Check for
updates

Sleeping Beauty Mouse Models of Cancer: Microenvironmental Influences on Cancer Genetics

Amy Guimaraes-Young, Charlotte R. Feddersen and Adam J. Dupuy*

Department of Anatomy and Cell Biology, Roy J. and Lucille A. Carver College of Medicine, University of Iowa, Iowa City, IA, United States

The Sleeping Beauty (SB) transposon insertional mutagenesis system offers a streamlined approach to identify genetic drivers of cancer. With a relatively random insertion profile, SB is uniquely positioned for conducting unbiased forward genetic screens. Indeed, SB mouse models of cancer have revealed insights into the genetics of tumorigenesis. In this review, we highlight experiments that have exploited the SB system to interrogate the genetics of cancer in distinct biological contexts. We also propose experimental designs that could further our understanding of the relationship between tumor microenvironment and tumor progression.

Keywords: sleeping beauty, insertional mutagenesis, mouse models of cancer, cancer, genetics

UTILIZATION OF SLEEPING BEAUTY TO MODEL TUMOR FORMATION IN MICE

Sleeping Beauty (SB) is a two-part DNA transposon system that has become an integral tool in identifying genetic drivers of cancers in mouse models. SB-induced cancer models have led to the discovery of novel cancer genes while comparative genomics have demonstrated the relevance of these models to human disease (1–8). In this review, we highlight characteristics of SB that make it uniquely equipped for unbiased *in vivo* forward genetic cancer screens and describe various ways in which SB cancer models are being used to address the impact of tumor microenvironment on cancer biology. Such approaches have provided insight into the multifaceted interplay between somatic mutations and other exposures that drive tumor progression. Collectively these experiments highlight the flexibility of SB mutagenesis in addressing complex genetic questions.

SB Origins and Optimization for Forward Genetic Screens

SB is derived from the Tc1/mariner superfamily of "cut-and-paste" transposable elements widely encoded across all animal kingdoms but silenced in vertebrates by the evolutionary accumulation of mutations. Site-directed mutagenesis of a consensus DNA sequence derived from fish species permitted synthesis of the active transposase enzyme and flanking recognition sequences mobilized by the enzyme (the transposon) (9). Functionality was found to be preserved in *trans* as a two-part system when co-transfection of two plasmids, one containing the transposase and the other containing a selectable cassette flanked by the transposase recognition sequences, resulted in successful integration of the selection cassette within chromosomal DNA (9). Indeed, the SB system demonstrated superior activity in mammalian cells relative to other transposons tested and, in 2001, several groups reported germline transmission and chromosomal transposition of SB in mice

harboring both the transposon and transposase in their genomes (i.e., double transgenic mice) (9–12). For use in forward genetic cancer screens, the SB system needed to be able to serve as an effective mutagen and achieve a mutational frequency sufficient to induce tumors.

The first SB transposons successfully used in insertional mutagenesis screens were T2Onc and T2Onc2 (13, 14). The constructs are similar: both contain the murine stem cell virus long terminal repeat (MSCV-LTR) promoter and a splice donor (SD) cassette to drive ectopic overexpression of downstream exons (mimicking oncogenic gain of function mutations). Gene trap components consisting of splice acceptors (SAs) on both DNA strands and a bidirectional polyadenylation signal (pA) to prematurely terminate gene transcription (mimicking tumor suppressor loss of function mutations) were also engineered into the constructs, permitting gene disruption independent of transposon reinsertion orientation (**Figure 1A**). Relative to T2Onc, T2Onc2 contains a larger SA fragment and an overall smaller size to increase transposition rate (14). Overall, the introduction of these components into the SB construct increased the likelihood of transcriptional interference upon reintegration of the mobilized transposon into gene-encoded chromosomal DNA.

Generation of the T2onc and T2Onc2 transgenic mice by pronuclear injection of linearized plasmid DNA yielded offspring with multiple copies of adjoining transposons, a phenomenon caused by homologous recombination between constructs, and subsequent concatemerized integration into chromosomal DNA (**Figure 2A**) (15). The result was an increased number of insertional events per cell when crossed with a transgenic mouse with an activating SB transposase (SBase) construct. T2Onc transgenic lines are "low copy number," with ~25 transposons per cell; T2Onc2 lines contain a "high copy number" of ~150–300 transposons per cell (13, 14). While transposon copy number contributes to increased frequency of transposon mutagenesis, alterations to the SBase component of the system also influence efficiency.

The first transgenic SBase mouse contained the originally identified transposase sequence (SB10) under the control of the chimeric chicken beta-actin and cytomegalovirus immediate early promoter sequences (CAGG) promoter (**Figure 1B**) (9, 13). This promoter was selected for its ubiquitous activity in mice (16). A shortcoming of transgenic models, however, is transgene silencing via methylation secondary to both concatemerization and positional effects depending on the chromosomal region into which the transgene inserts. This was of particular concern, as *in vitro* work demonstrated the influence of transposase expression levels on transposition rate (17). Indeed, transposition frequency in the double-transgenic low-copy T2Onc; *CAGGS-SB10* and high-copy T2Onc2; *CAGGS-SB10* mice were insufficient for tumorigenesis on a wildtype background, with transposase levels found to be quite low (13, 18). To address this, a knock-in SBase mouse (*RosaSB11*) was engineered by targeting an optimized transposase sequence (referred to as SB11) into the *Gt(ROSA)26Sor* locus (14). *SB11* is similar to *SB10* with the exception of several amino acid substitutions found to increase transposition activity (17). The *Gt(ROSA)26Sor* locus

was selected as prior research had demonstrated ubiquitous expression of genetic constructs targeted to this region (19). Crossing of the *RosaSB11* mouse to either the high-copy T2Onc2 or low-copy T2Onc mouse resulted in transposition-induced cancer with a predilection for lymphoid leukemias (14, 18). Interestingly, the low-copy T2Onc; *RosaSB11* mice demonstrated decreased embryonic lethality and a prolonged latency relative to the high-copy T2Onc2; *RosaSB11* offspring (18). Indeed, these early experiments demonstrated the ability of both components of the SB system to influence *in vivo* cancer models.

Since these initial publications, other mice have been engineered to expand the functionality of SB insertional mutagenesis, with 16 distinct SB constructs used to-date in *in vivo* cancer screens (**Figures 1A,B**). One limitation of the T2Onc2 mouse was its predisposition for leukemias despite ubiquitous SB activation in the animal (14). A possible explanation for this phenotype was the promoter used to drive oncogenic over-expression of downstream exons upon transposon reinsertion. Indeed, the MSCV-LTR promoter demonstrates greatest transcriptional activity in hematopoietic cells, potentially biasing the mutagenic rate in favor of blood tumors (20). To address this, the T2Onc3 transposon was generated by substituting the MSCV-LTR promoter for the CMV enhancer/chicken beta-actin (CAG) promoter (20). The T2Onc3 mouse demonstrated increased SB expression in epithelial cells and the formation of various solid tumors (20). Another limitation of the SB system was the high rate of embryonic lethality due to constitutive transposition induced by early, ubiquitous activation of the *RosaSB* transposase (14). As such, a *LoxP-stop-Lox* cassette was incorporated into the *RosaSB* allele (*RosaSbase^LSL*) rendering a *Cre*-inducible transposase (20). For the first time, SB activation could be induced in a tissue-specific manner by breeding with any of the *Cre*-transgenic mouse lines. Other SB design modifications have included the introduction of different transposon promoters, tissue-specific promoters for selective transposase activation, and transposase constructs with increased activity (3, 21–25). Mice harboring fusion constructs capable of mobilization by either SBase or the PiggyBac (PB) transposase (referred to as activating/inactivating transposons, or "ATPs") have also been generated (26). Collectively these models demonstrate the versatility of the SB system to model cancers.

Regulation of the SB Machinery

Mobilization of the SB transposon is a regulated process influenced by differences in SBase binding affinities, host cell factors, and DNA methylation status (recently reviewed elsewhere) (27). The transposase harbors both a catalytic and DNA binding domain, which is divided into two DNA binding subdomains, RED and PAI (28). Briefly, a pair of direct repeat (DR) binding sites in both inverted terminal repeat (ITR) ends of the transposon (collectively referred to as IR/DR elements) serve as transposase recognition sites for mobilization by the transposase. All four of these sites must be present for SB transposition to occur (28). Modifications to the primary DNA sequence of the ITR and/or transposase can influence transposition frequency (17, 29, 30).
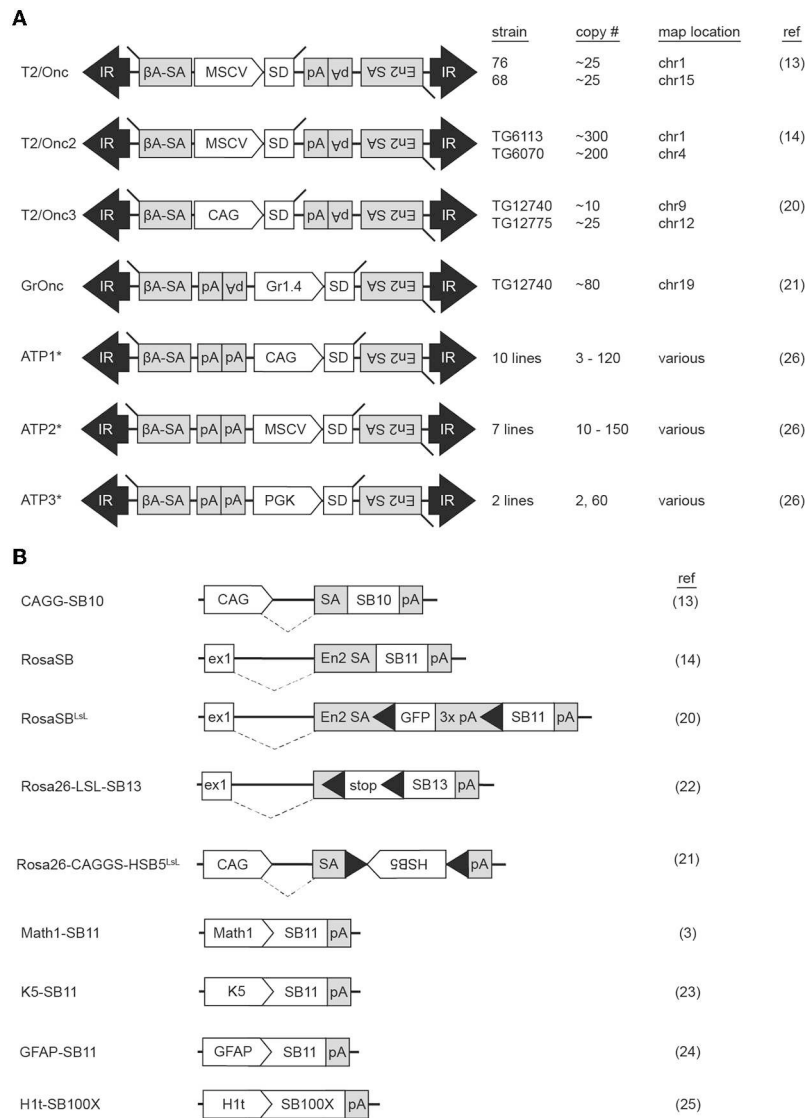
FIGURE 1 | Multiple Sleeping Beauty constructs have been engineered into mouse lines for use in forward genetic screens. Mouse strains harboring transposon **(A)** or transposase **(B)** constructs are available for use in forward genetic cancer screens. Transposition is initiated when a mouse harboring a transposon concatemer is crossed with a mouse harboring a transposase. For most transposon constructs, multiple lines have been developed, each harboring a different number of transposons on a different chromosome **(A)**. Transposase constructs incorporate various modifications to regulate location of enzyme activation **(B)**. ßA-SA, beta-actin splice acceptor; CAG, CAG promoter; En2 SA, engrailed-2 splice acceptor; ex1, Rosa26 exon 1; Gr1.4, Graffi1.4 murine leukemia virus LTR; IR, inverted repeat; MCSV, murine stem cell virus promoter; pA, polyadenylation signal; PGK, phosphoglycerate kinase promoter; SD, splice donor. Black triangles within transposase constructs represent Lox sequence sites recognized by Cre recombinase. *The ATP constructs incorporate both Sleeping Beauty and PiggyBac transposons.

In mammalian cells, the transcription factor HMG2L1 (also referred to as HMGXB4) aids transposase expression (31). Once transcribed, the transposase protein actively recruits the host cell High Mobility Group protein (HMGB1) to enhance its binding to an inner DR and facilitate DNA bending and complex formation with the inner DR of the ITR at the opposite end of the transposon (32). CpG methylation of transposon concatemers induces a condensed heterochromatin structure that further enhances this protein-DNA synaptic complex formation (33). Transposon excision then proceeds via DNA double strand breaks (32, 34). The transposase recruits Ku70 and DNA-dependent protein kinases that preferentially repair the excision sites via non-homologous end joining (NHEJ) (35). The transcription factor Miz-1 is also recruited, leading to down-regulation of Cyclin D1 and cell cycle arrest in G1, a phase favoring NHEJ (36). Upon repair, a 7-base footprint remains (**Figure 2B**) (35, 37). The exposed 3′-OH groups on the ends of the excised transposons are used in the transfer reaction upon transposon reintegration (9).
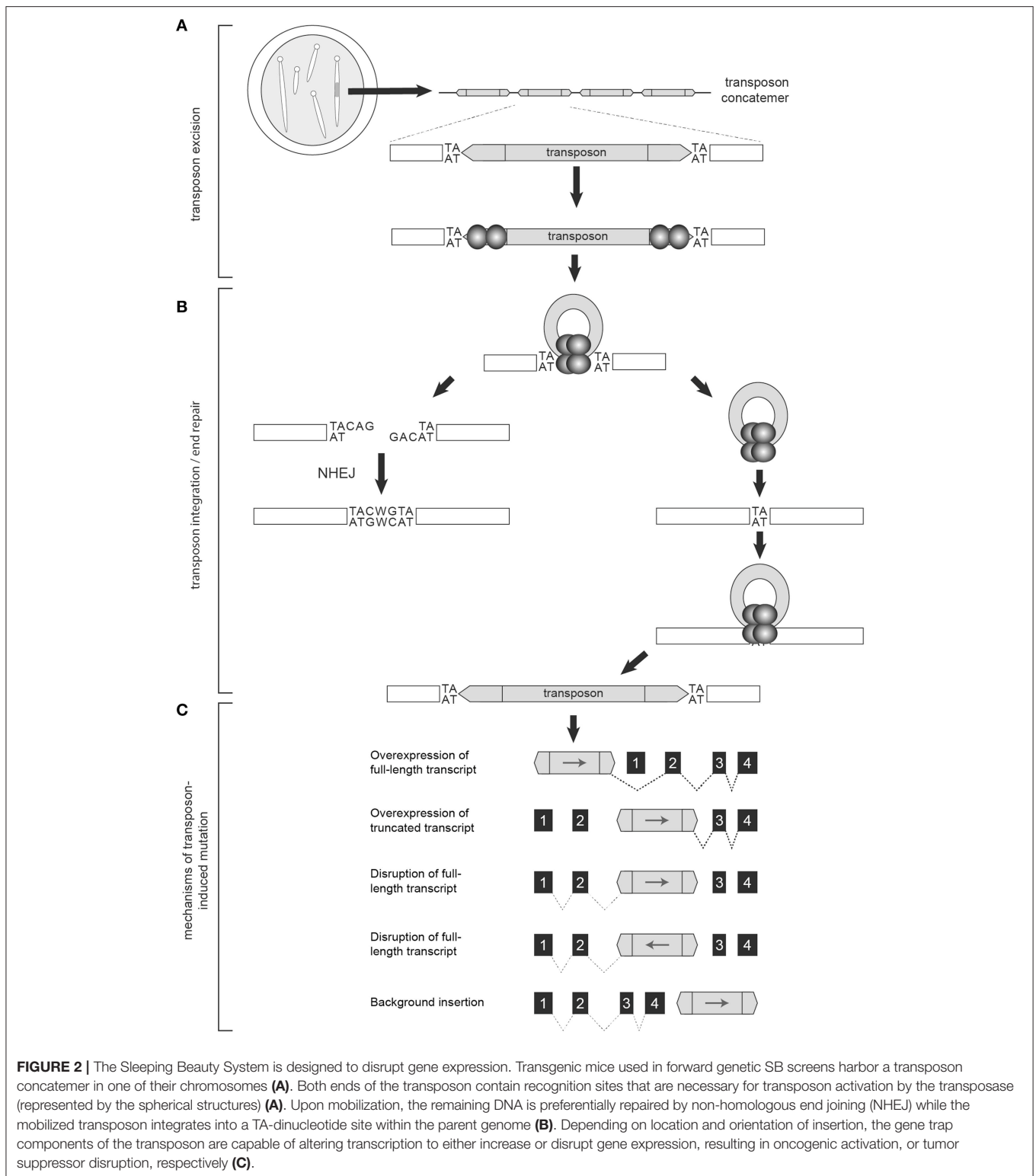
FIGURE 2 | The Sleeping Beauty System is designed to disrupt gene expression. Transgenic mice used in forward genetic SB screens harbor a transposon concatemer in one of their chromosomes **(A)**. Both ends of the transposon contain recognition sites that are necessary for transposon activation by the transposase (represented by the spherical structures) **(A)**. Upon mobilization, the remaining DNA is preferentially repaired by non-homologous end joining (NHEJ) while the mobilized transposon integrates into a TA-dinucleotide site within the parent genome **(B)**. Depending on location and orientation of insertion, the gene trap components of the transposon are capable of altering transcription to either increase or disrupt gene expression, resulting in oncogenic activation, or tumor suppressor disruption, respectively **(C)**.

SB transposons reintegrate into chromosomal DNA at TA-dinucleotide sites with preference for AT-repeat rich sequences, presumably due to optimized DNA bendability (**Figure 2B**) (9, 38). Additional TA dinucleotide sequences are added on either

end of the integrated transposon upon repair (9). Reintegration is highly efficient, with rates in murine embryonic stem cells as high as 75% (39, 40). There is some debate regarding the integration bias of the SBase enzyme. Several studies of unselected insertion

events detected no significant bias for insertion near specific chromatin marks or genomic features (40, 41). However, another report did find a slight bias for SB insertion events within transcribed regions (42). One clear source of bias occurs when SB is activated from within chromosomes resulting in the "local hopping" phenomenon—the tendency of SB transposons to reintegrate into the original donor chromosome (43). This bias can be removed by excluding reintegration sites that land within the donor chromosome from analyses.

Other Integrating Vector Systems

Other integrating vector systems commonly used in mouse models of cancer include retroviruses and the PiggyBac (PB) transposon system. Retroviral mutagenesis has the longest history of use in forward genetic cancer screens, however several drawbacks exist (44, 45). The cellular tropism of retroviruses used in mouse models confines them to the production of mammary and hematopoietic tumors (46, 47). These models also demonstrate a bias for oncogene activation over tumor suppressor inactivation (48, 49). Causative gene identification can be difficult due to the ability of retroviral enhancers to influence expression of genes hundreds of kilobases away from their integration sites (50, 51). Moreover, non-random integration patterns are commonly observed, with murine leukemia virus (MLV) exhibiting bias for transcription start sites (TSSs) of transcriptionally active genes (52). Similar to SB, PB is a "cut-and-paste" transposon with a unique integration motif ("TTAA" vs. "TA" for SB) that demonstrates activity across tissue types and is capable of candidate oncogene and tumor suppressor identification (41, 53, 54). However, as with retroviruses, PB shows preferential bias for TSSs and actively transcribed regions (55). Researchers have suggested combining SB and PB screens to maximize candidate tumor driver identification which would be readily feasible using mice harboring an ATP transposon (**Figure 1A**) (56). Indeed, in a model of pancreatic cancer using ATP1 activated by PB transposase (PBase), a rare subtype of pancreatic cancer not identified in two distinct models of SB-induced pancreatic cancer was observed (22, 57, 58). Currently, only limited reports are available on candidate tumor drivers identified in dual PB/SB screens in combination with PBase or SBase (56). It remains to be seen the extent to which these two systems may complement each other in the quest for tumor driver identification.

Interpreting SB Data

SB-induced tumors are harvested and the genomic DNA extracted. Since the transposon also functions as a "tag," mutated sites within DNA are easily identified. Candidate tumor drivers (historically referred to as Common Insertion Sites or "CISs") are called when a gene harbors integrations at a higher frequency than expected by chance alone. Put another way, SB insertions facilitating proliferation and malignant transformation are selected for and clonally expanded, resulting in non-random mutation profiles. To identify candidate tumor drivers, SB-genomic DNA junctions must be amplified, sequenced, mapped, and statistically evaluated. Variations at any step

of the process can influence sensitivity, quantification, and gene identification. This complicates comparisons between independent studies.

After tumor harvesting and DNA extraction, the tumor DNA is fragmented. Initial studies accomplished this using restriction enzyme digestion (59). DNA shearing via sonication is now preferred, however, as it results in uniformly sized fragments. SB-DNA junctions are then amplified using linker-mediated PCR (59–61). Next generation sequencing technology permits the simultaneous sequencing of products and has dramatically increased the ability to identify SB insertional events. Platforms commonly utilized today include the Roche 454 GS FLX Pyrosequencer (with an average read depth per sample in the thousands) and Illumina (with an average read depth per sample in the 100,000s). The combination of DNA fragmentation via sonication, which greatly reduces PCR amplification bias due to size variability, with the high read depths achievable on next generation sequencing platforms (primarily Illumina) permits semi-quantification of transposon reads (55, 59–61).

Upon read acquisition, sequences are aligned and mapped to the mouse genome (62). SB insertions mapping to the chromosome harboring the SB transposon concatemer are omitted from downstream analyses to account for local hopping. Consequently, studies utilizing only one SB mouse line do not achieve full genome coverage and genes facilitating tumorigenesis residing on the parent chromosome are missed. To overcome this limitation, investigators may elect to use two SB mouse lines in a screen. Omission of sequence reads mapping to the parent chromosome of one line are included in the reads from the other line and vice versa. It is important to note that mapping errors can result in erroneous candidate gene identification. As an example, *Sfi1* was identified as a candidate tumor driver in numerous screens, but was subsequently found to be a false call due to inaccurate mapping of reads to repetition-rich regions of the genome (63). It is inevitable that algorithms used to interpret the vast amount of genetic data generated from SB screens will continue to improve.

Various statistical models have been used to identify candidate tumor drivers among the mapped SB-DNA reads including: Monte Carlo (MC) simulation, Gaussian Kernel Convolution (GKC) method, two-dimensional Gaussian Kernel Convolution method (2DGKC), Poisson distribution, Poisson Regression Insertion Model (PRIM), and gene-centric common insertion site (gCIS) analysis (44, 60, 64–67). MC, GKC, and Poisson-based methods take into account the local density of TA dinucleotides (i.e., potential SB insertion sites) within genomic intervals of user-defined size to localize clusters of disruptions (44, 64, 67). The 2DGKC and PRIM methods incorporate parameters permitting identification of possible co-occurring mutations within individual tumors (65, 66). A limitation to these approaches is that identifying biologically meaningful windows can be challenging. In contrast, the gCIS approach examines transcribed regions of the genome and determines an expected SB insertion rate for each RefSeq gene based on the number of TA sites within the RefSeq unit and 10 kilobases of promoter sequence (60). Overlap in candidate tumor driver calls among statistical methods ranges from

60–90%, however the gCIS approach has been shown to have increased sensitivity relative to MC and GKC methods (60, 68). In addition, the pattern of insertion can predict the role the candidate gene may have in tumorigenesis. Regions of integration with the majority of transposons oriented upstream of a gene with the transposon promoter (e.g., MSCV, CAG) in the same direction as gene transcription often leads to increased oncogene expression. Insertions throughout the length of a gene with no orientation bias are suggestive of decreased gene expression, or loss of tumor suppressor activity (**Figure 2C**).

More recently, RNA from SB tumors has been used to further assess SB-induced genetic alterations (69). In their methods paper, investigators evaluated RNA sequencing (RNA-Seq) data in juxtaposition with traditional DNA insertion site data obtained from the same SB tumors. Regions of the genome with candidate insertion sites had observable changes in transcript levels when compared to tumors lacking those same insertions. Transposon-RNA fusion transcripts were also identified, many of which corresponded with DNA-derived data (53 and 71% of fusions were supported by restriction enzyme digested LM-PCR and sheared LM-PCR methods, respectively). Additionally, novel fusions and alterations in transcript expression not readily ascertained from genomic data were also identified (69). Overall, transcript analyses permitted direct identification of genes altered by SB and the nature of the disruption (i.e., increased expression, decreased expression, formation of a truncated transcript, etc.). While it is not uncommon for researchers to amplify individual fusion products to further investigate SB-induced alterations, large scale transcriptome analyses have hitherto not been performed. Given its utility, it is likely that the RNA-Seq technique will be incorporated into future SB forward screen analyses.

While different approaches to tumor preparation and bioinformatic analyses can interfere with side-by-side comparisons of independent studies, so too does the combination of different SB components (e.g., the pairing of T2Onc2, T2Onc3, or GrOnc with *RosaSB* or *RosaSbase*^{LSL}) in the generation of a model (**Figures 1A,B**). Biologically this makes sense, as temporal onset, location, and kinetics of transposition are all influenced by transposon and transposase selection (2, 13, 18, 20). Indeed, transposon selection is the largest determinant of variability in candidate tumor driver identification among SB models of the same cancer type (20, 70). It is precisely this diversity and flexibility that make the SB system an indispensable tool in studying cancer genetics. The SB insertional mutagenesis system has been used to model numerous forms of cancer types and led to the identification of innumerable candidate drivers, many of which have been subsequently validated (reviewed elsewhere) (71, 72). Two publicly available databases, the Candidate Cancer Gene Database (CCGD) and the Sleeping Beauty Cancer Driver Database (SBCDDB), allow for streamlined searching of these candidate genes (73, 74). In the following paragraphs we explore variables that, when present, have the potential to influence tumorigenesis as reflected by alteration of the SB tumor driver landscape.

TABLE 1 | Sensitizing mutations used in SB cancer screens.

Gene mutated	Tumor (or tissue) type and references
Apc	Intestinal (77, 78)
Blm	Glioma (24)
Braf	Melanoma (7, 75, 76)
Cadm1	Leukemia/CD3-positive T-Cell lymphoma (predominant) (79)
Ccne (cyclin E)	Leukemia (erythroleukemia and T-ALL) (80)
Cdh1	Breast (invasive lobular phenotype predominant) (81)
Csf	Glioma (24)
Ctnnb1	Breast (82)
Egfr	Peripheral nerve sheath (4)
Hras	Skin (non-melanoma) (23)
	Thyroid (poorly differentiated predominant) (83)
Jak2	leukemia (erythroleukemia predominant) (84)
Kras	Intestinal (78)
	Pancreatic adenocarcinoma (22, 57)
Myc	Liver (85)
Npm1c	Leukemia (AML predominant) (21)
p19arf	Gliomas (24, 86)
	Lung (87)
	Multiple tumor types (13)
Ptch	Medulloblastoma (3, 88, 89)
Pten	Lung (predominant) (87)
	Prostate (90)
	Breast (91)
	Liver (92)
	Medulloblastoma (93)
Rag2	Multiple tumor types (94)
Rassf1a	Leukemia/poorly differentiated lymphoma (predominant) (95)
Sav1	Liver (6)
Smad4	Intestinal (78)
	Gastric adenomas (predominant) (96)
Stat5b	B-ALL (97)
Tcl1	CLL (98)
Tgfbr2	Intestinal (99)
Trp53	Liver (1)
	Leukemia/CD3-positive T-Cell lymphoma (predominant) (100)
	Lung (87)
	Medulloblastoma (3, 88, 101)
	Osteosarcoma (8)
	Breast (102)
	Lymphoma, B cell (103)
	CNS-PNET (93)
Genes mutated (co-occurring)	**Tumor (or tissue) type and references**
BCR-ABL (translocation)	Leukemia (CML) (104)
Egfr and *Trp53*	Peripheral nerve sheath (4)
Etv6-RUNX1 (fusion)	Leukemia (BCP-ALL predominant) (105)
Etv6-RUNX1 (fusion) and *Pax5*	Leukemia (BCP-ALL predominant) (106)

BCP-ALL, B-cell precursor acute lymphoblastic leukemia; CNS-PNET, central nervous system primitive neuroectodermal tumor. AML, acute myeloid leukemia; B-ALL, B-cell acute lymphoblastic leukemia; CLL, chronic lymphocytic leukemia; CML, chronic myeloid leukemia.

SENSITIZING MUTATIONS IN SB MOUSE MODELS

Combining a predisposing (or sensitizing) mutation with the SB system can facilitate tumorigenesis. Sensitizing mutations are selected based on their known role as a tumor suppressor or oncogene in the cancer type being studied. In some cases, these mutations are required to induce cancer formation in an SB model (7, 22, 57, 75, 76). A comprehensive list of sensitizing mutations used in SB mouse models of cancer is presented in this review (**Table 1**). When tumorigenesis is observed in SB mice on both wildtype and sensitized backgrounds, investigators have the opportunity to compare differences in mutation profiles between the two cohorts. In this way, mutations driving tumor development on a wildtype background can be distinguished from genetic drivers co-occurring with a particular mutation. Such information allows for better tumor subtype stratification. Published comparative analyses conducted on screens of lymphomas, acute and chronic myeloid leukemias, osteosarcoma, and colorectal disease are discussed below.

Lymphomas

Lymphomas are cancers of the lymphatic tissues in the body. Three SB lymphoma publications studied SB-induced "lymphoma/leukemia" in the context of $Cadm1$-null, $Trp53$ homozygous and heterozygous mutant, and $Rassf1$-null backgrounds (79, 95, 100). $TP53$ and $RASSF1$ are both tumor suppressors known to be inactivated in human lymphomas (107–110). $CADM1$ belongs to the immunoglobulin superfamily of cell adhesion molecules. It was initially characterized as a tumor suppressor with loss of expression observed in various cancer types including hematologic malignancies (111, 112). More recently, cell surface expression of $CADM1$ has been described as an indicator of disease status and progression in some lymphoma subtypes (113, 114).

In the first of their studies published in close succession, tumors from 117 $Cadm1^{-/-}$; SB and 73 $Cadm1^{+/+}$; SB littermates were found to be predominantly CD3-positive T-cell lymphoma (79). SB mice on the $Cadm1$ null background had increased tumor multiplicity that developed with decreased latency (79). Ten genes including $Nr3c1$, the most frequently disrupted gene, were unique to the $Cadm1$-null background (79). Fifteen candidate genes were shared between the two cohorts while six genes were identified exclusively in $Cadm1^{+/+}$; SB tumors. The SB insertion pattern observed in $Nr3c1$, the gene that encodes the glucocorticoid receptor, is suggestive of its role as a tumor suppressor (79). Authors noted that three other genes identified exclusively in the $Cadm1$-null SB cohort of tumors ($St13$, $Ets1$, and $Csf3r$) encode proteins that regulate or interact with the glucocorticoid receptor (79). Indeed, impaired glucocorticoid signaling, via various mechanisms including deletion of $NR3C1$, has been associated with relapse and poor prognosis in pediatric acute lymphoblastic leukemia and blastic plasmacytoid dendritic cell neoplasm (115–117). These data demonstrate a selective advantage for combining impaired cell adhesion and disrupted glucocorticoid signaling in

lymphomagenesis. The mechanistic and clinical implications of this observation remain to be explored.

In the second forward screen performed in mice on a $Rassf1$-null background, 111 $Rassf1^{-/-}$; SB and 25 $Rassf1^{+/+}$; SB tumors were predominantly poorly differentiated lymphoma (95). Tumor latency was decreased in SB mice on the $Rassf1$-null background. Authors noted that their statistical analysis included data from an additional 101 SB mice on a wildtype background. These mice lacked the 129/Sv background introduced by breeding with $Rassf1$-null mice (95). Nevertheless, candidate drivers unique to the $Rassf1$-null background were identified including the transcription factor $Runx2$ (95). When present, Rassf1 alters Hippo signaling by influencing the proteins with which Yap1 complexes while Runx2 complexes with Yap1 directly (118–120). Authors posited that loss of both $RASSF1$ and $RUNX2$ exacerbates YAP1-TEAD complex formation leading to increased cellular proliferation and supported this hypothesis with in $vitro$ assays (95).

In the analysis of $Trp53$-related tumors, CD3-positive T-cell lymphoma was again identified as the predominant tumor type (100). Nine $Trp53^{-/-}$, 116 $Trp53^{+/-}$, and 36 $Trp53^{+/+}$ SB-induced tumors were analyzed. Eight of the nine candidate genes identified in the $Trp53^{-/-}$; SB group were unique to that group. Twelve genes were shared between $Trp53$ heterozygous and $Trp53$ wildtype SB cohorts. One gene, $Rapgef6$, was shared between both $Trp53^{-/-}$; SB and $Trp53^{+/-}$, SB cohorts but not detected in $Trp53^{+/+}$; SB tumors (100). $Jdp2$ disruption leading to overexpression was enriched in SB tumors maintaining $Trp53$ heterozygosity (100). Jdp2 directly binds the $Trp53$ promoter to represses its expression (100, 121). Thus, investigators concluded that $Jdp2$ mutation facilitates tumorigenesis in $Trp53$ heterozygotes by rendering a biallelic $Trp53$ loss-of-function phenotype.

In another model, T2Onc and RosaSBase^{LsL} with or without conditional $Trp53^{R270H}$ resulted in 65% of SB mice developing B cell lymphoid disease histologically compatible with follicular lymphoma or diffuse large B cell lymphoma (103). In an evaluation of 23 SB-only and 7 SB-$Trp53$ tumors, investigators identified 48 and 12 candidate tumor drivers within each group, respectively. It is difficult to draw conclusions on influence of the predisposing $Trp53$ mutation on mutational profile given the small sample size. Authors did, however, further evaluate the Ras-responsive element binding protein 1 ($Rreb1$) transcription factor identified in the SB-only tumors. $RREB1$'s influence on $KRAS$ expression increased RAS/MAPK signaling in $vitro$ and was found to be overexpressed in a subset of human diffuse large B-cell lymphomas (DLBCLs) (103).

Acute Myeloid Leukemia (AML)

Mutations involving nucleophosmin ($NPM1$) are observed in ~60% of cytogenetically normal AML tumors (122). Vassiliou et al. studied the impact of heterozygous expression of a "humanized" version of the most common mutation, $Npm1c^A$, on the transposon insertion profile of an SB model of AML (21). Tumors from 87 $Npm1c^{A/+}$; SB mice and 34 $Npm1c^{+/+}$; SB mice were assessed. Investigators found 75% (18/24) of candidate tumor drivers identified in the $Npm1c^{A/+}$; SB tumors to be

unique to that cohort (21). Cooperative mutations with $Npm1c^A$ included SB insertion patterns consistent with activation of *Csf2* (observed in 48% of tumors) and *Flt3*. Both granulocyte-macrophage colony stimulating factor (*CSF2*) and the tyrosine kinase receptor *FLT3* are frequently mutated in human AML with *FLT3* mutations commonly co-occurring with *NPM1* mutations (123–125). The internal tandem duplication mutation of *FLT3* (*FLT3-ITD*) is known to cause constitutive activation of the JAK/STAT pathway (126). The synergism of co-mutated *Npm1* and *Flt3-ITD* in AML tumorigenesis was further demonstrated *in vivo* and corresponded with a pronounced overall change in lymphoid progenitor cell gene expression (127, 128).

Chronic Myeloid Leukemia (CML)

The *BCR-ABL* fusion gene is generally considered a disease-defining mutation of CML. CML is a slow-growing tumor, with patients appropriately treated with a tyrosine kinase inhibitor targeting BCR-ABL remaining in the chronic phase for many years. Progression to the accelerated and blast crisis phases is characterized by acquisition of new chromosomal aberrations and increasing number of blasts. Once in blast crisis, CML behaves more like acute myeloid leukemia and median survival is ~12 months (129). Previously, mice engineered to express the disease-defining BCR-ABL translocation in hematopoietic stem cells were found to recapitulate chronic phase CML, with a small subset of tumors progressing to blast crisis (130). To further explore the genetics of CML progression, Giotopoulos et al. crossed the *BCR-ABL* mouse to mice harboring the GrOnc transposon and *Mx1-Cre* mediated activation of $RosaSbase^{LSL}$ SB system (*BCR-ABL; SB*) (104). The GrOnc construct contains the Graffi1.4 murine leukemia virus LTR, which preferentially promotes myeloid lineage cells (131). Authors also bred *SB*-only mice lacking the *BCR-ABL* translocation. Phenotypic and microarray gene expression comparisons between the *BCR-ABL* and *BCR-ABL; SB* cohorts were conducted. *BCR-ABL; SB* mice were found to have decreased survival, indicators of disease progression (increased terminal WBCs, decreased hemoglobin, increased spleen and liver weights), and changes in expression of genes previously implicated in CML progression when compared to *BCR-ABL*-only mice (104). Microscopic and immunophenotypic evaluation of tumors revealed all *BCR-ABL*-only mice to have disease that remained in the chronic phase. *SB*-only mice developed both lymphoid (26%) and myeloid (70%) acute leukemias. *BCR-ABL; SB* mice also developed primarily myeloid acute leukemia (85%). Interestingly, the *BCR-ABL; SB* cohort was the only group to manifest a continuum of CML progression, with tumors in the intermediate accelerated phase (10%) and blast crisis phase (5%) (104).

Candidate tumor driver analysis conducted on tumor DNA from 52 *BCR-ABL; SB* mice and 20 *SB*-only mice revealed 78/91 (86%) of candidate drivers identified in the *BCR-ABL; SB* tumors to be unique to that cohort (104). Authors noted that, among genes identified, several had already been implicated as potential drivers of CML and CML progression including: *Asxl1*, *Myb*, *Stat5b*, and *Pten* (104). Genes not previously associated with CML progression included: *Jak1*, *Flt3*, *Nf1*, *Erg*, and *Mll3* (104). Authors found the ETS-related gene (*Erg*) transcription factor to be most frequently disrupted, with an overall transposon insertion profile suggestive of oncogenic activation (104). To further assess its role *in vivo*, *ERG* was overexpressed using a retroviral vector in hematopoietic stem cells derived from *BCR-ABL*-expressing and wildtype mice. Cells were then transplanted into congenic recipient mice (104). The *ERG*; *BCR-ABL* mice developed acute leukemias phenotypically resembling the *BCR-ABL*; *SB* mice. Overexpression of *ERG* on a *BCR-ABL* background significantly decreased survival relative to *ERG* on a wildtype background and *BCR-ABL*-only mice. Authors noted that ERG has been implicated in poor prognosis and blast-phase transformation of other hematological diseases (104). As more data regarding chromosomal and gene expression changes in CML progression are obtained, clarity as to the clinical relevance of *ERG* in CML may begin to emerge.

Osteosarcoma

The third most common cancer among children and young adults, osteosarcomas harbor pronounced chromosomal aberrations (132). A limited number of genes are implicated in tumorigenesis, one of which is *TP53*. In an SB model of osteosarcoma, introduction of the $Trp53^{R290H}$ dominant negative allele accelerated tumorigenesis and increased tumor burden and penetrance relative to *SB*-only mice and *Trp53*-only mutant mice (8). Mice from both *SB* cohorts (with and without the *Trp53* mutant) developed liver and/or lung metastases (8). Comparative genomic hybridization and karyotyping revealed *SB*-only tumors had fewer genomic aberrations relative to *Trp53*; *SB* tumors. In a comparison of transposon insertions from 96 *Trp53*; *SB* and 23 *SB*-only osteosarcomas, Moriarity et al. observed overlap of the top hits in both cohorts including: *Pten*, *Eras*, and *Nf1* (8). Investigators validated the cooperative nature of the combined *Pten* and $Trp53^{R270H}$ mutations *in vivo* (8). Mutation of *TP53* and loss of *PTEN* are frequent derangements in human osteosarcoma, with loss of *PTEN* associated with a poorer clinical prognosis (133, 134). Indeed, in a recent whole exome phylogenetic analysis of osteosarcomas, mutation of *TP53* was determined to be an early event while loss of *PTEN* was associated with lung metastases (135).

Colorectal Disease

Starr et al. compared candidate tumor drivers of colorectal disease (predominantly adenomas) identified in heterozygous *Apc* (*Apc*; *SB*) model with those from *SB*-only tumors (77, 136). Investigators found three genes (fewer than 4%) identified in the *Apc*; *SB* screen to also be affected in the *SB*-only screen (77). The shared genes were *Apc* (resulting in biallelic inactivation in *Apc*; *SB* mice), *Nsd1*, and *Wac*. Authors noted 70% of candidate drivers from the *Apc*; *SB* tumors had one or more transposon insertions within the same locus in the *SB*-only tumors although these insertions did not occur at a rate high enough to be considered candidate drivers in *SB*-only tumors using their conservative statistical approach (77).

In another SB model of intestinal disease, Takeda et al. compared the candidate tumor drivers identified in tumors from mice with one of three predisposing mutations: $Kras^{G12D/+}$, $Smad4^{KO/+}$, or $Trp53^{R172H/+}$ with a pooled list of candidate

drivers identified in SB tumors from mice with somatic or germline *Apc* mutations (collectively referred to as *Apc*) (68, 78). Investigators found ~50% overlap in candidate genes identified in the three cohorts of tumors with the previously published *Apc* list (78). One hundred and eleven genes (~8% of all identified candidate tumor drivers) were disrupted in all four groups (78). While *Apc* disruption was the most prevalent event in *Apc*-, *Kras*^{G12D}-, and *Trp53*^{R172H}; *SB* tumors, only 32% of *Smad4*$^{KO/+}$; *SB* tumors demonstrated insertional disruption of *Apc* with the majority of tumors (79%) displaying biallelic inactivation of *Smad4* (78). Upregulation of Wnt/B-catenin signal activating R-spondin genes (*Rspo1* or *Rspo2*) via transposon insertion was also significantly enriched in *Smad4*$^{KO/+}$; *SB* tumors. While these comparisons suggest sensitizing mutations may uniquely influence tumor driver profiles, conclusions should be tempered by the fact that strain backgrounds varied among the mice, introducing additional genetic variability (78).

Most recently, SB mutagenesis was used to identify genes cooperating with loss of TGF-B in intestinal neoplasms (99). Tumors from SB mice with homozygous conditional inactivation of TGF-B receptor, type II (*Tgfbr2*) were compared with tumors from SB mice on a wildtype background. Authors found 34% (232/673) and 50% (187/372) of candidate genes identified in *Tgfbr2*; *SB* tumors to be unique to the presence of *Tgfbr2* inactivation, depending on statistical method used (99). Comparison of the two statistical approaches revealed overlap of 17 genes, with an enrichment in genes responsible for either Wnt/B-catenin or Hippo pathway signaling including *Lrp6*, *Ppp2r1a*, *Tcf7l2*, and *Yap1* (99). Given the role of SMAD4 in TGF-B signaling, candidate genes identified in *Smad4*$^{KO/+}$; *SB* tumors were compared to candidate genes in *Tgfbr2*; *SB* tumors with the expectation that significant overlap would be observed (78, 99). Indeed, authors found a 54% (243/449) overlap in the genes identified in the *Smad4*$^{KO/+}$; *SB* screen and the *Tgfbr2*; *SB* screen with a decrease in the number of independent tumors harboring *Apc* transposon insertions relative to wildtype (27% (35/130) vs. 45% (58/130), respectively) (99).

Collectively, these studies reveal the dramatic influence single gene modification can have on tumor gene recruitment and overall tumorigenesis. While some candidate genes appear in more than one context, the identification of candidate tumor drivers unique to particular mutant backgrounds suggests selective gene cooperation in the formation of these tumor subtypes.

OTHER FACTORS INFLUENCING TUMOR DEVELOPMENT IN SB MODELS

Manipulation of the tumor microenvironment can extend beyond the introduction of a sensitizing mutation. Known risk factors can be incorporated into study designs to interrogate their influences on tumor genetics. To date, investigators have used SB mouse models to explore genetic mediators of drug resistance, virus-associated disease, sex-biases, and immune function in cancer.

Therapeutic Interventions

Patients with melanoma positive for the *BRAF*^{V600E} point mutation have overactivation of the MAPK pathway and, consequently, inhibition of apoptosis and uncontrolled cellular proliferation. In advanced disease, patients are candidates for the targeted protein kinase inhibitor vemurafenib (PLX4720). The initial clinical response is typically quite strong, but patients rapidly relapse and experience aggressive disease progression (137). SB technology has been utilized to identify potential mediators of vemurafenib monotherapy resistance in melanoma (75). In the study, mice harboring a melanocyte-specific 4-hydroxytamoxifen (4-OHT)-inducible *Tyrosinase-CreER*^{T2} allele and inducible mutant *Braf* allele in the endogenous *Braf* locus (*Tyr-Cre*^{ERT2}; *Braf*$^{lsl-V618E/+}$) were mated with *SB* mice (offspring hereafter referred to as *Braf-SB* mice). After application of 4-OH tamoxifen and expression of oncogenic *Braf* and the *SB* system, *Braf-SB* mice were aged and monitored for tumor development. Upon melanoma formation, a subset of the mice were administered a diet containing vemurafenib. Tumors in these mice were observed to regress over a 1–4 week period followed by a period of stable disease and subsequent relapse marked by tumor regrowth/progression (75). Genomic DNA from treatment-naïve and vemurafenib-resistant melanomas was harvested and candidate tumor drivers identified. Statistical comparison of genes identified in treatment-naïve and vemurafenib-resistant melanomas revealed eight candidate tumor drivers enriched in treatment-resistant tumors: three known mediators of vemurafenib resistance (*Braf*, *Mitf*, and *Cdkn2a*) and five novel candidate genes (75). Authors validated one of these novel candidate genes, *ERas*, in a human cell line assay (75).

ERAS (embryonic stem cell-expressed Ras) encodes a constitutively active RAS-like protein that potentiates PI3K/AKT signaling. Crosstalk between MAPK and PI3K/AKT signaling is well-established, with PI3K/AKT upregulation a recognized mechanism by which vemurafenib resistance can occur [recently reviewed by (138)]. In their experiments, Perna et al. demonstrated the ability of ERAS to evade vemurafenib-mediated inhibition of the MAPK pathway and consequent dephosphorylation/activation of BAD, a proapoptotic protein, by promoting inactivation of BAD via the AKT pathway (75). Indeed, a phase I/II clinical trial of for the dual treatment of patients with *BRAF*-positive advanced melanoma with BKM120, a PI3K inhibitor, and vemurafenib had been initiated, although terminated early due to therapy-related toxicity (NCT01512251). Nevertheless, these data demonstrate the clinical relevance of SB screens in identifying possible genetic mediators of drug resistance and underscore their potential to inform future clinical research and treatment.

Environmental Exposures

Chronic HBV infection is known to increase the risk of hepatocellular carcinoma (HCC) and is attributable to half of all HCC cases (139, 140). Mice engineered to express the HBV surface antigen (HBsAg) develop hepatocellular inflammation, hepatic necrosis, regeneration, benign adenomas, and subsequent HCC (141). Two different SB models with liver-specific activation

of the SB system (one utilizing the T2Onc2 transposon, the other using T2Onc3) have been deployed in the presence of HBsAg activation to identify candidate drivers of tumor formation in the context of HBV-induced inflammation (6, 142). Although a detailed comparison was not made, Kodama et al. observed a high degree of concordance between candidate tumor drivers identified in SB tumors with HBsAg-induced inflammation and SB alone after pooling data from both T2Onc2 and T2Onc3 screens (142). Hepatic fibrosis and cirrhosis are common features of chronic HBV infection in humans. Indeed, 90% of all HCCs develop in the context of cirrhosis (143). It is important to note, however, that no overt fibrosis is observed in the HBsAg mouse model (6, 141).

To directly assess the influence of a fibrotic microenvironment on HCC formation, Riordan et al. evaluated tumor formation in the context of chemical-induced fibrosis in a liver-specific SB mutagenesis model ($Alb^{Cre/+}$; SB) (70). Two lines of T2/Onc3 transposon mice (TG12740 and TG12775) were used to permit full genome coverage. Chronic fibrosis was induced in 77 animals via intraperitoneal injection of carbon tetrachloride (CCl4). Tumors developed in both treated and untreated mice, permitting evaluation of the impact of fibrosis on tumor driver landscape. Twenty-one genes were identified as candidate tumor drivers in both cohorts including the most frequently mutated gene in both cohorts, $Rtl1$ (70). Certain drivers were more common in the CCl4-treated cohort of tumors including: $Gli2$, $Fign$, and Met, suggesting a propensity for their disruption in the context of fibrosis (70). Authors performed an $in\ vivo$ validation of $Gli2$, a transcription factor activated by Hedgehog pathway signaling, in the context of fibrosis (70). Authors also validated $Fign$, which was mutated in 8% (27/343) of CCl4 tumors and none of the SB-only tumors, in an $in\ vitro$ invasion assay (70). Interestingly, in a comparison of genes mutated in ≥2% of CCl4 tumors with those mutated in ≥2% of tumors from previously published HBsAg models, authors found 88% (50/57) overlap, suggesting an enrichment of these tumor drivers in the context of chronic hepatocellular injury.

Other investigators have assessed the impact of steatosis on hepatocellular disease (92, 144). Tschida et al. described an SB model of hepatic tumor formation in the context of alcohol-induced steatosis (144). SB mice harboring the T2Onc transposon and $Alb^{Cre/+}$-mediated induction of $RosaSbase^{LSL}$ were administration ethanol and a choline-deficient diet (144). In their model, 37/49 (76%) of SB mice receiving the treatment diet developed well-differentiated adenomas (81%) or HCC (19%) (144). Authors compared the transposon insertion profiles of tumors collected from 36 steatosis-induced SB mice (15 males and 21 females) to the insertion profile previously reported in tumors from SB mice harboring the same SB components on a $Trp53$-deficient background administered a normal diet (1). While this comparison is not ideal given the predisposing $Trp53$ mutation in one group and not the other, investigators chose to pursue $in\ vivo$ analyses of two candidate genes enriched in tumors from the steatosis cohort, $Prkaca$ and $Nat10$. Vectors expressing either constitutively active $Prkaca$ ($Prkaca^{L206R}$), fumarylacetoacetate hydrolase (Fah), and short hairpin RNA targeting $Trp53$ (sh$Trp53$) or Fah and sh$Trp53$

alone were introduced into Fah-deficient male and female mice on either the steatosis-inducing treatment diet or normal diet. Mice in both the treatment and normal diet groups injected with dual $Prkaca^{L206R}$ and sh$Trp53$ demonstrated higher tumor burden and penetrance relative to $Trp53$ alone (144). Moreover, a higher percentage of dual $Prkaca^{L206R}$/sh$Trp53$ tumors were histologically classified as HCC relative to sh$Trp53$ alone (14 and 0%, respectively). Among mice on the ethanol and choline-deficient steatosis-inducing diet, those receiving dual $Prkaca^{L206R}$/sh$Trp53$ demonstrated higher tumor penetrance and burden relative to $Prkaca^{L206R}$/sh$Trp53$ mice on normal diet, suggesting a tumorigenic role for $Prkaca^{L206R}$ in the context of hepatic steatosis (144). The same study design was executed using a $Nat10$ vector. Authors found no difference in tumor penetrance in $Nat10$/shTrp mice on treatment diet vs. normal diet.

Utilizing a similar study approach, Kodama et al. compared the genetic profiles of hepatocellular tumors developing in the context of steatosis induced by one of two mechanisms: high fat diet or homozygous $Pten$ loss (92). For their study, investigators utilized T2Onc2 mice with $Alb^{Cre/+}$-mediated induction of $RosaSbase^{LSL}$. Again, while utilization of the $Pten$ mouse strain introduced genetic variability between the two groups that is difficult to account for (the SB mice were crossed with the C;129S4-$Pten^{tm1Hwu}$/J strain), investigators found 10/30 (33%) of candidate tumor drivers making up ≥2% of total reads in tumors from the high fat diet group to overlap with those identified in the $Pten$ group. One of these genes, $Sav1$, had been previously used to predispose for cancer in an SB model but had not been identified as a candidate driver in an SB hepatocellular screen (6). Authors found $Sav1$ loss to accelerate hepatocellular carcinoma in the context of $Pten$-deficient steatotic liver $in\ vivo$ (92). These results were replicated by another group (145).

Indeed, biologically relevant connections between novel SB candidate hepatocellular tumor drivers and cellular responses to damage exist, strengthening the hypothesis that these genes mediate tumorigenesis in the context of hepatocellular injury. Candidate genes $Gli2$, $Fign$, and $Sav1$ are further discussed here.

$Gli2$ is a transcription factor activated by Hedgehog (Hh) signaling. While the Hh pathway is not active in normal hepatocytes, it can be induced in response to certain cellular stressors (146). In a mouse model of renal fibrosis, mice with myofibroblast-specific $Gli2$ deletion demonstrated cell cycle arrest and reduced fibrosis after ureteral obstruction (147). The same authors showed the ability of darinaparsin, a $GLI2$ inhibitor, to both prevent and ameliorate fibrosis in wildtype mice exposed to ureteral injury (147). The mechanism by which $Gli2$ may mediate hepatocellular cancer is less clear, although it is possible that overexpression is oncogenic by virtue of its ability to induce fibrosis and perpetuate cellular proliferation.

$Fign$ encodes an ATP-dependent microtubule severing protein. Hepatocytes rely heavily on the microtubule cytoskeleton to maintain polarity and traffic proteins to appropriate regions of the cell. Indeed, hepatocytes are uniquely multipolar, with each cell interfacing with multiple bile canaliculi and endothelial surfaces carrying sinusoidal or portal venous blood (148). Alpha-tubulin, one of the primary proteins making up the microtubule structure, is known to undergo

post-translational modifications upon exposure to ethanol metabolites (149). These modifications have been shown to decrease intracellular trafficking (149). In their review, Groebner and Tuma hypothesize that ethanol-induced microtubule modifications lead to altered lipid droplet transport, thereby contributing to the steatotic liver disease phenotype associated with ethanol consumption. *Fign* may function in a similar way, sufficiently disrupting microtubule stability to facilitate hepatic fibrosis and oncogenesis.

Sav1 is a scaffolding protein involved in the Hippo signaling pathway. Expression of *Sav1* leads to the phosphorylation and degradation of the transcription regulators Yap and Taz (150). Deletion of *Sav1* in the murine liver causes hepatomegaly and cancer after a latency of 1 year (145, 151). Mice with *Pten* deletion manifest a fatty liver phenotype prior to the development of cancer that also occurs after a prolonged latent period (92, 145, 152). Combined *Sav1* and *Pten* deletion not only accelerates HCC formation, but increases hepatic steatosis via synergistic activation of molecules downstream of Hippo and Pi3k signaling (145).

Divergences in findings among SB models of HCC underscore the significance of genetic and/or experimental variability in influencing tumorigenesis. Nevertheless, in assessing candidate tumor genes identified across SB models of hepatocellular disease, investigators have observed substantial overlap, suggesting an enrichment for particular tumor drivers in the context of chronic hepatocellular injury (70, 92). Further use of these models in the presence of specific types of hepatic injury (viral, alcoholic, fibrotic, steatotic) may increase the precision with which we are able to tease out differences and inform how we manage discrete patient populations.

Sex Bias

Male sex is a known risk factor for HCC across populations, with male to female incidence ratios averaging between 2:1 and 4:1 (153). Mouse models of HCC, including those utilizing the SB system, also demonstrate sexual dimorphism (1, 6, 20, 141). Keng et al. sought to identify sex-specific genetic drivers of HCC by comparing transposon insertion profiles of hepatic lesions occurring in a small cohort of male and female mice (154). Data from both sexes were pooled from mice harboring liver-specific activation of the SB system alone or in conjunction with *Trp53* inactivation for a total of 10 male and 9 female mice (154). Authors noted that while the majority of lesions harvested from males were classified as HCC, female nodules were histologically characterized as premalignant dysplastic lesions, adenomas, or HCC. *Egfr* was identified in nodules from all male mice evaluated (10/10), while only 22% of female mice (2/9) had tumors with *Egfr* identified as a candidate tumor driver (154). Genetic data from human HCC samples corroborated gender-specific findings, as tumors with polysomy of chromosome 7 (the location of human *EGFR*) displayed elevated *EGFR* mRNA expression and were disproportionately male (33:1 male to female ratio in the polysomy 7 subclass vs. 2:1 ratio in all other subclasses of human HCC evaluated) (154). Furthermore, authors performed an *in vivo* validation of EGFR by introducing expression vectors

of truncated *EGFR* and short hairpin RNA targeting *Trp53* into male and female *Fah*-deficient mice (154). Male mice were found to have a significantly increased number of hepatic lesions per mouse relative to female mice (154). Lesions in both sexes were categorized as HCCs by histological evaluation (154). Intriguingly, Tschida et al. reported a reduction in sexual dimorphism in their SB model of hepatic tumor formation in the context of steatosis, a trend also observed among human HCC cases when stratified by steatosis status (144).

Pronounced sex differences are known to exist in terms of hepatic response to toxins, including alcohol and viruses. These differences are, at least in part, attributable to differential expression of sex steroids and their influences on pathways involved in inflammation, lipid metabolism, and insulin response (155), (156–158). Indeed, crosstalk between steroids and EGFR is known to occur and baseline *Egfr* expression in normal mice differs between males and females (159). In a recent analysis of genetic data from The Cancer Genome Atlas (TCGA), male liver tumors were found to exhibit higher densities of single nucleotide variants (SNVs) and lower expression of mismatch repair (MMR) genes compared to female liver tumors (160). This trend was observed in several tumor types. Indeed, increased expression of *EGFR* mediates inhibition of MMR and facilitates error-prone DNA replication (161). As such, SB mouse models of cancer are able to recapitulate trends observed in human populations, reaffirming their potential utility in understanding tumorigenesis in humans.

Immune Function

In recent years, the concept of immunosurveillance as a force shaping tumor biology has emerged (162, 163). Therapies including immune checkpoint modulators, immune cell therapy, and immune-modifying agents are proving highly effective in inducing tumor regression in patients (164). Despite advances, genetic drivers of tumor immunogenicity and immune system evasion are poorly understood. Rogers et al. sought to characterize the impact of the adaptive immune system on tumorigenesis in an SB model of cancer (94). The *Rag2* deficient ($Rag2^{-/-}$) mouse is unable to produce mature B and T cells and has an increased incidence of spontaneous and carcinogen-induced cancers (165). Moreover, immunological control of tumor development has been demonstrated in $Rag2^{-/-}$ mice exposed to a chemical carcinogen (166). In their experiment, Rogers et al. bred immunocompromised ($Rag2^{-/-}$) and immunocompetent mice ($Rag2^{+/-}$) to mice with ubiquitously activated SB. Tumor latencies and multiplicities were similar in $Rag2^{-/-}$; *SB* and $Rag2^{+/-}$; *SB* mice (94). Authors offered various explanations for this including the possibility that SB-induced mutations, which cause altered gene expression rather than the creation of strong, highly mutated antigenic targets, result in less immunogenic tumors (94). Nevertheless, the genetic profiles of two out of three cancer types evaluated revealed stark differences. The E1 ubiquitin-conjugating enzyme *Uba1* was mutated exclusively in leukemias from immunocompromised mice. Clonal insertions were also observed in skin tumors and HCC from $Rag2^{-/-}$; *SB* mice

(94). HCC tumors from mice with intact adaptive immune systems ($Rag2^{+/-}$; SB) had insertions in $Rtl1$ at a higher frequency than HCCs harvested from immunocompromised mice (94).

That $Uba1$ was exclusively mutated in tumors that developed in the absence of an adaptive immune response suggests it is only an effective tumor driver in an immunocompromised microenvironment. A possible corollary to this is that $Uba1$ offers no tumorigenic advantage in an immunocompetent environment. It is feasible that mutation of $Uba1$, a ubiquitinase also known to be involved in protein folding and degradation, leads to the generation of innumerable antigenic peptides that facilitate clearance of the tumor cells in the presence of functional lymphocytes. Conversely, the predisposition for clonal selection of $Rtl1$ in an immunocompetent environment suggests it assists in immune evasion. $Rtl1$ is a retrotransposon-derived imprinted gene. It is expressed during development and is integral to maintaining the maternal-fetal interface of the placenta (167). Placental development involves marked cellular proliferation, invasion of the uterine wall, and successful evasion of the maternal immune system. Thus, clear biological and functional parallels between the placenta and cancers exist (168). Conceptually, it is within reason that $Rtl1$-mediated tumorigenesis may, at least in part, be due to a return to a fetal cell phenotype lacking immunogenic antigen expression.

These studies demonstrate that tumor microenvironments not only impact tumor incidence and kinetics, but also selection of driver mutations, supporting the hypothesis that complex genetic heterogeneity observed in human tumors could be explained by microenvironmental factors. Knowledge of context-specific genes driving tumorigenesis could guide treatment decisions in cancer care.

THE FUTURE OF SB MUTAGENESIS SCREENS

The utility of the SB system is manifest in its capacity to induce and tag mutations driving tumorigenesis in a targeted, streamlined, and unbiased manner. The simplicity of the SB system offers an ideal platform from which to make complex inquiries into cancer biology. Future modeling of SB cancers in unique genetic and microenvironmental milieus will provide a more sophisticated understanding of cancer genetics. Future directions in SB study design are explored below.

Therapeutic Interventions and Environmental Exposures

Despite epidemiological data establishing correlations between certain exposures, or "risk factors," and cancer, the mechanistic complexities mediating such risks render the biology of these relationships difficult to understand. As an example, GWAS studies suggest numerous alleles influence the link between obesity and increased risk of endometrial cancer (169, 170). Incorporation of obesity into an SB model of endometrial cancer by providing a high-fat diet during or prior to SB activation would permit comparison of insertional profiles

between obese and normal weight mice. Such interrogations could be conducted with a plethora of environmental factors including, but not limited to, UV exposure and skin cancers, pesticides and leukemias, or androgens and prostate cancer (**Figure 3A**). Multi-drug treatments and resulting relapses could also be systematically explored. To interrogate modulations in immune function, factors such as illness or pharmaceutical intervention (e.g., transplant anti-rejection drugs or HAART therapy) could be directly incorporated into experimental models to interrogate effects on tumor genetics (**Figure 3B**) (171, 172). Recently, the ability of experimental mice, housed in clean, low-pathogen facilities, to accurately recapitulate human immune responses has been brought into question (173). This is an intriguing idea in light of the enormous amount of scientific data emerging on the interplay among pathogens, the microbiome, and human physiology. Furthermore, preliminary studies suggest that pathogen exposure leads to maturation of the mouse immune response, with the genetic expression profile of peripheral mononuclear cells more closely resembling that of adult humans (174, 175). In the presence of such biologically relevant exposures, investigators may be able to better understand the biology of current treatment challenges and therapeutic roadblocks.
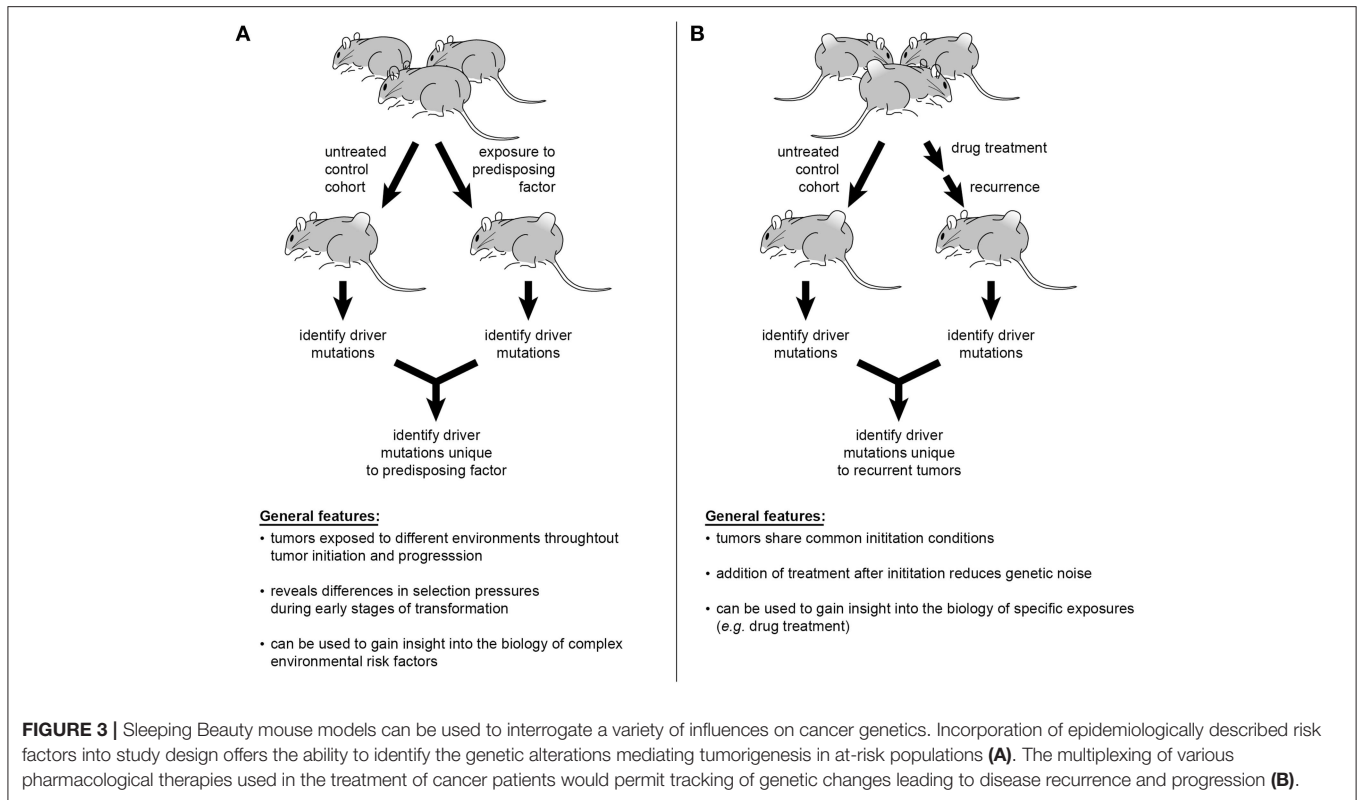
Inherent Traits

Inbred mouse strains allow scientists to control for confounding genetic variability in experiments. While essential to establishing causality of a specific mutation or exposure, it removes population diversity. The incidence of spontaneous and carcinogen-induced cancers is known to vary significantly among inbred strains, underscoring the importance of genetic background and polymorphisms on tumorigenesis (176–180). Exploiting polymorphisms in genetically engineered mouse models of cancer may facilitate characterization of cancer phenotypes observed in human populations (181). Twin studies have shown that, for some spontaneous cancers, heritability contributes to causation by 27–42%-higher than that accounted for by known single-gene familial cancer syndromes (182). Indeed, while influence of germline polymorphisms on cancer risk is recognized, the impact of germline alleles on driver mutation selection is far less clear. Since introduction of a single sensitizing mutation can influence candidate tumor driver selection (as discussed in the "Sensitizing mutations in SB mouse models" section above), we speculate that altering the entire genetic background could have profound effects on SB candidate tumor driver profiles. Experimental designs could be conducted using mice from different strains engineered to harbor the SB system and mutational landscape assessed. Differential findings in candidate tumor drivers could then be compared to known allelic variants between strains.

Genetics of Metastasis

Despite metastasis being the primary source of morbidity and mortality for many cancer types, the genetic drivers of metastasis are poorly understood. Moreover, the inherent complexity of tumors limits our ability to detect novel drivers of metastasis even when patient primaries and matched metastases are available for

FIGURE 3 | Sleeping Beauty mouse models can be used to interrogate a variety of influences on cancer genetics. Incorporation of epidemiologically described risk factors into study design offers the ability to identify the genetic alterations mediating tumorigenesis in at-risk populations **(A)**. The multiplexing of various pharmacological therapies used in the treatment of cancer patients would permit tracking of genetic changes leading to disease recurrence and progression **(B)**.

deep sequencing (183–185). In SB models, few shared transposon insertions is indicative of an early seeding event and independent tumor development while many shared insertions suggests clonal evolution of the primary tumor with subsequent dissemination. To date, comparative analyses have been conducted on metastatic lesions from SB models of medulloblastoma and osteosarcoma (3, 8). Wu et al. found little overlap between medulloblastoma primaries and metastases. The presence of known oncogenes such as *Notch2* and *Tert* in primary tumors and not metastatic lesions led authors to conclude that dissemination was an early event, occurring prior to the development of mutations in these genes (3). Retroviral vector assays have been used to validate *Eras*, *Lhx1*, *Ccrk*, and *Akt* as mediators of medulloblastoma dissemination/metastasis *in vivo* (3, 186). Moriarity et al. observed transposon insertional profiles in osteosarcoma that suggested a heterogeneous nature of metastasis (8). While 65% of genes (43/66) were unique to metastatic lesions, analysis of mice with three or more lesions suggested dissemination occurred at different points of tumorigenesis in different mice. Indeed, two mice appeared to have more than one metastatic seeding event (8). In future studies, the ability to turn off SB transposition,

such as through introduction of another recombinase within the SBase, could provide information on the temporality of metastasis and the genes driving it. Techniques such as laser microdissection or single cell sorting could also refine our understanding of tumor heterogeneity and metastasis (72, 187).

Ultimately, SB models of cancer are capable of interrogating microenvironmental factors in a multiplexed fashion that more closely approximates conditions under which human cancers develop and spread. SB experiments conducted to date reveal the tumor microenvironment to have pronounced ramifications on the genetic evolution of various tumor types. Continued application of the SB model to explore genetic heterogeneity of tumorigenesis promises to influence future therapeutic advancements in the treatment of human cancers.

AUTHOR CONTRIBUTIONS

AG-Y and CF performed the literature review. AD and AG-Y designed the figures. AG-Y wrote and all co-authors edited the manuscript.

REFERENCES

1. Keng VW, Villanueva A, Chiang DY, Dupuy AJ, Ryan BJ, Matise I, et al. A conditional transposon-based insertional mutagenesis screen for genes associated with mouse hepatocellular carcinoma. *Nat Biotechnol.* (2009) 27:264–74. doi: 10.1038/nbt.1526

2. Berquam-Vrieze KE, Nannapaneni K, Brett BT, Holmfeldt L, Ma J, Zagorodna O, et al. Cell of origin strongly influences genetic selection in a mouse model of T-ALL. *Blood.* (2011) 118:4646–56. doi: 10.1182/blood-2011-03-343947

3. Wu X, Northcott PA, Dubuc A, Dupuy AJ, Shih DJ, Witt H, et al. Clonal selection drives genetic divergence of metastatic medulloblastoma. *Nature.* (2012) 482:529–33. doi: 10.1038/nature10825

4. Rahrmann EP, Watson AL, Keng VW, Choi K, Moriarity BS, Beckmann DA, et al. Forward genetic screen for malignant peripheral nerve sheath tumor formation identifies new genes and pathways driving tumorigenesis. *Nat Genet.* (2013) 45:756–66. doi: 10.1038/ng.2641

5. Riordan JD, Keng VW, Tschida BR, Scheetz TE, Bell JB, Podetz-Pedersen KM, et al. Identification of rtl1, a retrotransposon-derived imprinted gene, as a novel driver of hepatocarcinogenesis. *PLoS Genet.* (2013) 9:e1003441. doi: 10.1371/journal.pgen.1003441

6. Bard-Chapeau EA, Nguyen AT, Rust AG, Sayadi A, Lee P, Chua BQ, et al. Transposon mutagenesis identifies genes driving hepatocellular carcinoma in a chronic hepatitis B mouse model. *Nat Genet.* (2014) 46:24–32. doi: 10.1038/ng.2847

7. Mann MB, Black MA, Jones DJ, Ward JM, Yew CC, Newberg JY, et al. Transposon mutagenesis identifies genetic drivers of Braf(V600E) melanoma. *Nat Genet.* (2015) 47:486–95. doi: 10.1038/ng.3275

8. Moriarity BS, Otto GM, Rahrmann EP, Rathe SK, Wolf NK, Weg MT, et al. A Sleeping Beauty forward genetic screen identifies new genes and pathways driving osteosarcoma development and metastasis. *Nat Genet.* (2015) 47:615–24. doi: 10.1038/ng.3293

9. 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. doi: 10.1016/S0092-8674(00)80436-5

10. Dupuy AJ, Fritz S, Largaespada DA. Transposition and gene disruption in the male germline of the mouse. *Genesis.* (2001) 30:82–8. doi: 10.1002/gene.1037

11. Fischer SE, Wienholds E, Plasterk RH. Regulated transposition of a fish transposon in the mouse germ line. *Proc Natl Acad Sci USA.* (2001) 98:6759–64. doi: 10.1073/pnas.121569298

12. Horie K, Kuroiwa A, Ikawa M, Okabe M, Kondoh G, Matsuda Y, et al. Efficient chromosomal transposition of a Tc1/mariner- like transposon Sleeping Beauty in mice. *Proc Natl Acad Sci USA.* (2001) 98:9191–6. doi: 10.1073/pnas.161071798

13. Collier LS, Carlson CM, Ravimohan S, Dupuy AJ, Largaespada DA. Cancer gene discovery in solid tumours using transposon-based somatic mutagenesis in the mouse. *Nature.* (2005) 436:272–6. doi: 10.1038/nature03681

14. 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. doi: 10.1038/nature03691

15. Bishop JO. Chromosomal insertion of foreign DNA. *Reprod Nutr Dev.* (1996) 36:607–18.

16. Okabe M, Ikawa M, Kominami K, Nakanishi T, Nishimune Y. 'Green mice' as a source of ubiquitous green cells. *FEBS Lett.* (1997) 407:313–9. doi: 10.1016/S0014-5793(97)00313-X

17. Geurts AM, Yang Y, Clark KJ, Liu G, Cui Z, Dupuy AJ, et al. Gene transfer into genomes of human cells by the sleeping beauty transposon system. *Mol Ther.* (2003) 8:108–17. doi: 10.1016/S1525-0016(03)00099-6

18. Collier LS, Adams DJ, Hackett CS, Bendzick LE, Akagi K, Davies MN, et al. Whole-body sleeping beauty mutagenesis can cause penetrant leukemia/lymphoma and rare high-grade glioma without associated embryonic lethality. *Cancer Res.* (2009) 69:8429–37. doi: 10.1158/0008-5472.CAN-09-1760

19. Soriano P. Generalized lacZ expression with the ROSA26 Cre reporter strain. *Nat Genet.* (1999) 21:70–1. doi: 10.1038/5007

20. Dupuy AJ, Rogers LM, Kim J, Nannapaneni K, Starr TK, Liu P, et al. A modified sleeping beauty transposon system that can be used to model a wide variety of human cancers in mice. *Cancer Res.* (2009) 69:8150–6. doi: 10.1158/0008-5472.CAN-09-1135

21. Vassiliou GS, Cooper JL, Rad R, Li J, Rice S, Uren A, et al. Mutant nucleophosmin and cooperating pathways drive leukemia initiation and progression in mice. *Nat Genet.* (2011) 43:470–5. doi: 10.1038/ng.796

22. Perez-Mancera PA, Rust AG, van der Weyden L, Kristiansen G, Li A, Sarver AL, et al. The deubiquitinase USP9X suppresses pancreatic ductal adenocarcinoma. *Nature.* (2012) 486:266–70. doi: 10.1038/nature11114

23. Quintana RM, Dupuy AJ, Bravo A, Casanova ML, Alameda JP, Page A, et al. A transposon-based analysis of gene mutations related to skin cancer development. *J Invest Dermatol.* (2013) 133:239–48. doi: 10.1038/jid.2012.245

24. Vyazunova I, Maklakova VI, Berman S, De I, Steffen MD, Hong W, et al. Sleeping Beauty mouse models identify candidate genes involved in gliomagenesis. *PLoS ONE.* (2014) 9:e113489. doi: 10.1371/journal.pone.0113489

25. Grandi FC, Rosser JM, Newkirk SJ, Yin J, Jiang X, Xing Z, et al. Retrotransposition creates sloping shores: a graded influence of hypomethylated CpG islands on flanking CpG sites. *Genome Res.* (2015) 25:1135–46. doi: 10.1101/gr.185132.114

26. Rad R, Rad L, Wang W, Cadinanos J, Vassiliou G, Rice S, et al. PiggyBac transposon mutagenesis: a tool for cancer gene discovery in mice. *Science.* (2010) 330:1104–7. doi: 10.1126/science.1193004

27. Narayanavari SA, Chilkunda SS, Ivics Z, Izsvak Z. Sleeping Beauty transposition: from biology to applications. *Crit Rev Biochem Mol Biol.* (2017) 52:18–44. doi: 10.1080/10409238.2016.1237935

28. Izsvak Z, Khare D, Behlke J, Heinemann U, Plasterk RH, Ivics Z. Involvement of a bifunctional, paired-like DNA-binding domain and a transpositional enhancer in Sleeping Beauty transposition. *J Biol Chem.* (2002) 277:34581–8. doi: 10.1074/jbc.M204001200

29. Cui Z, Geurts AM, Liu G, Kaufman CD, Hackett PB. Structure-function analysis of the inverted terminal repeats of the sleeping beauty transposon. *J Mol Biol.* (2002) 318:1221–35. doi: 10.1016/S0022-2836(02)00237-1

30. Zayed H, Izsvak Z, Walisko O, Ivics Z. Development of hyperactive sleeping beauty transposon vectors by mutational analysis. *Mol Ther.* (2004) 9:292–304. doi: 10.1016/j.ymthe.2003.11.024

31. Walisko O, Schorn A, Rolfs F, Devaraj A, Miskey C, Izsvak Z, et al. Transcriptional activities of the Sleeping Beauty transposon and shielding its genetic cargo with insulators. *Mol Ther.* (2008) 16:359–69. doi: 10.1038/sj.mt.6300366

32. Zayed H, Izsvak Z, Khare D, Heinemann U, Ivics Z. The DNA-bending protein HMGB1 is a cellular cofactor of Sleeping Beauty transposition. *Nucleic Acids Res.* (2003) 31:2313–22. doi: 10.1093/nar/gkg341

33. Yusa K, Takeda J, Horie K. Enhancement of Sleeping Beauty transposition by CpG methylation: possible role of heterochromatin formation. *Mol Cell Biol.* (2004) 24:4004–18. doi: 10.1128/MCB.24.9.4004-4018.2004

34. Wang Y, Pryputniewicz-Dobrinska D, Nagy EE, Kaufman CD, Singh M, Yant S, et al. Regulated complex assembly safeguards the fidelity of Sleeping Beauty transposition. *Nucleic Acids Res.* (2017) 45:311–26. doi: 10.1093/nar/gkw1164

35. Izsvak Z, Stuwe EE, Fiedler D, Katzer A, Jeggo PA, Ivics Z. Healing the wounds inflicted by sleeping beauty transposition by double-strand break repair in mammalian somatic cells. *Mol Cell.* (2004) 13:279–90. doi: 10.1016/S1097-2765(03)00524-0

36. Walisko O, Izsvak Z, Szabo K, Kaufman CD, Herold S, Ivics Z. Sleeping Beauty transposase modulates cell-cycle progression through interaction with Miz-1. *Proc Natl Acad Sci USA.* (2006) 103:4062–7. doi: 10.1073/pnas.0507683103

37. 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. doi: 10.1002/jgm.486

38. Vigdal TJ, Kaufman CD, Izsvak Z, Voytas DF, Ivics Z. Common physical properties of DNA affecting target site selection of sleeping beauty and other Tc1/mariner transposable elements. *J Mol Biol.* (2002) 323:441–52. doi: 10.1016/S0022-2836(02)00991-9

39. Luo G, Ivics Z, Izsvak Z, Bradley A. Chromosomal transposition of a Tc1/mariner-like element in mouse embryonic stem cells. *Proc Natl Acad Sci USA.* (1998) 95:10769–73. doi: 10.1073/pnas.95.18.10769

40. Liang Q, Kong J, Stalker J, Bradley A. Chromosomal mobilization and reintegration of Sleeping Beauty and PiggyBac transposons. *Genesis.* (2009) 47:404–8. doi: 10.1002/dvg.20508

41. Huang X, Guo H, Tammana S, Jung YC, Mellgren E, Bassi P, et al. Gene transfer efficiency and genome-wide integration profiling of Sleeping Beauty, Tol2, and piggyBac transposons in human primary T cells. *Mol Ther.* (2010) 18:1803–13. doi: 10.1038/mt.2010.141

42. de Jong J, Akhtar W, Badhai J, Rust AG, Rad R, Hilkens J, et al. Chromatin landscapes of retroviral and transposon integration profiles. *PLoS Genet.* (2014) 10:e1004250. doi: 10.1371/journal.pgen.1004250

43. Horie K, Yusa K, Yae K, Odajima J, Fischer SE, Keng VW, et al. Characterization of Sleeping Beauty transposition and its application

to genetic screening in mice. *Mol Cell Biol.* (2003) 23:9189–207. doi: 10.1128/MCB.23.24.9189-9207.2003

44. de Ridder J, Uren A, Kool J, Reinders M, Wessels L. Detecting statistically significant common insertion sites in retroviral insertional mutagenesis screens. *PLoS Comput Biol.* (2006) 2:e166. doi: 10.1371/journal.pcbi.0020166

45. Kool J, Berns A. High-throughput insertional mutagenesis screens in mice to identify oncogenic networks. *Nat Rev Cancer.* (2009) 9:389–99. doi: 10.1038/nrc2647

46. Callahan R. MMTV-induced mutations in mouse mammary tumors: their potential relevance to human breast cancer. *Breast Cancer Res Treat.* (1996) 39:33–44. doi: 10.1007/BF01806076

47. Mikkers H, Berns A. Retroviral insertional mutagenesis: tagging cancer pathways. *Adv Cancer Res.* (2003) 88:53–99. doi: 10.1016/S0065-230X(03)88255-6

48. Touw IP, Erkeland SJ. Retroviral insertion mutagenesis in mice as a comparative oncogenomics tool to identify disease genes in human leukemia. *Mol Ther.* (2007) 15:13–9. doi: 10.1038/sj.mt.6300040

49. Mattison J, van der Weyden L, Hubbard T, Adams DJ. Cancer gene discovery in mouse and man. *Biochim Biophys Acta.* (2009) 1796:140–61. doi: 10.1016/j.bbcan.2009.03.001

50. Zhang J, Markus J, Bies J, Paul T, Wolff L. Three murine leukemia virus integration regions within 100 kilobases upstream of c-myb are proximal to the 5′ regulatory region of the gene through DNA looping. *J Virol.* (2012) 86:10524–32. doi: 10.1128/JVI.01077-12

51. Coffin JM, HughesSH, Varmus HE (editors). *Retroviruses.* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press (1997).

52. Wu X, Li Y, Crise B, Burgess SM. Transcription start regions in the human genome are favored targets for MLV integration. *Science.* (2003) 300:1749–51. doi: 10.1126/science.1083413

53. Berry C, Hannenhalli S, Leipzig J, Bushman FD. Selection of target sites for mobile DNA integration in the human genome. *PLoS Comput Biol.* (2006) 2:e157. doi: 10.1371/journal.pcbi.0020157

54. Wang Y, Wang J, Devaraj A, Singh M, Jimenez Orgaz A, Chen JX, et al. Suicidal autointegration of sleeping beauty and piggyBac transposons in eukaryotic cells. *PLoS Genet.* (2014) 10:e1004103. doi: 10.1371/journal.pgen.1004103

55. Gogol-Doring A, Ammar I, Gupta S, Bunse M, Miskey C, Chen W, et al. Genome-wide profiling reveals remarkable parallels between insertion site selection properties of the MLV retrovirus and the piggyBac transposon in primary human CD4+ T cells. *Mol Ther.* (2016) 24:592–606. doi: 10.1038/mt.2016.11

56. Friedrich MJ, Rad L, Bronner IF, Strong A, Wang W, Weber J, et al. Genome-wide transposon screening and quantitative insertion site sequencing for cancer gene discovery in mice. *Nat Protoc.* (2017) 12:289–309. doi: 10.1038/nprot.2016.164

57. Mann KM, Ward JM, Yew CC, Kovochich A, Dawson DW, Black MA, et al. Sleeping Beauty mutagenesis reveals cooperating mutations and pathways in pancreatic adenocarcinoma. *Proc Natl Acad Sci USA.* (2012) 109:5934–41. doi: 10.1073/pnas.1202490109

58. Rad R, Rad L, Wang W, Strong A, Ponstingl H, Bronner IF, et al. A conditional piggyBac transposition system for genetic screening in mice identifies oncogenic networks in pancreatic cancer. *Nat Genet.* (2015) 47:47–56. doi: 10.1038/ng.3164

59. Largaespada DA, Collier LS. Transposon-mediated mutagenesis in somatic cells: identification of transposon-genomic DNA junctions. *Methods Mol Biol.* (2008) 435:95–108. doi: 10.1007/978-1-59745-232-8_7

60. Brett BT, Berquam-Vrieze KE, Nannapaneni K, Huang J, Scheetz TE, Dupuy AJ. Novel molecular and computational methods improve the accuracy of insertion site analysis in Sleeping Beauty-induced tumors. *PLoS ONE.* (2011) 6:e24668. doi: 10.1371/journal.pone.0024668

61. Koudijs MJ, Klijn C, van der Weyden L, Kool J, ten Hoeve J, Sie D, et al. High-throughput semiquantitative analysis of insertional mutations in heterogeneous tumors. *Genome Res.* (2011) 21:2181–9. doi: 10.1101/gr.112763.110

62. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* (2009) 10:R25. doi: 10.1186/gb-2009-10-3-r25

63. Riordan JD, Drury LJ, Smith RP, Brett BT, Rogers LM, Scheetz TE, et al. Sequencing methods and datasets to improve functional interpretation of sleeping beauty mutagenesis screens. *BMC Genomics.* (2014) 15:1150. doi: 10.1186/1471-2164-15-1150

64. Mikkers H, Allen J, Knipscheer P, Romeijn L, Hart A, Vink E, et al. High-throughput retroviral tagging to identify components of specific signaling pathways in cancer. *Nat Genet.* (2002) 32:153–9. doi: 10.1038/ng950

65. de Ridder J, Kool J, Uren A, Bot J, Wessels L, Reinders M. Co-occurrence analysis of insertional mutagenesis data reveals cooperating oncogenes. *Bioinformatics.* (2007) 23:i133–141. doi: 10.1093/bioinformatics/btm202

66. Bergemann TL, Starr TK, Yu H, Steinbach M, Erdmann J, Chen Y, et al. New methods for finding common insertion sites and co-occurring common insertion sites in transposon- and virus-based genetic screens. *Nucleic Acids Res.* (2012) 40:3822–33. doi: 10.1093/nar/gkr1295

67. Sarver AL, Erdman J, Starr T, Largaespada DA, Silverstein KA. TAPDANCE: an automated tool to identify and annotate transposon insertion CISs and associations between CISs from next generation sequence data. *BMC Bioinformatics.* (2012) 13:154. doi: 10.1186/1471-2105-13-154

68. March HN, Rust AG, Wright NA, ten Hoeve J, de Ridder J, Eldridge M, et al. Insertional mutagenesis identifies multiple networks of cooperating genes driving intestinal tumorigenesis. *Nat Genet.* (2011) 43:1202–9. doi: 10.1038/ng.990

69. Temiz NA, Moriarity BS, Wolf NK, Riordan JD, Dupuy AJ, Largaespada DA, et al. RNA sequencing of Sleeping Beauty transposon-induced tumors detects transposon-RNA fusions in forward genetic cancer screens. *Genome Res.* (2016) 26:119–29. doi: 10.1101/gr.188649.114

70. Riordan JD, Feddersen CR, Tschida BR, Jackson P, Keng VW, Linden MA, et al. Chronic liver injury alters driver mutation profiles in hepatocellular carcinoma. *Hepatology.* (2017) 67:924–39. doi: 10.1002/hep.29565

71. Mann MB, Jenkins NA, Copeland NG, Mann KM. Sleeping Beauty mutagenesis: exploiting forward genetic screens for cancer gene discovery. *Curr Opin Genet Dev.* (2014) 24:16–22. doi: 10.1016/j.gde.2013.11.004

72. Moriarity BS, Largaespada DA. Sleeping Beauty transposon insertional mutagenesis based mouse models for cancer gene discovery. *Curr Opin Genet Dev.* (2015) 30:66–72. doi: 10.1016/j.gde.2015.04.007

73. Abbott KL, Nyre ET, Abrahante J, Ho YY, 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.* (2015) 43:D844–848. doi: 10.1093/nar/gku770

74. 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:D1011–7. doi: 10.1093/nar/gkx956

75. Perna D, Karreth FA, Rust AG, Perez-Mancera PA, Rashid M, Iorio F, et al. BRAF inhibitor resistance mediated by the AKT pathway in an oncogenic BRAF mouse melanoma model. *Proc Natl Acad Sci USA.* (2015) 112:E536–545. doi: 10.1073/pnas.1418163112

76. Karreth FA, Tay Y, Perna D, Ala U, Tan SM, Rust AG, et al. *In vivo* identification of tumor- suppressive PTEN ceRNAs in an oncogenic BRAF-induced mouse model of melanoma. *Cell.* (2011) 147:382–95. doi: 10.1016/j.cell.2011.09.032

77. Starr TK, Scott PM, Marsh BM, Zhao L, Than BL, O'Sullivan MG, 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) 108:5765–70. doi: 10.1073/pnas.1018012108

78. Takeda H, Wei Z, Koso H, Rust AG, Yew CC, Mann MB, et al. Transposon mutagenesis identifies genes and evolutionary forces driving gastrointestinal tract tumor progression. *Nat Genet.* (2015) 47:142–50. doi: 10.1038/ng.3175

79. van der Weyden L, Arends MJ, Rust AG, Poulogiannis G, McIntyre RE, Adams DJ. Increased tumorigenesis associated with loss of the tumor suppressor gene Cadm1. *Mol Cancer.* (2012) 11:29. doi: 10.1186/1476-4598-11-29

80. Loeb KR, Hughes BT, Fissel BM, Osteen NJ, Knoblaugh SE, Grim JE, et al. Insertional mutagenesis using the Sleeping Beauty transposon system identifies drivers of erythroleukemia in mice. *Sci Rep.* (2019) 9:5488. doi: 10.1038/s41598-019-41805-x

81. Kas SM, de Ruiter JR, Schipper K, Annunziato S, Schut E, Klarenbeek S, et al. Insertional mutagenesis identifies drivers of a novel oncogenic pathway in invasive lobular breast carcinoma. *Nat Genet.* (2017) 49:1219–30. doi: 10.1038/ng.3905

82. Chen L, Jenjaroenpun P, Pillai AM, Ivshina AV, Ow GS, Efthimios M, et al. Transposon insertional mutagenesis in mice identifies human breast cancer susceptibility genes and signatures for stratification. *Proc Natl Acad Sci USA.* (2017) 114:E2215–e2224. doi: 10.1073/pnas.1701512114

83. Montero-Conde C, Leandro-Garcia LJ, Chen X, Oler G, Ruiz-Llorente S, Ryder M, et al. Transposon mutagenesis identifies chromatin modifiers cooperating with Ras in thyroid tumorigenesis and detects ATXN7 as a cancer gene. *Proc Natl Acad Sci USA.* (2017) 114:E4951–e4960. doi: 10.1073/pnas.1702723114

84. Tang JZ, Carmichael CL, Shi W, Metcalf D, Ng AP, Hyland CD, et al. Transposon mutagenesis reveals cooperation of ETS family transcription factors with signaling pathways in erythro-megakaryocytic leukemia. *Proc Natl Acad Sci USA.* (2013) 110:6091–6. doi: 10.1073/pnas.1304234110

85. O'Donnell KA, Keng VW, York B, Reineke EL, Seo D, Fan D, et al. A Sleeping Beauty mutagenesis screen reveals a tumor suppressor role for Ncoa2/Src-2 in liver cancer. *Proc Natl Acad Sci USA.* (2012) 109:E1377–1386. doi: 10.1073/pnas.1115433109

86. Bender AM, Collier LS, Rodriguez FJ, Tieu C, Larson JD, Halder C, et al. Sleeping beauty-mediated somatic mutagenesis implicates CSF1 in the formation of high-grade astrocytomas. *Cancer Res.* (2010) 70:3557–65. doi: 10.1158/0008-5472.CAN-09-4674

87. Dorr C, Janik C, Weg M, Been RA, Bader J, Kang R, et al. Transposon mutagenesis screen identifies potential lung cancer drivers and CUL3 as a tumor suppressor. *Mol Cancer Res.* (2015) 13:1238–47. doi: 10.1158/1541-7786.MCR-14-0674-T

88. Genovesi LA, Ng CG, Davis MJ, Remke M, Taylor MD, Adams DJ, 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. doi: 10.1073/pnas.1318639110

89. Lastowska M, Al-Afghani H, Al-Balool HH, Sheth H, Mercer E, Coxhead JM, et al. Identification of a neuronal transcription factor network involved in medulloblastoma development. *Acta Neuropathol Commun.* (2013) 1:35. doi: 10.1186/2051-5960-1-35

90. Ahmad I, Mui E, Galbraith L, Patel R, Tan EH, Salji M, et al. Sleeping Beauty screen reveals Pparg activation in metastatic prostate cancer. *Proc Natl Acad Sci USA.* (2016) 113:8290–5. doi: 10.1073/pnas.1601571113

91. Rangel R, Lee SC, Hon-Kim Ban K, Guzman-Rojas L, Mann MB, Newberg JY, et al. Transposon mutagenesis identifies genes that cooperate with mutant Pten in breast cancer progression. *Proc Natl Acad Sci USA.* (2016) 113:E7749–e7758. doi: 10.1073/pnas.1613859113

92. Kodama T, Yi J, Newberg JY, Tien JC, Wu H, Finegold MJ, et al. Molecular profiling of nonalcoholic fatty liver disease-associated hepatocellular carcinoma using SB transposon mutagenesis. *Proc Natl Acad Sci USA.* (2018) 115:E10417–e10426. doi: 10.1073/pnas.1808968115

93. Beckmann PJ, Larson JD, Larsson AT, Ostergaard JP, Wagner S, Rahrmann EP, et al. Sleeping beauty insertional mutagenesis reveals important genetic drivers of central nervous system embryonal tumors. *Cancer Res.* (2019) 79:905–17. doi: 10.1158/0008-5472.CAN-18-1261

94. Rogers LM, Olivier AK, Meyerholz DK, Dupuy AJ. Adaptive immunity does not strongly suppress spontaneous tumors in a Sleeping Beauty model of cancer. *J Immunol.* (2013) 190:4393–9. doi: 10.4049/jimmunol.1203227

95. van der Weyden L, Papaspyropoulos A, Poulogiannis G, Rust AG, Rashid M, Adams DJ, et al. Loss of RASSF1A synergizes with deregulated RUNX2 signaling in tumorigenesis. *Cancer Res.* (2012) 72:3817–27. doi: 10.1158/0008-5472.CAN-11-3343

96. 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:E2057–2065. doi: 10.1073/pnas.1603223113

97. Heltemes-Harris LM, Larson JD, Starr TK, Hubbard GK, Sarver AL, Largaespada DA, et al. Sleeping Beauty transposon screen identifies signaling modules that cooperate with STAT5 activation to induce B-cell acute lymphoblastic leukemia. *Oncogene.* (2016) 35:3454–64. doi: 10.1038/onc.2015.405

98. Zanesi N, Balatti V, Riordan J, Burch A, Rizzotto L, Palamarchuk A, et al. A Sleeping Beauty screen reveals NF-kB activation in CLL mouse model. *Blood.* (2013) 121:4355–8. doi: 10.1182/blood-2013-02-486035

99. Morris SM, Davison J, Carter KT, O'Leary RM, Trobridge P, Knoblaugh SE, 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) 140:853–63. doi: 10.1002/ijc.30491

100. van der Weyden L, Rust AG, McIntyre RE, Robles-Espinoza CD, del Castillo Velasco-Herrera M, Strogantsev R, et al. Jdp2 downregulates Trp53 transcription to promote leukaemogenesis in the context of Trp53 heterozygosity. *Oncogene.* (2013) 32:397–402. doi: 10.1038/onc.2012.56

101. Koso H, Tsuhako A, Lyons E, Ward JM, Rust AG, Adams DJ, et al. Identification of FoxR2 as an oncogene in medulloblastoma. *Cancer Res.* (2014) 74:2351–61. doi: 10.1158/0008-5472.CAN-13-1523

102. Suarez-Cabrera C, Quintana RM, Bravo A, Casanova ML, Page A, Alameda JP, et al. A Transposon-based analysis reveals RASA1 is involved in triple-negative breast cancer. *Cancer Res.* (2017) 77:1357–68. doi: 10.1158/0008-5472.CAN-16-1586

103. Rahrmann EP, Wolf NK, Otto GM, Heltemes-Harris L, Ramsey LB, Shu J, et al. Sleeping beauty screen identifies RREB1 and other genetic drivers in human B-cell lymphoma. *Mol Cancer Res.* (2019) 17:567–82. doi: 10.1158/1541-7786.MCR-18-0582

104. Giotopoulos G, van der Weyden L, Osaki H, Rust AG, Gallipoli P, Meduri E, et al. A novel mouse model identifies cooperating mutations and therapeutic targets critical for chronic myeloid leukemia progression. *J Exp Med.* (2015) 212:1551–69. doi: 10.1084/jem.20141661

105. van der Weyden L, Giotopoulos G, Rust AG, Matheson LS, van Delft FW, Kong J, et al. Modeling the evolution of ETV6-RUNX1-induced B-cell precursor acute lymphoblastic leukemia in mice. *Blood.* (2011) 118:1041–51. doi: 10.1182/blood-2011-02-338848

106. van der Weyden L, Giotopoulos G, Wong K, Rust AG, Robles-Espinoza CD, Osaki H, et al. Somatic drivers of B-ALL in a model of ETV6-RUNX1; Pax5(+/-) leukemia. *BMC Cancer.* (2015) 15:585. doi: 10.1186/s12885-015-1586-1

107. Lu TX, Young KH, Xu W, Li JY. TP53 dysfunction in diffuse large B-cell lymphoma. *Crit Rev Oncol Hematol.* (2016) 97:47–55. doi: 10.1016/j.critrevonc.2015.08.006

108. Xu P, Liu X, Ouyang J, Chen B. TP53 mutation predicts the poor prognosis of non-Hodgkin lymphomas: Evidence from a meta-analysis. *PLoS ONE.* (2017) 12:e0174809. doi: 10.1371/journal.pone.0174809

109. Sakhdari A, Ok CY, Patel KP, Kanagal-Shamanna R, Yin CC, Zuo Z, et al. TP53 mutations are common in mantle cell lymphoma, including the indolent leukemic non-nodal variant. *Ann Diagn Pathol.* (2019) 41:38–42. doi: 10.1016/j.anndiagpath.2019.05.004

110. Murray PG, Qiu GH, Fu L, Waites ER, Srivastava G, Heys D, et al. Frequent epigenetic inactivation of the RASSF1A tumor suppressor gene in Hodgkin's lymphoma. *Oncogene.* (2004) 23:1326–31. doi: 10.1038/sj.onc.1207313

111. Fu L, Gao Z, Zhang X, Tsang YH, Goh HK, Geng H, et al. Frequent concomitant epigenetic silencing of the stress-responsive tumor suppressor gene CADM1, and its interacting partner DAL-1 in nasal NK/T-cell lymphoma. *Int J Cancer.* (2009) 124:1572–8. doi: 10.1002/ijc.24123

112. Paulsson K, An Q, Moorman AV, Parker H, Molloy G, Davies T, et al. Methylation of tumour suppressor gene promoters in the presence and absence of transcriptional silencing in high hyperdiploid acute lymphoblastic leukaemia. *Br J Haematol.* (2009) 144:838–47. doi: 10.1111/j.1365-2141.2008.07523.x

113. Nakahata S, Morishita K. CADM1/TSLC1 is a novel cell surface marker for adult T-cell leukemia/lymphoma. *J Clin Exp Hematop.* (2012) 52:17–22. doi: 10.3960/jslrt.52.17

114. Yuki A, Shinkuma S, Hayashi R, Fujikawa H, Kato T, Homma E, et al. CADM1 is a diagnostic marker in early-stage mycosis fungoides: multicenter study of 58 cases. *J Am Acad Dermatol.* (2018) 79:1039–46. doi: 10.1016/j.jaad.2018.06.025

115. Kuster L, Grausenburger R, Fuka G, Kaindl U, Krapf G, Inthal A, et al. ETV6/RUNX1-positive relapses evolve from an ancestral clone and frequently acquire deletions of genes implicated in glucocorticoid signaling. *Blood.* (2011) 117:2658–67. doi: 10.1182/blood-2010-03-275347

116. Grausenburger R, Bastelberger S, Eckert C, Kauer M, Stanulla M, Frech C, et al. Genetic alterations in glucocorticoid signaling pathway components are associated with adverse prognosis in children with relapsed ETV6/RUNX1-positive acute lymphoblastic leukemia. *Leuk Lymphoma.* (2016) 57:1163–73. doi: 10.3109/10428194.2015.1088650

117. Emadali A, Hoghoughi N, Duley S, Hajmirza A, Verhoeyen E, Cosset FL, et al. Haploinsufficiency for NR3C1, the gene encoding the glucocorticoid receptor, in blastic plasmacytoid dendritic cell neoplasms. *Blood.* (2016) 127:3040–53. doi: 10.1182/blood-2015-09-671040

118. Halder G, Johnson RL. Hippo signaling: growth control and beyond. *Development.* (2011) 138:9–22. doi: 10.1242/dev.045500

119. Piccolo S, Dupont S, Cordenonsi M. The biology of YAP/TAZ: hippo signaling and beyond. *Physiol Rev.* (2014) 94:1287–312. doi: 10.1152/physrev.00005.2014

120. Sebio A, Lenz HJ. Molecular Pathways: Hippo Signaling, a Critical Tumor Suppressor. *Clin Cancer Res.* (2015) 21:5002–7. doi: 10.1158/1078-0432.CCR-15-0411

121. Piu F, Aronheim A, Katz S, Karin M. AP-1 repressor protein JDP-2: inhibition of UV-mediated apoptosis through p53 down-regulation. *Mol Cell Biol.* (2001) 21:3012–24. doi: 10.1128/MCB.21.9.3012-3024.2001

122. Walker A, Marcucci G. Molecular prognostic factors in cytogenetically normal acute myeloid leukemia. *Expert Rev Hematol.* (2012) 5:547–58. doi: 10.1586/ehm.12.45

123. Kornblau SM, McCue D, Singh N, Chen W, Estrov Z, Coombes KR. Recurrent expression signatures of cytokines and chemokines are present and are independently prognostic in acute myelogenous leukemia and myelodysplasia. *Blood.* (2010) 116:4251–61. doi: 10.1182/blood-2010-01-262071

124. Levis M. FLT3 mutations in acute myeloid leukemia: what is the best approach in 2013? *Hematol Am Soc Hematol Educ Program.* (2013) 2013:220–6. doi: 10.1182/asheducation-2013.1.220

125. Boddu P, Kantarjian H, Borthakur G, Kadia T, Daver N, Pierce S, et al. Co-occurrence of FLT3-TKD and NPM1 mutations defines a highly favorable prognostic AML group. *Blood Adv.* (2017) 1:1546–50. doi: 10.1182/bloodadvances.2017009019

126. Choudhary C, Brandts C, Schwable J, Tickenbrock L, Sargin B, Ueker A, et al. Activation mechanisms of STAT5 by oncogenic Flt3-ITD. *Blood.* (2007) 110:370–4. doi: 10.1182/blood-2006-05-024018

127. Mupo A, Celani L, Dovey O, Cooper JL, Grove C, Rad R, et al. A powerful molecular synergy between mutant Nucleophosmin and Flt3-ITD drives acute myeloid leukemia in mice. *Leukemia.* (2013) 27:1917–20. doi: 10.1038/leu.2013.77

128. Dovey OM, Cooper JL, Mupo A, Grove CS, Lynn C, Conte N, et al. Molecular synergy underlies the co-occurrence patterns and phenotype of NPM1-mutant acute myeloid leukemia. *Blood.* (2017) 130:1911–22. doi: 10.1182/blood-2017-01-760595

129. Jain P, Kantarjian HM, Ghorab A, Sasaki K, Jabbour EJ, Nogueras Gonzalez G, et al. Prognostic factors and survival outcomes in patients with chronic myeloid leukemia in blast phase in the tyrosine kinase inhibitor era: cohort study of 477 patients. *Cancer.* (2017) 123:4391–402. doi: 10.1002/cncr.30864

130. Koschmieder S, Gottgens B, Zhang P, Iwasaki-Arai J, Akashi K, Kutok JL, et al. Inducible chronic phase of myeloid leukemia with expansion of hematopoietic stem cells in a transgenic model of BCR-ABL leukemogenesis. *Blood.* (2005) 105:324–34. doi: 10.1182/blood-2003-12-4369

131. Voisin V, Barat C, Hoang T, Rassart E. Novel insights into the pathogenesis of the Graffi murine leukemia retrovirus. *J Virol.* (2006) 80:4026–37. doi: 10.1128/JVI.80.8.4026-4037.2006

132. NCI. *General Information About Osteosarcoma and Malignant Fibrous Histiocytoma (MFH) of Bone.* (2018). Available online at: https://www.cancer.gov/types/bone/hp/osteosarcoma-treatment-pdq#link/_1802_toc (accessed August 27, 2018).

133. Robl B, Pauli C, Botter SM, Bode-Lesniewska B, Fuchs B. Prognostic value of tumor suppressors in osteosarcoma before and after neoadjuvant chemotherapy. *BMC Cancer.* (2015) 15:379. doi: 10.1186/s12885-015-1397-4

134. Sayles LC, Breese MR, Koehne AL, Leung SG, Lee AG, Liu HY, et al. Genome-informed targeted therapy for osteosarcoma. *Cancer Discov.* (2019) 9:46–63. doi: 10.1158/2159-8290.CD-17-1152

135. Xu H, Zhu X, Bao H, Wh Shek T, Huang Z, Wang Y, et al. Genetic and clonal dissection of osteosarcoma progression and lung metastasis. *Int J Cancer.* (2018) 143:1134–42. doi: 10.1002/ijc.31389

136. Starr TK, Allaei R, Silverstein KA, Staggs RA, Sarver AL, Bergemann TL, et al. A transposon-based genetic screen in mice identifies genes altered in colorectal cancer. *Science.* (2009) 323:1747–50. doi: 10.1126/science.1163040

137. Flaherty KT, Puzanov I, Kim KB, Ribas A, McArthur GA, Sosman JA, et al. Inhibition of mutated, activated BRAF in metastatic melanoma. *N Engl J Med.* (2010) 363:809–19. doi: 10.1056/NEJMoa1002011

138. Tolcher AW, Peng W, Calvo E. Rational approaches for combination therapy strategies targeting the MAP kinase pathway in solid tumors. *Mol Cancer Ther.* (2018) 17:3–16. doi: 10.1158/1535-7163.MCT-17-0349

139. Parkin DM. The global health burden of infection-associated cancers in the year 2002. *Int J Cancer.* (2006) 118:3030–44. doi: 10.1002/ijc.21731

140. Yang JD, Roberts LR. Hepatocellular carcinoma: a global view. *Nat Rev Gastroenterol Hepatol.* (2010) 7:448–58. doi: 10.1038/nrgastro.2010.100

141. Chisari FV, Klopchin K, Moriyama T, Pasquinelli C, Dunsford HA, Sell S, et al. Molecular pathogenesis of hepatocellular carcinoma in hepatitis B virus transgenic mice. *Cell.* (1989) 59:1145–56. doi: 10.1016/0092-8674(89)90770-8

142. Kodama T, Bard-Chapeau EA, Newberg JY, Kodama M, Rangel R, Yoshihara K, et al. Two-step forward genetic screen in mice identifies Ral GTPase-activating proteins as suppressors of hepatocellular carcinoma. *Gastroenterology.* (2016) 151:324–37.e312. doi: 10.1053/j.gastro.2016.04.040

143. Seitz HK, Stickel F. Risk factors and mechanisms of hepatocarcinogenesis with special emphasis on alcohol and oxidative stress. *Biol Chem.* (2006) 387:349–60. doi: 10.1515/BC.2006.047

144. Tschida BR, Temiz NA, Kuka TP, Lee LA, Riordan JD, Tierrablanca CA, et al. Sleeping Beauty insertional mutagenesis in mice identifies drivers of steatosis-associated hepatic tumor. *Cancer Res.* (2017) 77:6576–88. doi: 10.1158/0008-5472.CAN-17-2281

145. Jeong SH, Kim HB, Kim MC, Lee JM, Lee JH, Kim JH, et al. Hippo-mediated suppression of IRS2/AKT signaling prevents hepatic steatosis and liver cancer. *J Clin Invest.* (2018) 128:1010–25. doi: 10.1172/JCI95802

146. Hu L, Lin X, Lu H, Chen B, Bai Y. An overview of hedgehog signaling in fibrosis. *Mol Pharmacol.* (2015) 87:174–82. doi: 10.1124/mol.114.095141

147. Kramann R, Fleig SV, Schneider RK, Fabian SL, DiRocco DP, Maarouf O, et al. Pharmacological GLI2 inhibition prevents myofibroblast cell-cycle progression and reduces kidney fibrosis. *J Clin Invest.* (2015) 125:2935–51. doi: 10.1172/JCI74929

148. Treyer A, Musch A. Hepatocyte polarity. *Compr Physiol.* (2013) 3:243–87. doi: 10.1002/cphy.c120009

149. Groebner JL, Tuma PL. The altered hepatic tubulin code in alcoholic liver disease. *Biomolecules.* (2015) 5:2140–59. doi: 10.3390/biom5032140

150. Meng Z, Moroishi T, Guan KL. Mechanisms of Hippo pathway regulation. *Genes Dev.* (2016) 30:1–17. doi: 10.1101/gad.274027.115

151. Lu L, Li Y, Kim SM, Bossuyt W, Liu P, Qiu Q, et al. Hippo signaling is a potent *in vivo* growth and tumor suppressor pathway in the mammalian liver. *Proc Natl Acad Sci USA.* (2010) 107:1437–42. doi: 10.1073/pnas.0911427107

152. Horie Y, Suzuki A, Kataoka E, Sasaki T, Hamada K, Sasaki J, et al. Hepatocyte-specific Pten deficiency results in steatohepatitis and hepatocellular carcinomas. *J Clin Invest.* (2004) 113:1774–83. doi: 10.1172/JCI20513

153. El-Serag HB, Rudolph KL. Hepatocellular carcinoma: epidemiology and molecular carcinogenesis. *Gastroenterology.* (2007) 132:2557–76. doi: 10.1053/j.gastro.2007.04.061

154. Keng VW, Sia D, Sarver AL, Tschida BR, Fan D, Alsinet C, et al. Sex bias occurrence of hepatocellular carcinoma in Poly7 molecular subclass is associated with EGFR. *Hepatology.* (2013) 57:120–30. doi: 10.1002/hep.26004

155. Ma WL, Lai HC, Yeh S, Cai X, Chang, C. Androgen receptor roles in hepatocellular carcinoma, fatty liver, cirrhosis and hepatitis. *Endocr Relat Cancer.* (2014) 21:R165–82. doi: 10.1530/erc-13-0283

156. Guy J, Peters MG. Liver disease in women: the influence of gender on epidemiology, natural history, and patient outcomes. *Gastroenterol Hepatol.* (2013) 9:633–9.

157. Nguyen-Lefebvre AT, Horuzsko A. Kupffer cell metabolism and function. *J Enzymol Metab.* (2015) 1:101.

158. Kurt Z, Barrere-Cain R, LaGuardia J, Mehrabian M, Pan C, Hui ST, et al. Tissue-specific pathways and networks underlying sexual

dimorphism in non-alcoholic fatty liver disease. *Biol Sex Differ.* (2018) 9:46. doi: 10.1186/s13293-018-0205-7

159. Wang L, Xiao J, Gu W, Chen H. Sex difference of egfr expression and molecular pathway in the liver: impact on drug design and cancer treatments? *J Cancer.* (2016) 7:671–80. doi: 10.7150/jca.13684

160. Li CH, Haider S, Shiah YJ, Thai K, Boutros PC. Sex differences in cancer driver genes and biomarkers. *Cancer Res.* (2018) 78:5527–37. doi: 10.1158/0008-5472.CAN-18-0362

161. Ortega J, Li JY, Lee S, Tong D, Gu L, Li GM. Phosphorylation of PCNA by EGFR inhibits mismatch repair and promotes misincorporation during DNA synthesis. *Proc Natl Acad Sci USA.* (2015) 112:5667–72. doi: 10.1073/pnas.1417711112

162. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell.* (2011) 144:646–74. doi: 10.1016/j.cell.2011.02.013

163. Mittal D, Gubin MM, Schreiber RD, Smyth MJ. New insights into cancer immunoediting and its three component phases–elimination, equilibrium and escape. *Curr Opin Immunol.* (2014) 27:16–25. doi: 10.1016/j.coi.2014.01.004

164. Lesokhin AM, Callahan MK, Postow MA, Wolchok JD. On being less tolerant: enhanced cancer immunosurveillance enabled by targeting checkpoints and agonists of T cell activation. *Sci Transl Med.* (2015) 7:280sr281. doi: 10.1126/scitranslmed.3010274

165. Shankaran V, Ikeda H, Bruce AT, White JM, Swanson PE, Old LJ, et al. IFNgamma and lymphocytes prevent primary tumour development and shape tumour immunogenicity. *Nature.* (2001) 410:1107–11. doi: 10.1038/35074122

166. Koebel CM, Vermi W, Swann JB, Zerafa N, Rodig SJ, Old LJ, et al. Adaptive immunity maintains occult cancer in an equilibrium state. *Nature.* (2007) 450:903–7. doi: 10.1038/nature06309

167. Sekita Y, Wagatsuma H, Nakamura K, Ono R, Kagami M, Wakisaka N, et al. Role of retrotransposon-derived imprinted gene, Rtl1, in the feto-maternal interface of mouse placenta. *Nat Genet.* (2008) 40:243–8. doi: 10.1038/ng.2007.51

168. Macaulay EC, Chatterjee A, Cheng X, Baguley BC, Eccles MR, Morison IM. The genes of life and death: a potential role for placental-specific genes in cancer: active retrotransposons in the placenta encode unique functional genes that may also be used by cancer cells to promote malignancy. *Bioessays.* (2017) 39:1700091. doi: 10.1002/bies.201700091

169. Chen X, Xiang YB, Long JR, Cai H, Cai Q, Cheng J, et al. Genetic polymorphisms in obesity-related genes and endometrial cancer risk. *Cancer.* (2011) 118:3356–64. doi: 10.1002/cncr.26552

170. Delahanty RJ, Beeghly-Fadiel A, Xiang YB, Long J, Cai Q, Wen W, et al. Association of obesity-related genetic variants with endometrial cancer risk: a report from the Shanghai Endometrial Cancer Genetics Study. *Am J Epidemiol.* (2011) 174:1115–26. doi: 10.1093/aje/kwr233

171. Perrett CM, Harwood CA, McGregor JM, Karran P. Carcinogenic mechanisms related to immunosuppressive therapy. *Cancer Treat Res.* (2009) 146:123–32. doi: 10.1007/978-0-387-78574-5_11

172. Brugnaro P, Morelli E, Cattelan F, Petrucci A, Panese S, Eseme F, et al. Non-AIDS definings malignancies among human immunodeficiency virus-positive subjects: epidemiology and outcome after two decades of HAART era. *World J Virol.* (2015) 4:209–18. doi: 10.5501/wjv.v4.i3.209

173. Tao L, Reese TA. Making mouse models that reflect human immune responses. *Trends Immunol.* (2017) 38:181–93. doi: 10.1016/j.it.2016.12.007

174. Beura LK, Hamilton SE, Bi K, Schenkel JM, Odumade OA, Casey KA, et al. Normalizing the environment recapitulates adult human immune traits in laboratory mice. *Nature.* (2016) 532:512–6. doi: 10.1038/nature17655

175. Reese TA, Bi K, Kambal A, Filali-Mouhim A, Beura LK, Burger MC, et al. Sequential infection with common pathogens promotes human-like immune gene expression and altered vaccine response. *Cell Host Microbe.* (2016) 19:713–9. doi: 10.1016/j.chom.2016.04.003

176. Manenti G, Gariboldi M, Fiorino A, Zedda AI, Pierotti MA, Dragani TA. Pas1 is a common lung cancer susceptibility locus in three mouse strains. *Mamm Genome.* (1997) 8:801–4. doi: 10.1007/s003359900582

177. Bilger A, Bennett LM, Carabeo RA, Chiaverotti TA, Dvorak C, Liss KM, et al. A potent modifier of liver cancer risk on distal mouse chromosome 1: linkage analysis and characterization of congenic lines. *Genetics.* (2004) 167:859–66. doi: 10.1534/genetics.103.024521

178. Woodworth CD, Michael E, Smith L, Vijayachandra K, Glick A, Hennings H, et al. Strain-dependent differences in malignant conversion of mouse skin tumors is an inherent property of the epidermal keratinocyte. *Carcinogenesis.* (2004) 25:1771–8. doi: 10.1093/carcin/bgh170

179. Meunier C, Kwan T, Turbide C, Beauchemin N, Gros P. Genetic control of susceptibility to carcinogen-induced colorectal cancer in mice: the Ccs3 and Ccs5 loci regulate different aspects of tumorigenesis. *Cell Cycle.* (2011) 10:1739–49. doi: 10.4161/cc.10.11.15817

180. Bult CJ, Blake JA, Smith CL, Kadin JA, Richardson JE, Mouse Genome Database Group. Mouse Genome Database (MGD) 2019. *Nucleic Acids Res.* (2019) 47:D801–6. doi: 10.1093/nar/gky1056

181. Hunter KW. Mouse models of cancer: does the strain matter? *Nat Rev Cancer.* (2012) 12:144–9. doi: 10.1038/nrc3206

182. Lichtenstein P, Holm NV, Verkasalo PK, Iliadou A, Kaprio J, Koskenvuo M, et al. Environmental and heritable factors in the causation of cancer–analyses of cohorts of twins from Sweden, Denmark, and Finland. *N Engl J Med.* (2000) 343:78–85. doi: 10.1056/NEJM200007133430201

183. Lawrence MS, Stojanov P, Mermel CH, Robinson JT, Garraway LA, Golub TR, et al. Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature.* (2014) 505:495–501. doi: 10.1038/nature12912

184. Hong MK, Macintyre G, Wedge DC, Van Loo P, Patel K, Lunke S, et al. Tracking the origins and drivers of subclonal metastatic expansion in prostate cancer. *Nat Commun.* (2015) 6:6605. doi: 10.1038/ncomms7605

185. Paik PK, Shen R, Won H, Rekhtman N, Wang L, Sima CS, et al. Next-generation sequencing of stage IV squamous cell lung cancers reveals an association of PI3K aberrations and evidence of clonal heterogeneity in patients with brain metastases. *Cancer Discov.* (2015) 5:610–21. doi: 10.1158/2159-8290.CD-14-1129

186. Mumert M, Dubuc A, Wu X, Northcott PA, Chin SS, Pedone CA, et al. Functional genomics identifies drivers of medulloblastoma dissemination. *Cancer Res.* (2012) 72:4944–53. doi: 10.1158/0008-5472.CAN-12-1629

187. Copeland NG, Jenkins NA. Harnessing transposons for cancer gene discovery. *Nat Rev Cancer.* (2010) 10:696–706. doi: 10.1038/nrc2916