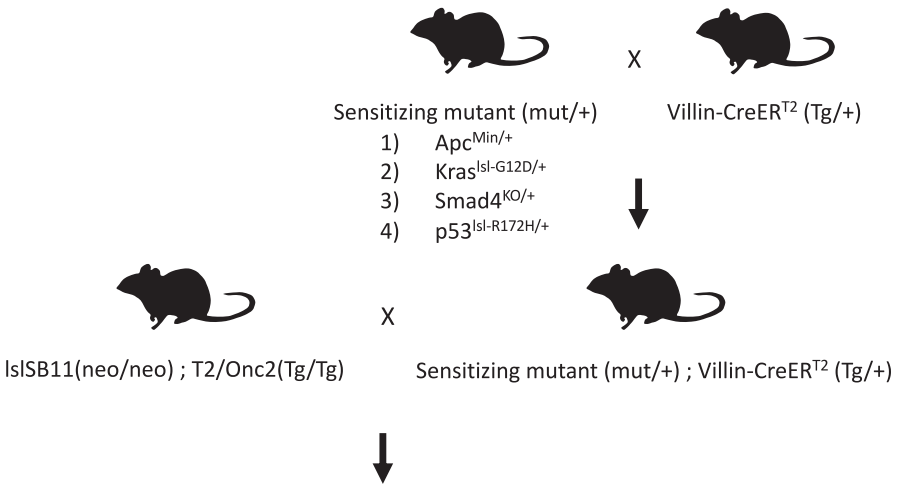

FIGURE 4 The mating strategy used to generate 4 different quadruple compound mice. Heterozygous mice for each sensitizing mutation were crossed with *Villin-CreERT2* transgenic mice to generate double compound mutant mice, which were subsequently crossed with mice homozygous for the *SB* transposase (lslSB11) and the T2/Onc2 transposon concatamer. Note that the T2/Onc2 transposon concatamer contains hundreds of copies of the T2/Onc2 transposon, all linked together at a single site in the genome. The *ERT2* cassette makes it possible to control the activity of *Cre* recombinase by tamoxifen addition[24]

Smad4KO:SB tumors. In Smad4KO:SB tumors, the most frequently mutated gene was *Smad4* itself. Transposons were specifically inserted in the 1 remaining wild-type *Smad4* allele present in these mice to induce bi-allelic inactivation of *Smad4* in tumor cells. Insertions in *Rspo1* and *Rspo2* were also enriched in Smad4KO:SB tumors. R-spondins are secreted proteins that bind to Lgr4/5/6 cell-surface receptors to enhance Wnt signaling.[47] R-spondins are necessary to maintain intestinal stem cells and are therefore one of the essential components for intestinal organoid culture.[48] In Smad4KO:SB tumors, insertions in *Rspo1* and *Rspo2* were primarily located in 5′ promoter regions, in the same transcriptional orientation, and resulted in overexpression of the *Rspo1/2* gene. This finding is consistent with evidence that Wnt signaling activation is observed in nearly all the CRC. It still remains unknown why R-spondins were selectively overexpressed in *Smad4*-deficient tumors, but some studies have shown that R-spondins induce apoptosis through activation of TGFβ signaling.[49,50] Inhibition of TGFβ signaling by bi-allelic *Smad4* inactivation in the environment of abundant R-spondins may be a strategy for tumor cells to escape apoptosis. However, *SB* mutagenesis on the homozygous *Tgfbr2* knockout background (a gene encoding the type II TGFβ receptor) did not enrich for *SB* insertions in *Rspo1* or *Rspo2*.[25] Therefore, there may be a specific cooperation between overexpression of R-spondins and loss of *Smad4* in intestinal tumor development.

Collectively, these data showed that a pre-existing mutation in a pre-malignant cancer cell can induce an intrinsic selection pressure that influences the nature of the mutation that is selected next. Whereas mutations occur randomly, their selection does not. In other words, the cancer-causing mutation that occurs first, influences the nature of the mutation that is selected next, which in turn, influences the mutation that is selected third, and so on down the line. To put this another way, the order in which mutations accumulate in tumor cells over time is not random. Mutations occur randomly but the selection of mutations does not. This helps to explain why, in human CRC, the mutations largely occur in a non-random order, with *APC* mutations occurring first, *KRAS* mutations second, followed by *SMAD4* and then *TP53* mutations.[51] These studies also demonstrated the power of *SB* mutagenesis for understanding how cancer genomes evolve during cancer progression.

6 | *SB* SCREENS IN THE STOMACH

Gastric cancer is one of the leading causes of cancer-related deaths worldwide. Approximately 10% of human gastric cancers contain loss-of-function mutations in *SMAD4*,[52] consistent with studies showing that *Smad4* knockout mice developed gastric tumors.[44] At the time when *SB* screening in the stomach was performed, a mouse line expressing Cre specifically in gastric epithelial cells was not available. Therefore, whole body *SB* mutagenesis on the *Smad4* knockout background was performed. Analysis of 66 gastric tumors identified 941 CCGs, which included genes involved in Wnt signaling, Akt signaling, TGFβ signaling or regulation of tight junctions. As expected, the most highly mutated gene was *Smad4*, which presumably resulted in bi-allelic inactivation of *Smad4*. Interestingly, no insertions in *Rspo1* or *Rspo2* were observed in this screen, which was in sharp contrast with the results obtained from *SB* screens in the intestine on the *Smad4* mutant background. Instead, mutations were identified in other upstream Wnt signaling components, such as *Lrp6*. These results showed that tissue type also matters for tumor evolution.

7 | CANCER ORIGINS

Based upon these results and combined with studies of others, we recently suggested that mutations of selected oncogenes or tumor suppressor genes function initially in different tissue-specific stem cells to clonally increase the number of cells in the tissue, producing a small begin tumor.[51] We further proposed that these initial or truncal mutations are themselves, tissue-specific, with each different type of stem cell responding to only a limited set of truncal mutations. A good example of this can be found in the type of cancers that occur in patients with Li-Fraumeni syndrome, who have germline mutations in *TP53*. In these patients, tumors derived from ectodermal and mesodermal tissues are observed at a high excess risk (250-500-fold) in childhood (1-15 y of age) and middle age (20-50 y of age), whereas tumors derived from endodermal tissues occur at lower excess risk (2-4-fold) and older ages (50-70 y of age).[53] This reason for this is thought to be that inherited *TP53* mutations are truncal mutations in ectodermal-derived and mesodermal-derived cancers but are progression mutations in endodermal derived cancers.

8 | CONCLUSIONS

Sleeping Beauty mutagenesis screens have been performed in many different mouse tissues, including the GI tract. Genes mutated in human and mouse are then identified and assumed to be cancer genes based upon their evolutionary conservation in cancer development. Other genes that are mutated in mouse tumors, but not in human tumors, could also represent new cancer genes that escaped detection in human tumors due to a variety of different reasons. In both cases, experimental validation will be necessary to confirm their oncogenic ability. One high-throughput method for validating CRC tumor suppressor genes makes use of CRISPR-Cas9-mediated gene knockouts in benign tumor-derived intestinal organoids.[54] In the end, the hope is that the identification of novel driver genes by *SB* mutagenesis will help to identify new therapeutic targets for the treatment of CRC.

Sleeping Beauty mutagenesis can also be applied to understanding the mechanisms for drug resistance. This is important as acquiring resistance to anti-cancer drugs is one of the major problems in the clinic. *BRAF* inhibitors such as PLX4720 exert anti-tumor responses in ~80% of *BRAF* mutated tumors, however, resistant tumors quickly

emerge. One study used *SB* mutagenesis to model melanoma in mice and obtain PLX4720-resistant tumors. Eight genes, including *Braf*, *Mitf* and *Eras* were significantly more mutated by *SB* in PLX4720-resistant tumors,[28] highlighting the power of SB mutagenesis for identifying genes involved in drug resistance. Another example is the identification of resistance mechanisms to p53-MDM2 inhibition by *PiggyBac* transposon mutagenesis. MDM2 degrades p53, therefore compounds inhibiting the interaction between p53 and MDM2 restore p53 function. However, tumors treated with p53-MDM2 inhibitors relapsed very rapidly. Interestingly, genes mutated by *SB* in these resistance tumors were significantly enriched for genes that function in the *p53* pathway, including *p53* itself and the *p53* family members *p63* and *p73*.[55] These studies may one day help us to predict the future clinical course of patients treated with these inhibitors.

Most CCGs identified in *SB* screens are usually identified in more than 1 screen and 1 type of cancer. Therefore, the comparison of CCGs identified in different cancer types can help in understanding the complexity of cancer. To share the data obtained from different SB screens, databases have been developed,[56,57] such as the Sleeping Beauty Cancer Driver DataBase (SBCDDB, http://sbcddb.moffitt.org) or The Candidate Cancer Gene Database (http://ccgd-starrlab.oit.umn.edu/index.html). By keying-in the name of your gene of interest, you can identify the type of cancers carrying *SB* mutations in your gene, including the frequency of mutations in your gene, the patterns of transposon insertions in your gene, and studies in which your gene was identified. The results are cumulative and, as more data become available, the value of *SB* mutagenesis for cancer studies becomes increasingly more powerful.

ACKNOWLEDGMENTS

This work was supported by Grant-in-Aid for Scientific Research (20H03522), and the MSD Foundation.

ORCID

Haruna Takeda https://orcid.org/0000-0002-1643-2524

REFERENCES

1. Gao J, Aksoy BA, Dogrusoz U, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal*. 2013;6:pl1.
2. Cerami E, Gao J, Dogrusoz U, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov*. 2012;2:401-404.
3. Copeland NG, Jenkins NA. Harnessing transposons for cancer gene discovery. *Nat Rev Cancer*. 2010;10:696-706.
4. Weber J, Braun CJ, Saur D, Rad R. In vivo functional screening for systems-level integrative cancer genomics. *Nat Rev Cancer*. 2020;20:573-593.
5. Akagi K, Suzuki T, Stephens RM, Jenkins NA, Copeland NG. RTCGD: retroviral tagged cancer gene database. *Nucleic Acids Res*. 2004;32:D523-D527.
6. Kool J, Berns A. High-throughput insertional mutagenesis screens in mice to identify oncogenic networks. *Nat Rev Cancer*. 2009;9:389-399.
7. Largaespada DA. Transposon mutagenesis in mice. *Methods Mol Biol*. 2009;530:379-390.
8. 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.
9. Collier LS, Carlson CM, Ravimohan S, Dupuy AJ, Largaespada DA. Cancer gene discovery in solid tumours using transposon-based somatic mutagenesis in the mouse. *Nature*. 2005;436:272-276.
10. Mann KM, Ward JM, Yew CC, et al. Sleeping Beauty mutagenesis reveals cooperating mutations and pathways in pancreatic adenocarcinoma. *Proc Natl Acad Sci USA*. 2012;109:5934-5941.
11. Perez-Mancera PA, Rust AG, van der Weyden L, et al. The deubiquitinase USP9X suppresses pancreatic ductal adenocarcinoma. *Nature*. 2012;486:266-270.
12. Mann MB, Black MA, Jones DJ, et al. Transposon mutagenesis identifies genetic drivers of Braf(V600E) melanoma. *Nat Genet*. 2015;47:486-495.
13. Keng VW, Villanueva A, Chiang DY, et al. A conditional transposon-based insertional mutagenesis screen for genes associated with mouse hepatocellular carcinoma. *Nat Biotechnol*. 2009;27:264-274.
14. Bard-Chapeau EA, Nguyen AT, Rust AG, et al. Transposon mutagenesis identifies genes driving hepatocellular carcinoma in a chronic hepatitis B mouse model. *Nat Genet*. 2014;46:24-32.
15. Kodama T, Bard-Chapeau EA, Newberg JY, et al. Two-step forward genetic screen in mice identifies Ral GTPase-activating proteins as suppressors of hepatocellular carcinoma. *Gastroenterology*. 2016;151(324–337):e312.
16. Rangel R, Lee SC, Hon-Kim Ban K, et al. Transposon mutagenesis identifies genes that cooperate with mutant Pten in breast cancer progression. *Proc Natl Acad Sci USA*. 2016;113:E7749-E7758.
17. Suárez-Cabrera C, Quintana RM, Bravo A, et al. A Transposon-based analysis reveals RASA1 is involved in triple-negative breast cancer. *Can Res*. 2017;77:1357-1368.
18. Kas SM, de Ruiter JR, Schipper K, et al. Insertional mutagenesis identifies drivers of a novel oncogenic pathway in invasive lobular breast carcinoma. *Nat Genet*. 2017;49:1219-1230.
19. Genovesi LA, Ng CG, Davis MJ, et al. Sleeping Beauty mutagenesis in a mouse medulloblastoma model defines networks that discriminate between human molecular subgroups. *Proc Natl Acad Sci USA*. 2013;110:E4325-4334.
20. Morrissy AS, Garzia L, Shih DJ, et al. Divergent clonal selection dominates medulloblastoma at recurrence. *Nature*. 2016;529:351-357.
21. Starr TK, Allaei R, Silverstein KA, et al. A transposon-based genetic screen in mice identifies genes altered in colorectal cancer. *Science*. 2009;323:1747-1750.
22. Starr TK, Scott PM, Marsh BM, et al. 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.
23. March HN, Rust AG, Wright NA, et al. Insertional mutagenesis identifies multiple networks of cooperating genes driving intestinal tumorigenesis. *Nat Genet*. 2011;43(12):1202-1209.
24. Takeda H, Wei Z, Koso H, et al. Transposon mutagenesis identifies genes and evolutionary forces driving gastrointestinal tract tumor progression. *Nat Genet*. 2015;47:142-150.
25. Kuwahara R, Tomita R, Ogawa N, et al. Crystallization and hardening of poly(ethylene-co-vinyl acetate) mouthguards during routine use. *Sci Rep*. 2017;7:44672.
26. Takeda H, Rust AG, Ward JM, Yew CC, Jenkins NA, Copeland NG. Sleeping Beauty transposon mutagenesis identifies genes that cooperate with mutant Smad4 in gastric cancer development. *Proc Natl Acad Sci USA*. 2016;113:E2057-E2065.
27. Rad R, Rad L, Wang W, et al. PiggyBac transposon mutagenesis: a tool for cancer gene discovery in mice. *Science*. 2010;330:1104-1107.
28. Yoshida J, Akagi K, Misawa R, Kokubu C, Takeda J, Horie K. Chromatin states shape insertion profiles of the piggyBac, Tol2 and

Sleeping Beauty transposons and murine leukemia virus. *Sci Rep*. 2017;7:43613.
29. Kemp JR, Longworth MS. Crossing the LINE toward genomic instability: LINE-1 retrotransposition in cancer. *Front Chem*. 2015;3:68.
30. 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;91:501-510.
31. Dupuy AJ, Rogers LM, Kim J, et al. A modified sleeping beauty transposon system that can be used to model a wide variety of human cancers in mice. *Can Res*. 2009;69:8150-8156.
32. Wood LD, Parsons DW, Jones S, et al. The genomic landscapes of human breast and colorectal cancers. *Science*. 2007;318:1108-1113.
33. Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature*. 2012;487:330-337.
34. Seshagiri S, Stawiski EW, Durinck S, et al. Recurrent R-spondin fusions in colon cancer. *Nature*. 2012;488(7413):660-664.
35. Giannakis M, Mu XJ, Shukla SA, et al. Genomic correlates of immune-cell infiltrates in colorectal carcinoma. *Cell Rep*. 2016;15:857-865.
36. el Marjou F, Janssen KP, Chang BH, et al. Tissue-specific and inducible Cre-mediated recombination in the gut epithelium. *genesis*. 2004;39:186-193.
37. Moser AR, Pitot HC, Dove WF. A dominant mutation that predisposes to multiple intestinal neoplasia in the mouse. *Science*. 1990;247:322-324.
38. Oshima M, Oshima H, Kitagawa K, Kobayashi M, Itakura C, Taketo M. Loss of Apc heterozygosity and abnormal tissue building in nascent intestinal polyps in mice carrying a truncated Apc gene. *Proc Natl Acad Sci USA*. 1995;92:4482-4486.
39. Kinzler KW, Vogelstein B. Lessons from hereditary colorectal cancer. *Cell*. 1996;87:159-170.
40. Su LK, Kinzler KW, Vogelstein B, et al. Multiple intestinal neoplasia caused by a mutation in the murine homolog of the APC gene. *Science*. 1992;256:668-670.
41. Pinto D, Clevers H. Wnt, stem cells and cancer in the intestine. *Biol Cell*. 2005;97:185-196.
42. Bos JL. ras oncogenes in human cancer: a review. *Can Res*. 1989;49:4682-4689.
43. Haigis KM, Kendall KR, Wang Y, et al. Differential effects of oncogenic K-Ras and N-Ras on proliferation, differentiation and tumor progression in the colon. *Nat Genet*. 2008;40:600-608.
44. Takaku K, Miyoshi H, Matsunaga A, Oshima M, Sasaki N, Taketo MM. Gastric and duodenal polyps in Smad4 (Dpc4) knockout mice. *Can Res*. 1999;59:6113-6117.
45. Takaku K, Oshima M, Miyoshi H, Matsui M, Seldin MF, Taketo MM. Intestinal tumorigenesis in compound mutant mice of both Dpc4 (Smad4) and Apc genes. *Cell*. 1998;92:645-656.
46. Olive KP, Tuveson DA, Ruhe ZC, et al. Mutant p53 gain of function in two mouse models of Li-Fraumeni syndrome. *Cell*. 2004;119:847-860.
47. de Lau W, Barker N, Low TY, et al. Lgr5 homologues associate with Wnt receptors and mediate R-spondin signalling. *Nature*. 2011;476:293-297.
48. Sato T, Vries RG, Snippert HJ, et al. Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature*. 2009;459:262-265.
49. Zhou X, Geng L, Wang D, Yi H, Talmon G, Wang J. R-Spondin1/LGR5 activates TGFbeta signaling and suppresses colon cancer metastasis. *Can Res*. 2017;77:6589-6602.
50. Lahde M, Heino S, Hogstrom J, et al. Expression of R-Spondin 1 in Apc(Min/+) mice suppresses growth of intestinal adenomas by altering Wnt and transforming growth factor beta signaling. *Gastroenterology*. 2021;160:245-259.
51. Levine AJ, Jenkins NA, Copeland NG. The roles of initiating truncal mutations in human cancers: the order of mutations and tumor cell type matters. *Cancer Cell*. 2019;35:10-15.
52. Cancer Genome Atlas Research Network. Comprehensive molecular characterization of gastric adenocarcinoma. *Nature*. 2014;513:202-209.
53. Levine AJ. p53: 800 million years of evolution and 40 years of discovery. *Nat Rev Cancer*. 2020;20:471-480.
54. Takeda H, Kataoka S, Nakayama M, et al. CRISPR-Cas9-mediated gene knockout in intestinal tumor organoids provides functional validation for colorectal cancer driver genes. *Proc Natl Acad Sci USA*. 2019;116:15635-15644.
55. Chapeau EA, Gembarska A, Durand EY, et al. Resistance mechanisms to TP53-MDM2 inhibition identified by in vivo piggyBac transposon mutagenesis screen in an Arf(-/-) mouse model. *Proc Natl Acad Sci USA*. 2017;114:3151-3156.
56. Newberg JY, Mann KM, Mann MB, Jenkins NA, Copeland NG. SBCDDB: Sleeping Beauty Cancer Driver Database for gene discovery in mouse models of human cancers. *Nucleic Acids Res*. 2018;46:D1011-d1017.
57. Abbott KL, Nyre ET, Abrahante J, Ho Y-Y, Isaksson Vogel R, Starr TK. The Candidate Cancer Gene Database: a database of cancer driver genes from forward genetic screens in mice. *Nucleic Acids Res*. 2014;43:D844-D848.
58. Ireland H, Kemp R, Houghton C, et al. Inducible Cre-mediated control of gene expression in the murine gastrointestinal tract: effect of loss of beta-catenin. *Gastroenterology*. 2004;126:1236-1246.
59. Shibata H, Toyama K, Shioya H, et al. Rapid colorectal adenoma formation initiated by conditional targeting of the Apc gene. *Science*. 1997;278:120-123.
60. Tuveson DA, Shaw AT, Willis NA, et al. Endogenous oncogenic K-ras(G12D) stimulates proliferation and widespread neoplastic and developmental defects. *Cancer Cell*. 2004;5:375-387.
61. Morris SM, Davison J, Carter KT, et al. Transposon mutagenesis identifies candidate genes that cooperate with loss of transforming growth factor-beta signaling in mouse intestinal neoplasms. *Int J Cancer*. 2017;140:853-863.
62. Chytil A, Magnuson MA, Wright CV, Moses HL. Conditional inactivation of the TGF-beta type II receptor using Cre:Lox. *Genesis*. 2002;32:73-75.

How to cite this article: Takeda H, Jenkins NA, Copeland NG. Identification of cancer driver genes using *Sleeping Beauty* transposon mutagenesis. *Cancer Sci*. 2021;112:2089–2096. https://doi.org/10.1111/cas.14901