CHAPTER 2. MATERIALS AND METHODS

2.1 Cell culture

2.1.1 Mouse ES cell culture

The XY AB1 and AB2.2 cell line were derived from a black agouti 129 Sv embryo described previously (Evans *et al.* 1981). The AB2.2 (129 S7/SvEv$^{Brd\text{-}Hprt^{b\text{-}m2}}$) cell line has an inactivated *Hprt* gene (on the X chromosome) mutated by a recombinant retrovirus (Kuehn *et al.* 1987). These cells and their derivatives were cultured on irradiated (60 Gray) mitotically inactivated monolayer SNL 76/7 STO cells (McMahon *et al.* 1990). The SNL 76/7 cell line is G418^R and expresses leukaemia inhibitory factor (LIF) to maintain ES cells in an undifferentiated state (Williams *et al.* 1988).

ES cells were cultured at 37°C with 5% CO_2 in M15 media (Knockout™ D-MEM (Invitrogen 10829-018) supplemented with 15% foetal bovine serum (Hyclone; CatNo. CH30160.03; LotNo. CPCO285), 2 mM L-glutamine, 50 units/ml penicillin, 40 mg/ml streptomycin and 0.1 mM b-mercaptoethanol (b-ME)). In cell culture, M15 media with supplements was renewed daily if not specified.

2.1.2 *Blm*-deficient ES cell line

The cell line I used in this project is called NGG5-3 (Figure 2-1), which has two targeted *Blm* alleles *Blm*(m1/m3) (Luo *et al.* 2000). The NGG5-3 cell line is derived from AB2.2, an *Hprt*-deficient ES cell line (Kuehn *et al.* 1987). NGG5-3 is cultured on irradiated (60 Gray) SNL 76/7 STO feeder cells. The *Blm*^{tm1Brd} allele contained a *loxP*-flanked *PGK-neo* cassette, which was removed by Dr Ge Guo to generate the *Blm*^{tm4Brd} allele using Cre-*loxP* recombination (Guo 2004).

To enable a screen for genes involved in mismatch repair in *Blm*-deficient cells, the *Hprt* gene product has to be present in the cell line. Thus, an *Hprt* mini-gene was introduced into both autosomal copies of the growth differentiation factor 9 (*Gdf9*) gene loci by sequential gene targeting. This resulted in a null allele of *Gdf9* missing exon 2. As *Blm*-deficient cells have a high rate of LOH, heterozygous alleles may be lost during cell replication. Therefore, two copies of *Hprt* were targeted to *Gdf9*, ensuring that the remaining cells are *Hprt* positive. The *Gdf9* gene has only two exons and is required for sex development in female mice. Importantly, the expression of *Gdf9* should be destroyed and the absence of *Gdf9* does not have any effect on ES cells (Dong *et al.* 1996). The first allele was targeted using Dr Jinwen

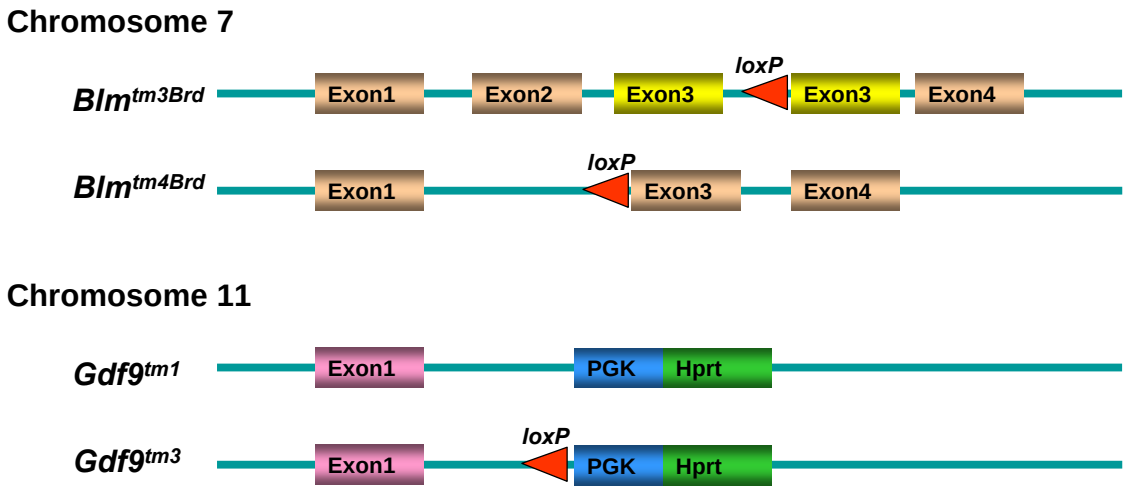


Figure 2-1 Genotype of the *Hprt*$^{+/+}$ *Blm*-deficient ES cell line NGG5-3

In the cell line NGG5-3, both copies of the *Blm* and *Gdf9* genes are null. The *Blm*^{tm3Brd} has all of the endogenous *Blm* exons plus an extra copy of exon 3; the *Blm*^{tm4Brd} allele lacks exon 2. Both mutations cause a frame shift, thus they are believed to be null alleles for *Blm*. Both *Gdf9*^{tm1} and *Gdf9*^{tm3} lack exon 2 of *Gdf9*. The two copies of *PGK-Hprt* mini-genes cause the NGG5-3 cell line to be sensitive to 6TG and resistant to HAT.

Dong's vector. The second targeting vector was modified by Dr Ge Guo so that it had a *loxP*-flanked *PGK-neo* cassette inserted in front of *PGK-Hprt*. The *Neo* selection marker was removed by Cre after the second targeting resulting in the *Gdf9*^{tm3} allele (Guo 2004).

2.1.3 *Dnmt1*-deficient ES cell line

A *Dnmt1*-deficient ES cell line was derived from the previously established *Dnmt1* gene knockout cell line (Guo *et al.* 2004). From the reported *Dnmt1*$^{tm1/tm2}$ cell (Guo *et al.* 2004), a cell line was generated by Dr Wei Wang by removing the neomycin cassette through Cre-*loxP* recombination, which left two alleles with the same structure (*Dnmt1*$^{tm2/tm2}$, Figure 2-2). These have a *loxP* site in both alleles and exons 2–4 deleted. The deletion of exons 2–4 removes a 359-bp coding region, resulting in a frame shift mutation, which can be expected to be a null mutation (Guo 2004). I have targeted one allele of the *Gdf9* locus in this cell line with an *Hprt* mini-gene using the same targeting vector, which produced *Hprt*-proficiency in the *Blm*-deficient cell line NGG5-3 (Dong *et al.* 1996).

2.1.4 Chemicals and media used for ES cell culture

M15 medium: 500 mL Knockout™ D-MEM (Invitrogen 10829-018) is supplemented with 15% foetal bovine serum (Hyclone; CatNo. CH30160.03; LotNo. CPCO285), 2 mM L-Glutamine, 50 units/ml Penicillin, 40 mg/ml Streptomycin and 0.1 mM b-mercaptoethanol (b-ME). Storage: at 4°C.

Foetal bovine serum: Hyclone; CatNo. CH30160.03; LotNo. CPCO285. Storage: at -20°C.

Blasticidin: Blasticidin S HCl (Invitrogen, Cat. No. R210-01), 1000× stock (5 mg/ml) was made in Phosphate Buffered Saline (PBS). After mixing, the 1000× stock solution was sterilized by filtering through a 0.2 µm syringe filter. Storage: at –20°C.

FIAU: 1-(2'-deoxy-2'-fluoro-β-D-arabinofuranosyl)-5-iodouracil, 1000× stock
(200 µM) was made in PBS and 5 M NaOH was added dropwise until it was dissolved. After mixing, the 1000× stock solution was filter sterilized through a
0.2 µm syringe filter. Storage: at –20°C.

G418: Geneticin (Invitrogen, Cat. No. 10131) was bought as a sterile stock solution containing 50 mg/ml active ingredient. Storage: at –20°C.

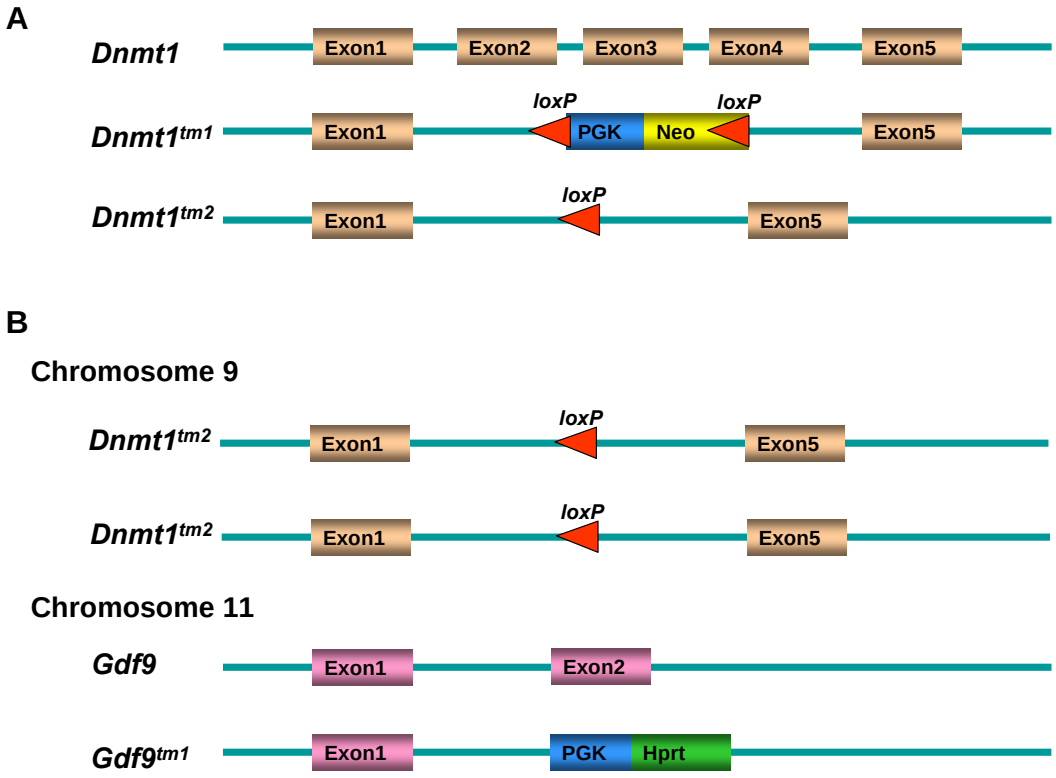


Figure 2-2 Structure of the *Dnmt1* locus in the targeted cell lines

A. Targeting strategy of the *Dnmt1* gene A *loxP*-flanked *PGK-neo* cassette replaces exons 2–4 of the *Dnmt1* gene to generate allele tm1 through homologous recombination. The allele tm2 is generated from tm1 through Cre-*loxP* recombination. **B**. Schematic structure of the *Dnmt1* locus and the *Gdf9* locus in the *Dnmt1*-deficient *Hprt*-proficient cell line.

Puromycin: ($C_{22}H_{29}N_7O_5$.2HCL, Sigma) 1000× stock (3 mg/ml) was made in MiliQ water. After mixing, the 1000× stock solution was filter sterilized through a 0.2 µm syringe filter. Storage: at -20°C.

HAT: a liquid sterile mixture of sodium Hypoxanthine, Aminopterin and Thymidine (50×, Invitrogen, Cat. No. 21060-017). HAT supplement contains 5 mM Hypoxanthine, 20 µM Aminopterin and 0.8 mM Thymidine. Storage: -5°C to -20°C.

HT: a sterile mixture of sodium Hypoxanthine and Thymidine (100×, Invitrogen, Cat. No.11067030). HT supplement contains 10 mM Hypoxanthine and 1.6 mM Thymidine.

6TG: (2-amino-6-mercaptopurine (Sigma)) 5 mM stock was made in PBS and 5 M NaOH was added dropwise until it dissolved. Storage: -20 °C.

PBS (Phosphate Buffered Saline): PBS is prepared in 10 litre quantities, aliquoted into sterile disposable bottles, and stored at room temperature. Mix NaCl (80.0 g), KCl (2.0 g), $Na_2HPO_4.7H_2O$ (10.72 g) and KH_2PO_4 (2.0 g) in 8 Litre of Milli-Q water, then bring the total volume to 10 L. Adjust the PH to 7.2 with a saturated solution of $Na_2HPO_4.7H_2O$. Add phenol red until a peach colour is achieved.

GPS (Glutamine-Penicillin-Streptomycin) 100×Stock Solution: Thaw the 100 ml bottle(s) of 200 mM Glutamine (filtered to be sterile, stored at -20°C). From each bottle, aseptically remove 10 ml of the 200 mM Glutamine to a 15 ml tube. Add the following to the tube and dissolve by mixing: Penicillin (300.0 mg), Streptomycin (500.0 mg). After mixing, filter sterilizes the GPS solution through a 0.2 µm syringe filter back into the original 100 ml bottle of Glutamine. Mix thoroughly. The resulting 100× solution is: 200 mM L-Glutamine; 4,950 U/ml Penicillin; and 500 mg/ml Streptomycin. Aliquot from the original bottle into 25 ml Nunc Universals, label with "100× GPS" and the date, and store at 4°C.

Trypsin Stock solution: This is prepared in 5 L quantities, filter-sterilized, aliquoted into 50 ml tubes, and stored -20°C. Mix the following in Milli-Q water, then bring the total volume to 5 L: NaCl (35.00 g), D-Glucose (5.00 g), $Na_2HPO_4.7H_2O$ (0.90 g), KCl (1.85 g), KH_2PO_4 (1.20 g), EDTA (2.00 g), Trypsin (12.50 g, Invitrogen, Cat. No. 840-7250IL) and Tris Base (15.00 g); these ingredients generally take 30-60 minutes to dissolve thoroughly. Adjust the pH from its initial 8.71 to pH 7.6 with HCl. Add phenol red until a pink colour is achieved.

BME (β-Mercaptoethanol) Stock Solution (100×, 10^{-2} M): Add 72 µL of 14 M β -Mercaptoethanol to 100 ml of PBS and mix well. Filter-sterilize it using a 0.22 µm filter and aseptically decant into 50 ml tubes. Storage: 4 °C.

Gelatin (0.1%): The gelatin is used to treat (coat) tissue culture plasticware so the cells will adhere better to the plate surface. It is prepared in 4 litre quantities, sterilized by autoclaving, aliquoted into sterile disposable bottles, and stored at room temperature. To 4 L of Milli-Q water, add 4.0 g of Gelatin (Sigma; from porcine skin, 300-bloom; cell culture tested). Add phenol red until a bright yellow colour is achieved; swirl to evenly distribute. Autoclave on liquid cycle.

Cell staining buffer: methylene blue 2%, 70% ethanol.

2.1.5 Passaging ES cells

When ES cells were 70-90% confluent on feeder plates, media was replaced 2-4 hours before passaging them. ES cell were washed twice with PBS and trypsin solution added to the plate (50 uL for a well of 96-well plate; 150 uL for a well of 24-well plate; 0.3 mL for a well of 6-well plate; 2 mL for a 10 cm plate). After incubation at 37°C for 10-15 minutes, an appropriate volume (1–5 times volume of trypsin solution used) of M15 was added. Cells were pipetted up and down vigorously to breakup the clumps and re-plate at ratio of 1:3 or 1:4. Cells usually took 2–4 days to reach confluence.

2.1.6 Freezing ES cells

Cells are best frozen at 70–80% confluence. Cells were reefed by M15 media 2–4 hours before they were frozen. Briefly, ES cells were suspended by trypsin according to passaging protocol. Suspended cells were harvested by centrifuge (1000rpm 3 min) and resuspended in one volume of M15 and an equal volume of freezing medium. Usually cells were frozen at density about 10^7 per mL. Freezing media has 60% Knockout™ D-MEM (Invitrogen 10829-018), 20% foetal bovine serum (FBS) and 20% DMSO (Dimethyl Sulfoxide). Cells can be frozen in 96-well / 24-well plates (wells were covered by 0.2 µm filtered mineral oil) or Cryovials at -80°C. If in Cryovials, the temperature decrease is controlled to approximate 1°C per minute. Then Cryovials can be transferred into liquid nitrogen freezer the second day.

2.1.7 Electroporation of DNA into ES cells

DNA used for electroporation was usually prepared by the standard alkaline lysis mini-prep method or by QIAGEN Plasmid Maxi Kit (Cat. No.12163). DNA used for gene-targeting was linearized by enzyme digestion with an appropriate enzyme (NEB) under the conditions recommended by the manufacturers. Before electroporation, DNA was precipitated by ethanol and air-dried in a tissue culture (TC) hood. 5–20 ug of DNA (0.5–1.0 μg/μL) was used in appropriate experiments.

Mouse ES cell electroporation was conducted as described before (Ramirez-Solis *et al.* 1993). Briefly, 70–90% confluent ES cells were fed 2–4 hours before media was removed. The cells were washed with PBS twice and 2 mL of trypsin was applied to a 90 mm plate, which was incubated at 37°C for 15 minutes. Then 8 mL M15 media was added and the cells were pipetted repeatedly to make a suspension of single cells. This mixture was centrifuged at 1000 rpm for 3 minutes and the supernatant was removed. Cells were suspended in PBS at the concentration of 1×10^7 cells/mL. Cells then were transferred into 0.4 cm gap cuvette (Bia-Rad) after mixing with 5–20 ug DNA. The gene pulser (Bio-Rad) was configured at 230 V, 500 μF. After electroporation, cells were plated onto 10 cm feeder plate. Selective drugs were usually applied with M15 media one day after electroporation.

2.1.8 Picking ES cell colonies

Colonies need 8–10 days to grow up from a single cell. They were washed twice by PBS immediately before picking. With a P20 Pipetman®, individual colonies were picked into round bottom wells in 96-well format, which were filled with 50 uL trypsin. Colonies were incubated at 37°C for 10–15 minutes then cell clumps were broken up into single cells by pipetting in 50 uL M15 medium and transferred into 96-well flat bottom feeder plates and cultured in normal ES cell culture conditions.

2.1.9 Production of recombinant retrovirus

The *PuroΔtk* (Chen, Y. T. *et al.* 2000) containing recombinant MMuLV retrovirus was produced by the B4-5 virus producing cell line, a gift from Wei Wang (Wang and Bradley 2007). *PuroΔtk* is a bifunctional fusion protein between puromycin N-acetyltransferase (Puro) and a truncated Herpes Simplex Virus type 1 thymidine kinase (*Δtk*). Murine embryonic stem (ES) cells transfected with *PuroΔtk* become resistant to puromycin and sensitive to 1-(-2-deoxy-2-fluoro-1-beta-D-arabino-furanosyl)-5-iodouracil (FIAU) (Chen, Y. T. *et al.* 2000). B4-5 cells were cultured in M15 ES cell media until they are >80% confluent.

Supernatant was collected after 24 hours (containing MMuLV recombinant virus) was filtered (0.2 μm) before it was used to infect ES cells. Filtered MMuLV containing supernatant can be stored at -80°C. Virus was added in the media then infected cells were selected in puromycin media. Virus titer is calculated as puromycin resistant colonies generated per mL of virus containing supernatant used. The virus titer will vary when different numbers of cells are used for the infection. About 600 infected clones were obtained when 60 μL virus containing supernatant was used to infect 3×10^6 cells. Thus, the virus titer is 10^4 CFU/mL.

2.1.10 Recombination Mediated Cassette Exchange

Recombination mediated cassette exchange (RMCE) is a technology, in which a DNA fragment (cassette) can be introduced into a host genome (exchange) at a specific site through Cre-*loxP* recombination (Figure 2-3). Bacterial Artificial Chromosomes (BACs) can be used for this purpose (Liu *et al.* 2006). Considering the average size of many BACs (~200 kb), it is possible to carry most genes. Transforming a BAC into a mutant genome by RMCE provides a practical way to complement a mutant gene's function.

To introduce single copy BAC DNA into genome, an efficient acceptor site and a selection scheme are needed. The discovery of heterospecific *lox* sites makes it possible to transform a DNA fragment into a defined location in the mouse genome with a specific orientation (Baer *et al.* 2001). A single positive selection can be used to select for BAC insertions, but this does not rule out random integrations of the BAC. RMCE will normally give single copy insertions. Thus, multicopy of the transformed exogenous gene is avoided.

Heterospecific *lox* sites targeted to the X-linked hypoxanthine phosphoribosyltransferase (*Hprt*) locus provide a positive-negative selection scheme. The *Hprt* gene was the first locus mutated in ES cells (Kuehn *et al.* 1987) and it was found to exert a minimal influence on transgene expression (Heaney *et al.* 2004). *Hprt* has also been used as a defined locus to accept a single copy transgene by homologous recombination (Bronson *et al.* 1996; Heaney *et al.* 2004). When an exogenous gene is transformed into *Hprt* locus and destroys *Hprt* gene transcription, cells are resistant to 6TG. The combination of this negative selection for loss of *Hprt* function and positive seletion for the exogenous fragment provide a straightforward selection system to isolate single copy BAC insertions at a specific site. Using RMCE, functional copies of human α Globin and human cardiac sodium channels were used to replace the mouse genes (Liu *et al.* 2006; Wallace *et al.* 2007).

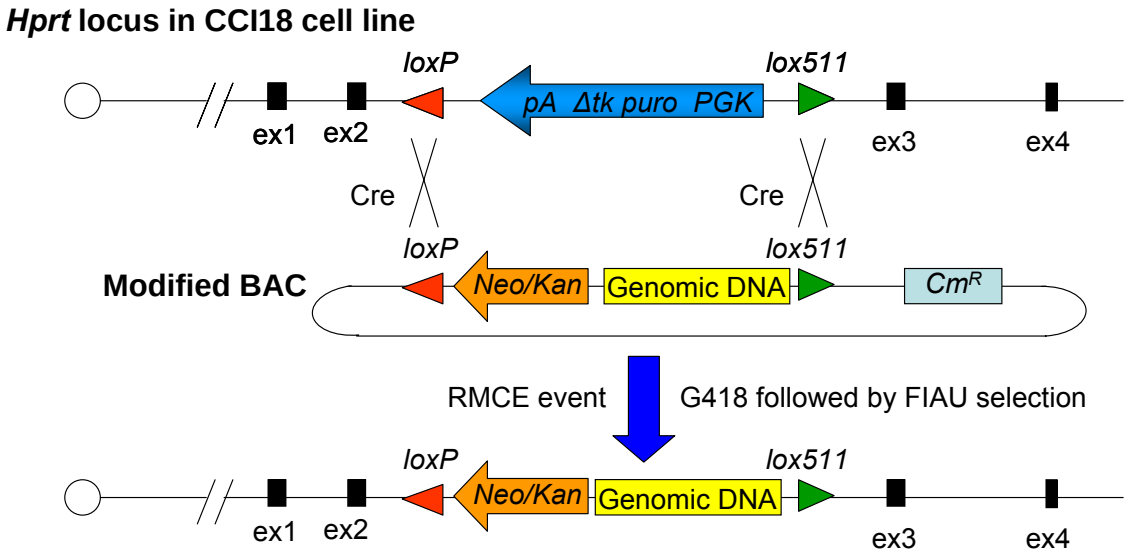


Figure 2-3 The principle of Recombination Mediated Cassette Exchange

Genomic DNA insert from mouse BAC libraries, such as RPCI-23, 24 can be inserted into *Hprt* locus of CCI 18 mouse ES cell line through site directed recombination. BAC: Bacteria Artificial Chromosome. Cm^R: Chloramphenicol antibiotic resistance gene. *Neo/Kan*: *Neo*mycin/Kanamycin resistance gene. *PGK*: phophoglycerate kinase promoter. *PuroΔtk*: a positive-negative selection marker, which is resistant to puromycin and sensitive to FIAU (Chen, Y. T. *et al.* 2000). Ex: Exon.

Dr Prosser in Prof. Bradley's laboratory introduced *PGK-PuroΔtk-pA* cassette flanked by *loxP* and *lox511* (Hoess *et al.* 1986) sites into *Hprt* intron 2 of AB1 mouse ES cell line (McMahon *et al.* 1990). The resulting unpublished cell line CCI18 has the capacity to accept BAC DNA into *Hprt* locus through Cre mediated site directed recombination – recombination mediated cassette exchange (Prosser 2007 unpublished).

Most BAC libraries can be adapted in a single overnight step for RMCE. Mouse (C57BL/6J) BAC libraries RPCI-23, 24 have *loxP* and *lox511* sites, located at either side of genomic DNA insert. Before BACs are electroporated into ES cells, a neomycin/kanamycin resistance gene (*Neo/Kan*) is inserted into the BAC by recombineering. The locus used for this purpose, the *SacBII* gene, is located between the two recombination sites: *loxP* and *Lox511*, thus the *Neo/Kan* gene will be transferred into the *Hprt* locus of CCI18 cell line. After RMCE, positive selection (G418) for the presence of *neomycin* gene and a negative selection (FIAU) for the loss of *tk* transcript can isolate single copy BAC integration events at 100% efficiency. If RMCE is used to introduce a BAC with a functional gene into a mutant cell with a homozygous mutation at the same gene, the incoming gene will complement the mutant gene function, leads to a normal phenotype.

2.2 Gamma Irradiation

Cells were irradiated in Gammacell® 1000 Elite (MDS Nordion). This irradiation generates gamma rays from Caesium^{137} (Cs^{137}). Due to the half life of the isotope, the actual dose decreases with time. However, the half life of Cs^{137} is 30.07 years so the decrease in dose rate is minimal. Doses of ionizing radiation are usually measured by the unit Gray (Gy). 1 Gy corresponds to 1 J/kg of absorbed energy. The conversion factor with the unit "rad" used in the older scientific literature is 1 Gy = 100 rads = 100 cG. The actual dose received by cells using Gammacell® 1000 Elite is at a rate of approximately 12 cG per second.

2.3 DNA methods

2.3.1 Solutions used in molecular experiments

TE: 10 mM Tris-Cl, pH 7.5, 1 mM EDTA in dH_2O.

T0.1E: 10 mM Tris-Cl, pH 7.5, 0.1 mM EDTA in dH_2O.

Cell lysis buffer: 100 mM NaCl, 50 mM Tris pH 7.5, 10 mM EDTA (pH 8.0), 0.5% SDS.

Before use, add Proteinase K to 0.5–1.0 mg/mL.

10× TAE solution: Tris 400 mM, EDTA-Na_2-salt 10 mM and Acetic acid 200 mM, are resolved in dH_2O, adjusted to pH 8.3, autoclave sterilized.

Depurination buffer: 0.25 M HCl in dH_2O.

Denaturation buffer: 0.4 M NaOH, 1 M NaCl in dH_2O.

Neutralization buffer: 0.5 M Tris-HCl (PH 7.4), 1 M NaCl.

Southern blot Hybridization Buffer: PerfectHyb™Plus Hybridization Buffer, Sigma, CatNo.H7033.

Southern blot wash buffer: 0.2×SSC, 0.1% SDS.

2.3.2 PCR

PCR reactions were carried out using Platinum® Taq DNA polymerase (Invitrogen Cat.No.10966-018) if not specified. The reaction mixture were prepared by 5μL PCR buffer (10×, 200 mM Tris-HCl (pH 8.4), 500 mM KCl), 1.5 μL $MgCl_2$ (original 50mM stored at -20°C, final 1.5 mM), 0.4 μL dNTP mixture (original 25 mM stored at -20°C, PH8.0 adjusted by 2N NaOH, final 0.2 mM), 0.5-1 μL Platinum® Taq DNA polymerase, 5–100 ng DNA template, 0.5 μL each of upper and lower primers (original 100 μM stored in water at -20°C, final 1 μM) and appropriate volume of Mili-Q water to bring a final volume to 50 uL.

Reactions were usually designed as: 94°C 2 minutes to denature templates, 25–40 cycles of [94°C 30sec; 60–70°C annealing 45sec; 72°C extension time], followed by 72°C 2 minutes, 4°C 5 minutes. Annealing temperature depends on the primer design. Extension time ensures 1minute for every 1 kb target fragment.

2.3.3 Primers

Lower case usually indicates designed restriction enzyme sites. But in primers zk198.3, 198.4, 199.1 and 199.2, lower case indicates *lox* sites.
zk102.1: 5’- accgCTCGAGTTTTATGGACAG – 3’
zk107: 5’ - CGgaattcAATTCAGAAGAACT - 3’

zk111.1: 5’ – TCCcccgggCCTTTTGTAACCTTTCATAT – 3’

zk111.2: 5’ – GactagtAAGGTGGCATACTTCTCACC – 3’

zk112.1: 5’ – GGggtaccTTTATATCTCCCCGAACCCT – 3’

zk112.2: 5’ – CCatcgatGCAAGTCATACGACTTAGCTAC – 3’

zk113.1: 5’ – CGgaattcAGGGTTGAGTTTCCCTCATC – 3’

zk113.2: 5’ – CCCaagcttGTGTAAGTGTAAGTGCTGTG – 3’

zk114.1: 5’ – GactagtCTCACATTCAAGCCTTGAGCC – 3’

zk114.2: 5’ – GCtctagaCCCTTCCATAGACTCATTTGTG -3’

zk136.1: 5’ – AGCCATcccgggCTTGTTACAACAAAAAGAATGAAA – 3’

zk136.2: 5’ – CCATactagtTGTACACCATATTTTAATAAGTAT – 3’

zk137.1: 5’ – CCTAactagtAGAGGTTGTGCACCGTCATGTGGG – 3’

zk137.2: 5’ – CCAAtctagaCCCCCACTGACGTTAGACGGAATC – 3’

zk142.1: 5’ – CTTTATTCTTGCCATCTGGCTATC – 3’

zk142.2: 5’ – CTATGTAGAACAAAGAAAAAATCA – 3’

zk175.1: 5’ – tttggtaccAAAGTGGTTGAGTAGAAGGGCTTC – 3’

zk175.2: 5’ – tttctcgagCAGGACGTGAATTCTCTGTGCAAA – 3’

zk176.1: 5’ – tttggatccACATGAAGAAGGTCGGTCCGCAGA – 3’

zk176.2: 5’ – tttccgcggCAGAGAAGCTTCTAGGTGTGCCAG – 3’

zk184.1-3’*Blm*TV probe: 3’ – TGAGACAGGATCTTTGACCAGTTTGGCTAG – 3’

zk184.2-3’*Blm*TV probe: 3’ – AACTTGAAGTAACTGAAGTATGTTTTCAT – 3’

zk189: 5’ – CCCCGACTGCATCTGCGTGTTCCAATTC – 3’

zk198.3: 5’ – ATCataacttcgtatagcatacattatacgaagttatC – 3’

zk198.4: 5’ – TCGAGataacttcgtataatgtatgctatacgaagttatGAT – 3’

zk199.1: 5’ – AATTCataacttcgtataatgtatactatacgaagttatG – 3’

zk199.2: 5’ – GATCCataacttcgtatagtatacattatacgaagttatG – 3’

zk212.2-*Hprt*: 5’ – AGAAACATGCCAAGCATGATGGTGTGTCTCA – 3’

zk212.3-*Hprt*: 5’ – TGAACTGTCAAATTGTCCTCCAACAAGT – 3’

zk215-RC-pCMV: 5’ – ATGGGCGGGGGTCGTTGGGCGGTCAGC – 3’

zk216-RC-pCMV: 5’ – GAAATCCCCGTGAGTCAAACCGCTATCCA – 3’

zk217-BSD: 5’ – GGACGGTGCCGACAGGTGCTTCTC – 3’

zk218-BSD: 5’ – TTCGTGAATTGCTGCCCTCTGGTTATGTGT – 3’

zk219-RC-puro: 5’ – CGAGGCGCACCGTGGGCTTGTA – 3’

zk237 *Blm*: 5’ – TAAAGCTAGGCGGGTGTAGCCATGGCGTCT – 3’

2.3.4 Transformation of DNA into *E. coli* strains

A 5 mL *E. coli* (DH10B/EL350/EL250/DY380) (Liu *et al.* 2003) culture in LB media was

produced from a single colony, by shaking at 180 rpm in a 15 mL tube at 30°C (37°C for DH10B strain) overnight. The next day, this was diluted 1:10 or 1:20 to an OD_{600}=0.6–0.8 in 20 mL. The bacteria were allowed to grow for another 1–2 hours. Cells were harvested from LB media by centrifugation at 4000 rpm for 5 min at 0°C then were suspended in 800 µL ice-cold water. Cells were transferred to a 1.5 mL tube on ice followed by another centrifugation at 4000 rpm for 4 min at 0°C. The supernatant was discarded then cell pellets were washed again with ice-cold water followed by another centrifugation. Final cell pellets were suspended in 50 µL ice-cold water. 100 ng BAC (1 µL)/1.0 ng plasmid DNA/20 ng (all quantified by photospectrometer) linearized DNA fragment was mixed well with *E. coli* cells. Linearized DNA fragment was prepared by enzyme digestion followed by agarose gel purification. The mixture was transferred into an electroporation cuvette (1 mm gap). The electroporation machine was set as 1750 V, 25 uF, 200 Ω. The pulse time span is usually 4.0 ms. After electroporation, 1.0 mL LB was added then cells were incubated at 30°C (37°C for DH10B strain) for 1 hour. Cells (20 µL–1mL depends on efficiency of the experiment) were spread on appropriate plate for clonal growth.

2.3.5 Recombineering

2.3.5.1 Principle

The recombineering technique uses homologous recombination as a molecular manipulation tool in *E. coli* (Copeland *et al.* 2001; Yu *et al.* 2000; Zhang *et al.* 1998) and facilitates construction of DNA vectors. Using this technique, targeting vectors can be constructed from BACs within a week. The recombineering system is based on the recombination system from bacteriophage λ to recombine DNA fragments in *E.coli*. Bacteriophage λ contains a recombination system, the Red operon, which utilizes two proteins (Figure 2-4): redα (exo) and redβ (bet). Exo is an exonuclease, which digest linear dsDNA in a 5’ to 3’ direction (Carter *et al.* 1971; Matsuura *et al.* 2001). When the nucleotides are removed, long 3’ ssDNA tails are exposed, which are possibly as long as half the length of the original duplex DNA (Carter *et al.* 1971; Hill *et al.* 1997). Then λ Beta promotes the annealing of the complementary DNA strand, which stably binds to 3’ single-strand DNA (Figure 2-5) (Radding *et al.* 1971), which is greater than 35 nucleotides in length (Mythili *et al.* 1996). It is known that beta protein specifically protects single-stranded DNA from digestion by DNase, thus exogenous DNA cannot be destroyed in *E. coli* (Muniyappa *et al.* 1986). If 3’ ssDNA is annealed to the replication fork, it initiates the recombination process.

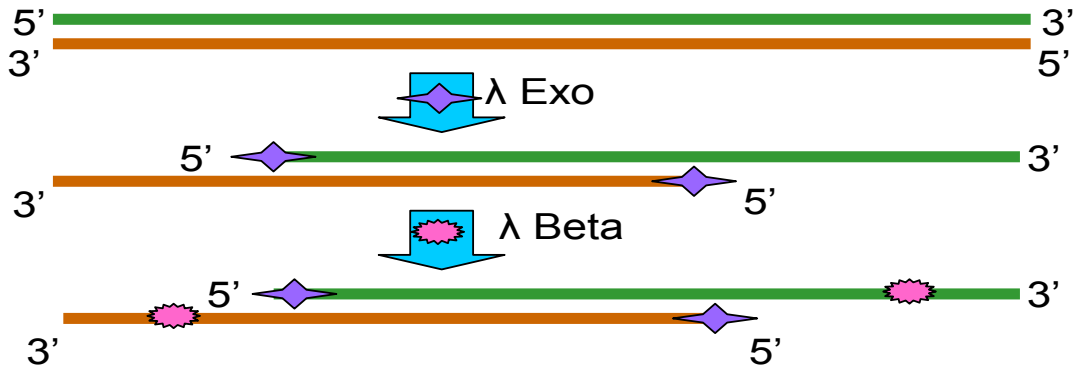


Figure 2-4 The basic functions of Bacteriophage lambda proteins: Exo and Beta

An Exo protein (purple star) digests the 5’ nucleotides of linear dsDNA. The protein Beta (red) binds to the 3’ ssDNA.

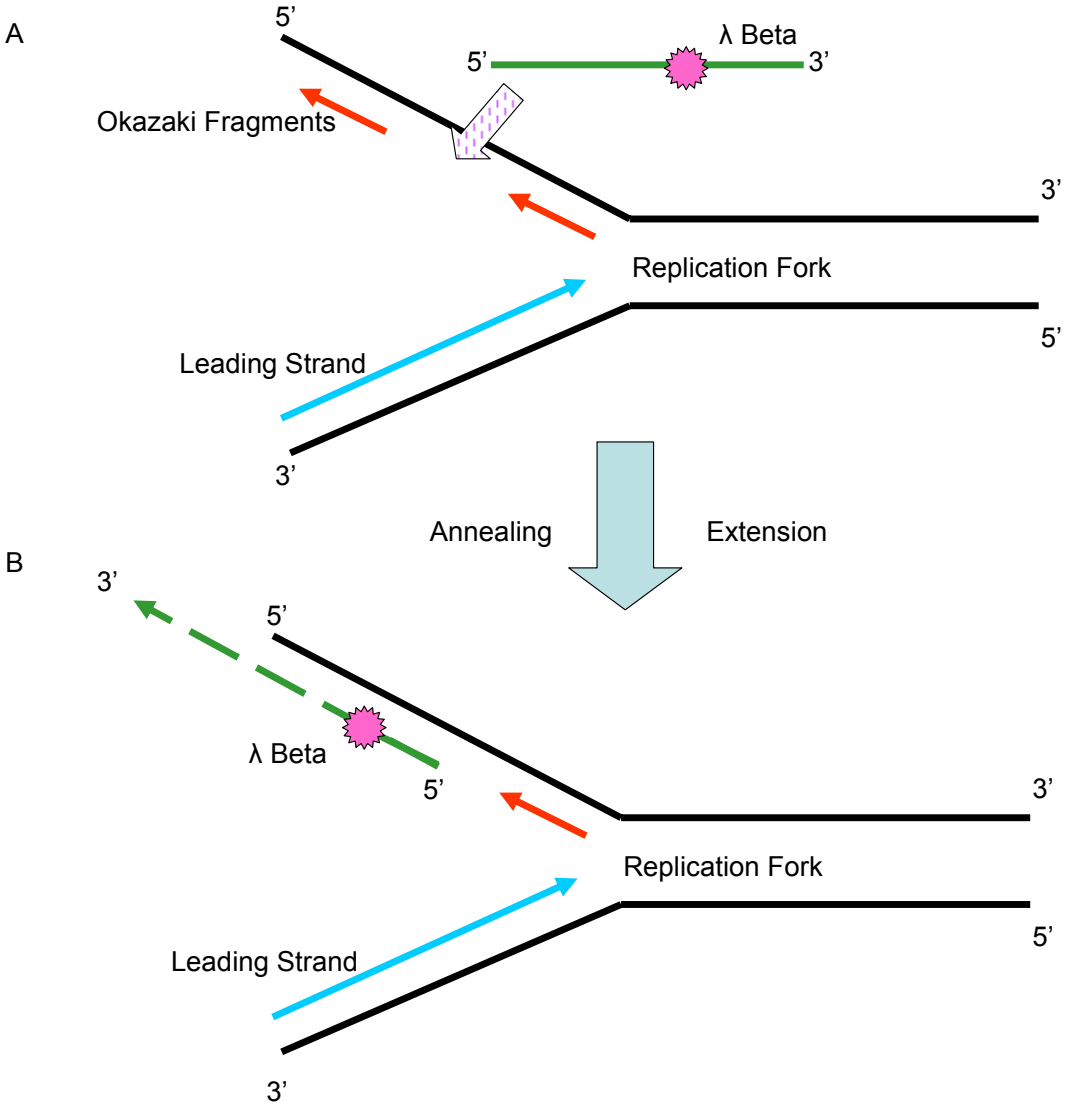


Figure 2-5 The annealing of the ssDNA to the replication fork by lambda Beta protein

A. A protein Beta anneals a 3’ ssDNA tail of a linearized DNA fragment (green line) to the complement lagging strand gap in the replication fork. **B**. This annealed strand starts recombination between linearized DNA fragment and genome DNA (black line).

To make use of recombineering in *E. coli*, Daiguan Yu integrated λ prophage with the λ Red system genes exo, beta and gam into *E. coli* genome. These genes were under the control of powerful λ P_L promoter and the temperature sensitive repressor cl857 (Yu *et al.* 2000). The λ prophage protein expression is tightly regulated at the temperature range from 32°C to 42°C by the repressor cl857 (Figure 2-6). The λ gam protein inhibits the helicase activity and some recombination activities of *Escherichia coli* RecBCD enzyme (Murphy 1991).

Applying the above principle, linearized plasmids can be used to retrieve genomic DNA from Bacterial Artificial Chromosome (BAC) through gap repair in *E. coli* strains containing the integrated λ prophage (Figure 2-7) (Lee *et al.* 2001). A BAC of interest is first transformed into the engineered *E. coli* strain. Transformed single colonies are isolated and confirmed by enzyme digestion and agarose gel electrophoresis. Retrieval plasmids are constructed with two short (50 bp) homology arms (left arm and right arm), which are homologous with the flanking sequence of the relevant fragment of genomic DNA on the BAC. A transformed *E. coli* strain is cultured at 30°C overnight followed by a heating shock at 42°C for 15 minutes to promote λ protein Exo, Beta and Gam expression. These proteins enable recombination between the homologous arms. The result of this gap repair process is to subclone a large genomic DNA fragment from the BAC into a high copy plasmid.

2.3.5.2 Recombineering protocol

Strains EL350/EL250/DY380 were cultured as described in 2.3.4. Before harvesting, cells were transferred into a 42°C water bath for 15min with shaking (180 rpm). Then cells were placed on ice for 5 min to cool. Cells were washed and electroporated as described in 2.3.4.

To retrieve genomic DNA from a BAC, 1–2 μL of linearized and purified retrieval vector (200-500ng) was used. For mini-targeting the retrieval product, 1–2 μL agarose gel purified DNA fragment (20–100 ng) was used in the electroporation.

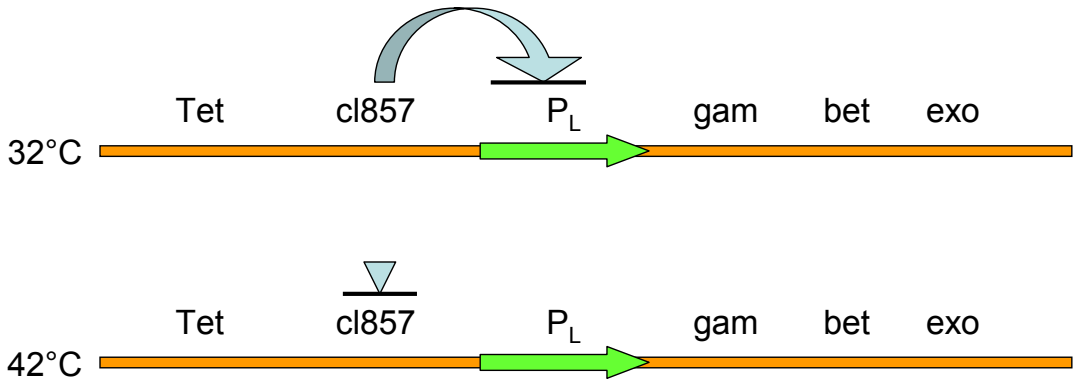


Figure 2-6 The structure and function of the defective lambda prophage

The defective λ prophage contains a Tet selection marker, a cl857 repressor, a PL promoter and λ protein gam, bet and exo. When cells are cultured at 32°C, the cl857 repressor blocks the PL promoter thus repressing the transcription of gam, bet and exo. At 42°C, this repressor is inactivated and the PL promoter initiates expression of gam, beta and exo genes. Tet: tetracycline resistance gene serves as a selection marker.

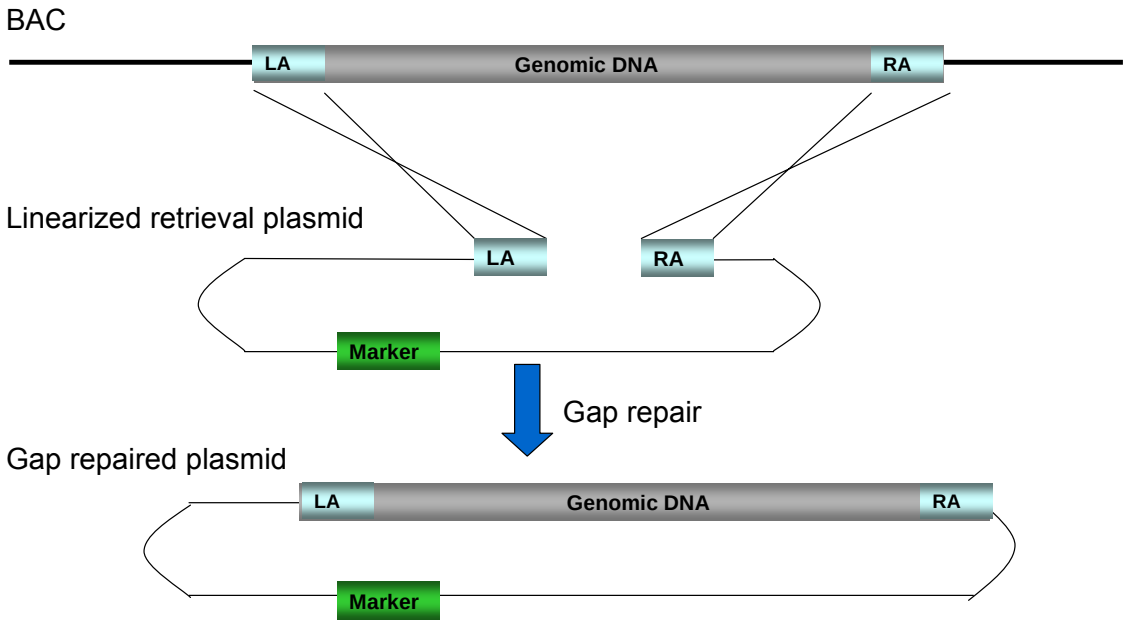


Figure 2-7 Gap repair process

Through the function of λ proteins Exo, Beta and Gam, the gap repair process can recover large fragments of genomic DNA from BACs into high copy number plasmids. The process requires the presence of homology arms (LA and RA, indicates Left Arm and Right Arm, respectively).

2.3.6 Vectors

2.3.6.1 pZK5 – *PuroΔtk/Kanamycin* cassette

As most plasmids carry Ampicillin resistance marker (*Amp*), the *PuroΔtk* cassette was modified to contain a Kanamycin resistance gene (*Kan*) to make the cassette selectable in *E.coli*. The *PuroΔtk* fragment was from the vector YTC-37 (Chen, Y. T. *et al.* 2000).

A Kanamycin resistance cassette was generated by PCR using the pCR 2.1-TOPO vector (Kanamycin resistant, Invitrogen) as template and primers zk107 and zk102.1. The sequences of these oligo nucleotides are shown in 2.3.3. These two primers contain a homologue of pCR2.1-TOPO at 3' and a restriction enzyme digestion site at 5'. Amplified DNA was digested with *Eco* RI and *Xho* I following gel purification. This product was ligated into pBluescript II (pBSII) at the *Xho* I/*Eco* RI sites to produce pZK4 (data not shown).

A *Cla* I-*PuroΔtk-Xho* I and the *Xho* I-*Kan-Eco* RI fragment (from pZK4) were ligated into pBSII to generate vector pZK5 (Figure 2-8). The cassette between *Cla* I and *Eco* RI sites contained the fused positive / negative selection marker *PuroΔtk* for eukaryotic cells and a Kanamycin selection marker (*Kan*) for *E. coli*.

2.3.6.2 pZK9 – Vector to retrieve gene desert DNA from BAC

The retrieval vector pZK9 for the chromosome 6 gene desert targeting was constructed using a pBS *DT-A* backbone (Figure 2-9), which contained the *MC1 DT-A* (Diphtheria toxin A-fragment) cassette (Yagi *et al.* 1990). My colleague Dr Wei Wang inserted a *Kpn* I-*MC1-DT-A-SV40pA-Sph* I fragment into pBS SK(+) *Cla* I-*Eco* RI backbone fragment after both were blunted by T4 polymerase.

The retrieval left arm (PCR product 111) and right arm (PCR 114) were generated by PCR with primers 111.1 and 111.2 for product 111, primers 114.1 and 114.2 for product 114 (Figure 2-10). The sequences of these oligo nucleotides are shown in 2.3.3. The PCR product 111 contained 462 bp chromosome 6 sequence from 10435434 to 10435895 bp and product 114 contained 391 bp chromosome 6 sequence from 10443570 to 10443960 bp. They were digested with *Sma* I/*Spe* I and *Spe* I/*Xba* I respectively and cloned into pBS *DT-A*.

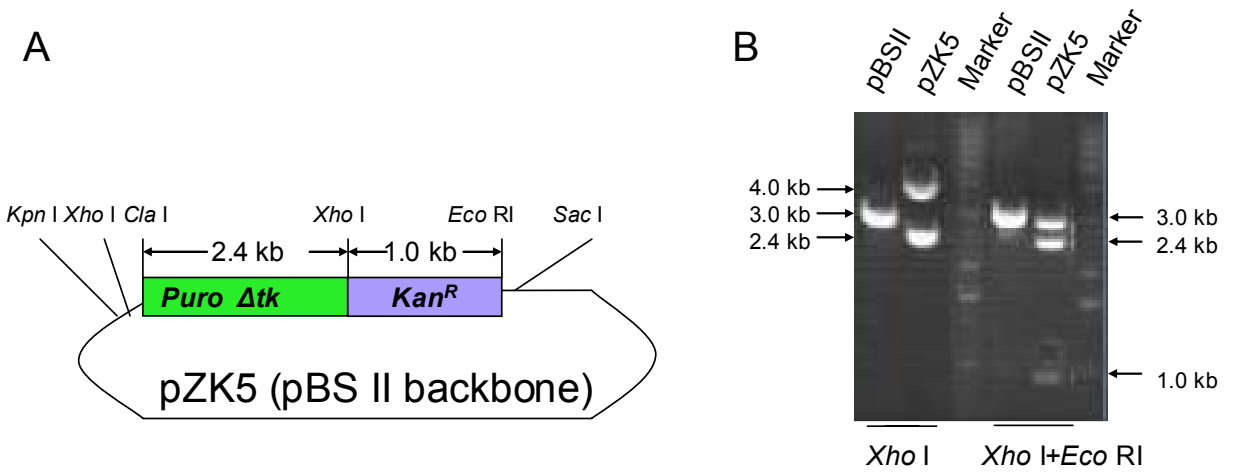


Figure 2-8 **Vector pZK5 – *PuroΔtk/Kanamycin* cassette**

A. Structure of pZK5; **B**. Enzyme digestions confirmed pZK5 was correct as designed. *Xho* I digestion showed pZK5 had 4kb and 2.4kb fragments. *Xho* I+*Eco* RI double digestion showed pZK5 had 1.0kb *Kan*, 2.4kb *PuroΔtk* and 3.0kb pBSII vector.

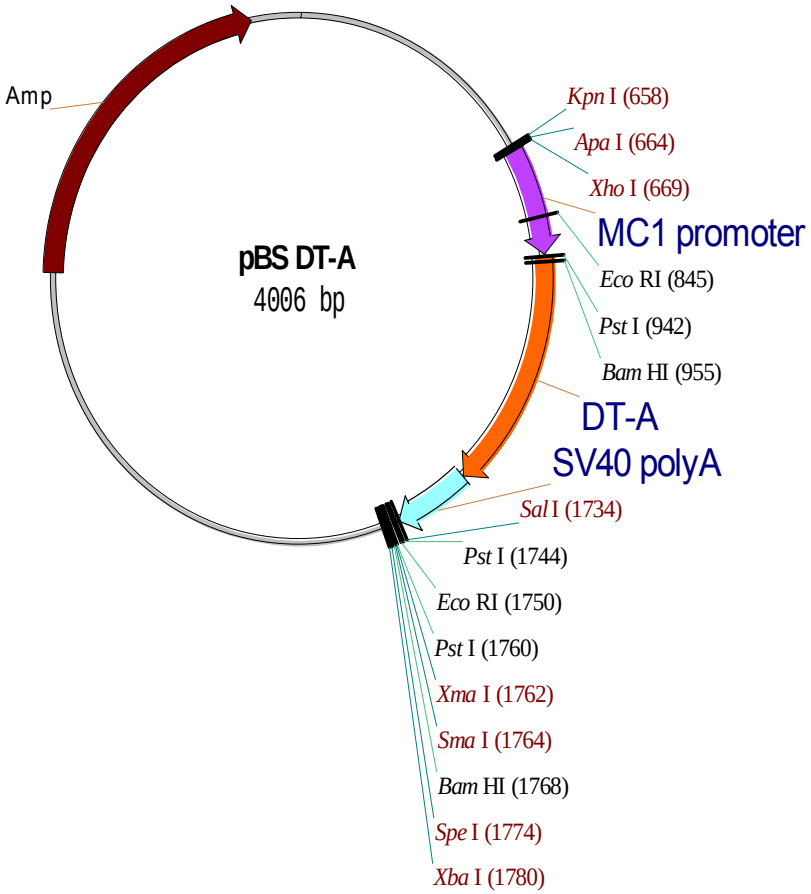


Figure 2-9 **Structure of pBS *DT-A* plasmid**

This vector was used to construct retrieval vectors to retrieve genomic DNA from BACs. The *DT-A* gene was driven by the *MC1* promoter. The SV40 (Simian virus 40) polyA signal was used.

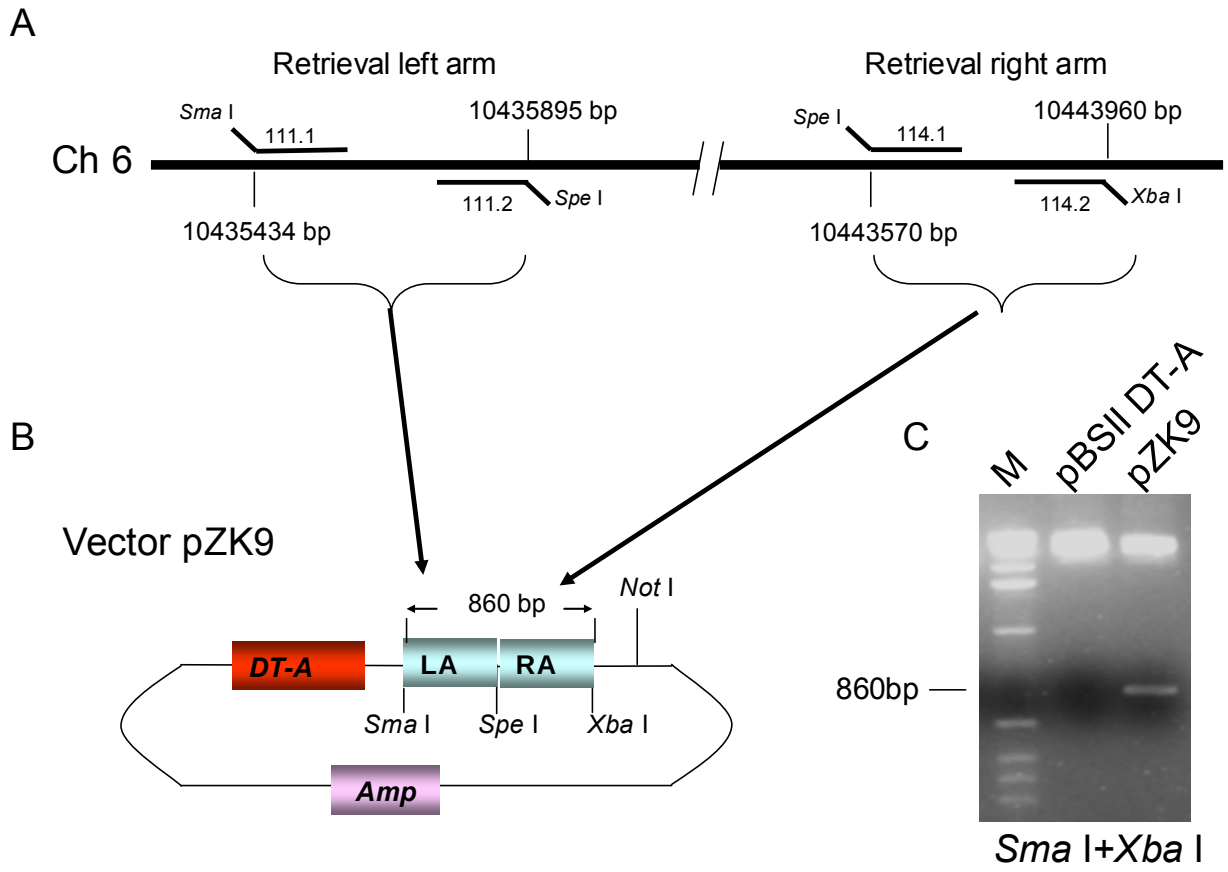


Figure 2-10 Vector pZK9 – to retrieve gene desert genomic DNA from a BAC

A. Position of primers to amplify retrieval homology arms. Part of mouse chromosome 6 is shown here. Primer 111.1 and 111.2 were used to amplify the left retrieval arm. Primer 114.1 and 114.2 were used to amplify the right retrieval arm. Coordinates of start and end point of homology arms are shown in base pairs (bp). *Sma* I, *Spe* I and *Xba* I enzyme sites were synthesized in the 5' of the primers. **B**. Structure of pZK9. pZK9 contained a retrieval left homology arm (LA) and a retrieval right homology arm (RA). The vector contains a unique *Not* I site that will be used to linearize pZK9 prior to targeting experiment. **C**. Enzyme digestion confirmed the pZK9 structure. An ethidium bromide stained agarose gel showing the 860 bp LA/RA fragment cut with *Sma* I/*Xba* I.

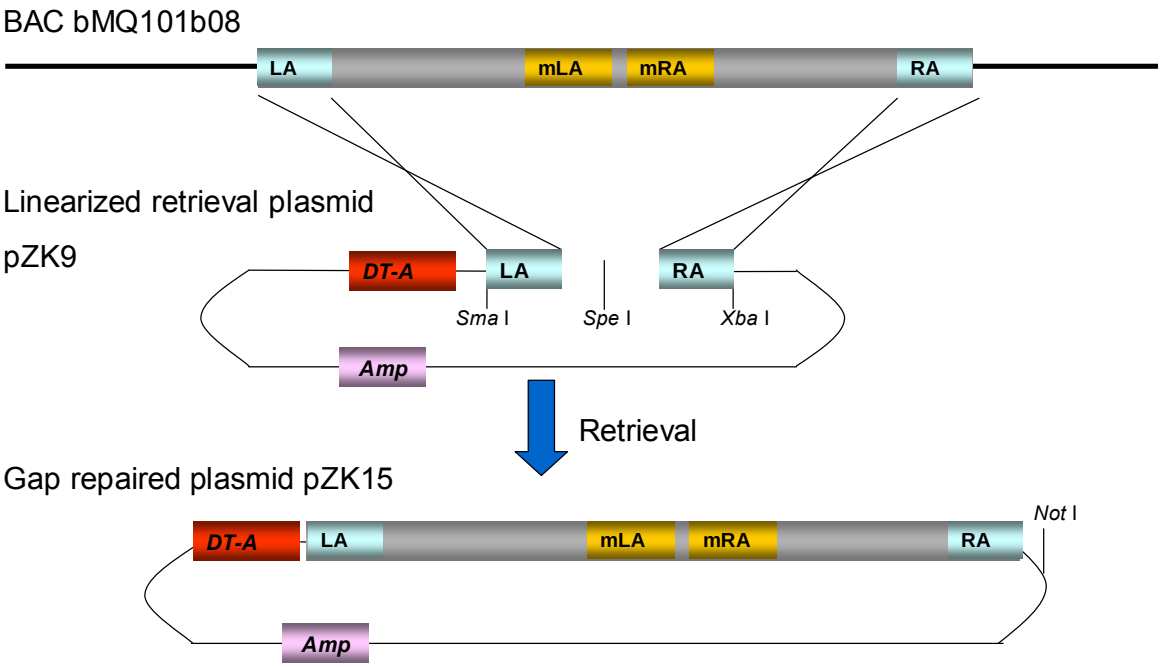
2.3.6.3 pZK15 – Retrieving the homology required for targeting from a BAC

To generate a targeting vector, a large piece of genomic DNA is required for the homology arms. The retrieval plasmid pZK9 has small regions of homology and is designed to recover an 8.5 kb piece of genomic DNA from the BAC bMQ101b08. The BAC was introduced into the EL350 bacteria strain in advance. Transformation was selected in LB agar plates with chlorophenical as the BAC bMQ101b08 contained the chlorophenical marker gene. Correct clones were confirmed by enzyme digestion. To accomplish the retrieval product, pZK9 was linearized with *Spe* I and electroporated into the EL350 bacteria strain. After linearized pZK9 was electroporated into the bacteria, cells were recovered in liquid LB at 30 °C for 45 minutes then placed on Ampicillin agar plates. Ampicillin resistant colonies were picked and the plasmids in those colonies were extracted. Correct pZK9 candidates were identified by a series of enzyme (*Spe* I, *Cla* I, *Eco* RV, *Not* I, *Hind* III and *Sma* I) digestion (Figure 2-11).

2.3.6.4 pZK8 - Chromosome 6 gene desert mini-targeting vector

The mini-targeting vector pZK8 was designed to recombine with the retrieved genomic fragment to generate the chromosome 6 gene desert target vector (Figure 2-12). Two mini targeting homology arms mLA (249 bp) and mRA (286 bp) were amplified using the PCR primer pairs zk112.1 with 112.2 and zk113.1 with 113.3, respectively. The sequences of these oligo nucleotides are shown in 2.3.3. The amplified products were digested with *Kpn* I*/Cla* I and *Eco* RI*/Hind* III, respectively before they were cloned into either side of the *PuroΔtk/Kanamycin* cassette of pZK5.

A.



B.

Expected restriction pattern of pZK15

Restriction Enzyme	Cut	Fragments (kb)
Spe I	3	0.47, 4.2, 7.8
Cla I	1	12.5
Eco RV	3	0.3, 2.3, 9.9
Not I	1	12.5
Hind III	1	12.5
Sma I	2	4, 8.5

C.

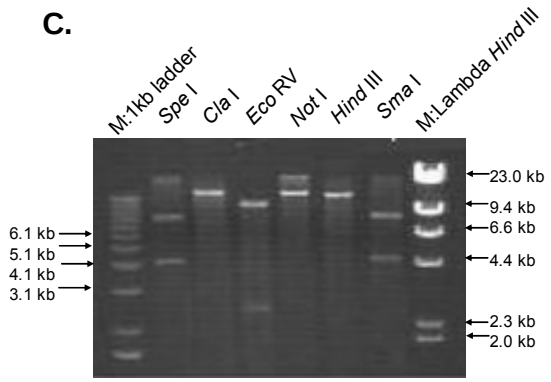


Figure 2-11 Generation of retrieval product pZK15

A. Schematic process of gap repair – retrieving DNA from the BAC bMQ101b08 to form pZK15; **B**. In silico prediction of restricted fragments of pZK15; **C**. Confirmation of one clone of pZK15.

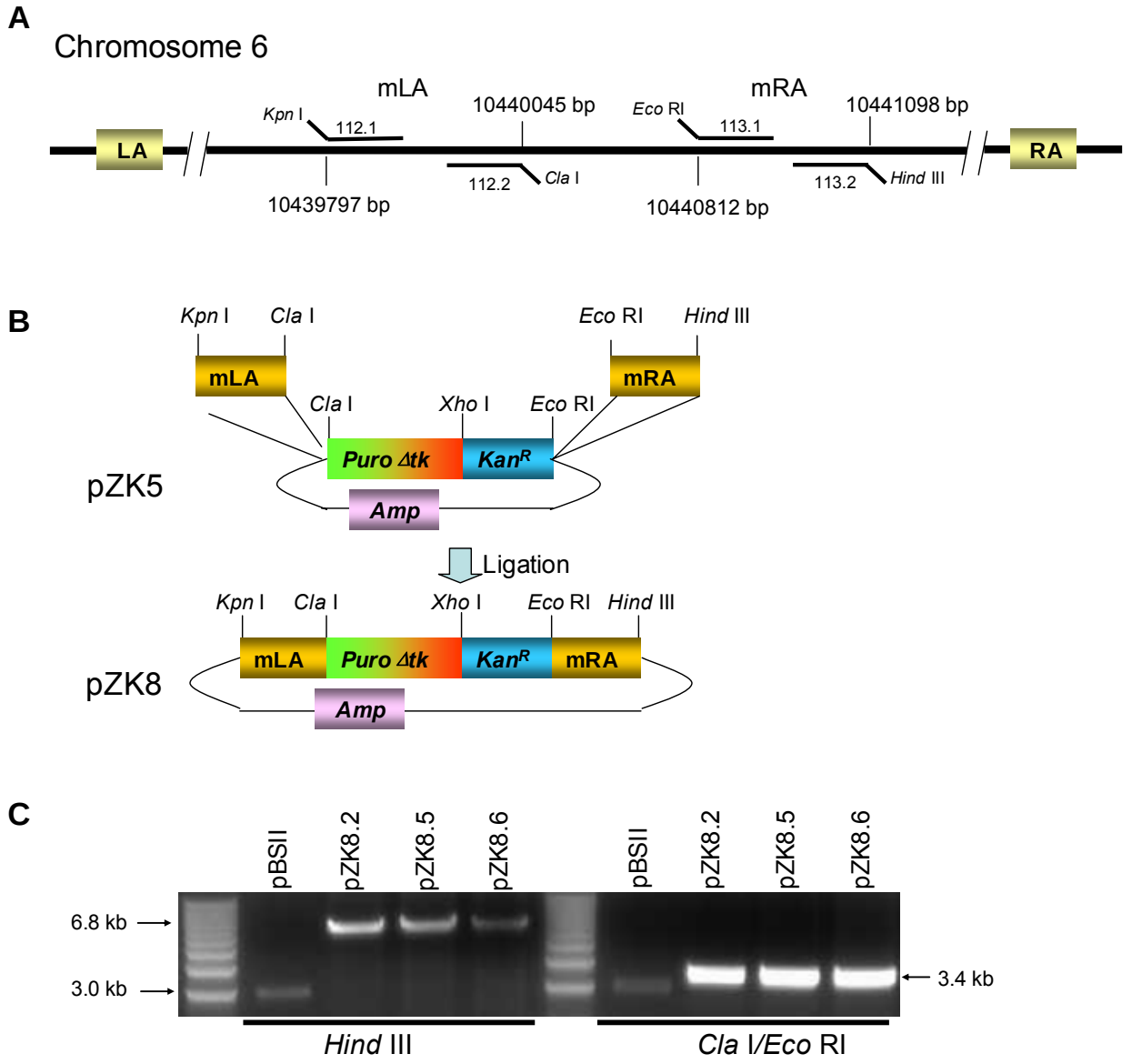


Figure 2-12 Structure of pZK8 – gene desert mini targeting vector

A. Genomic DNA between retrieval left arm (LA) and right arm (RA) of chromosome 6 is shown here. Mini targeting homology arms mLA and mRA were amplified by primer pairs 112.1 with 112.2 and 113.1 with 113.2. Enzyme sites were included in the 5' sequence of primers. Genomic coordinates are shown as base pair number on chromosome 6. **B**. *Kpn* I-mLA-*Cla* I and *Eco* RI-mRA-*Hind* III fragments were cloned one after another into either side of *PuroΔtk/Kanamycin* cassette on pZK5, resulting in pZK8, the gene desert mini targeting vector. **C**. *Hind* III and *Cla* I/*Eco* RI digestion confirmed correct plasmid structure of clones 2, 5 and 6. *Cla* I/*Eco* RI digestion generated two 3.4 kb fragments**.**

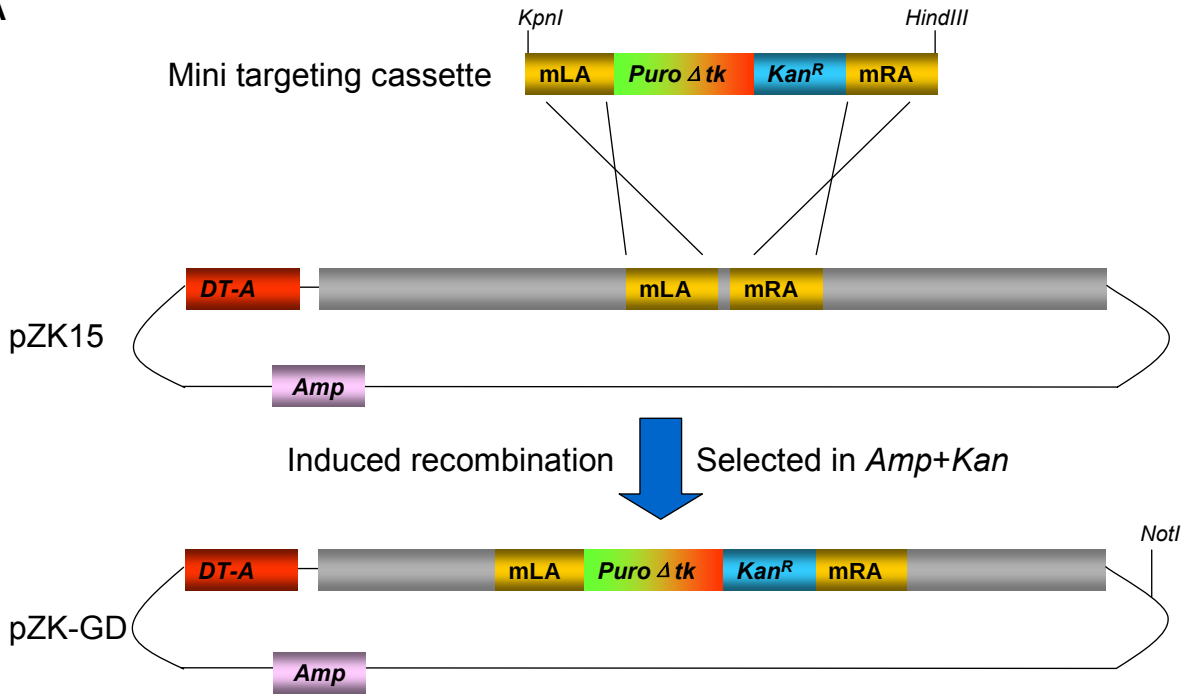
2.3.6.5 Construction of gene desert targeting vector pZK-GD

To generate pZK-GD (Figure 2-13), the mini targeting cassette needed recombine with the retrieved homologous DNA fragment contained in pZK15. The EL350 strain containing pZK15 was cultured in 42°C for 15 minutes and cooled in ice-cold water prior to the transformation of the mini targeting cassette. This temperature shift induces λ protein Exo, Beta and Gam expression and facilitates recombination. Vector pZK8 was digested with *Kpn* I and *Hind* III. The 3.9 kb mini targeting fragment was separated from the vector on an agarose gel. DNA was purified and the fragment was transformed into the EL350 *E. coli* strain containing pZK15 with the retrieved homology arms.

After transformation the *E. coli* cells were plated on Ampicillin/Kanamycin plates. EL350 *E. coli* colonies contained a mixture of the original plasmids and the recombined product. To recover the product, a plasmid mini-prep was performed. The product was re-transformed into DH10B *E. coli* strain and re-plated in Ampicillin/Kanamycin plates.

The correct clones were identified by restricted enzyme digests with *Bam* HI, *Eco* RI, *Eco* RV, *Hind* III, *Pst* I, *Sma* I and *Spe* I (Figure 2-14). The agarose gel pattern was compared to that predicted in silico. This confirmed that pZK-GD clone 31, 34, 35, 103, 105 had the correct structure.

A



B

Restriction Enzyme	Fragments (kb)	
	pZK15	**pZK-GD**
BamHI	2.2, 2.7, 7.5	1.6, 2.2, 2.7, 8.6
EcoRI	0.9, 2.5, 2.8, 6.2	0.035, 0.90, 2.5, 5.5, 6.2
EcoRV	0.3, 2.3, 9.9	0.104, 0.30, 1.24, 1.37, 2.3, 9.8
HindIII	12.5	6.3, 8.9
NotI	12.5	15.2
PstI	0.016, 0.60, 0.69, 0.80, 1.75, 8.7	0.016, 0.60, 0.68, 0.80, 1.0, 1.2. 1.7, 2.1, 6.9
SmaI	0.007, 4.3, 8.2	0.007, 0.9. 1.5, 4.3, 8.4
SpeI	0.47, 4.2, 7.8	0.5, 1.2, 5.6, 7.9

Figure 2-13 Generation of the chromosome 6 gene desert targeting vector - pZK-GD

A. Recombination at the mini-homology arms (mLA and mRA) between the targeting cassette and retrieval product vector pZK15 resulted in the gene desert targeting vector pZK-GD. pZK-GD is resistant to Ampicillin and Kanamycin. **B**. In silico prediction of restriction fragments of pZK-GD compared to pZK15.

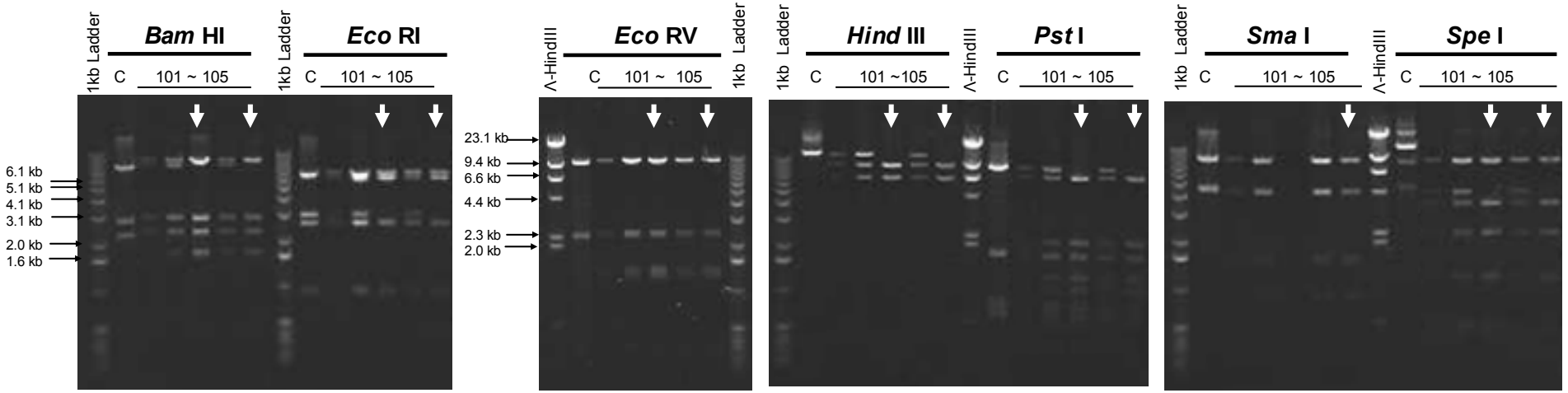


Figure 2-14 Restriction pattern confirmed correct pZK-GD plasmids

Clones 101–105 were digested with *Bam* HI, *Eco* RI, *Eco* RV, *Hind* III, *Pst* I, *Sma* I and *Spe* I. Clones 103 and 105 (white arrow) have the correct restriction fragments compared to the control plasmid pZK15.1 (C).

2.3.6.6 RMCE plasmid with the *Bsd* cassette

A vector was designed (Figure 2-15) to replace the positive/negative selection marker *PuroΔtk* in the BAC acceptor cell line CCI18 with a Blasticidin resistance gene (*Bsd*). The *pCMV-EM7-Bsd-SV40pA* fragment, from pCMV/Bsd (Invitrogen) was isolated by *Xho* I/*Eco* RI double digestion and ligated into pBSII SK+. *LoxP* and *Lox511* adapters were ligated into the vector. The *loxP* adapter was produced by annealing oligo nucleotides zk198.3 and zk198.4. The *Lox511* adapter was produced by annealing oligo nucleotides zk199.1 and zk199.2. The sequences of these oligo nucleotides are shown in 2.3.3. Plasmid pZK33.3, pZK33.4 were confirmed to be correct clones by sequence analysis.

2.3.6.7 Retrieve part of *Blm* gene for targeting from a BAC

To construct a *Blm* gene-targeting vector, homology arms need to be retrieved from a BAC containing the gene. The retrieval vector pZK21 (Figure 2-16) was designed to retrieve the *Blm* gene genomic DNA from the 129S5 BAC bMQ-436e20 by recombineering. The left arm (LA, 500 bp) was amplified from this BAC by the primers ZK136.1 and ZK136.2. And the right arm (RA, 660 bp) was amplified from the same BAC by the primers ZK137.1 and ZK137.2. The sequences of these oligo nucleotides are shown in 2.3.3. The fragments were cut with *Sma* I/*Spe* I (LA) and *Spe* I/*Xba* I (RA), gel purified and ligated in a 3-way reaction with pBS *DT-A* cut with *Sma* I/*Xba* I. This resulting plasmid (pZK21) can be linearized with *Spe* I to facilitate retrieval of the *Blm* gene from the BAC.

2.3.6.8 Retrieval product with *Blm* gene exons 2 and 3 (pZK19)

The recovery of the 10 kb homology arm required for targeting is achieved by gap repair recombinerring. The retrieval vector pZK21 was linearized with *Spe* I and purified by gel purification. The mouse BAC bMQ-436e20 was transferred into EL350 cells first. The cells were isolated on LB agar plates with chlorophenical. Correct clones were resistant to chlorophenical as the BAC bMQ-436e20 contained the chlorophenical marker gene. These clones were confirmed by enzyme digestion. One correct clone was cultured in LB at 30°C and the recombination protiens - λ phage proteins Exo, Beta and Gam were induced to express for 15 minutes at 42°C. These bacteria cells were prepared to be electroporation competent cells then electroporated with 20ng linearized pZK21 plasmid. Cells were recovered in LB at 30°C for 45 minutes then selected in Ampicillin agar plates. The recovered colonies may contain 10 kb recovered *Blm* gene DNA. Plasmid minipreps were performed to candidate clones. Then these clones were screened by the digestion of *Bam* HI*, Eco* RI*, Spe* I and *Xba* I. Correct clones, for example clone 1, 2, 4 and 6 were named

pZK19 and were used to generate homology arms of the *Blm* gene-targeting vector (Figure 2-17).

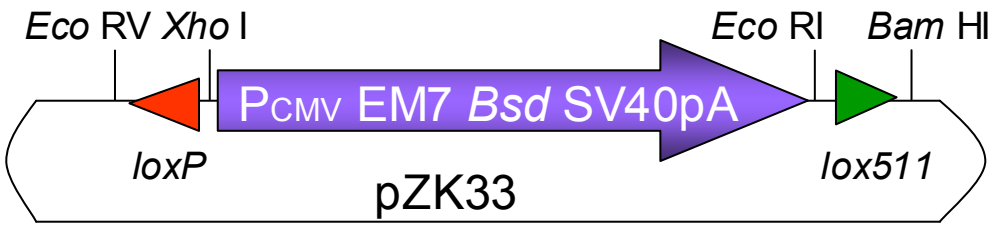


Figure 2-15 **Construction of the RMCE vector pZK33**

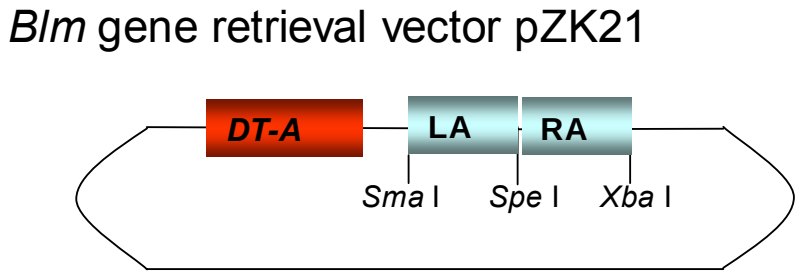


Figure 2-16 ***Blm* gene retrieval vector pZK21**

Retrieval left arm (LA) and right arm (RA) were ligated into pBS *DT-A*. The *Spe* I site was designed to linearize this plasmid for retrieval. *DT-A*: Diphtheria toxin A-fragment.

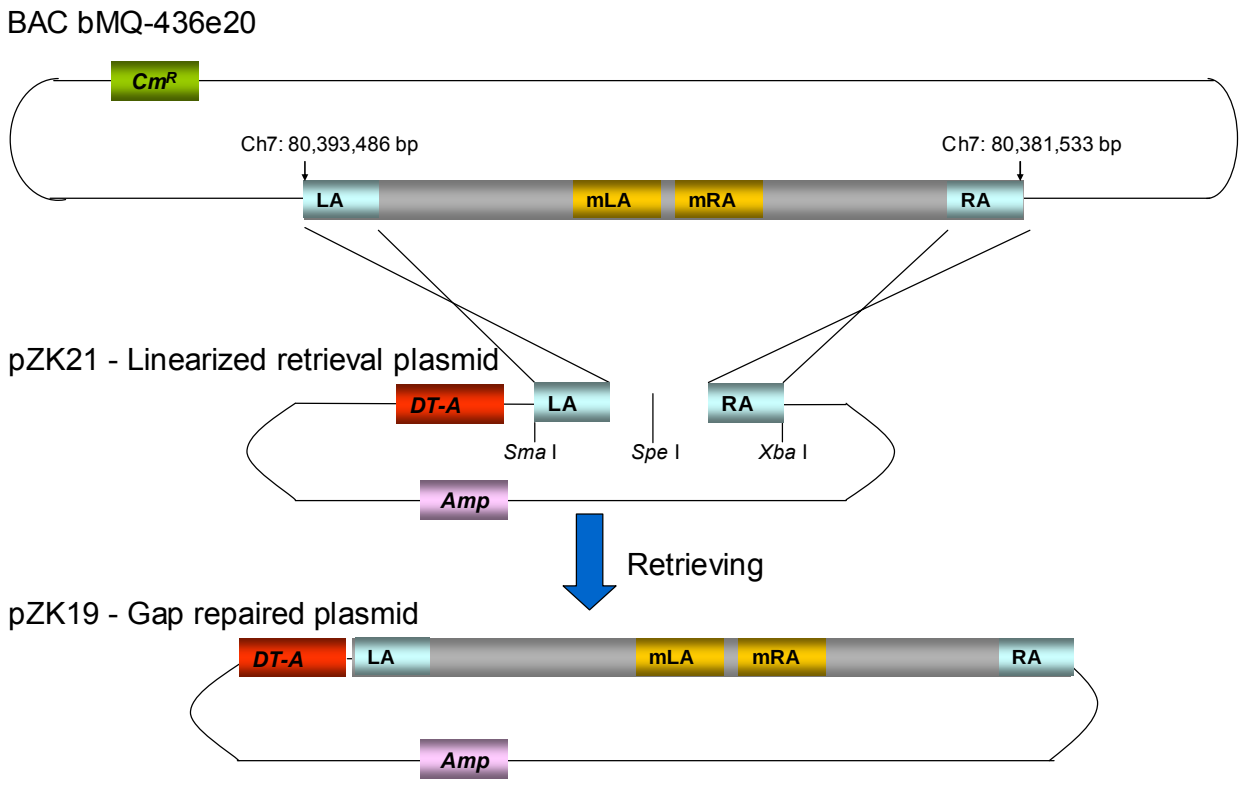


Figure 2-17 Retrieving *Blm* gene-targeting arms from a BAC

Spe I linearized retrieval vector pZK21 was designed to recombine with the BAC bMQ-436e20 and retrieve a fragment of homologous sequence from the BAC.

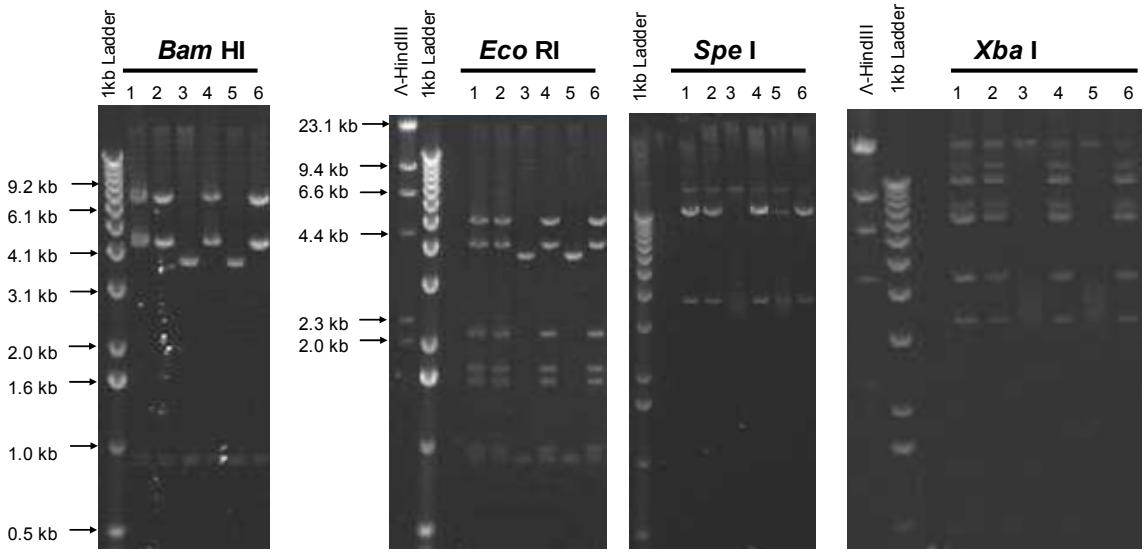


Figure 2-18 Enzyme digestion to identify the correct recombinants

Candidate plasmids of pZK19 were prepared (n=16, 6 are shown here) and digested with several enzymes. Electronic analysis of the vector and genomic sequence (Table 2-1) revealed that clones 1, 2, 4, 6 (shown) and 7–10, 12–16 were correct.

2.3.6.9 *Blm* gene mini targeting vector (pZK31)

To generate the final gene-targeting vector, a *Blm* gene mini targeting cassette is needed. This cassette was designed to contain *PGK/EM7/Neo-MC1-tk* cassette flanked by *FRT* sites. This vector was designed to replace the exons 2 and 3 of the gene with the *PGK/EM7/Neo-MC1-tk* cassette. *PGK* is a eukaryotic promoter and *EM7* is a prokaryotic promoter. Both *PGK* and *EM7* can initiate the expression of the *Neo* cassette but provide different resistance. *PGK-neo* generates resistance to G418 in eukaryotic cells and *EM7-neo* generates resistance to Kanamycin in bacteria. After targeting the *Blm* gene, the *PGK/EM7/Neo-MC1-tk* cassette can be recycled by Flpe then the same gene-targeting vector can be used to target the second allele.

The mini-targeting vector was designed (Figure 2-19) to target the fragment retrieved in pZK19 to generate the final *Blm* gene-targeting vector. The mini left homology arm (mLA) was amplified with primers zk175.1 and zk175.2. The mini right homology arm (mRA) was amplified with primers zk176.1 and zk176.2. The sequences of these oligo nucleotides are shown in 2.3.3. mLA and mRA were cut with *Kpn* I/*Xho* I (LA) and *Bam* HI/*Sac* II (RA), gel purified and ligated into pCOI5 cut with *Kpn* I/ *Sac* II, which contains the *PGK/EM7/Neo-MC1-tk* cassette flanked by *FRT* sites (Haydn Prosser's unpublished construct), through a four-way ligation and selected by Kanamycin.

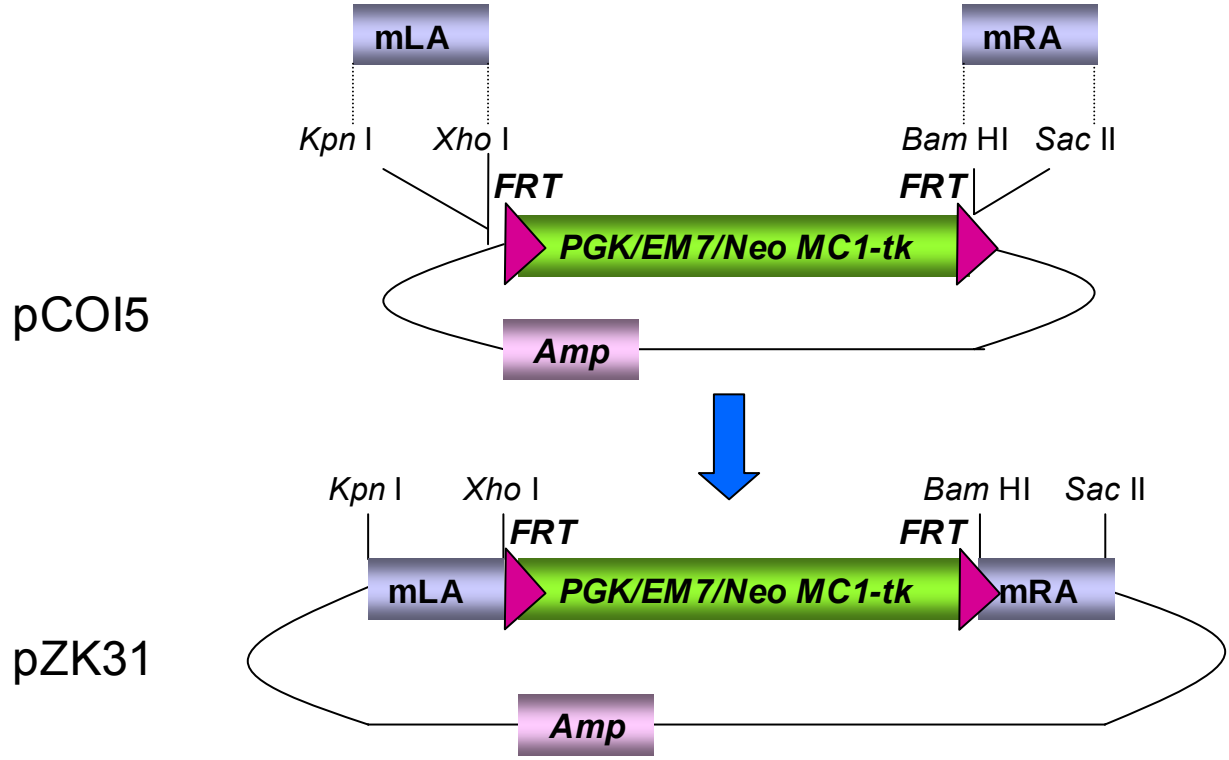


Figure 2-19 Generation of the *Blm* gene mini targeting vector - pZK31

The *PGK/EM7/Neo* and *MC1-tk* cassettes are flanked by two *FRT* sites. The mini targeting left arm (mLA) and mini targeting right arm (mRA) were inserted into pCOI5 backbone through a 4-way ligation into the *Kpn* I and *Xho* I site and *Bam* HI and *Sac* II site, respectively.

2.3.6.10 *Blm* gene-targeting vector (pZK30)

To assemble the *Blm* targeting vector (Figure 2-20), the mini targeting cassette (from pZK31) was isolated by digestion with *Kpn* I/*Sac* II followed by gel purification and electroporated into the EL350 recombinogenic *E. coli* containing the *Blm* homology fragment of pZK19. The *PGK/EM7/neo-MC1-tk* cassette replaced exon 2 and 3 DNA of *Blm* gene and resulted in Kanamycin resistant colonies. The produced plasmid contains a left homology arm (LA, 5 kb) and right homology arm (RA, 4.6 kb) to facilitate homologous recombination of *Blm* gene in ES cells.

The *Blm* targeting vector pZK30 can be selected in Ampicillin and Kanamycin. However, EL350 cells could also contain both of the non-recombined original plasmids, which would also give Ampicillin and Kanamycin resistance. To isolate the recombinant pZK30 plasmid, these plasmids were diluted and retransfected into *E. coli* strain DH10B followed by Ampicillin and Kanamycin selection. Candidate Ampicillin and Kanamycin resistant clones were screened by restriction enzyme digestion to identify those with the correct pattern, which was predicted by electronic sequence analysis (Table 2-1). Several clones were correct. Figure 2-21 shows the correct subclones: pZK30.3.1–30.3.12 and pZK30.4.1–30.4.9.

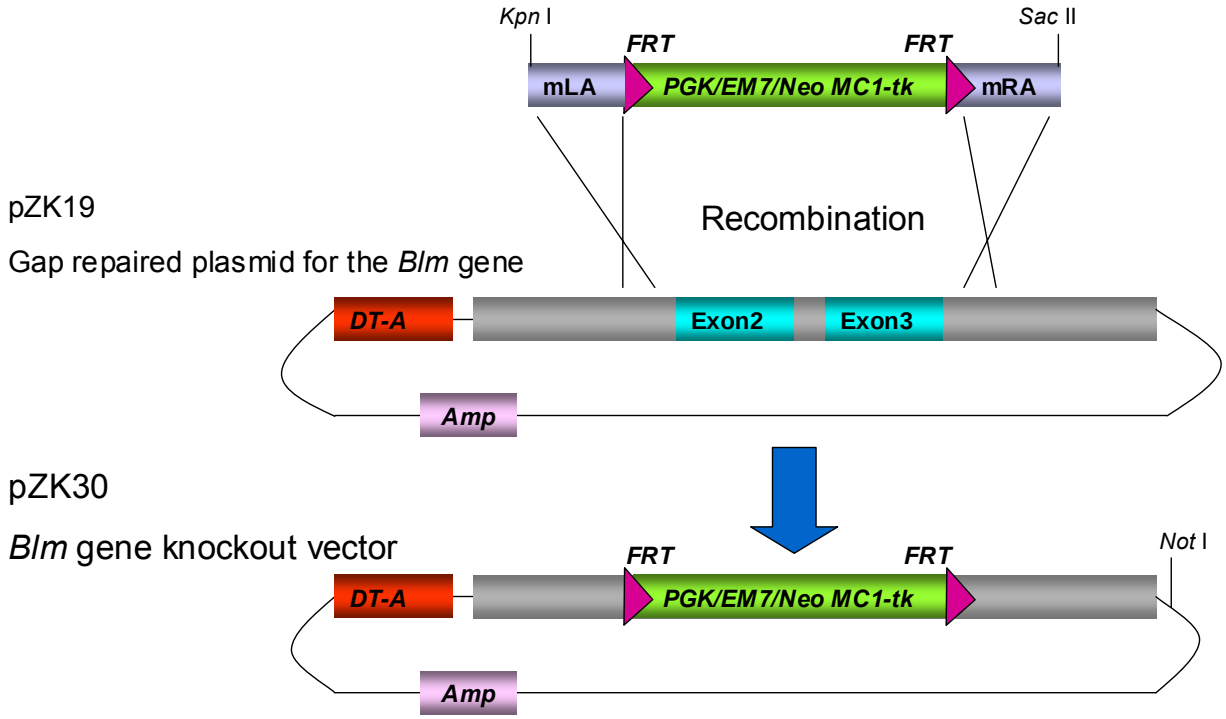


Figure 2-20 Generation of *Blm* targeting vector pZK30

The mini targeting cassette recombines with the *Blm* gene homology fragment in pZK19 to form the *Blm* gene knockout vector pZK30. pZK30 can be selected on Ampicillin and Kanamycin agar plates. The *Not* I site will be used to linearize this gene-targeting vector.

Table 2-1 Electronic sequence analysis of pZK19 and pZK30

Restriction Enzyme	Cuts		Fragments (kb)	
	pZK19	pZK30	pZK19	pZK30
Bam HI	4	5	0.914, 4.248×2, 6.584	0.914, 4.248×2, 3.589, 4.541
Eco RI	7	7	0.905, 0.954, 1.5, 1.7, 2.1, 4.0, 4.9	0.047, 0.905, 0.954, 1.7, 4.0, 4.9, 5.1
Hind III	4	5	0.722, 1.8, 4.0, 9.5	0.722, 1.2, 1.8, 4.4, 9.5
Not I	1	1	15.99	17.54
Xba I	4	7	0.963, 3.3, 4.4, 7.3	0.388, 0.963, 1.5, 1.7, 2.4, 3.3, 7.3

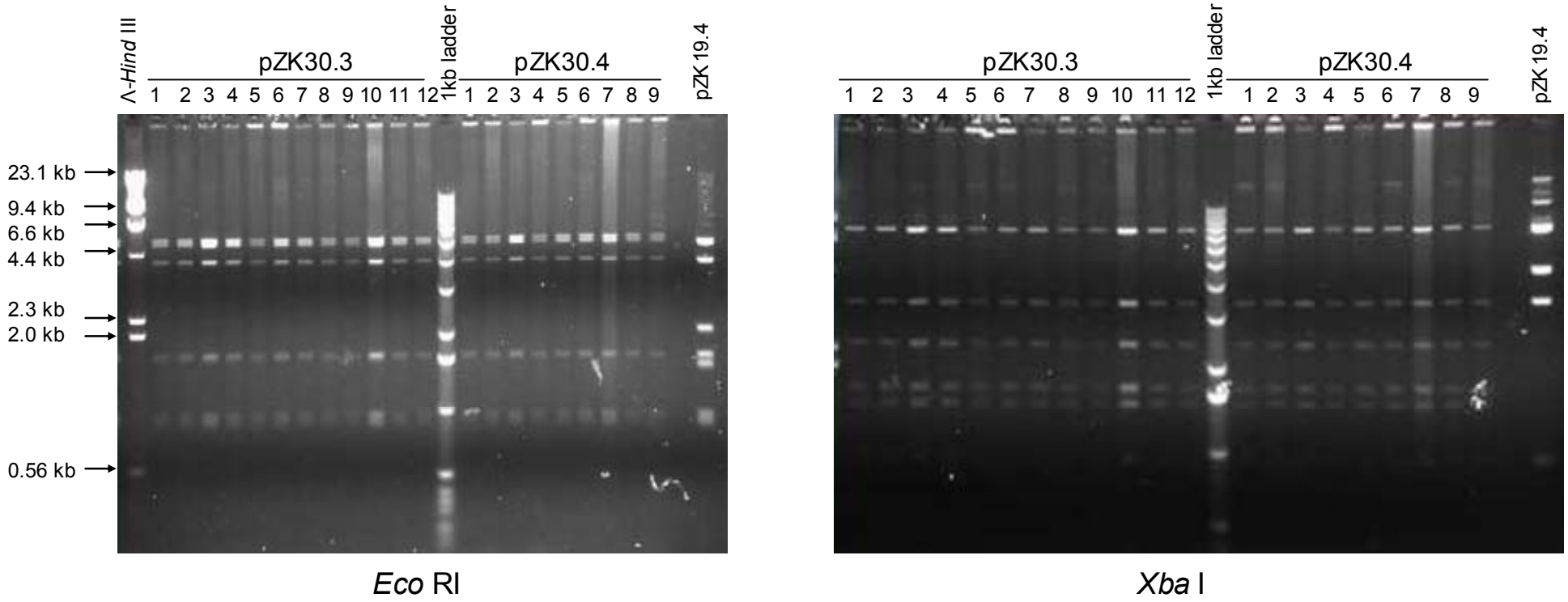


Figure 2-21 Screens for the pZK30 by enzyme digestion

pZK30.3 and pZK30.4 subclones were digested with *Eco* RI and *Xba* I. Subclones 1–12 of pZK30.3 and subclones 1–9 of pZK30.4 were correct. The *Blm* gene retrieval product pZK19.4 served as the control.

2.3.7 Genomic DNA isolation from tissue culture (6-well plate)

Cells in 6-well plates produce a large quantity of DNA, which is sufficient for use in Southern blot or array CGH. When preparing samples of genomic DNA, materials and solutions reserved for genomic use were used to diminish plasmid contamination. These include: Eppendorf tubes, proteinase K, dH_2O, and T0.1E. Genomic pipette tips from which the tips have been snipped off were used to prevent shearing the DNA and keep the average size in excess of approximate 30 kb.

6-well plates were washed twice with PBS before they were lysated by 1 ml of Cell Lysis Buffer. Then the plate was shaken at room temperature at 15-60 rpm for 10 minutes until all the cells detached. Then cell lysate was transferred into 15 ml falcon tubes to be shaken gently at 55°C for 3 hours to overnight. Then two times volume of 100% ethanol was added into the tubes to precipitate DNA. Tubes were inverted a few times to mix well. DNA was spooled onto flame-sealed micropipettes and rinsed once in 70% ethanol, once in 100% ethanol followed by an air dry. Finally, genomic DNA was dissolved in 0.6 mL T0.1E in a sterile Eppendorf tube overnight at room temp or a couple of hours at 55°C if it is hard to dissolve. Genomic DNA can be stored at 4°C for further analysis. Typically, the yield from a confluent 3 cm plate was 100 ug DNA.

2.3.8 Genomic DNA isolation from 96-well plates

Cells cultured in 96-well plates can be used for extracting DNA for Southern blot. When cells were confluent in wells, they were washed twice with PBS before adding 50 µL cell lysis buffer containing freshly prepared 0.5–1 mg/ml Proteinase K. Cell lysates were kept in a humidity chamber at 55°C overnight then 100 µL 100% ethanol was added to each well to precipitate the genomic DNA, which were allowed to stand at room temperature for a minimum of 30 minutes followed by 3 washes with 70% ethanol and 3 washes with 100% ethanol. DNA was left to dry in a 55°C oven for 5 minutes then it was ready to be dissolved for endonucleases digestion. Each well can generate 1–3 µg DNA.

2.3.9 Endonuclease digestion for Southern blotting

10 units of the appropriate enzyme were used in a total reaction volume of 40 µL with the proper reaction buffer according to the instruction from New England Biolabs (NEB). 2–5 µg genomic DNA was digested for each sample and digestion usually lasted overnight at 37°C to fully cut genomic DNA.

2.3.10 Southern blot hybridization

2.3.10.1 Probes

Probes were either PCR amplified from templates or excised from plasmids. They were all separated from other unwanted products by agarose electrophoresis and purified by QIAquick PCR Purification Kit™ (Cat.No. 28104) following manufacture's instructions.

PuroΔtk probe: 1.2 kb DNA fragment from *Pst* I digested plasmid YTC37 (Chen, Y. T. *et al.* 2000).
Δtk probe: 1.0 kb DNA fragment from *Bam* HI/*Xba* I double digested plasmid YTC37 (Chen, Y. T. *et al.* 2000).

2.3.10.2 Southern blotting

1–10 µg genomic DNA was digested with the appropriate restriction enzyme overnight and the DNA fragments were separated by electrophoresis on a 0.8% agarose gel in 1×TAE solution. Following electrophoresis, the gel was sequential soaked in Depurination Buffer for 5–10 minutes and denaturation buffer for 30–60 minutes with gentle agitation. A capillary blot was set up according to the standard protocols with Hybond XL membrane (GE Healthcare RPN203S) and denaturation buffer as the transfer buffer. After an overnight transfer, the membrane was neutralized by neutralization buffer for 5-10 minutes and baked for 1 hour at 80°C.

2.3.10.3 Probe preparation

10–30 ng DNA probe was labelled by Prime-It® II Random Primer Labelling Kit (Stratagene, Cat. No.300385) according to the manufacturer's instructions. Specific activities of ≥1×10^9 dpm/µg were achieved with DNA probes ranging in size from 400 bp to 10 kb. Briefly, 25 ng DNA probe and 10 µL random oligonucleotide primers were mixed in a final volume of 34 µL in dH_2O. Probe and primers were heat at 100°C for 5 minutes and cooled down at room temperature. 10 µL 5× dCTP buffer (provided in kit), 5 µL $dCTP^{32}$ (5 µCi) and 1 µL $Exo^{(-)}$ Klenow (5U/µL, final concentration 0.1U/µL) were added to probe and primers to make 50 µL mixture. The mixture was incubated at 37°C for 10 minutes and purified by ProbeQuant G-50 Micro Columns (GE Healthcare, Cat.No. 27-5335-01). Probes were then denatured at 100°C for 10 minutes and chilled on ice for 5 minutes before use.

2.3.10.4 Hybridization and membrane washes

The blot was pre-hybridized at 65°C for 1 hour with 1 mg single-strand salmon testes DNA (Sigma, Cat. No.D9156) in 10–15 mL hybridization buffer. The labelled probe then was added into the hybridization tube and incubated in a rotating oven at 65°C for 4 hours or overnight. The blot was rinsed by wash buffer at 65°C for 15 minutes and at room temperature for 10 minutes and exposed to X-ray film (Fuji) overnight at -80°C before development.

2.3.11 Array comparative genomic hybridization

2.3.11.1 Principle

Irradiation mutagenesis generates large amplifications and deletions on chromosomes. To identify these, a method to detect genomic DNA copy number variations is needed. Array comparative genomic hybridization (CGH) has become a powerful approach to detect genome-wide chromosomal imbalances. The principle behind this technology is the use of differentially fluorescently labelled samples of DNA and a reference, which are hybridized simultaneously to nucleotide probes on a glass slide (Figure 2-22). The signal of two fluorochromes at each probe location will depend on the copy number of DNAs. Through this method, any gain or loss of DNA sequences, such as deletions, duplications and amplifications can be identified in the whole genome.

The resolution of array CGH ranges from 50bp to 10 Mb. BACs, PACs, PCR amplified DNA/cDNA and oligo-nucleotides are all able to be used as probes in array CGH (Chung *et al.* 2004; Kallioniemi *et al.* 1992; Selzer *et al.* 2005).

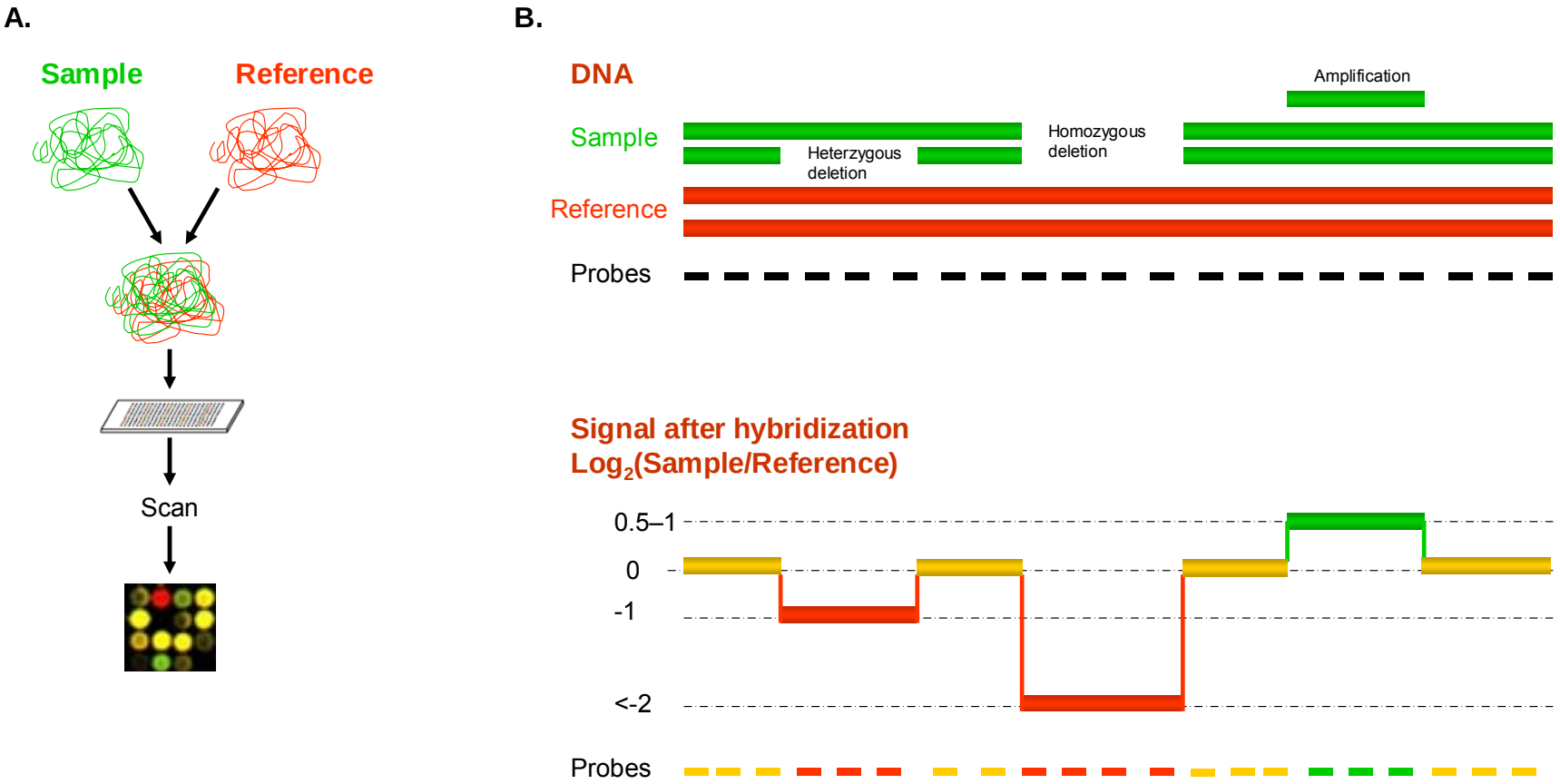


Figure 2-22 The principle of array comparative genome hybridization

A. The schematic view of the pocedures to conduct array CGH. Sample DNA and reference DNA labelled by green and red fluorescence are hybridized with probes arranged on a glass slide. Copy number variation is observed by the ratio of the two probes on the array. **B**. Example of Log_2 signal associated with various copy number alterations. Lack of one or two copies in the sample DNA (green) is reflected by lack of a green signal with the probes, while if a region of the genome is amplified in the sample, the green signal will be stronger than the red control. The $base_2$ logarithm of the signal intensity is used to reflect these alterations.

2.3.11.2 Reagents

For all data the following arrays were used: mWGTP1.1.1 (mouse whole genome BAC tilepath array, contains 2 copies of all probes) and mWGTP1.2.1 (mouse whole genome BAC tilepath array, contains 1 copy of all probes). One copy of all probes includes 18,294 tile path BACs. The original protocol has been reported previously by Chung *et al* (Chung *et al.* 2004)

BioPrime labelling kit (Invitrogen, Cat. 18094-011)

2.5× random primers solution: 125 mM Tris-HCl (pH 6.8), 12.5 mM MgCl2, 25 mM 2-mercaptoethanol, 750 µg/ml oligodeoxyribonucleotide primers (random octamers).

10× dNTP mix: 1 mM dCTP, 2 mM dATP, 2 mM dGTP, 2 mM dTTP in TE buffer, stored at -20

Exo-Klenow Fragment: 40 U/µl, Klenow fragment in 100 mM Potassium Phosphate (pH 7.0), 1 mM DTT, and 50% Glycerol.

1 mM Cy3-dCTP (Amersham, Cat. PA53021)

1 mM Cy5-dCTP (Amersham, Cat. PA55021)

NucAway spin columns (Ambion, Cat. 10070)

Hybridisation buffer : ULTRAHYB (Ambion, Cat. 8670). Aliquots stored at -20°C. Remove and heat to 70ºC prior to use, making sure that all components are in solution.

3 M NaAc pH 5.2 AMBION (Cat. 9740): store at room temperature.

mouse Cot-1 DNA (Invitrogen, Cat. 18440-016): 1 mg/ml in 10 mM Tris-HCl (pH 7.4), 1 mM EDTA. Store at -20°C.

100% Ethanol, 80% Ethanol

Yeast tRNA (Invitrogen, Cat. 15401-029): Store at -20°C.

Cysteamine: Sigma, 0.2 M in Water, filtered to be sterile, pre-warmed at 37°C before use.

2.3.11.3 Random labelling DNA for array CGH

500 ng DNA sample was mixed with 60 µl 2.5× random primers solution to make up final volume of 130.5 µl with water. This DNA/primer mixture was denatured at 100°C for 10 minutes followed by immediately cooling on ice. Labelling reaction (150 µl) was made by adding 15 µl 10× dNTP, 1.5 µl Cy3 or Cy5-labelled dCTP (GE Healthcare, PA53021 and PA55021, 1 mM, final 10 µM) and 3 µl Klenow enzyme on ice (final 0.8 U/µL). Then it was mixed well and kept at 37°C overnight.

To purify labelled DNA, ProbeQuant G-50 Micro Columns (GE, Cat.No. 27-5335-01) were used. The pre-loaded columns were spun at 4300rpm for 1.5 min before loading. ~74 µL labelling reaction per column (2 columns per labelling) were loaded carefully followed by spinning for 2 min at 4300rpm (750 g) to collect the labelled DNA. It then was assessed by a Nanodrop in the “Micro-array” mode to measure Cy3 and Cy5 labelled DNA concentration by detecting the fluorescence signal. Proper parameter will be: Cy3: ~7 pmol/uL, DNA 370 ng; Cy5: ~5 pmol/uL, DNA 350 ng.

2.3.11.4 Hybridization

The following mixes per array were prepared and placed at -70°C for at least 30 min to precipitate fluorescence labelled DNA.

In 2 mL tube:

Cy3-labelled DNA	~180 µl (final ~37 ng/µL)
Cy5-labelled DNA	~180 µl (final ~35 ng/µL)
mouseCot1 DNA	135 µl (final 0.075 µg/µL)
Yeast tRNA	6 µl (final 0.33 µg/µL)
3 M NaAc pH 5.2 (Ambion #9740)	50 µl (final 83 mM)
100% ethanol	1250 µl

Then labelled DNA was centrifuged for 30min at 15000 rpm at 4°C. Precipitated DNA was washed once by 600 µl 80% ethanol followed by a quick centrifuge at 13000 rpm for 5min. Trace volumes of ethanol were pipetted off and a purple pellet was left in the tube. After air drying for approx. 5 min at 70°C, 10ul of Cysteamine and 35µl (or 70µl for the higher resolution arrays) of pre-warmed hybridisation buffer (Ambion) were applied to each pellet

and incubated for 15 min at 70°C with frequent mixing during the first 5 minutes.

The tube was incubated at 37°C for 1–2.5 hours. A gentle pulse spin before 37°C incubation is recommended. Then the contents (~40 µL) in the tube were applied onto a coverslip (24 x 36 mm) and this was placed on the warmed array. The array was hybridized with labelled DNA in slide booster at 37°C for 16–24 hours. 15 uL AdvaSon™ Coupling Liquid for SlideBooster hybridizations was added below the array to distribute DNA on the array evenly. Water was used to maintain humidify in the hybrid booster.

2.3.11.5 Array washing

After the arrays were hybridized they were rinsed in 1×PBS (with 0.8 µM cysteamine) to removed coverslips. Then they were sequentially washed in the following buffers to achieve an increasing stringency. All solutions supplemented with 10 mL 0.2M cysteamine in 2500 mL as an antioxidant.

10 min at room temperature in PBS with 0.1% Tween-20
30 min at 55°C in 0.1X SSC
15 min at room temperature in 1XPBS
2 min at room temperature in PBS with 0.05% Tween-20
30 sec–1 min at room temperature in water

The arrays were centrifuged at 2200 rpm for 3 minutes to dry then they were ready to be scanned.

2.4 RNA methods

2.4.1 RNA isolation

Total RNA was extracted by RNAqueous™ Kit (Ambion) following the manufacturer's instructions. Briefly, cells were cultured in a 24-well plate without feeder cells until confluent. 700 µL lysis/binding solution was added to each well to lyse cells. An equal volume (700 µL) of 64% ethanol was added to the lysates. These mixtures were applied to filter cartridges separately followed by a centrifugation at 10,000 g for 30 seconds. The flow-through was discarded. A first wash was conducted by applying 700 µL wash solution 1 to a filter cartridge followed by a centrifugation at 10,000 g for 30 seconds. The flow-through was discarded. Filter cartridges were then washed twice with wash solutions 2 and 3 with the same wash procedure. Finally, filter cartridges were put into a new collection tube. 50 µL of

elution solution was applied followed by centrifugation at 10,000 g for 30 seconds. The concentration of the eluted RNA was measured by a spectrophotometer and stored at -80°C.

2.4.2 Expression arrays - RNA isolation and sample labelling

RNA was isolated using the RNAqueous™ Kit (Ambion). RNA samples were amplified and labelled with biotinylated uridine triphosphate (UTP) using the Illumina® TotalPrep RNA Amplification Kit (Cat No.IL1791) following the manufacture's instruction (illustrated in Figure 2-23). Briefly, total RNA was reverse transcribed to produce the first strand complementary DNA (cDNA) using oligo d(T) primers with a T7 RNA polymerase binding sequence (T7 promoter) at the primer's 5' end. Then the RNA strand of the RNA-DNA hybrid duplex is degraded by RNaseH. The second strand of cDNA is synthesized by DNA polymerase I (Pol I) and T4 DNA polymerase. A purification step is applied to remove degraded RNA, excess primers, enzymes and salts. Finally, antisense RNA is transcribed and amplified by T7 RNA polymerase. Biotinylated uridine triphosphate (UTP) is used to detect the signal from the complementary RNA. Amplified cRNA is purified to remove enzymes and salts. The biotin labeled cRNA is hybridized with the probes on Illumina® Mouse-6 Expression BeadChip. Each sample was hybridized individually. Wash procedures are applied to remove non-complementary hybridized cRNA. Cyanine3 (Cy3, wavelength 532 nm) conjugated streptavidin is used to visualize the biotin labelled cRNA.

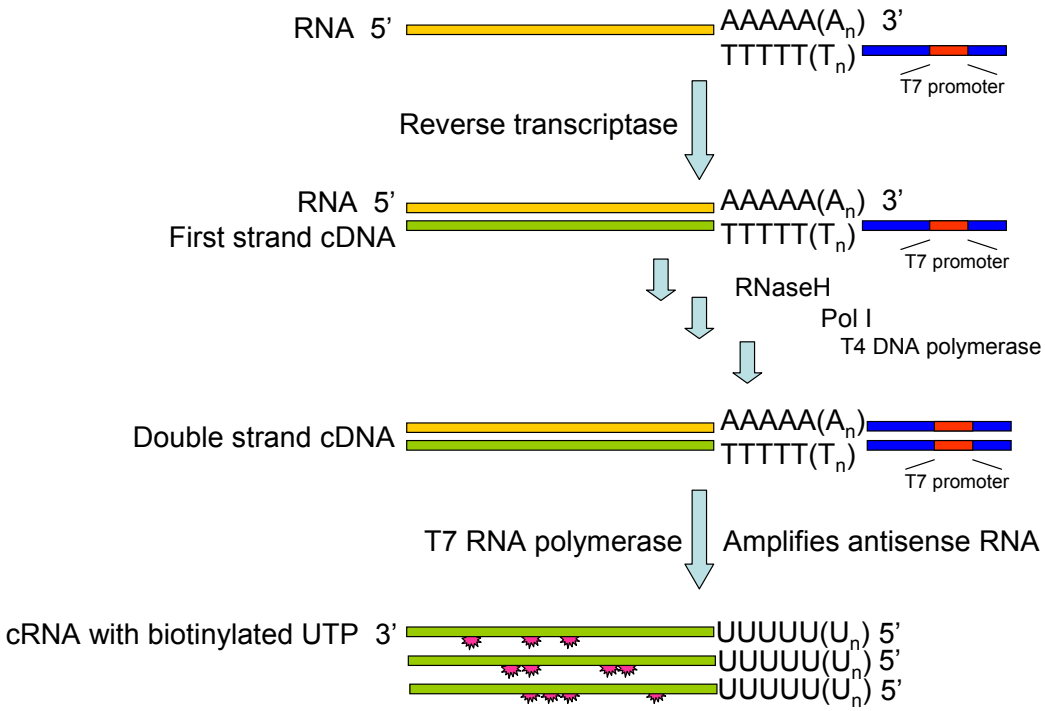


Figure 2-23 RNA amplification and labelling for expression array

Total RNA is reverse transcribed to produce the first strand cDNA from oligo d(T) primers containing a T7 promoter sequence at the 5' end. The RNA strand in the RNA-DNA hybrid duplex is degraded with RNaseH. DNA polymerase I (Pol I) and T4 DNA polymerase then synthesize the second trand cDNA with T7 promoter sequence at the 3' end. Double-strand cDNA is purified to remove RNA, primers, enzymes and salts before antisense RNA is transcribed *in vitro* by T7 RNA polymerase. Uridine triphosphate (UTP) used in the *in vitro* transcription is labeled by biotin (red star). Biotinylated UTP incorporated in the amplified complementary RNA (cDNA) is visualized on the array by using Cy3 conjugated streptavidin. This figure is modified from Van Gelder *et al*. (Van Gelder *et al.* 1990).

2.5 Statistical analysis of arrays

2.5.1 Array CGH

Reciprocal fluorescence array CGH was mainly conducted to compare 6TG-resistant mutant and reference sample (*Blm*-deficient NGG5-3 cell line). DNA of NGG5-3 was obtained from the passage immediately before the starting point of generating the mutation library (retrovirus tagging). Following stardard washing protocols arrays were scanned using an Agilent Technologies DNA microarray scanner. Images were analysed using BlueFuse software. Within an experiment, Log_2(ratio) between -0.29 and +0.29 for the signal from any BAC probe was regarded as no copy number change. When more than one reciprocal fluorescence array CGH were analysed in a batch, raw data was used without excluding any BAC probes. Thus, more copy number changes may be detected in a single mutant genome. However, a large scale normalization analysis has been conducted, generating a better standard to distinguish if a BAC probe was amplified, deleted or its copy number remains the same.

2.5.2 Expression array

2.5.2.1 Detection P-value

A whole genome expression array, Illumina® mouse-6 beadchip was used to attain genome-wide transcriptional information from all types of cells analysed. A detection P-value was calculated for probes on all Illumina® arrays using BeadStudio (version 2.3.41) software. The detection P-value provides a measure of the probability of a measured signal being due to hybridization with Cy3 labelled complementary RNA rather background non-specific binding. There are >700 negative control probes on Illumina® whole genome bead arrays. These negative control probes do not have any specific target transcripts in the mouse transcriptome. Any signals detected with these probes are regarded as non-specific binding. The detection P-value is calculated by ranking the signal of the negative control probes with the signals of other probes. Thus, the lower a detection P-value, the higher probability there is that the linked probe has bound specifically.

2.5.2.2 Combined P-value and transcripts' presence

A combined P-value was obtained for each probe to indicate signal consistency between replicates. AB1 and AB2.2 wild type cells had 2 biological replicates; each mutant as well as *Blm*-deficient cell and *Dnmt1*-deficient cell, had 3 biological replicates. All RNA samples for

replicate samples were isolated independently and array hybridizations were conducted separately. Analysis of each array generated a detection P-value for each probe.

The combined P-value of a probe is a measure of the probability that the signal for a particular probe is a background signal. Thus, if a specific probe's combined P-value is 0.05, it means there is 5% probability that this probe has the same signal level as background. A combined P-value of 0.05 was set to distinguish the transcriptional status of a probe's target sequence. A combined P-value between 0 and 0.05 (exclusive) is interpreted as indicating that the probe sequence is present in the RNA of replicates. A combined P-value between 0.05 (inclusive) and 1 means the relevant probe is absent in the RNA of replicates. This criterion to distinguish presence and absence of transcripts can be raised if more confidence in the expression data is needed.

2.5.2.3 Comparison analysis

Comparison analysis revealed the statistic significance of the relative expression the fold change between two cell lines. Within each pair of comparison, the expression level of each probe is compared between the replicates of one cell line and the replicates of another cell line. The adjusted P-value (adj.P.val) is used to indicate the probability that the two groups of replicates for a given probe are expressed at the same transcriptional level. The lower the adjusted P value, the lower is the probability that the two groups are expressed at the same level. Five thousand probes with the lowest adjusted P-value in each pair of comparison were selected for analysis. Comparison analysis does use the detection P-value and the combined P-value as parameters.