

# Isolation of homozygous mutant mouse embryonic stem cells using a dual selection system

Yue Huang<sup>1,2,\*</sup>, Stephen J. Pettitt<sup>1</sup>, Ge Guo<sup>1</sup>, Guang Liu<sup>2</sup>, Meng Amy Li<sup>1</sup>, Fengtang Yang<sup>1</sup> and Allan Bradley<sup>1,\*</sup>

<sup>1</sup>The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK and <sup>2</sup>State Key Laboratory of Medical Molecular Biology, Institute of Basic Medical Sciences, Department of Medical Genetics, Peking Union Medical College & Chinese Academy of Medical Sciences, Dong Dan San Tiao 5, Beijing 100005, China

Received June 20, 2011; Revised August 10, 2011; Accepted October 7, 2011

## ABSTRACT

Obtaining random homozygous mutants in mammalian cells for forward genetic studies has always been problematic due to the diploid genome. With one mutation per cell, only one allele of an autosomal gene can be disrupted, and the resulting heterozygous mutant is unlikely to display a phenotype. In cells with a genetic background deficient for the Bloom's syndrome helicase, such heterozygous mutants segregate homozygous daughter cells at a low frequency due to an elevated rate of crossover following mitotic recombination between homologous chromosomes. We constructed DNA vectors that are selectable based on their copy number and used these to isolate these rare homozygous mutant cells independent of their phenotype. We use the *piggyBac* transposon to limit the initial mutagenesis to one copy per cell, and select for cells that have increased the transposon copy number to two or more. This yields homozygous mutants with two allelic mutations, but also cells that have duplicated the mutant chromosome and become aneuploid during culture. On average, 26% of the copy number gain events occur by the mitotic recombination pathway. We obtained homozygous cells from 40% of the heterozygous mutants tested. This method can provide homozygous mammalian loss-of-function mutants for forward genetic applications.

## INTRODUCTION

Mammalian cell lines provide a convenient model for mammalian cell biology, particularly in high-throughput

applications where using mice is not feasible. Forward genetic screens using various cell lines have identified genes required for many cellular processes (1). Gain-of-function screens, where genes are overexpressed or ectopically expressed, have been successful for many phenotypes (2,3). However, loss-of-function screens using mutant cells that lack expression of a particular gene are harder to conduct in diploid mammalian cells, as in most cases both alleles of a gene must be knocked out to see a phenotype. This difficulty has meant that loss-of-function screens have not been applied as widely as in yeast, *Drosophila* or *Caenorhabditis elegans*, where homozygous mutants can easily be obtained on a genome-wide scale (4,5).

This problem of obtaining functional null mutants can be solved using a number of technologies, each with limitations. First, a cell line that is functionally hemizygous can be used. Chinese hamster ovary cells are one such cell type in which many loci are functionally hemizygous and these cells have been used to isolate, for example, X-ray sensitive mutants, with a single random hit (6). However, the extent of hemizyosity in the genome of these cells is unknown and it is likely that not all genes are accessible. A near-haploid human leukemia cell line has also been described, which can be used to isolate loss-of-function mutants in the haploid portion of the genome with a single hit (7,8). Further derivatives of this cell line (9), and the recent report of a haploid ES cell line (10), show promise for screens in the future.

RNA interference (RNAi) screens have also been used in mammalian cells (11). These act at the mRNA level and thus the zygoty of the gene is irrelevant. Many RNAi screens have been performed in mammalian cell lines, and the reagents are applicable to a variety of cell types enabling many phenotypes to be studied. Common problems encountered are incomplete

\*To whom correspondence should be addressed. Tel: +44 1223 496998; Fax: +44 1223 494714; Email: [abradley@sanger.ac.uk](mailto:abradley@sanger.ac.uk)  
Correspondence may also be addressed to Yue Huang. Tel: +86 10 65296462; Fax: +86 10 65105089; Email: [huangyue@pumc.edu.cn](mailto:huangyue@pumc.edu.cn)

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

© The Author(s) 2011. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

knockdown at the protein level, and off-target effects where transcripts other than the one predicted are affected, which can lead to variable results (12,13).

The most robust way to make loss-of-function mutations in a relatively normal cell line would be serially target both alleles in mouse embryonic stem (ES) cells. Although resources of targeting vectors and targeted heterozygous mutant ES cells are increasing in size [www.knockoutmouse.org and ref. (14)], this is still a very time-consuming task on the scale required for genome-wide genetic screens of homozygous mutations. We have previously used ES cells deficient for the Bloom syndrome gene [*Blm*, ref. (15)] to obtain homozygous mutants. *Blm*-deficient cells (referred to here as *Blm* cells for simplicity) have an increased frequency of crossing over following mitotic recombination relative to wild-type cells (Figure 1A). In practice, this means that cells carrying a heterozygous mutation segregate homozygous mutations at division at a low rate of the order of  $10^{-4}$  events/locus/cell/division.

Screens of mutant libraries made in *Blm* cells have been successful for phenotypes where null mutants are selectable, for example resistance to 6-thioguanine (mismatch repair mutants), aerolysin (glycosylphosphatidylinositol anchor synthesis mutants) or retroviral infection (16–18). Reporter systems can also be used to make the phenotype artificially selectable (19,20). The requirement for a selectable phenotype is due to the fact that each potentially interesting homozygous cell in the library is outnumbered

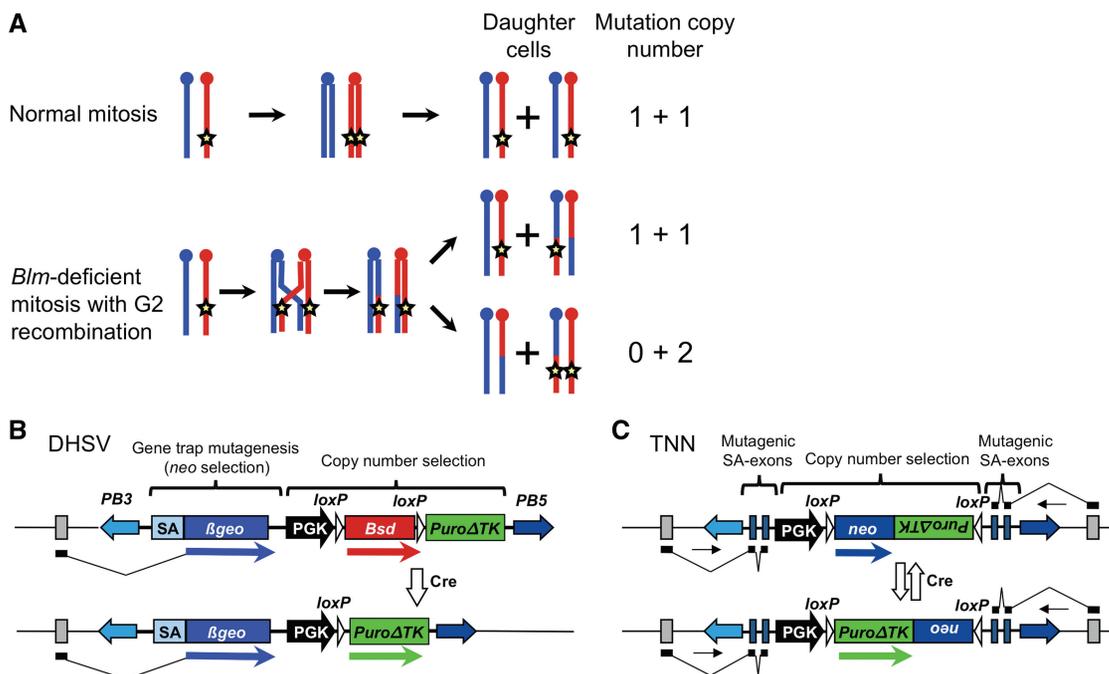
by the order of one thousand cells heterozygous for the insertion, which are unlikely to display a loss-of-function phenotype. We were therefore interested in extending this method to other non-selectable phenotypes by increasing the proportion of homozygous cells in the library.

We present here a method to isolate homozygous cells from these libraries independent of their phenotype. We use the *piggyBac* (*PB*) transposon (21) to cause loss-of-function insertion mutations and also to deliver a selection construct that carries two drug resistance genes but can express only one at a time. The expressed gene can be switched using Cre recombinase, and only cells with two copies of this construct can acquire resistance to both corresponding drugs simultaneously. This allows selection for the increase in copy number of the mutation that occurs after loss of heterozygosity (LOH), and thus for homozygous mutants. We have isolated homozygous mutants in 48 genes using this method, which should be applicable on a larger scale to generate clonal libraries of ES cells with null mutations for genetic screens.

## MATERIALS AND METHODS

### Construction of the gene trap vectors

The *PB* transposon vector, which contains 313 bp of the 5' inverted terminal DNA repeat (TR) and 235 bp of the 3' inverted TR, has been described previously (19). For the deletion homozygosity selection vector (DHSV), the *PB* 5'



**Figure 1.** Copy number selection for recovery of homozygous mutants. (A) Mechanism of LOH in *Blm*-deficient cells. Blue and red chromosomes represent homologs. Crossing over between homologous chromosomes is normally suppressed in mitosis (top). In the absence of *Blm*, crossovers can occur in G2 phase, forming recombinant chromatids. If these segregate to different daughter cells, LOH occurs distal to the point of crossover. If the cells carry a heterozygous mutation (represented by the star), this becomes homozygous after LOH, and the copy number of the mutation increases from one in the starting cell to two in the homozygous daughter. (B and C) Design of vectors for copy number-based selection. The vectors DHSV (B) and TNN (C) are shown integrated into an intron, with grey boxes representing exons of the disrupted gene; note that TNN can trap genes transcribed in either direction. SA, splice acceptor; *PB5* and *PB3*, *PB* repeats.

and 3' TRs were polymerase chain reaction (PCR) amplified and cloned upstream and downstream of the selection cassettes. The *SA-βgeo-bpA* gene trap cassette was derived from RGTV1 (18). The *Bsd-bpA* fragment was derived from pYTC77 (22). The *PGK-loxP-Bsd-bpA-loxP-PuroΔtk-bpA* fragment was derived from pYTC86 (22).

The TNN vector was assembled using a *neo* gene and SV40 polyA from pcDNA3 (Invitrogen), and *puΔTK-bpA* from pYTC86. The mutagens are pairs of terminal exons with their preceding introns from the genes *Ccdc107* and *Dom3z*, and were amplified by PCR and cloned in opposite orientations. These exon pairs were chosen for small size, constitutive splicing, lack of protein domains and potential to cause frameshift and premature termination. The QuikChange site-directed mutagenesis kit (Stratagene) was used to introduce premature stop codons in the native reading frame. This is predicted to cause non-sense-mediated decay of any fusion transcripts produced, thus increasing the chance that insertions toward the 3'-end of genes, which may not disrupt functional domains, will be mutagenic. The final vectors were confirmed by sequencing and full sequence of these vectors is available from GenBank under accession numbers HQ683721 (DHSV1) and HQ683722 (TNN).

#### ES cell culture and generation of gene trap mutations

Blm-deficient (m3/m4) ES cells (NN5), originally derived from the AB2.2 ES cell line, were cultured on a layer of  $\gamma$ -irradiated SNL76/7 feeder cells as described previously (18). Five million cells expanded from a single-cell clone of NN5 were co-electroporated with 0.5 or 1  $\mu$ g of each gene trap vector and 10  $\mu$ g of a PBase expression plasmid (pCAG-PBase). After electroporation, the cells were cultured on 90 mm feeder plates for 3 days without selection, then trypsinized and replated onto new feeder plates to reduce mosaicism in individual clones, before applying G418 selection (180  $\mu$ g/ml, Invitrogen).

#### Transient expression of Cre plasmid and Cre protein transduction

Cre was introduced into ES cells by various methods such as electroporation, lipofection and transduction with cell-permeable Cre protein (23,24). For electroporation, 20  $\mu$ g of pCAGGS-Cre plasmid was electroporated into  $1.0 \times 10^7$  ES cells, one-tenth of the electroporated cells were plated on one 90 mm feeder cell plate. Puromycin (3 mg/l, Sigma) and blasticidin S (10 mg/l, Invitrogen) selection was initiated 48 h later and continued for 7 days until individual ES cell clones were visible. For Lipofection, 1  $\mu$ g of pCAGGS-Cre plasmid and 3  $\mu$ l Lipofectamine 2000 (Invitrogen) were applied to ES cells cultured in one well of 24-well feeder plate (50% confluent) in accordance with instructions from the manufacturer. Two days later, the cells were replated on six-well plates and drug selection was conducted as above. Cre transduction was performed as described (23,24). In brief, one well of 50% confluent ES cells from a 24-well feeder plate were cultured in 200  $\mu$ l ES cell medium (without serum and antibiotics) containing 1 mM HTNC

(His<sub>6</sub>-Tat-NLS-Cre) protein produced as previously described (24) for 8 h. After treatment, 1.5 ml fresh ES cell medium (without antibiotics) was added to each well. Forty-eight hours later, cells were replated on a six-well plate and drug selection was applied.

For some experiments, an inducible Cre-ERT2 protein stably expressed from the *Rosa26* locus was used. The *Rosa26* locus was targeted as previously described (25) using a vector provided by Adams, D. and Jonkers, J. (unpublished data). To induce recombination, cells were treated with 1 mM 4-hydroxytamoxifen overnight.

#### Splinkerette-PCR to identify transposon insertion sites

Isolation of the transposon–chromosome junction was performed using the Splinkerette-PCR method as described (26). Briefly, genomic DNA was isolated from ES cell colonies on 96- or 24-well plates. Two to three micrograms of DNA was digested with Sau3AI and ligated with the corresponding Splinkerette adaptors HMSp-Sau3AI (generated by annealing Splinkerette oligos HMSpBb-Sau3AI with HMSpAa). A first-round of PCR was carried out with Splinkerette primer HMSp1 and *PB* transposon primers PB5'-1 or PB3'-1. Then 1% volume of the PCR product was directly used for second-round nested PCR that was carried out with Splinkerette primer HMSp2 and the *PB* transposon primers PB5'-2 or PB3'-2. The nested PCR products were purified by High Pure 96 UF cleanup Plate (Roche) and were used for sequencing with the primer PB5'-seq or PB3'-seq separately. All nucleotide positions are from National Center for Biotechnology Information (NCBI) mouse build 37.

#### Genomic PCR and reverse transcription PCR

All primers for genomic PCR and reverse transcription-PCR (RT-PCR) are described in Supplementary Table S1.

#### Karyotyping and Fluorescence *in situ* hybridisation of metaphase chromosome spreads of ES cells

Metaphase chromosome spreads were prepared from the ES cells cultured on 6-well gelatin-treated plates according to a standard protocol (27). Hybridization was performed by using appropriate labeled mouse whole chromosome paints or a specific DHSV vector probe (27,28).

#### Adriamycin treatment

*Atm*<sup>+/+</sup> (NN5), *Atm*<sup>+/-</sup> and *Atm*<sup>-/-</sup> ES cells were plated on  $\gamma$ -irradiated SNL76/7 feeder layer in a series of six-well plates at a density of 600–800 cells/well. Twenty-four hours later, cells were treated with graded concentration of adriamycin (Sigma) for 7 h and then replaced with fresh ES cell medium. The surviving ES cell colonies in each well were counted after 7–9 days of culture.

## RESULTS

### Copy number-based selection vectors

We considered that the copy number of the mutagenesis vector could be used to discriminate homozygous and heterozygous cells, and thus to select for homozygous mutants independently of phenotype. If the initial mutagenesis is limited to one mutation per cell, then homozygous mutant cells will have two mutations (at allelic positions). We therefore designed reporter constructs that are selectable based on the number of expressed copies in the cell. These constructs contain only one promoter and coding sequence for two different resistance genes. The two genes are expressed mutually exclusively; the expressed gene can be switched by the action of Cre recombinase. This is achieved either by deletion of one of the genes (vector DHSV, Figure 1B) or inversion when the genes are arranged tail-to-tail (vector TNN; Figure 1C). DHSV contains tandemly arranged *bsd* and *puroATK* genes (29) that confer blasticidin and puromycin resistance, respectively. TNN contains *puroATK* and a *neo* gene, which confers resistance to the drug G418, in the opposite orientation.

Cells bearing a single copy of such a construct can only be resistant to one drug at a time. However, in the case where cells have two or more copies of the construct, transient treatment with Cre will result in some cells with both forms of the construct, expressing both resistance genes. These cells will be resistant to both drugs. Thus, these constructs are predicted to allow stringent selection against cells with only one copy (i.e. heterozygous mutants) and to recover a subset of cells with two or more copies (homozygous mutants).

### Steps to increase the number of mutable loci

As well as the copy number selection cassette, each vector contains mutagenic elements designed to disrupt splicing. The constructs are flanked by TR sequences of the *PB* transposon for efficient delivery into the genome of mammalian cells. *PB* is highly active in ES cells and exhibits wide genome coverage (30,31). Moreover, ~50% of *PB* insertions occur in genes, making it an ideal mutagenesis vector (32). The DHSV vector contains a conventional selectable gene trap based on the  *$\beta$ -geo* gene, which combines G418 resistance with a *lacZ* reporter (33). Selectable gene trap vectors ensure that each insertion is in a gene, although there is a requirement for this gene to be expressed at the time of mutagenesis. *PB* appears to have wider genome coverage than the retroviruses used for gene trap mutagenesis—a previous study isolating components of the mismatch repair pathway showed that *PB* mutagenesis found more known components of the pathway than retroviral mutagenesis (34). As an option to further increase coverage, the TNN vector includes two non-selectable gene trap cassettes. These consist of pairs of exons and splice acceptors, containing multiple premature stop codons in all reading frames that are predicted to cause non-sense-mediated decay of fusion transcripts (35). These were designed to maximize the chance of causing a

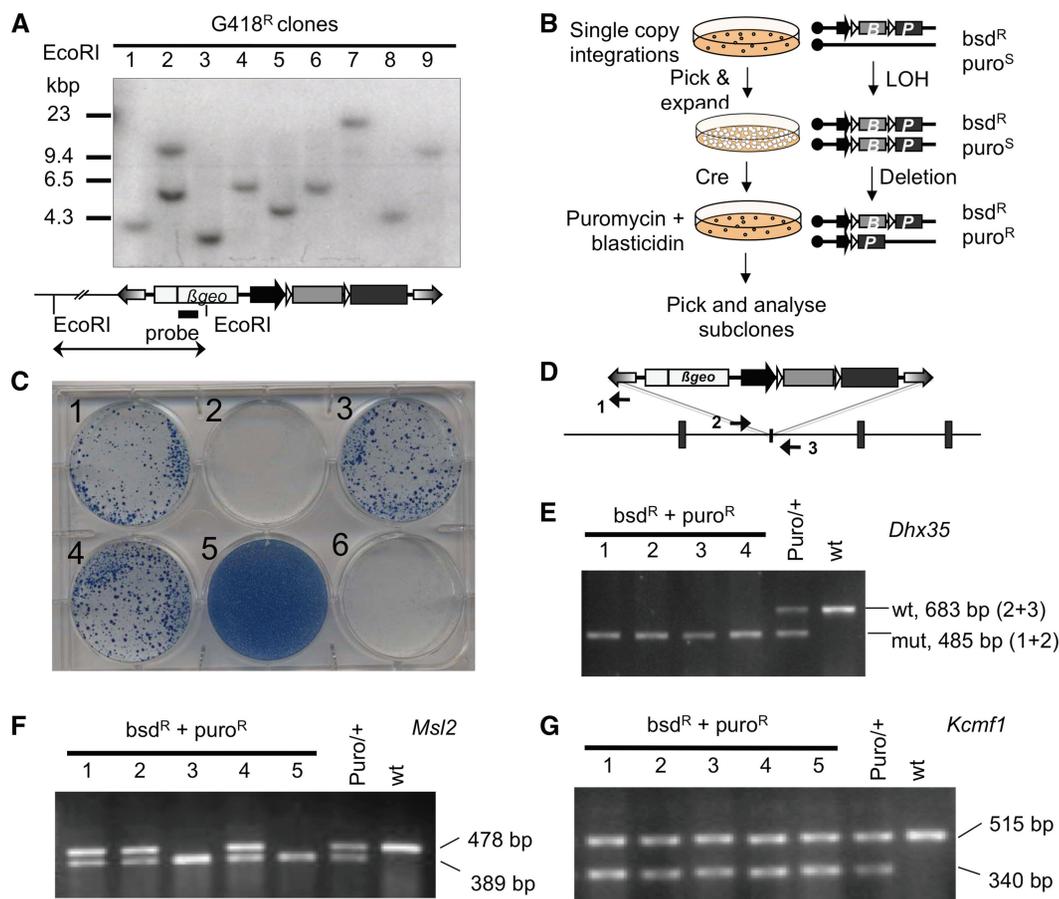
mutation, given that *PB* insertions frequently occur in genes, without requiring sustained expression of the gene trap coding sequence.

### Homozygous mutants recovered using selection for copy number increase

We applied these constructs to the selection of rare homozygous mutants that arise in cultures of heterozygously mutated *Blm*-deficient cells. Using transfection conditions in which a *PB* transposase expression plasmid is cotransfected with a limiting amount of transposon donor plasmid, we obtained hundreds of clonal *Blm*-deficient mouse ES cell lines. The majority of these contain a single-copy integration of the transposon (Figure 2A). Selection with both drugs at this point did not yield any double-resistant subclones (data not shown). For each cell line, we mapped the transposon integration site by Splinkerette-PCR. After culturing these cell lines for several passages, we could isolate double drug-resistant subclones after providing Cre by transfection of an expression plasmid, transduction with cell-permeable Cre protein or activation of an inducible ERT2-Cre transgene (Figure 2B). Some cell lines produced very large numbers of double-resistant subclones, comparable to the number of cells plated; these lines contained two or more non-allelic transposon insertions (Figure 2C, well 5) and were not analyzed further. Some lines did not yield double-resistant cells. However, for most lines we obtained varying numbers of double-resistant subclones, of the order of one per  $10^2$ – $10^4$  cells plated (Figure 2C). We genotyped these double-resistant subclones to determine whether these represented homozygous mutants. We assessed presence of the wild-type and mutant alleles by PCR using primers designed to flank the transposon integration site, along with a common primer that hybridizes to one transposon terminus and extends outwards (Figure 2D). For some cell lines, all double-resistant subclones showed the expected result for a homozygous mutant—amplification of only the PCR product from the mutant allele, and loss of the wild-type allele (Figure 2E). However, for most cell lines we observed some double-resistant subclones from which a wild-type PCR product could be amplified (Figure 2F), and in some cases all double-resistant subclones retained a wild-type locus as assessed in this PCR assay (Figure 2G).

### Double-resistant clones that retain a wild-type locus are aneuploid

Further investigation of these subclones that retained the wild-type locus by Southern blot revealed that although there were two allelic copies of the transposon present, these were accompanied by one or more extra wild-type chromosomes. In some cases, the cells had a tetraploid karyotype (Figure 3A–D), while in others only the chromosome with the insertion site was in excess (Figure 3E). These data show that LOH is not the only pathway for copy number increase in *Blm*-deficient cells, and that acquisition of aneuploidy is the major alternative pathway. Across all cell lines that we investigated, the average proportion of genuine homozygous mutant



**Figure 2.** Isolation of homozygous mutants by selection for copy number increase. (A) Transposition using limiting amounts of transposon donor plasmid results in most G418 resistant subclones having one transposon per cell. (B) Selection scheme, showing predicted number and type of transposon present at each stage. (C) Typical result of double-selection stage. Well 5 has two non-allelic transposon integrations to begin with and thus all cells can become double resistant. (D) Genotyping scheme—primer positions (1–3) are shown relative to the transposon integration site. (E) All double-resistant subclones (derived from an integration in *Dhx35*) fail to amplify the wild-type PCR product. (F and G) Examples of double-resistant populations in which some (F) and all (G) subclones have a wild-type allele.

subclones isolated was 26% (Table 1). As the rate of LOH is low even in *Blm*-deficient ES cells ( $4.2 \times 10^{-4}$  events/locus/cell/division), this represents a 25- to 250-fold enrichment for homozygous mutants, which would be expected to form only 0.1–1% of the culture prior to double selection

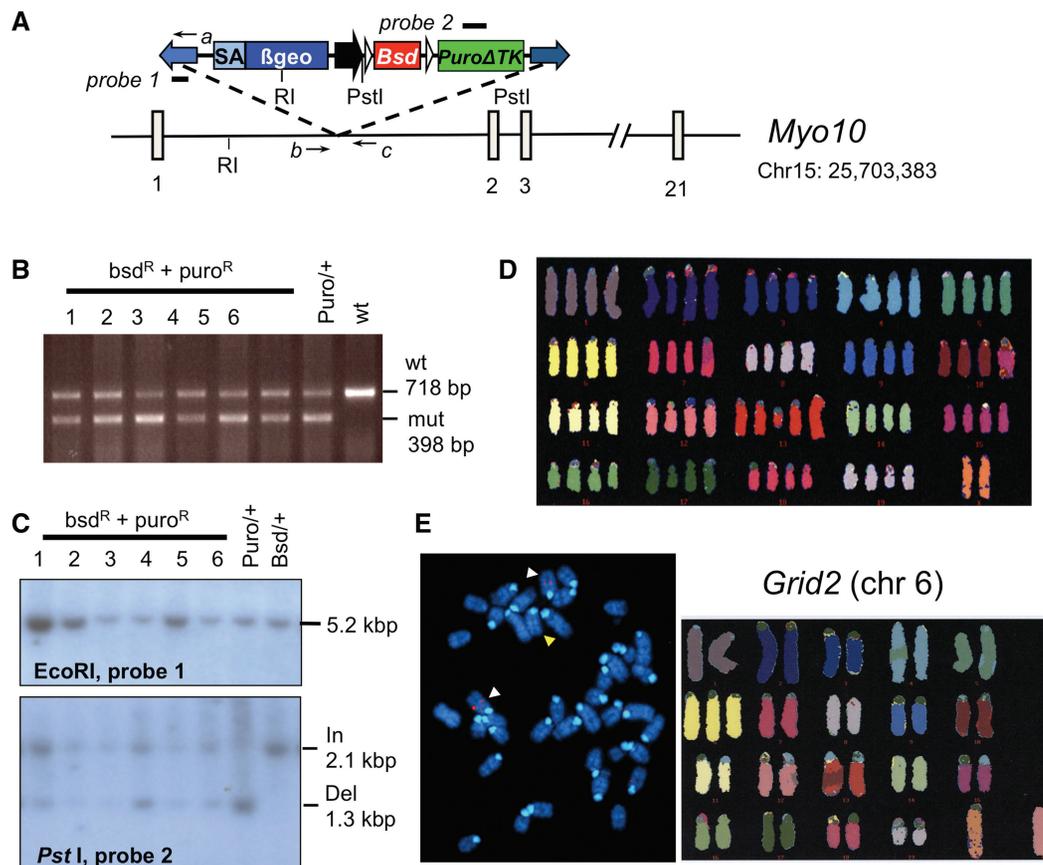
### LOH rate at different loci

As the process of LOH depends on a recombination event between the locus in question and the centromere, it is possible that it would be difficult to obtain homozygous mutants for loci close to the centromere. To address this, we measured the rate of LOH at three different loci on chromosome 11, one of which was a targeted insertion of a PGK-*puro* $\Delta$ TK cassette at the *Gdf9* locus, where previous measurements of the LOH rate in *Blm*-deficient ES cells have been made (15). We isolated other chromosome 11 insertions of TNN in the *puro* orientation (Figure 4A). In all of these cases, the wild-type cell that results from crossover and loss of the *puro* $\Delta$ TK cassette can be isolated by selection in 1-(2-deoxy-2-fluoro-1-d-arabinofuranosyl)-5-iodouracil; 200 nM (FIAU). For

each locus we set up replicate cultures starting from a single cell, allowed them to expand to  $7.5 \times 10^5$  cells and selected cells with LOH events in FIAU (Figure 4B). The rate of LOH was calculated by the fluctuation analysis using the method of medians (36), and did show an increase with distance from the centromere (Table 2). The rates were lower than previously measured using HPRT for negative selection, even for *Gdf9* (previously measured as  $4.2 \times 10^{-4}$ ). This may be related to the use of  $\Delta$ TK for selection in this case. We found that the LOH rate varied by a factor of four over the 20–96 Mb range tested. As 92% of genes lie >20 Mb from the centromere [Figure 4C, ref. (37)], the probability of crossover is not likely to limit the recovery of homozygous mutants for the vast majority of loci.

### The mutagenesis vectors cause null mutations

Homozygotes were readily isolated using both the deletion (DHSV) and inversion (TNN) vectors. For one insertion of DHSV in the *Atm* gene (Figure 5A), which encodes the ataxia telangiectasia-mutated kinase involved in the DNA damage response (38,39), we obtained three homozygous



**Figure 3.** Abnormal karyotype of clones that retain a wild-type allele. (A) Structure of *Myo10* locus analyzed in (B–D). (B) PCR genotyping as in Figure 2D. All double-resistant subclones retain the wild-type allele. (C) All these subclones have a single-transposon integration site (top) but contain both pre- (In: insertion) and post-Cre (Del: deleted) forms of the transposon (bottom). (D) Spectral karyotype showing tetraploidy in a double-resistant subclone. (E) An example of a less severe aberration—only chromosome 6, which bears the insertion at the *Grid2* locus, is present in excess. Left, FISH showing three copies of chromosome 6 with (white arrows) and without (yellow arrow) transposon insertions (red dots). Right, spectral karyotype.

**Table 1.** Number and proportion of homozygous mutant clones isolated using each vector

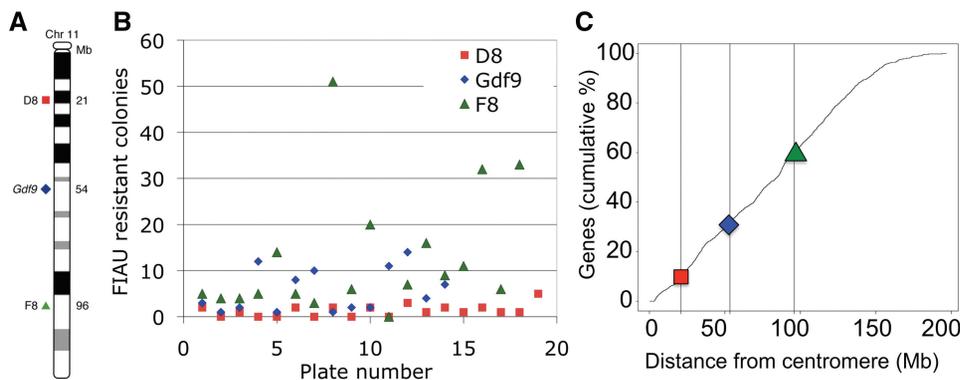
Vector	Clones analyzed	Clones with homozygotes	Clones with homozygotes (%)	Mean hom:het ratio in subclones (%) <sup>a</sup>
DHSV	89	36	40	23
TNN	19	12	63	34
Combined	108	48	44	26

<sup>a</sup>Includes clones for which no homozygous subclones were isolated.

subclones after double selection (Figure 5B). Although, the vector was oriented with the splice acceptor in the opposing direction, and is therefore trapping another promoter in this instance, *Atm* transcription was effectively disrupted (Figure 5C). Mutant subclones showed sensitivity to the DNA damaging drug adriamycin, as expected for an *Atm* null mutant. In all other cases that we tested, both the DHSV and TNN vectors effectively disrupted transcription of genes that they were inserted into as assessed by RT-PCR (Figure 6). These data show that mutations generated with DHSV or TNN cause null mutations and therefore can be used to screen non-selectable phenotypes.

## DISCUSSION

We have developed a method for enriching libraries of *Blm*-deficient ES cells carrying single-copy transposon mutations for homozygous mutants. The method is based on selection for two copies of the transposon construct in the rare homozygous mutants, which is achieved by a switchable drug selection cassette. Using our method of single-copy mutagenesis, clonal expansion, cassette switching by Cre recombinase and selection for two copies, we obtained double-drug resistant populations enriched in homozygous mutants. The proportion of homozygous mutants increased from a predicted 0.1% in the expanded unselected population to a measured

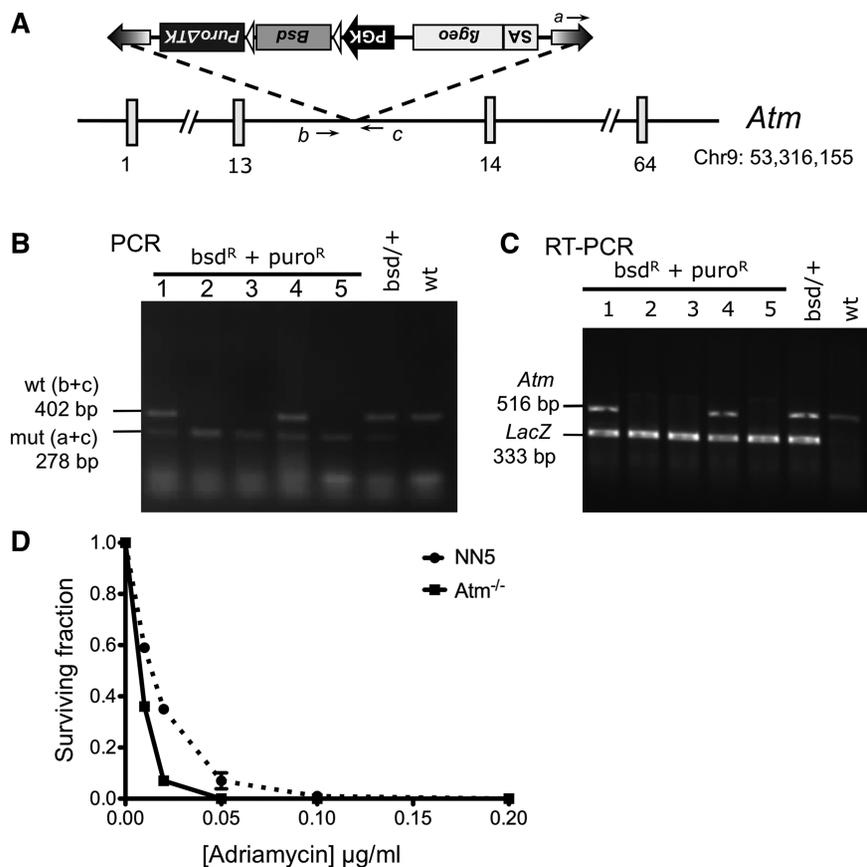


**Figure 4.** Measurement of LOH rate at different loci. (A) Position of loci investigated on chromosome 11. (B) Number of FIAU resistant colonies obtained (one-tenth of each culture was plated). Data are shown in the original random order, and the rate calculation shown in Table 2. (C) Cumulative frequency plot of all Vega-curated gene start sites (37). The positions of the loci investigated here are marked.

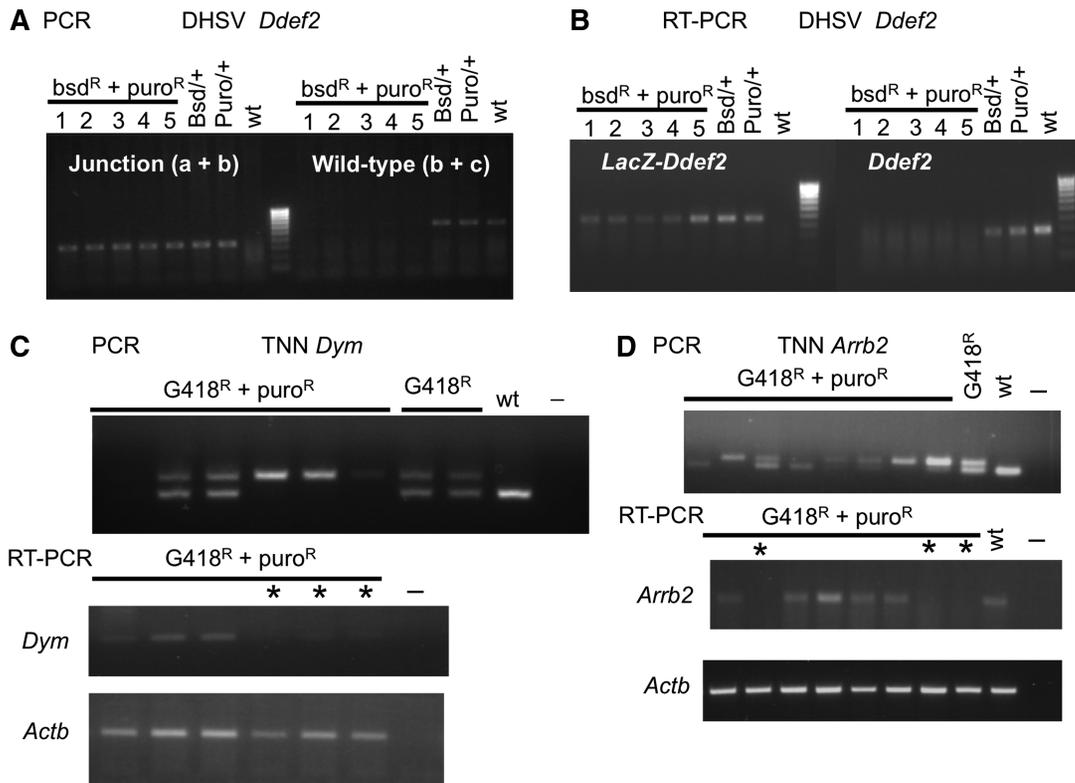
**Table 2.** Calculation of the rate of LOH at three loci on chromosome 11 by fluctuation analysis

Locus	Mb	Cultures	Median FIAU <sup>R</sup> /culture	Plating efficiency	Mean cells/culture	LOH events/culture ( $m_h$ ) <sup>a</sup>	LOH rate ( $10^{-5}$ events/cell/division)
D8	21	19	1	0.35	753 000	7.53	0.99
<i>Gdf9</i>	54	14	3.5	0.38	762 000	18.56	2.5
F8	96	18	6.5	0.27	739 000	40.95	5.5

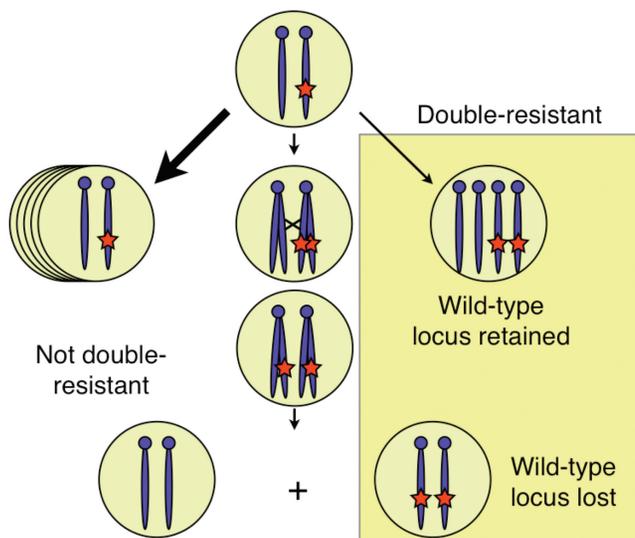
<sup>a</sup>See ref. (34) for details of calculation. FIAU<sup>R</sup>, FIAU-resistant colonies.



**Figure 5.** Insertion of DHSV at *Atm* causes DNA damage sensitivity. (A) Structure of mutant *Atm* locus. (B) Triple primer PCR genotyping using primers a, b and c shown in (A). Clones 2, 3 and 5 lack the wild-type band, consistent with being homozygous mutants. (C) RT-PCR showing loss of *Atm* transcript. Internal primers to *LacZ* were included as a positive control. (D) *Atm* mutants are hypersensitive to adriamycin. Mean surviving fraction (colonies) is plotted. NN5,  $n = 2$ ; *Atm*<sup>-/-</sup>,  $n = 3$ . Error bars show SEM.



**Figure 6.** Robust disruption of transcription in homozygous mutants. (A) PCR genotyping of DHSV insertion in *Ddef2* using primers for mutant (Junction, left) and wild-type alleles. All double-resistant subclones are homozygous mutants. (B) RT-PCR showing amplification of fusion transcript (left) and loss of wild-type transcript (right) in double-resistant subclones. (C and D) Insertions of the TNN vector also disrupt transcription. PCR and RT-PCR genotyping is shown for insertions in *Dym* (C) and *Arrb2* (D). Asterisks indicate homozygous mutant subclones, all of which have lost wild-type transcript expression. Minus indicates no template control, *Actb* is an internal control.



**Figure 7.** Pathways for copy number gain in *Blm*-deficient cells. Most cell divisions do not result in copy number increase (left). The two pathways of mitotic recombination (middle) and aneuploidy acquisition (right) result in copy number gain and contribute to the double-resistant population.

average of 26%, depending on the vector used, in the double-resistant population. This increased proportion means that homozygous mutants can be obtained by picking relatively small numbers of subclones. It is possible that this number is an underestimate where clones were genotyped by absence of wild-type locus by PCR, as this may be very sensitive to contamination either between clones or from wild-type feeder cells.

Genotyping of subclones from the double-drug resistant populations revealed two classes of mutants. We found that while some subclones had lost the wild-type locus, consistent with a homozygous mutant, others retained a wild-type locus that could be amplified by PCR. Further analysis of some of these clones revealed numerical chromosome abnormalities, either genome wide (tetraploidy) or affecting the chromosome bearing the transposon insertion (trisomy). We conclude that our method faithfully selects for two copies of the construct, and there are two routes by which *Blm* cells can increase mutation copy number—the predicted LOH pathway and by numerical chromosome instability (Figure 7). Often we isolated similar numbers of these two classes of double-resistant subclone, suggesting that each of these events occurs at a similar low rate. In some cases, this may be exacerbated by a growth advantage linked to trisomy, particularly for chromosomes 8, 11 and 16 (40–42).

However, we did successfully isolate homozygous mutants from insertions on these chromosomes.

The system of selection for copy number increase in *Blm*-deficient ES cells could be applied more broadly to obtain large collections of mutants for study, for example in genetic screens. Using the procedure described above, we isolated homozygous mutants for 48 out of 108 cell lines tested (Table 1 and Supplementary Tables S1 and S2). We screened relatively small numbers of subclones (usually six or fewer); therefore it is possible that homozygous mutants could be isolated for some of the remaining 60 cell lines by screening further subclones.

We were unable to isolate homozygous mutants for some clones, either because no double-resistant cells were obtained or because all double-resistant subclones analyzed retained the wild-type locus. This could be simply because the events leading to double resistance are rare and may not occur in every expanded clone, or because only a chromosome instability event and not an LOH event occurs. Another explanation could be that the gene mutated is essential for ES cell survival or growth, and therefore homozygous cells do not survive or are at a growth disadvantage compared with wild-type retaining cells. This is likely to be the case for some clones that we tested, for example *Faf1*, a mutant of which has been reported to cause early embryonic lethality and for which no homozygotes were isolated using our system (43). It should also be noted that when the DHSV vector is used, double-resistant cells may not be recovered if Cre is very efficient and causes deletion in both copies of the resistant cassette. In the case of TNN, the theoretical chance of isolating a cell with two copies is 50% if the Cre reaction reaches an effective equilibrium—this represents the probability that both copies are in different orientations. However, we did not notice an overall difference in the number of double-resistant clones isolated with each vector in practice.

The *Blm* mutant alleles used here are constitutive, rather than inducible, mutations. Since *Blm* is a key regulator of homologous recombination (44), there is cause for concern that a constitutive mutation may result in ongoing genome instability. We have not observed high levels of gross chromosomal instability, for example genuine homozygotes isolated from cell lines that also gave aneuploid subclones usually have a normal diploid karyotype (at the microscopic level). An inducible *Blm* allele (17) is an obvious potential improvement to the system, and a similar system has been developed using such an allele (45). Even in this case, *Blm* function must be still disrupted to allow LOH, hence an inducible allele will not improve genome stability during the homozygote isolation process, although it would allow for formal separation of the effects of the transposon-induced mutations from the *Blm*-deficient genetic background.

Even better validation of mutations could come from resources of targeted mouse ES cell knockouts and targeting vectors, which are increasing in size to the point where obtaining a heterozygous mutant in a gene of interest will become trivial. These ES cell lines can be used to generate homozygous mutants after disruption of the remaining wild-type allele using the vectors provided (14,46). Given

the ease of obtaining mutants from the public resources in a low-passage cell line with a ‘clean’ genetic background, this represents the best option for further functional study. Mammalian screening methods such as the one we describe here are an excellent complement to these resources, as they provide rapid functional information to prioritize study of the available mutants.

In conclusion, we describe a system using *PB* transposon constructs that cause null mutations with wide genome coverage in mouse ES cells that are a versatile model mammalian cell type. The mutations generated are stable, but reversion can be induced and selected for by a second round of transposition to allow phenotype rescue experiments (26). This gives a robust link between genotype and phenotype, representing a key advantage compared to RNAi-based screens where causally linking the phenotype to the targeted gene can be problematic. Moreover, other mutations or reporter constructs can easily be introduced into the *Blm*-deficient cells to create customized libraries for more sophisticated screens. By providing easy access to elusive homozygous mammalian mutant cells, our method significantly expands the scope of genetic screens and future functional studies in mammalian cells.

## ACCESSION NUMBERS

GenBank: HQ683721, HQ683722.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Tables 1–3.

## ACKNOWLEDGEMENTS

The authors thank Xiaozhong Wang (Northwestern University) for providing the *PB* transposon vector and *PB* transposase expression plasmid (pCAG-PBase); Pentao Liu and Leopold Parts for comments on this manuscript; Qi Liang, Frances Law and Hiroko Yusa for help with tissue culture.

## FUNDING

Wellcome Trust (79643 to A.B.); Royal Society (to Y.H.); National Natural Science Foundation of China (30971670 to Y.H. and G.L.); National Basic Research Program of China (2011CB965203 to Y.H. and G.L.). Funding for open access charge: Wellcome Trust.

*Conflict of interest statement.* None declared.

## REFERENCES

1. Grimm, S. (2004) The art and design of genetic screens: mammalian culture cells. *Nat. Rev. Genet.*, **5**, 179–189.
2. Feng, Y., Broder, C.C., Kennedy, P.E. and Berger, E.A. (1996) HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science*, **272**, 872–877.

3. Goldfarb, M., Shimizu, K., Perucho, M. and Wigler, M. (1982) Isolation and preliminary characterization of a human transforming gene from T24 bladder carcinoma cells. *Nature*, **296**, 404–409.
4. St Johnston, D. (2002) The art and design of genetic screens: *Drosophila melanogaster*. *Nat. Rev. Genet.*, **3**, 176–188.
5. Jorgensen, E.M. and Mango, S.E. (2002) The art and design of genetic screens: *Caenorhabditis elegans*. *Nat. Rev. Genet.*, **3**, 356–369.
6. Doetschman, T., Gregg, R.G., Maeda, N., Hooper, M.L., Melton, D.W., Thompson, S. and Smithies, O. (1987) Targeted correction of a mutant HPRT gene in mouse embryonic stem cells. *Nature*, **330**, 576–578.
7. Carette, J.E., Guimaraes, C.P., Varadarajan, M., Park, A.S., Wuethrich, L., Godarova, A., Kotecki, M., Cochran, B.H., Spooner, E., Ploegh, H.L. *et al.* (2009) Haploid genetic screens in human cells identify host factors used by pathogens. *Science*, **326**, 1231–1235.
8. Kotecki, M., Reddy, P.S. and Cochran, B.H. (1999) Isolation and characterization of a near-haploid human cell line. *Exp. Cell Res.*, **252**, 273–280.
9. Carette, J.E., Raaben, M., Wong, A.C., Herbert, A.S., Obernosterer, G., Mulherkar, N., Kuehne, A.I., Kranzusch, P.J., Griffin, A.M., Ruthel, G. *et al.* (2011) Ebola virus entry requires the cholesterol transporter Niemann-Pick C1. *Nature*, **477**, 340–343.
10. Leeb, M. and Wutz, A. (2011) Derivation of haploid embryonic stem cells from mouse embryos. *Nature*, **479**, 131–134.
11. Boutros, M. and Ahringer, J. (2008) The art and design of genetic screens: RNA interference. *Nat. Rev. Genet.*, **9**, 554–566.
12. Jackson, A.L., Bartz, S.R., Schelter, J., Kobayashi, S.V., Burchard, J., Mao, M., Li, B., Cavet, G. and Linsley, P.S. (2003) Expression profiling reveals off-target gene regulation by RNAi. *Nat. Biotechnol.*, **21**, 635–637.
13. Bushman, F.D., Malani, N., Fernandes, J., D'Orso, I., Cagney, G., Diamond, T.L., Zhou, H., Hazuda, D.J., Espeseth, A.S., König, R. *et al.* (2009) Host cell factors in HIV replication: meta-analysis of genome-wide studies. *PLoS Pathogens*, **5**, e1000437.
14. Skarnes, W.C., Rosen, B., West, A.P., Koutourakis, M., Bushell, W., Iyer, V., Mujica, A.O., Thomas, M., Harrow, J., Cox, T. *et al.* (2011) A conditional knockout resource for the genome-wide study of mouse gene function. *Nature*, **474**, 337–342.
15. Luo, G., Santoro, I.M., McDaniel, L.D., Nishijima, I., Mills, M., Youssoufian, H., Vogel, H., Schultz, R.A. and Bradley, A. (2000) Cancer predisposition caused by elevated mitotic recombination in Bloom mice. *Nat. Genet.*, **26**, 424–429.
16. Wang, W. and Bradley, A. (2007) A recessive genetic screen for host factors required for retroviral infection in a library of insertionally mutated Blm-deficient embryonic stem cells. *Genome Biol.*, **8**, R48.
17. Yusa, K., Horie, K., Kondoh, G., Kouno, M., Maeda, Y., Kinoshita, T. and Takeda, J. (2004) Genome-wide phenotype analysis in ES cells by regulated disruption of Bloom's syndrome gene. *Nature*, **429**, 896–899.
18. Guo, G., Wang, W. and Bradley, A. (2004) Mismatch repair genes identified using genetic screens in Blm-deficient embryonic stem cells. *Nature*, **429**, 891–895.
19. Trombly, M.I., Su, H. and Wang, X. (2009) A genetic screen for components of the mammalian RNA interference pathway in Bloom-deficient mouse embryonic stem cells. *Nucleic Acids Res.*, **37**, e34.
20. Guo, G., Huang, Y., Humphreys, P., Wang, X. and Smith, A. (2011) A PiggyBac-based recessive screening method to identify pluripotency regulators. *PLoS One*, **6**, e18189.
21. Fraser, M.J., Ciszczon, T., Elick, T. and Bauser, C. (1996) Precise excision of TTA-specific lepidopteran transposons piggyBac (IFP2) and tagalong (TFP3) from the baculovirus genome in cell lines from two species of Lepidoptera. *Insect Mol. Biol.*, **5**, 141–151.
22. Chen, Y., Liu, P. and Bradley, A. (2004) Inducible gene trapping with drug-selectable markers and Cre/loxP to identify developmentally regulated genes. *Mol. Cell Biol.*, **24**, 9930–9941.
23. Joshi, S.K., Hashimoto, K. and Koni, P.A. (2002) Induced DNA recombination by Cre recombinase protein transduction. *Genesis*, **33**, 48–54.
24. Peitz, M., Pfannkuche, K., Rajewsky, K. and Edenhofer, F. (2002) Ability of the hydrophobic FGF and basic TAT peptides to promote cellular uptake of recombinant Cre recombinase: a tool for efficient genetic engineering of mammalian genomes. *Proc. Natl Acad. Sci. USA*, **99**, 4489–4494.
25. Vooijs, M., Jonkers, J. and Berns, A. (2001) A highly efficient ligand-regulated Cre recombinase mouse line shows that Cre recombination is position dependent. *EMBO Rep.*, **2**, 292–297.
26. Li, M.A., Pettitt, S.J., Yusa, K. and Bradley, A. (2010) Genome-Wide Forward Genetic Screens in Mouse ES Cells. *Methods Enzymol.*, **477**, 217–242.
27. Hernandez, D., Phillips, I.R. and Shephard, E. (2006) In: Phillips, I.R. and Shephard, E. (eds), *Methods in Molecular Biology*, Vol. 320. Humana Press, Totwa, NJ, pp. 321–327.
28. Rabbitts, P., Impey, H., Heppell-Parton, A., Langford, C., Tease, C., Lowe, N., Bailey, D., Ferguson-Smith, M. and Carter, N.P. (1995) Chromosome specific paints from a high resolution flow karyotype of the mouse. *Nat. Genet.*, **9**, 369–375.
29. Chen, Y.T. and Bradley, A. (2000) A new positive/negative selectable marker, puDeltat, for use in embryonic stem cells. *Genesis*, **28**, 31–35.
30. Ding, S., Wu, X., Li, G., Han, M., Zhuang, Y. and Xu, T. (2005) Efficient transposition of the piggyBac (PB) transposon in mammalian cells and mice. *Cell*, **122**, 473–483.
31. Wang, W., Lin, C., Lu, D., Ning, Z., Cox, T., Melvin, D., Wang, X., Bradley, A. and Liu, P. (2008) Chromosomal transposition of PiggyBac in mouse embryonic stem cells. *Proc. Natl Acad. Sci. USA*, **105**, 9290–9295.
32. Liang, Q., Kong, J., Stalker, J. and Bradley, A. (2009) Chromosomal mobilization and reintegration of Sleeping Beauty and PiggyBac transposons. *Genesis*, **47**, 404–408.
33. Friedrich, G. and Soriano, P. (1991) Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. *Genes Dev.*, **5**, 1513–1523.
34. Wang, W., Bradley, A. and Huang, Y. (2009) A piggyBac transposon-based genome-wide library of insertionally mutated Blm-deficient murine ES cells. *Genome Res.*, **19**, 667–673.
35. Hentze, M.W. and Kulozik, A.E. (1999) A perfect message: RNA surveillance and nonsense-mediated decay. *Cell*, **96**, 307–310.
36. Jones, M.E., Thomas, S.M. and Rogers, A. (1994) Luria-Delbrück fluctuation experiments: design and analysis. *Genetics*, **136**, 1209–1216.
37. Wilming, L.G., Gilbert, J.G., Howe, K., Trevanion, S., Hubbard, T. and Harrow, J.L. (2008) The vertebrate genome annotation (Vega) database. *Nucleic Acids Res.*, **36**, D753–D760.
38. Shiloh, Y. (2003) ATM and related protein kinases: safeguarding genome integrity. *Nat. Rev. Cancer*, **3**, 155–168.
39. Xu, Y. and Baltimore, D. (1996) Dual roles of ATM in the cellular response to radiation and in cell growth control. *Genes Dev.*, **10**, 2401–2410.
40. Sugawara, A., Goto, K., Sotomaru, Y., Sofuni, T. and Ito, T. (2006) Current status of chromosomal abnormalities in mouse embryonic stem cell lines used in Japan. *Comp. Med.*, **56**, 31–34.
41. Liu, X., Wu, H., Loring, J., Hormuzdi, S., Disteche, C.M., Bornstein, P. and Jaenisch, R. (1997) Trisomy eight in ES cells is a common potential problem in gene targeting and interferes with germ line transmission. *Dev. Dyn.*, **209**, 85–91.
42. Liang, Q., Conte, N., Skarnes, W.C. and Bradley, A. (2008) Extensive genomic copy number variation in embryonic stem cells. *Proc. Natl Acad. Sci. USA*, **105**, 17453–17456.
43. Adham, I.M., Khulan, J., Held, T., Schmidt, B., Meyer, B.I., Meinhardt, A. and Engel, W. (2008) Fas-associated factor (FAF1) is required for the early cleavage-stages of mouse embryo. *Mol. Hum. Reprod.*, **14**, 207–213.
44. Hickson, I. (2003) RecQ helicases: caretakers of the genome. *Nat. Rev. Cancer*, **3**, 169–178.
45. Horie, K., Kokubu, C., Yoshida, J., Akagi, K., Isotani, A., Oshitani, A., Yusa, K., Ikeda, R., Huang, Y., Bradley, A. *et al.* (2011) A homozygous mutant embryonic stem cell bank applicable for phenotype-driven genetic screening. *Nature Methods*, 23 October (doi:10.1038/nmeth.1739; epub ahead of print).
46. Tate, P.H. and Skarnes, W.C. (2011) Bi-allelic gene targeting in mouse embryonic stem cells. *Methods*, **53**, 331–338.