Published in final edited form as:

Science. 2010 November 19; 330(6007): 1104–1107. doi:10.1126/science.1193004.

PiggyBac Transposon Mutagenesis: A Tool for Cancer Gene Discovery in Mice

Roland Rad¹, Lena Rad¹, Wei Wang¹, Juan Cadinanos^{1,2}, George Vassiliou¹, Stephen Rice¹, Lia S. Campos¹, Kosuke Yusa¹, Ruby Banerjee¹, Meng Amy Li¹, Jorge de la Rosa², Alexander Strong¹, Dong Lu¹, Peter Ellis¹, Nathalie Conte¹, Fang Tang Yang¹, Pentao Liu¹, and Allan Bradley^{1,#}

¹The Wellcome Trust Sanger Institute Genome Campus Hinxton/Cambridge CB10 1SA United Kingdom

²Instituto de Medicina Oncológica y Molecular de Asturias (IMOMA) Avda. Richard Grandío s/n Oviedo, 33193 Spain

Abstract

Transposons are mobile DNA segments that can disrupt gene function by inserting in or near genes. Here we show that insertional mutagenesis by the *PiggyBac* transposon can be used for cancer gene discovery in mice. *PiggyBac* transposition in genetically engineered transposon/transposase mice induced cancers whose type (hematopoietic versus solid) and latency were dependent on the regulatory elements introduced into transposons. Analysis of 63 hematopoietic tumors revealed the unique qualities of *PiggyBac* for genome-wide mutagenesis and discovered many cancer genes not identified in previous retroviral or *Sleeping Beauty* transposon screens, including *Spic*, which encodes a PU.1-related transcription factor, and *Hdac7*, a histone deacetylase gene. *PiggyBac* and *Sleeping Beauty* have different integration preferences. To maximize the utility of the tool, we engineered 20 mouse lines to be compatible with both transposases in constitutive, tissue- or temporal-specific mutagenesis. Mice with different transposon types, copy numbers and chromosomal locations support wide applicability.

Genetic screening in higher organisms has been hampered for decades by the lack of efficient insertional mutagenesis tools. Retroviruses have been used for cancer gene discovery in mice, but their application has been limited to the study of hematopoietic and mammary tumors due to viral tropism for these tissues (1). DNA transposons, which are the key insertional mutagens in lower organisms, were inactivated in vertebrate genomes millions of years ago. Only recently have new transposons been engineered to be active in mammalian cells, a development that provides opportunities for their use as genetic tools in higher organisms (2). *Sleeping Beauty* (*SB*), a TC1/mariner transposon, was reconstructed from dormant elements in fish genomes and optimized to transpose in multiple cell types (3), including mouse embryonic stem cells (4). Further improvements of *SB* led to its successful application for somatic mutagenesis in mice (5, 6). Another transposon, *PiggyBac* (*PB*) from the cabbage looper moth, was recently engineered to be highly active in mammalian cells (7) and has been shown to have biological properties distinct from those of *SB* (2, 7-9). *PB* can move larger DNA fragments (allowing complex transgene designs to be incorporated into the transposon) and it has a weaker tendency for local hopping *in vitro* (which makes it an attractive candidate for genome-wide mutagenesis). Furthermore, in contrast to *SB*, *PB* does not leave undesired footprint mutations after transposition. Finally, *PB* and *SB* have different integration preferences.

#Correspondence: a Bradley@sanger.ac.uk.

To deploy *PB* for genetic screening in mice, we generated *PB* transposase knockin-mice (*RosaPB*) and 19 mouse lines carrying different types of transposons (Figures 1 and S1(10)). All transposons possess *PB* and *SB* inverted terminal repeats (ITRs), allowing mobilization with both transposases. Promoter/enhancer elements, a splice donor, bidirectional SV40 polyAs and two splice acceptors were introduced in between the ITRs to allow gain or loss of function mutations, depending on the transposon orientation and its spatial relation to genes. Transgenic mice were generated with three variants of these bi-functional “activating/inactivating” transposons (ATP1, ATP2, ATP3), which carry different promoter/enhancer elements (CAG, MSCV, PGK). The chromosomal location and copy number of transposons were determined for each line by FISH and qPCR (Figures S2 and S3). Nineteen transgenic mouse lines were established, with transposon arrays on 14 different chromosomes and 2 to 150 transposon copies (Figure 1).

To analyze the biology of *PB* transposition *in vivo*, we intercrossed *RosaPB* mice with 14 *ATP* lines and assessed nearly 900 progeny. When “low copy” *ATP* lines (< 20 copies) were used, double-transgenic offspring were born at Mendelian frequency or close to it, whereas “high copy” *ATP* mice (e.g. *ATP1-H39*) did not produce live-born double-positive progeny (Figure 2A). Double-transgenic *RosaPB;ATP1-H39* embryos were however identified, suggesting extensive transposition-induced embryonic lethality (Figure S4). Less pronounced was the effect of the transposon type and donor location on embryonic lethality, although weak positional effects could be observed for some lines (e.g. *ATP1-H5*). Further characteristics of *PB* mobilization *in vivo* are shown in Figure S4.

We next investigated whether *PB*-based mutagenesis induced cancer in the mice. We monitored 393 animals from multiple cohorts and found that beginning at 2 months of age, double-transgenic *RosaPB;ATP* mice, but not single-transgenic animals, developed cancers. The tumor type and latency were profoundly influenced by the transposon type. Mobilization of ATP2 transposons resulted in highly penetrant hematopoietic cancer formation. More than 90% of *RosaPB;ATP2* mice developed aggressive leukemias and lymphomas (Figure 2B and S5). In contrast, *RosaPB;ATP1* animals had almost exclusively solid tumors, including sarcomas and various carcinomas with poor differentiation and metastasis (Figure 2C and S6; Tables S1 and S2).

To explore whether the lack of hematopoietic carcinogenesis in ATP1 mice is related to tissue-specificity of the CAG promoter, we generated reporter mice expressing DsRed constitutively under control of CAG from the *Hprt* locus. This knock-in line has two major advantages over conventional transgenic reporter mice: (i) Promoters inserted into *Hprt* are not subject to epigenetic silencing and drive gene expression in all tissues (11); (ii) Only one copy of the promoter/reporter is used, which accurately reflects the effect of one ATP1-CAG-transposon. The analysis of *Hprt-Cag-DsRed* mice revealed robust DsRed expression in most organs but not in the spleen (Figure 2E), suggesting low CAG activity in the hematopoietic compartment.

Finally, we analyzed *RosaPB;ATP3* mice, which developed solid and/or hematopoietic tumors, with a high percentage of mice having both (Figure 2D and S7). Further characterization of *PB*-induced tumors is shown as supplementary information, including immunohistochemical subtype analysis of hematologic cancers (Figure S8), evidence for lack of generalized transposon-induced genomic instability (Figure S9), proof of clonality of insertions in tumors (Figure S10), analysis of transposase expression and transposon copy numbers in tumors (Figure S11) and the distribution pattern of different transposon types in genes (Table S3).

SB has a strong tendency for local hopping, with 32-45% of all integrations mapping to the donor chromosome in *SB*-induced tumors (12, 13). To examine the potential of *PB* for genome-wide mutagenesis, we cloned more than 50,000 transposon integration sites in 63 hematopoietic tumors originating from 3 transposon mouse lines. A total of 5,590 non-redundant insertions were identified. Figure 3 and table S4 show that *PB* integrations were uniformly distributed across the genome and correlated with the size and the “TTAA” frequency (the *PB* integration site) of individual chromosomes. There was a slightly increased insertion density on chromosomes 2, 9, 11 and 19 in all three groups of mice, which might be related to the high gene density (Table S5) or the high number of oncogenic insertions (Figure S12) on those chromosomes. On donor chromosomes the number of integrations was as expected or only slightly higher and local hopping was confined to a small region (1-3Mb) around the donor locus (Figure 3). Together these results demonstrate the unique qualities of *PB* for genome-wide mutagenesis.

To discover candidate cancer genes, we used a statistical framework based on Gaussian Kernel convolution (14), which identified 72 unique common integration sites (CISs) at 67 independent loci (Figure 4; Table S6). Figure S12 visualizes the global transposon integration pattern for each chromosome at a 1Mb resolution and indicates all significant CIS. A search in the Retrovirus-Tagged-Cancer-Gene-Database (15) and in recent large-scale insertional mutagenesis studies (13, 16) revealed, as expected, that some of these CIS (or genes within CIS) have also been identified in earlier retroviral or *SB* screens (Figure 4A). Remarkably however, 42% of the CISs have not been reported before our study. Given that thousands of murine hematologic cancers have been generated and analyzed earlier using retroviral or *SB* mutagenesis, this high number of novel CIS was unexpected and underscores the potential of *PB* for genetic screening.

The integration pattern in oncogenes suggests a high selective pressure for insertions in specific regions where transposons cause gene activation. All *Nras* or *Evi1* integrations, for example, were upstream of the translation start site and transposons were sense-oriented, supporting gene expression from the unidirectional MSCV promoter (Figure 4B and S13). Some oncogenes had more complex integration patterns. *Notch1*, for example, had one insertion cluster in intron 27, leading to overexpression of an intracellular Notch1 domain like in human T-ALL with t(7;9) translocations (17), and a second cluster of truncating integrations in intron 2, leading to a “decoupled” intracellular domain that can be expressed from a cryptic *Notch1* promoter (18).

In contrast to oncogene insertions, hits in tumor suppressor genes (TSGs) were randomly distributed throughout genes (e.g. *Ikzf1* or *Pten*; Figure 4B) and were either sense- or antisense-oriented, suggesting that gene disruption is the disease-causing mechanism. In some tumors multiple integrations in TSGs occurred, leading to trapping of both alleles (Figure S14). *Pten* was frequently hit in this screen but not in previous retroviral screens, highlighting the different integration preference of these insertional mutagens.

Examples of novel candidate cancer genes with promoter insertions identified in this screen are the PU.1-related transcription factor *Spic*, the histone deacetylase *Hdac7*, the Wnt pathway component *Bcl9* or the cell cycle regulatory phosphatase *Cdc14* (Figure 4 and S13). *Spic* was hit in 9 tumors, which were all MPO-positive myeloid leukemias. In all tumors transposons were sense-oriented and located within a narrow region (760bp) upstream of the translation start site. *MSCV-Spic* fusion transcripts were confirmed by sequencing and *Spic* overexpression was shown by qRT-PCR (Figure 4C). Little is known about the biological function of *Spic*, but a recent study has suggested its role in myelomonocytic development (19), which further supports our findings. *HDAC7* is specifically expressed in CD4/CD8-double positive human thymocytes (20), but its role in

hematopoietic tumorigenesis has not been studied so far. The beta-catenin co-factor Bcl9 is a rare translocation partner in human hematopoietic malignancies (21), and a recent study revealed frequent amplifications of a chromosomal region that includes *BCL9* in various human malignancies (22).

The mouse is a major model organism for the study of human physiology and disease. Our study adds an efficient and versatile tool kit for genetic screening in mice that can be exploited for a wide range of applications, including constitutive, tissue-specific or spatio-temporally controlled somatic mutagenesis using conditional and tamoxifen-inducible *PB* (23). In tissue-specific screens we were able to induce aggressive metastatic solid tumors, which in principle will facilitate future studies of tumor initiation, progression and metastasis, as well as the genetic basis of various histological subtypes of individual cancers. *PB* can be used alone for genome-wide mutagenesis, but the *ATP* lines described here with donor loci on 14 different chromosomes can also be used in combination with *SB* (e.g., to exploit its local hopping tendency (24) for chromosome-specific screens in known cancer susceptibility regions). These transposon tool kits in mice have the potential to increase the speed and efficiency with which cancer genes and oncogenic signalling networks are discovered, and should likewise facilitate functional analysis of these genes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the Research support facility team, B.L. Ng, B. Fu and Y. Hooks for technical assistance; N. Carter and C. Lopez-Otin for experimental resources; L. Wessels, J. ten Hoeve and J. de Ridder for software used in data analysis and H. Prosser for reagents. The work was supported by the Wellcome Trust. R.R. and J.R. are recipients of a fellowship from the German Research Foundation and the Fundacion Maria-Cristina-Masaveu-Peterson, respectively.

Reference List

1. Kool J, Berns A. Nat. Rev. Cancer. 2009; 9:389. [PubMed: 19461666]
2. Ivics Z, et al. Nat. Methods. 2009; 6:415. [PubMed: 19478801]
3. Ivics Z, Hackett PB, Plasterk RH, Izsvak Z. Cell. 1997; 91:501. [PubMed: 9390559]
4. Luo G, Ivics Z, Izsvak Z, Bradley A. Proc. Natl. Acad. Sci. U. S. A. 1998; 95:10769. [PubMed: 9724779]
5. Dupuy AJ, et al. Nature. 2005; 436:221. [PubMed: 16015321]
6. Collier LS, et al. Nature. 2005; 436:272. [PubMed: 16015333]
7. Ding S, et al. Cell. 2005; 122:473. [PubMed: 16096065]
8. Wang W, et al. Proc. Natl. Acad. Sci. U. S. A. 2008; 105:9290. [PubMed: 18579772]
9. Liang Q, Kong J, Stalker J, Bradley A. Genesis. 2009; 47:404. [PubMed: 19391106]
10. Supporting material available on Science online.
11. Bronson SK, et al. Proc. Natl. Acad. Sci. U. S. A. 1996; 93:9067. [PubMed: 8799155]
12. Starr TK, et al. Science. 2009; 323:1747. [PubMed: 19251594]
13. Collier LS, et al. Cancer Res. 2009; 69:8429. [PubMed: 19843846]
14. De Ridder J, et al. PLoS. Comput. Biol. 2006; 2:e166. [PubMed: 17154714]
15. Akagi K, et al. Nucleic Acids Res. 2004; 32:D523. [PubMed: 14681473]
16. Uren AG, et al. Cell. 2008; 133:727. [PubMed: 18485879]
17. Ellisen LW, et al. Cell. 1991; 66:649. [PubMed: 1831692]
18. Hoemann CD, et al. Mol. Cell Biol. 2000; 20:3831. [PubMed: 10805726]
19. Kohyama M, et al. Nature. 2009; 457:318. [PubMed: 19037245]

20. Dequiedt F, et al. *Immunity*. 2003; 18:687. [PubMed: 12753745]
21. Willis TG, et al. *Blood*. 1998; 91:1873. [PubMed: 9490669]
22. Beroukhi R, et al. *Nature*. 2010; 463:899. [PubMed: 20164920]
23. Cadinanos J, Bradley A. *Nucleic Acids Res*. 2007; 35:e87. [PubMed: 17576687]
24. Keng VW, et al. *Nat. Methods*. 2005; 2:763. [PubMed: 16179923]

Fig. 1. Mouse lines carrying the genetic components of the transposon systems

(A) *RosaPB* and *RosaSB* knock-in mice express *PiggyBac* or *Sleeping-Beauty* transposase under control of the constitutively active *Rosa26* promoter. (B) Transposon design and transposon mouse lines. *Left*: Three transposon constructs were designed, which differ in their promoter/enhancer. All three transposons have *PB* as well as *SB* inverted terminal repeats (ITRs) and can therefore be mobilized with both transposases. *Right*: Three types of transposon mouse lines (*ATP1*, *ATP2* and *ATP3*) were generated using these constructs. A total of 19 *ATP* lines were developed and are listed in the tables on the right. For each line the chromosome harbouring the transposon donor locus and the transposon copy number in the donor concatemer are indicated. C β ASA, Carp β -actin splice acceptor; En2SA, *Engrailed-2* exon-2 splice acceptor; SD, *Foxf2* exon-1 splice donor; pA, bidirectional SV40 polyadenylation signal; CAG, CMV enhancer and chicken beta-actin promoter; MSCV, murine stem cell virus LTR; PGK, phosphoglycerate-kinase promoter;

Fig. 2. Embryonic lethality, tumor spectrum and latency in different ATP mouse lines upon *PiggyBac* mobilization

(A) Embryonic lethality in offspring from *RosaPB*;*ATP* intercrosses. 14 *ATP* mouse lines were used. Each line's transposon copy number is indicated below the graph. The number of mice born in each colony is shown above the graph. Bars represent double transgenic progeny as a percentage of expected. (B; C; E) Tumor spectrum and latency in different *ATP* mouse lines upon *PB* transposon mobilization. The percentage of mice with hematopoietic or solid tumors is indicated for *RosaPB*;*ATP2* (B), *RosaPB*;*ATP1* (C) and *RosaPB*;*ATP3* mice (E). Kaplan-Meier survival curves are shown for indicated genotypes. (D) CAG promoter-driven DsRed expression in mouse tissues. *Left columns*: red fluorescence detection in organs from wild-type (WT) and *hprt*^{CAG-DsRed} (R) male mice. H, heart; M, skeletal muscle; K, kidney; B, brain; T, testicle; IN, small intestine; L, lung; S, spleen. *Right columns*: Western blot analysis of the same organs using a DsRed-specific antibody (r) or control antibodies (t, alpha-tubulin; a, beta-actin).

Fig. 3. PB integration pattern in tumors from different ATP2 mouse lines

(A) Analysis of tumors from *RosaPB;ATP2-S1* mice (50 tumors; 4379 insertions). The lower graph compares the percentage of transposon insertions to the size and TTAA frequency on individual chromosomes. The donor chromosome and locus were identified by FISH and are indicated by blue arrows. The upper graph shows the distribution of insertions on Chr17 at a 1 Mb resolution. The number of expected (exp) and observed insertions outside (odl) and inside (idl) a 3MB interval surrounding the donor locus are indicated. (B) Results from *RosaPB;ATP2-S2* mice (7 tumors; 594 insertions; donor Chr12), and *RosaPB;ATP2-H32* mice (6 tumors; 618 insertions; donor Chr2) are presented accordingly.

Figure 4. Identification of common integration sites in *PiggyBac*-induced hematopoietic tumors
 (A) Pie chart representing all CIS identified: Some were reported in earlier screens using retroviruses (RV; 36%), SB (4%) or both (18%). 42% have not been identified in earlier screens. (B) Insertion pattern in known oncogenes and tumor suppressor genes. Arrows indicate individual insertions and their orientation (red/blue, sense- or antisense-orientation of the transposon's promoter to genes, respectively). (C) Promoter insertions in *Spic* in 9 tumors. MSCV-*Spic* fusion transcripts were detected by RT-PCR and sequencing. *Spic* expression was analyzed by qRT-PCR and normalized to *Gapdh* expression. Tumors with *Spic* integrations are myeloid leukemias, as shown by immunohistochemical myeloperoxidase staining (left, spleen; right lymph node from one mouse). Scale bars, 25 μ m. ***p<0.001