Chapter Four - Generation of a *Blm*-deficient mouse ES cell line with a single copy of the mutagenic *piggyBac* transposon

1. Introduction

A *Blm*-deficient ES cell line with an intact *Hprt* locus is a very useful reagent. It can be used to conduct genome-wide recessive screens when coupled with *piggyBac*-mediated intra-genomic mobilisation using *Hprt* as a donor locus to facilitate the enrichment for PB re-mobilisation (Chapter 3). However, the existing *Blm*-knockout ES cell line was generated in the *Hprt*-deficient AB2.2 cell line (Chapter 2). In addition, a reporter system is required for my screens in order to select for mutants in the miRNA biogenesis and effector pathways (Chapter 1 and 5). Therefore, a new *Blm*-deficient ES cell line has been designed and generated to incorporate both requirements.

2. Results

2.1. Generation of a new *Blm*-deficient mouse ES cell line

The gene-targeting strategy used to create the *Blm*-null mutation was based on the design of the Blm^{tm1Brd} allele of the *Blm*-knockout ES cells, using a positive drug selection cassette to replace the start codon-containing exon by the replacement gene-targeting method (Luo et al., 2000). An eGFP and Bsd-resistant gene co-expression cassette driven by the human ubiquitin C promoter (Huc-eGFP-IRES-Bsd) was introduced into the *Blm* locus as the selection marker to replace exon 2 of the *Blm* gene Figure 4-1b. The eGFP-IRES-Bsd is constitutively expressed under the Huc promoter and the coding regions of eGFP and Bsd are connected by IRES, which allows the two coding regions to be transcribed from a single mRNA. Both alleles of *Blm* were targeted sequentially to give rise to the *Blm*-null mouse ES cell line.

The pBlmMLTV2 targeting vector (Figure 4-1a) was constructed based on a C57BL/6 BAC covering the *Blm* gene, and the detailed vector contrusction is shown in Chapter 2. The targeting vector was linearised with *Pme*I and electroporated into $1x10^7$ JM8.F6 cells. The cells were selected under Blasticidin for seven days and 48 colonies were picked and screened

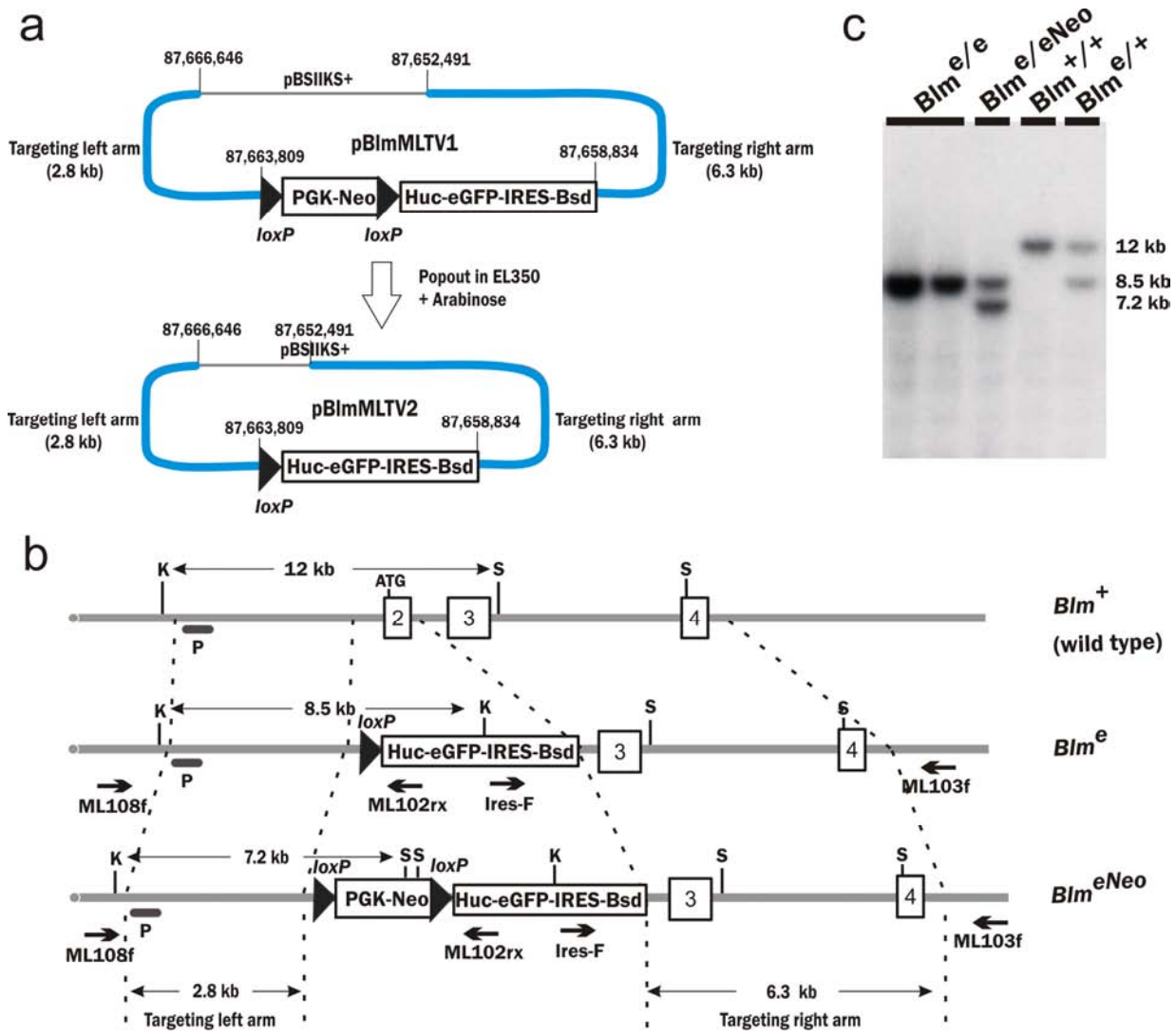
with long-range PCR (with primers ML108f and ML102rx) to detect homologous recombination on the short targeting arm end for correctly targeted clones. The targeting efficiency was 18 %. Long-range PCR (with primers Ires-F and ML103f) was also conducted to confirm the homologous recombination on the long targeting arm end. The correctly targeted cells are named *Blm*$^{e/+}$.

The six of the correctly targeted clones were subjected to karyotyping analysis to ensure the ES cell clones were not aneuploid during the targeting and subcloning processes. The cells were incubated with Colcemid (100μg/ml final concentration) for two hours to accumulate cells with cell-cycle arrested at metaphase before the metaphase preparation (Chapter 2 for detailed protocols). For each ES cell clone, ten metaphases were examined for the total chromosome numbers. Wild-type cells with a diploid genome should contain 40 chromosomes at metaphase of the cell cycle, and the most common aneuploidty includes the loss of Y chromosome (the cells used are male) to give rise to 39 chromosomes, and trisomy for chromosome 8 or 11, to give rise to 41 chromosomes. All six *Blm*$^{e/+}$ clones examined had 40 chromosomes per metaphase.

One of the clones with the correct karyotype was selected for the second round targeting with the *Pme*I-linearised targeting vector pBlmMLTV1. The electroporated cells were selected using G418 for eight days and 96 colonies were picked. The efficiency of the second allele targeting is expected to be slightly lower, as the targeting of the second allele is only 50 % of all the correctly targeted events. Therefore, a larger number of colonies were screened for the correctly targeted clones than the first allele targeting. Long-range PCR (with primers ML108f and ML102rx) was conducted to screen for colonies with both alleles targeted, as both alleles can be amplified with the primer set. Two PCR products with different sizes can be observed. The second allele targeting efficiency was 17 %. The double-allele targeted ES cell clones (*Blm*$^{e/eNeo}$) were expanded and were subjected to karyotyping analysis. Out of the six clones analysed, no aneuploidity was observed.

One correctly targeted and karyotypically normal clone was expanded and the PGK-Neo cassette was popped out by transient transfection of a Cre expression plasmid (pCAG-Cre). $3x10^6$ cells were electroporated with 25µg of pCAG-Cre plasmid, and the cells were plated onto a 90 mm plate. Three days post-electroporation in non-selective medium, the cells were trypsinised and replated in duplicates at the density of 1,000 cells per 90 mm plate. The next day, G418 selection was initiated on one plate and no selection was added to the other plate. The number of G418 resistant colonies provided a background level for the clones without the cassette popped out. Twelve colonies were picked from the 90 mm plate without any selection and were subjected to PCR analysis to detect the presence of the genomic junction with the cassette removed. A third of the colonies picked were positive for the cassette popout. After removal of the PGK-Neo cassette, both alleles of the *Blm* locus are identical, giving rise to the final *Blm*-deficient ES cells, $Blm^{e/e}$ (Figure 4-1b). Southern blotting was also conducted using an external locus-specific probe to further confirm the locus structure after each step of the manipulation, shown in Figure 4-1c.

Figure 4-1: Double-allele targeting of the *Blm* locus of the JM8.F6 ES cell line.



a, Two targeting vectors designs for the sequential *Blm* targeting. pBlmMLTV2 was derived from pBlmMLTV1 with the PGK-Neo cassette deleted in EL350 *E coli* strain in which Cre expression is induced by *L*-arabinose. The genomic coordinates (NCBI Build 37) shown on the targeting vector are the coordinate of the start and base pair taken for the targeting arm. b, Allele structures of the *Blm* locus. The wild-type allele was targeted with pBlmMLTV2 to give $Blm^{e/+}$ cells. $Blm^{e/+}$ cells were further targeted with pBlmMLTV1 to give $Blm^{e/eNeo}$ cells. Finally, the PGK-Neo cassette was popped out to give the final $Blm^{e/e}$ cells. The locations of the long-range PCR primers used for the initial genotype screening are shown. P, DNA probe for Southern blotting; K, *Kpn*I site; S, *Spe*I site. c, Southern blot confirming the genotypes of all the intermediate and final *Blm*-deficient ($Blm^{e/e}$) ES cells.

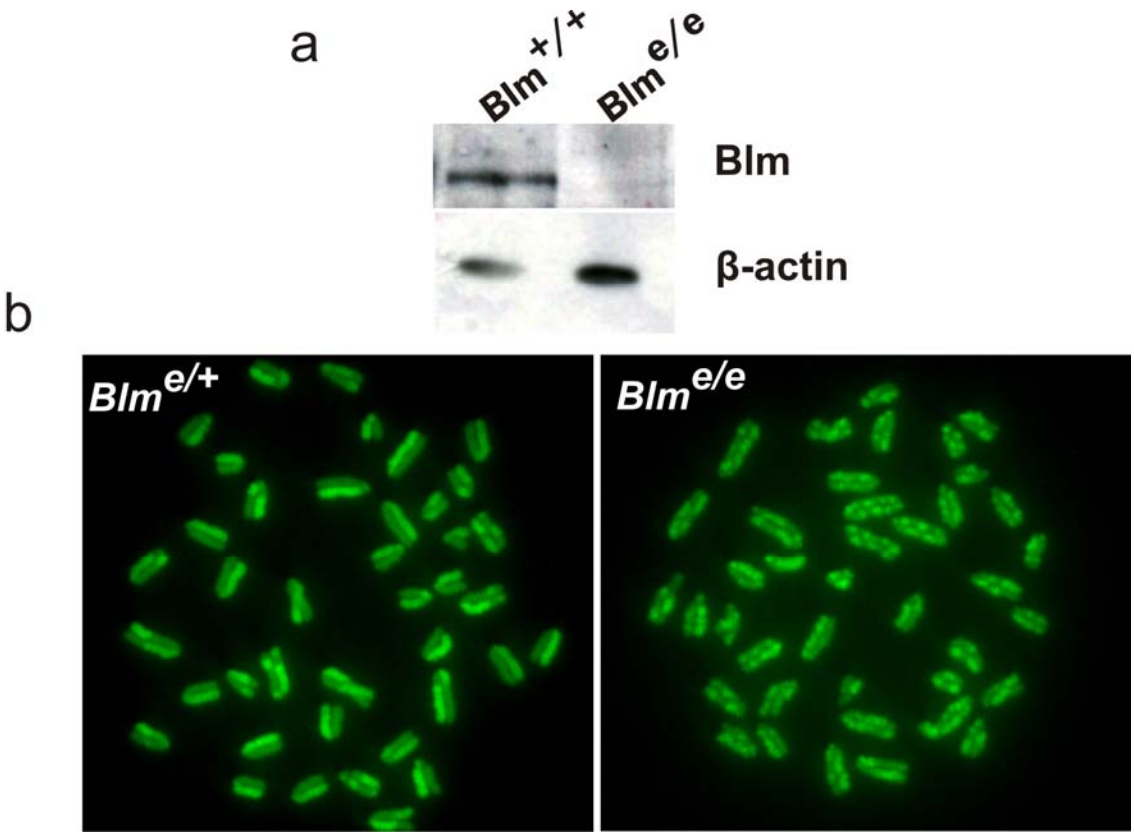
2.2. Phenotypic assessments of the *Blm*$^{e/e}$ cell line

Although the targeted replacement of exon 2 should give rise to a null mutant of *Blm*, it is important to assess the absence of Blm protein and to confirm *Blm*-deficiency phenotypically. The Blm protein level of the *Blm*$^{e/e}$ cell line was assessed by Western blotting, Figure 4-2a. No protein product was detected in the *Blm*$^{e/e}$ cells, suggesting that the targeting completely abolished *Blm* expression.

One of the main characteristics of *Blm*-deficient mammalian cells is hyper-recombination between homologous chromosomes and between the sister chromosomes. Sister chromatid exchange, SCE, can be visualised on metaphase preparation of cells with the sister chromatids differentially labelled with BrdU. Both of the heterozygous and homozygous *Blm* mutants (*Blm*$^{e/+}$ and *Blm*$^{e/e}$) were treated with Brdu in two consecutive cell cycles (approx. 34 hours). Metaphase spreads were prepared which were stained with Acridine orange for SCE visualisation. BrdU is a thymidine analogue, which can be incorporated into the newly synthesised DNA strand during S phase. After the first round of DNA replication in the presence of BrdU, the sister chromatids are equally labelled, with only the newly synthesized strand of the DNA containing BrdU. In the second round of DNA synthesis, the two sister chromatids are differentially labelled. One sister chromatid contains both strands of the DNA labelled with BrdU, while the other sister chromatid which inherited the original template strand (without BrdU labelling) has only one strand of the DNA labelled with BrdU. Acridine orange binds to dsDNA and emits green fluorescence upon binding. BrdU can quench the Acridine orange fluorescence, thus the sister chromatid with both DNA strands labelled with BrdU appears dimmer than the one with only single DNA strand labelled.

In *Blm*$^{e/+}$ cells, most of the chromosomes do not have SCE, occasionally one exchange per chromosome can be observed. However, in *Blm*$^{e/e}$ cells, SCE occurs in almost all chromosomes, and four or even more exchanges per chromosome are commonly observed. Figure 4-2b shows the typical images of metaphase spreads of *Blm*$^{e/+}$ and *Blm*$^{e/e}$ cells . Taken together, the newly generated ES cells with both *Blm* alleles targeted show the cellular characteristics of Blm deficiency.

Figure 4-2: Functional validation of the new *Blm*-deficient ES cell line *Blm*$^{e/e}$.



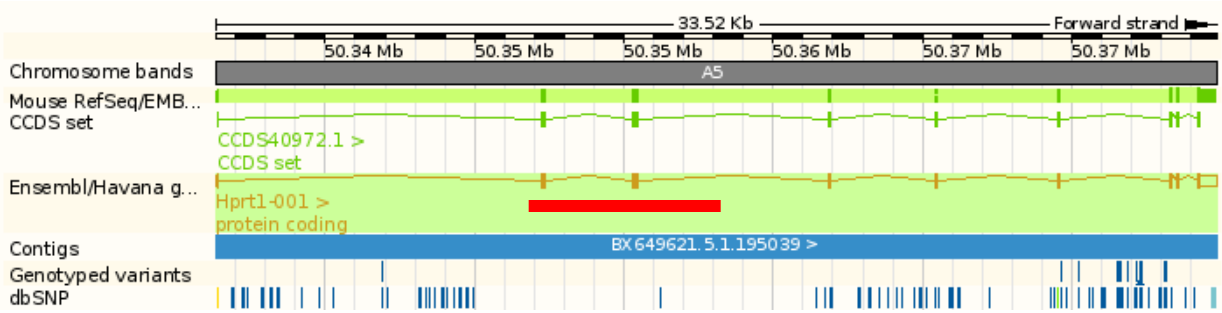
a, Western blot of wild-type (*Blm*$^{+/+}$) and *Blm*-deficient (*Blm*$^{e/e}$) ES cells. b, Sister chromatid exchange analysis of heterozygous (*Blm*$^{e/+}$) and homozygous (*Blm*$^{e/e}$) *Blm* mutant cells. The sister chromatids are differentially labelled with BrdU and visualised by staining the slides with acridine orange. A SCE is an abrupt point of colour exchange on both sister chromatids.

2.3. Gene targeting of the mutagenic PB to the *Hprt* locus in *Blm*$^{e/e}$ cells

With the *Hprt*-proficient *Blm*-deficient ES cell line established, the mutagenic *piggyBac* transposon was introduced into the *Hprt* locus to generate a *Blm*-deficient ES cell line with a single-copy PB transposon. Two cell lines were generated with the PB transposon targeted in two different locations within the *Hprt* locus, as local surrounding sequence may affect the PB excision efficiency. The designs of both of the targeted alleles are shown in Figure 4-3a,e.

The first cell line was made by introducing my mutagenic PB within the intron 2 of the *Hprt* locus (BlmHprtPBin2 cell line). The targeting vector was constructed with homologous arms derived from a 129 strain-based *Hprt* intron 2 targeting vector obtained from Haydn Prosser (Prosser *et al.*, 2008). Strain-specific single-nucleotide polymorphisms (SNPs) introduced by the targeting arms to the ES cell genome can be present in exons and introns. The genomic region where the targeting arms reside does not show any SNPs in the SNP database (dbSNP), suggesting that this region is highly conserved (Figure 4-3). Therefore introduction of 129 strain-based DNA fragment into the C57BL/6 ES cell genome is unlikely to introduce any variation which may affect Hprt function. The PB recognition site "TTAA" was introduced immediately adjacent to the two PBITRs within the targeting arms to ensure the efficient excision of the PB from the donor site. The mutagenic PB was placed in such an orientation that the *Dom3z* mutagen unit should mediate endogenous Hprt trapping in correctly targeted clones, thus the cells should be HAT sensitive.

Figure 4-3: Genetic variations with the *Hprt* targeting region.



Ensembl screen shot of the entire Hprt gene. The red line highlights the region covering both of the targeting arms. The coordinates are based on NCBI Build37.

The linearised HprtTVPB targeting vector was electroporated into the $1x10^7$ *Blm*$^{e/e}$ cells and puromycin selection was initiated 24 hours post-electroporation and colonies were picked after seven days. 6-TG selection was not used for the isolation of correctly targeted clones, as homologous recombination can occur in a portion of the daughter cells of the electroporated parental cells. Under 6-TG selection, the correctly targeted cells can be cross killed by their neighbouring untargeted sister cells due to the toxic metabolite sharing. Long range PCR was

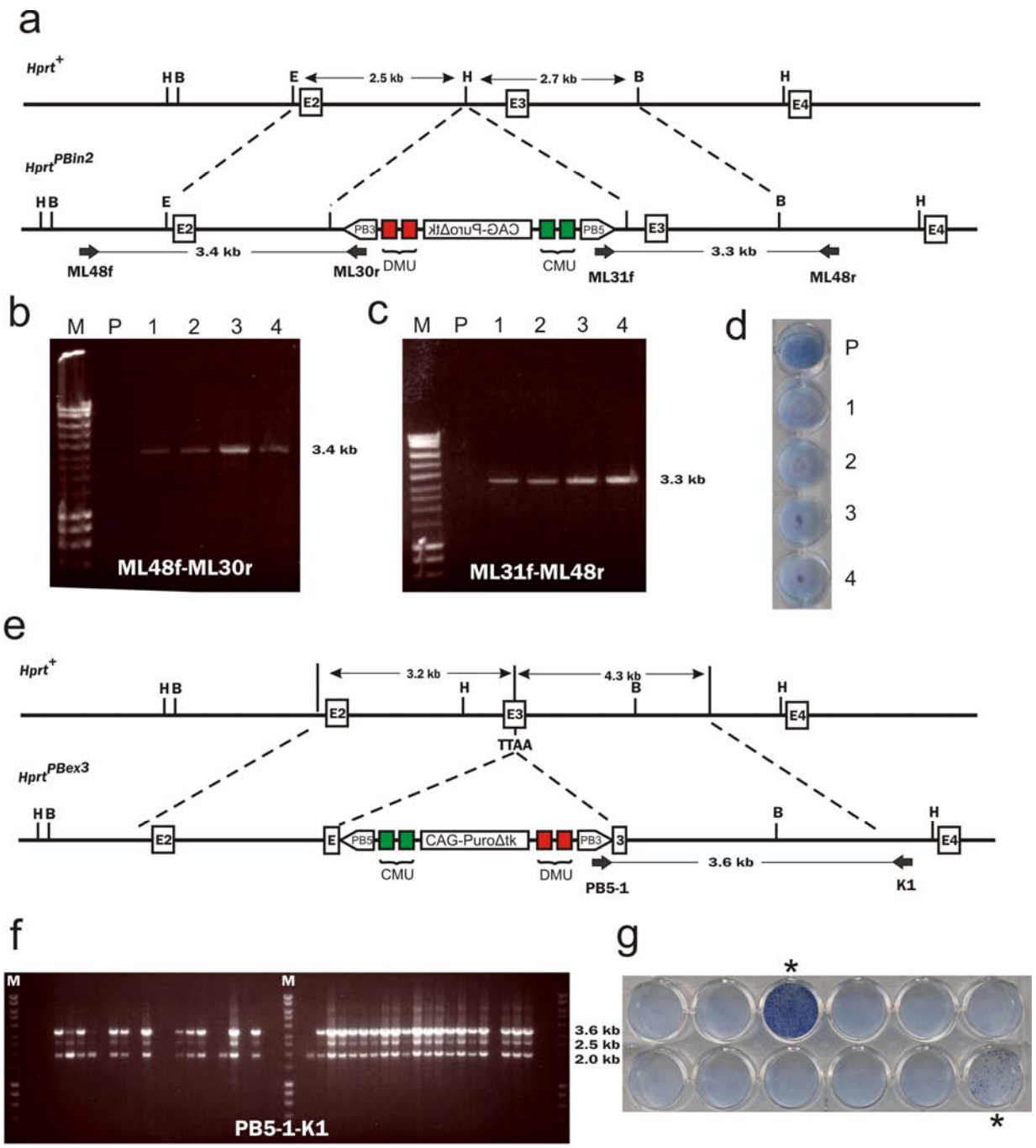
conducted to detect the homologous recombination of the short targeting arm to screen for targeting-positive clones (primers ML48f and ML30r), and the homologous recombination at the other end of the targeting arm was also confirmed by long-range PCR (primers ML48r and ML31f), Figure 4-4b,c. In correctly targeted clones, the *Dom3z* mutagenic unit should trap the *Hprt* expression, thus cells should be sensitive to HAT, whereas random integration of the targeting vectors will not cause *Hprt* inactivation. Therefore, HAT sib-selection was conducted on the PCR positive clones to further validate the gene targeting functionally, Figure 4-4d.

The second cell line was made by introducing the PB into a TTAA site within exon 3 of the *Hprt* gene (BlmHprtPBEx3 cell line), Figure 4-4e. C57BL/6N BAC DNA was used to construct the targeting vector, HprtTVE3PB. The gene targeting procedure was conducted as described above, 29 out of 48 clones were positive for the targeting events screened by the Long-range PCR (with primers PB5-1 and K1) detecting the homologous recombination of the short targeting arm, Figure 4-3f. The HAT selection was also conducted to confirm *Hprt* inactivation by targeted insertion of the PB transposon, Figure 4-4g. Most of the clones show complete sensitivity to HAT, suggesting that the clones are unlikely to contain mixed untargeted cells.

2.4. PB-remobilisation assessment of the newly generated cell lines

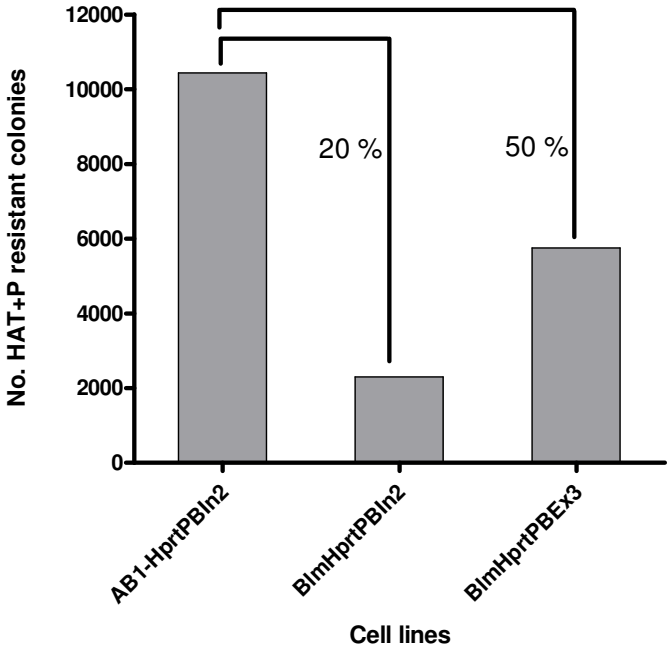
BlmHprtPBIn2 and BlmHprtPBEx3 cell lines were assessed for the efficiency of PB remobilisation. Twenty five µg of CMV-HyPBase plasmid was electroporated into $1x10^7$ cells for both cell lines as well as the positive control, AB1- $Hprt^{PBin2}$, which contains an identical targeting allele of the mutagenic PB transposon in intron 2 of the *Hprt* locus, shown in Figure 4-4a. A tenth of each electroporation was plated on a 90 mm plate and HAT and puromycin double selection was commenced 24 hours post-electroporation to select for PB excision and reintegration, respectively. The colony numbers were counted after ten days selection and are shown in Figure 4-5.

Figure 4-4: *Hprt* targeting of the mutagenic PB transposon.



a, The *Hprt* intron 2 targeted allele with the PB transposon. b,c, long-range PCR confirmations of the targeted clones 1-4. P, parental untargeted cells. d, HAT selection on the targeted clones with the parental untargeted cells as the control. e, The *Hprt* exon 3 targeting with the PB transposon. f, Long-range PCR screening of the homologous recombination on the shorter targeting arm end with PB5-1 and k1 primers. The larger molecular weight PCR product is the correct size whereas others are non-specific PCR products. g, HAT selection on 12 randomly selected PCR-positive clones. Two clones (HAT resistant, marked with *) contain non-targeted cell mixtures. M, Bio-Rad hyperladder I marker.

Figure 4-5: PB re-mobilisation efficiency assessments.



The PB re-mobilisation efficiency was compared among three cell lines with AB1-HprtPBIn2, BlmHprtPBIn2 and BlmHprtPBEx3. AB1-HprtPBIn2 was Blm proficient with PB targeted into intron 2 of the *Hprt* locus. BlmHprtPBIn2 and BlmHprtPBEx3 were derived from Blm-deficient ES cells. PB was targeted to the intron 2 of the *Hprt* locus in BlmHprtPBIn2, while PB was targeted to the intron 3 of the *Hprt* in BlmHprtPBEx3. The number of HAT and puromycin resistant colonies plotted were the total number obtained per $1x10^7$ cells. The percentage shown in the graph represents the percentage of HAT and puromycin double resistant colonies in the *Blm*-deficient cell lines *vs* the positive control. P, puromycin.

Compared to the positive control, the PB remobilisation efficiency in the two newly generated *Blm*-deficient cell lines was lower. Similar experiment was done with HAT selection alone to measure the excision efficiency and the results were similar to that obtained with HAT and puromycin double selection. PB transposon targeting in the BlmHprtPBIn2 cell line was identical to the AB2.2-HprtPBIn2 cell line, however, PB remobilisation efficiency in the BlmHprtPBIn2 and BlmHprtPBEx3 cell lines were only 20 % and 50 % of AB2.2-HprtPBIn2 cell line, respectively.

3. Discussion

In this chapter, I have described the generation of a *Blm*-deficient but otherwise wild-type C57BL/6N ES cell line. The proficiency of wild-type *Hprt* expression permits the use of this locus as a donor site for PB mediated genome-wide re-mobilisation, previously described in Chapter 3. The *Blm*-null ES cell line was generated by sequential gene targeting of a selection cassette to replace the start codon-containing exon 2 of the *Blm* gene, a strategy that has been used previously to generate the Blm-knockout ES cells (Luo *et al.*, 2000; Guo, 2004). The selection cassette used here was an essential part of the reporter strategy for the miRNA biogenesis and effector pathway screening design. The detailed features of this reporter strategy are described in Chapter 5.

The new Blm-deficient ES cells ($Blm^{e/e}$) were confirmed to be null since Blm protein expression could not be detected. In addition, Blm deficiency was further assessed functionally by accessing the frequency of the SCE. A highly elevated level of SCE was observed in these *Blm*-deficient ES cells compared to their heterozygous counterpart. Therefore, my *Blm*-deficient ES cell line is a true *Blm* loss-of-function mutant.

The mutagenic PB transposon was introduced into the *Hprt* locus of the $Blm^{e/e}$ cells in two independent loci, and both mutagenic units can mediated efficient *Hprt* trapping and gave rise to HAT sensitive targeted clones. This data confirmed further of the mutagenicity of my PB transposon. PB re-mobilisation efficiency was compared in these newly generated cell lines.

PB re-mobilisation in the AB2.2-HprtPBIn2 cell line is five times more efficient than in the BlmHprtPBIn2 cell line. The major difference between these two cell lines is their genetic backgrounds. The AB2.2-HprtPBIn2 cell line was derived from the 129 SvEvBrd strain, whereas the BlmHprtPBIn2 cell line was derived from C57BL/6 strains. Moreover, the AB2.2-HprtPBIn2 cell line is *Blm*-proficient whereas the BlmHprtPBIn2 cells are *Blm*-null. The PB transposition is host independent, as highly efficient transposition can occur in a wide range of species (Chapter One). Thus the difference in genetic background may not be the cause for the

reduction in PB transposition efficiency. Thus the observed difference may be due to *Blm* deficiency, although the connection between the DNA transposon excision and Blm has not yet been investigated.

Blm is known to be involved in DNA double-stranded break (DSB) repair using the local sequence microhomology, and Blm may function in unwinding the DNA at the breakpoint in order to allow the local microhomology search to repair the DSB (Langland et al., 2002). Although DSB can be repaired via other system such as non-homologous end joining, the loss of Blm may partially compromise the DSB repair efficiency upon PB transposon excision. Cell-cycle arrest or apoptosis may be induced in cells without efficient DSB repair, thus contributing to the reduction in colony number. If PB undergoes fast transposition kinetics with several cycles of excision and reintegration occurring per cell, cells harbouring rounds of DSB generated by PB remobilisation are more prone to cell death. Such selection pressure in the *Blm*-deficient background may also enrich cells with less PB intra-genomic mobilisation, and possibly enhance the local hopping events. In yeast, the absence of both Sgs1 (Blm homologue in yeast) and Exo1 causes a pronounced hypersensitivity to DSB inducing agents and the cells have compromised DSB resection, DNA damage-mediated signalling and strongly impaired homologous recombination-mediated repair (Gravel et al., 2008). BlmHprtPBIn2 and AB1-HprtPBIn2 cell lines generated here are very useful reagents for the future investigation into the PB transposon mediated DSB and Blm-mediated DSB repair. Research into the relation between the two could potentially reveal biological insights into the mechanisms of the mammalian DNA repair systems which are used to eliminate DNA transposition-mediated genomic insults.

PB re-mobilisation from the BlmHprtPBEx3 cell line performed three times more efficient than in BlmHprtPBIn2 cell line. This difference may be due to differences in the excision efficiency, as the PB transposon was introduced into slightly different genomic sequence contexts. Therefore, it is possible that AT-rich local DNA sequence context may favour the PB excision. Furthermore, local methylation status and chromatin structure may also influence the PB excision efficiency.

With respect to the genome-wide insertional mutagenesis, the BlmHprtPBEx3 cell line is a better choice of the two due to its higher PB remobilisation efficiency, Figure 4-5. Despite the reduction in PB remobilisation efficiency observed in *Blm*-deficient ES cells, per electroporation of $1x10^7$ of BlmHprtPBEx3 cells, approximately 6,000 HAT and puro double resistant colonies can be generated representing the PB transposon excised from the *Hprt* locus and reintegrated elsewhere in the genome, respectively, Figure 4-5. Based on the previous estimate, 47 % PB insertions are in genes and 80 % of which land in actively transcribed genes in ES cells, assuming no local PB hopping phenomenon is present (Liang et al., 2009). Thus 40,000 heterozygous mutants can be generated in merely seven electroporations in BlmHprtPBEx3 cells, to cover approximately 15,000 PB insertions in actively transcribed genes in ES cells. Therefore in terms of generating a genome-wide mutant library, such a cell line is sufficient enough to cover the whole ES cell expressed genome, assuming half of the genome is transcribed in ES cells. Therefore, the *Blm*-deficient cell line with a mutagenic PB transposon residing in the *Hprt* locus constitutes a very simple, useful and efficient means to generate genome-wide insertional mutagenesis with a tightly controlled mutagen copy number. Libaries of mutants generated using such a cell line can support recessive genetic screens for phenotypically selectable pathways, such as phenotypes involving viral, toxin and drug resistance (Yusa *et al.*, 2004a; Wang and Bradley, 2007; Wang *et al.*, 2008a). This cell line is also highly versatile, as any genetic modification can be engineered into the cell line prior to library generation to suit screen-specific requirements for probing phenotypes which are not directly selectable. This aspect is demonstrated in the latter part of this thesis by introducing reporter systems to probe the miRNA biogenesis and effector pathway genes.