Chapter 4. Modelling *Tel-AML1* oncogenic translocation using knockin mice and transposon-mediated insertional mutagenesis

4.1 Introduction

Chromosome translocation is an event which results in exchange of genetic material between the arms of two non-homologous chromosomes. Occasionally this is accompanied by expression of a fusion chimeric protein produced at the translocation join point. Chromosome translocations can be classified into two types. Balanced chromosome translocation is the exchange of genetic materials between two chromosomes without the gain or loss of genetic information (**Figure 4-1 A**). In contrast, unbalanced chromosome translocation results in the gain or loss of genetic information through unequal chromosome exchange. Chromosome translocation has been suggested to be the cause of many types of genetic disorders including cancer, infertility and Downs syndrome (130-132). In particular, balanced chromosomal translocations have been identified as predisposing events in many types of haematological malignancies (133). For example, the ETS domain encoding genes are involved in several chromosomal arrangements (134-136).

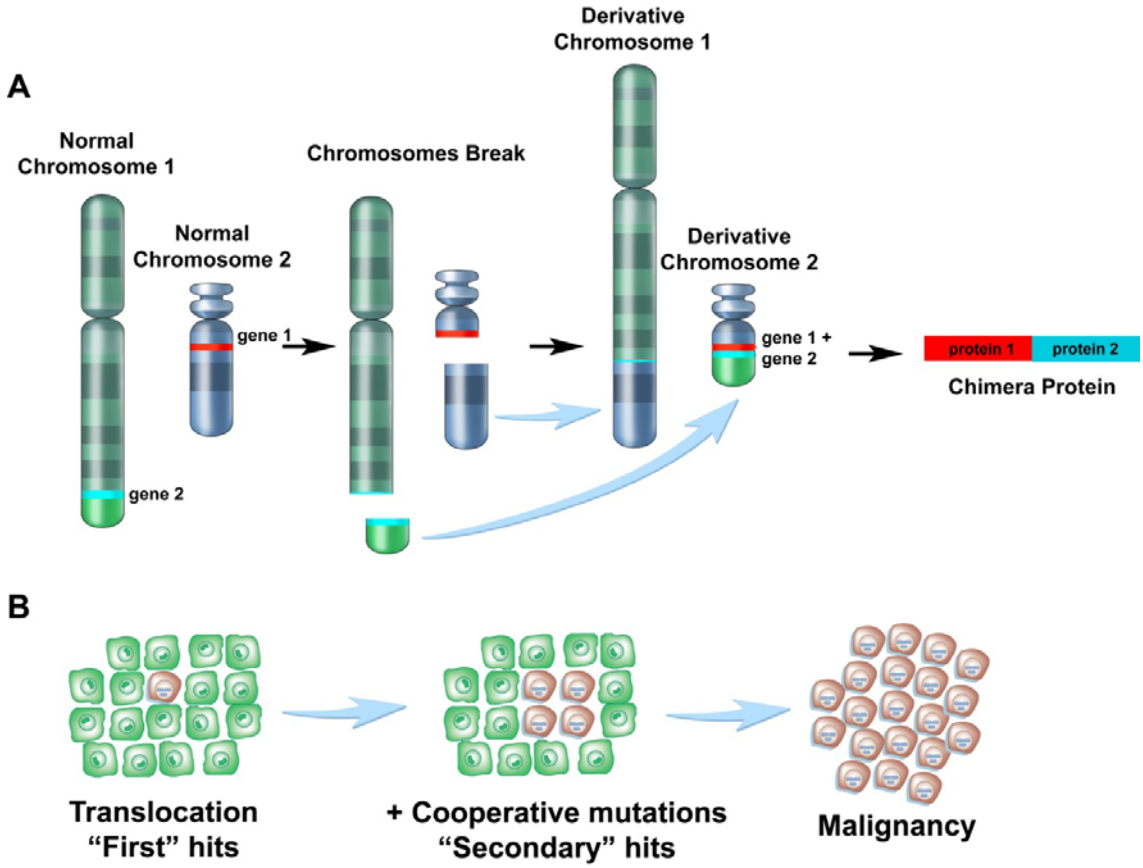


Figure 4-1. Balanced chromosome translocation and cancer cell malignancy initiated by chromosome translocation.

(A) Schematic cartoon of balanced chromosome translocation. During chromosome translocation, the chromosome breakpoints recombine to form chimeric chromosomes and express a chimeric fusion protein from fusion junction site. (B) The progression of cancer cell from harbouring a chromosome translocation to full malignancy.

The t(12;21)(p13;q22) translocation is the most common chromosomal translocation in paediatric cancers, occurring in approximately 25 % of cases of childhood pro-B cell acute lymphoblastic leukaemia (cALL) (137). The rearrangement results in the in-frame fusion of the 5' terminal of the ETS transcription factor TEL (also known as ETV6), to almost the entirety of the AML1 gene (also known as RUNX1). *AML1* encodes one of the DNA binding subunits of the core binding factor (CBF) and is related to the Drosophila gene *RUNT*. AML1 has been shown to play a role in regulating lymphoid and myeloid development (138). Clinical studies have found that the TEL-AML1 translocation occurs *in utero*, followed by a protracted time delay for leukaemia to develop (139). The disease has been recently identified to originate from a $CD34^{+}CD38^{-}$/low$CD19^{+}$ rare blood cell population (140).

The human *TEL* gene encodes a 452 amino acid ETS family transcription factor which was isolated as a fusion partner with the β chain of PDGF Receptor in t(5;12) chronic myelomonocytic leukaemia (135). In addition to its C terminus 85 amino acid ETS DNA binding domain, the N terminus of TEL contains a conserved interaction domain, which is responsible for oligomerization, and also for maintaining transcriptional activity (141). Apart from TEL-AML fusion-mediated malignancies, the TEL gene is involved in several 12p13 chromosomes translocations associated with a range of human malignancies, both as a N-terminal and a C-terminal fusion partner (142). AML1, like TEL, has been found to be rearranged with a number of different genes in leukaemogenic translocations including ETO, ETO-related MTG16 and EVI1 (143-145).

Although the formation of balanced chromosomal translocations is a frequent event in the pathogenesis of human malignancy, it is commonly believed the eventual formation of cancer normally requires additional mutations, or *'*secondary hits*'* (146) (**Figure 4-1 B**). In t(12;21) cALL, the *TEL-AML1* fusion was shown to occur in haematopoietic cells of a strikingly high proportion of live births (147). However, only a small fraction of these neonates go on to develop leukaemia (approximately 1/100), often with a long latent period, suggesting that although the *TEL-AML1* fusion may be acting as an initiating mutation, it is not sufficient to cause the disease. In addition, animal studies where the *TEL-AML1* fusion was expressed under the control of the IGH enhancer failed to cause leukaemia in mouse (148), indicating that additional mutations might be required to cooperate with the *TEL-AML1* fusion for cALL to develop. Research into the nature of mutations that can cooperate with *TEL-AML1* to cause cALL has yielded interesting results. In particular, it appears that the most common such mutation is a complete or partial deletion of the second allele of *TEL* which occurs in

approximately 70 % of cases (149). In addition, it has also been proposed that deletion or mutation of *Paired-Box-Containing Gene 5* (*PAX5)* participates with *TEL-AML1* in the induction of cALL (150) but again this has not been experimentally verified. The co-operating mutations that cause the most aggressive treatment refractory forms of cALL are yet to be identified.

Identification of mutations that cooperate with the *TEL-AML1* in the development of cALL remains far from complete, but will ultimately help to understand the mechanism of this disease and to find new treatments. One technique that can help to identify cooperative mutations is *in vivo* transposon-mediated insertional mutagenesis. In lower organisms and cell cultures, transposon systems have been routinely used for genetic manipulation and mutagenesis screens (55) (59,60). Recent publications have indicated transposons can also be engineered as mutagens in higher eukaryotes (47). The well characterized *Tc1*-like transposon *Sleeping Beauty* (SB) system consists of two components: the transposase, the enzyme responsible for mobilisation ('jumping'), and the transposon, the actual mobilised piece of DNA. It was recently shown that the SB transposon is an effective somatic insertional mutagen for studying oncogenesis and for identifying novel candidate cancer genes in mice (88,89). This was achieved through generation of two independent transgenic mouse lines, one harbouring chromosomal concatamers of the transposon DNA, and the other expressing SB transposase under the control of a ubiquitous promoter. The crossing of transposase and transposon mice yielded experimental mice in which the transposon is mobilised in the soma. The tumours were found to harbour transposon integration sites in both novel genes, as well as known human cancer genes (88,89).

4.2 Aims and summary of the project

In this chapter I plan to generate a *Tel-AML1* knockin mouse to model childhood pro-B cell acute lymphoblastic leukaemia (cALL), one of the most common paediatric cancers in human patients. The first stage of this project is to generate the model in mouse ES cells and validate the ES cell line *in vitro* before testing them *in vivo*. The second stage of the project is to model the cancer initiation and progression in the *Tel-AML1* mouse models, and combing this model with the *Sleeping Beauty* transposon system to identify *'*secondary hits*'*. I then aim to identify candidate genes and pathways underlining the cancer progression and formation process. Specifically the aims of this project are:

1. Validate the *TEL-AML1* knockin system i*n vitro* in ES cell for fusion transcripts and protein expression.

2. Cross *Tel-AML1* knockin mice with *T2/Onc* transposon mice (contains *Sleeping Beauty* transposon array) to initiate transposon mutagenesis and monitor the tumour formation by tumour watch.

3. Validate the tumour types using histology and FACS analysis

4. Combine the *TEL-AML1* system with the *Sleeping Beauty* transposon system to identify genes that when either inactivated or over-expressed represent *'*secondary hits*'* and cooperate with expression of the *TEL-AML1* to form ALL in mice.

A mouse model of *Tel-AML1* was generated by knocking-in the human *AML1* cDNA into the locus of *Tel* by DNA engineering (**Figure 4-2**), allowing expression of the fusion protein from the endogenous *Tel* promoter constitutively. A cDNA sequence encoding the *Sleeping Beauty* transposase was also knocked into the *Tel* locus after an internal ribosomal entry site (IRES), allowing deployment of transposon-mediated mutagenesis for screening the 'secondary hits'. The *loxP* sequences flanking the knockin cassette were originally designed to make a 'conditional' targeting construct. As the constitutive expressing *Tel-AML1* ES cells are viable and could be transmitted through the germ line, the conditional version was not used in this study. A backup model for *TEL-AML1* mouse model was also generated, where *TEL-AML1* was knocked into the Rosa 26 locus under a ubiquitous promoter. This mouse model will be further discussed in the results section.

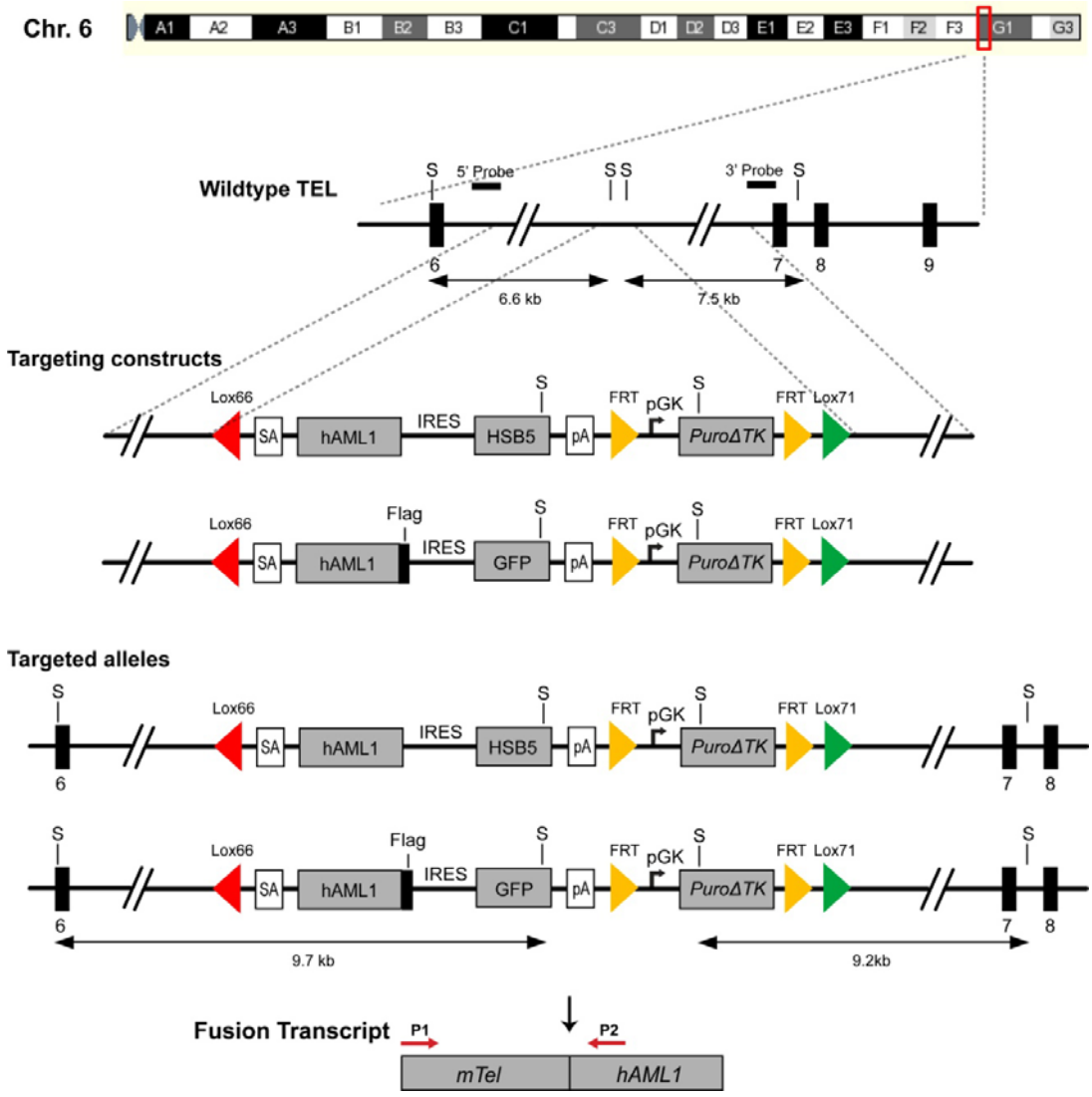


Figure 4-2. Schematic diagrams of *Tel-AML1* targeting constructs and targeted alleles

Both the SB transposon and GFP versions targeting constructs and targeted alleles are shown. The GFP version targeting construct is modified from the original SB version construct by adding a Flag tag sequence to the C-terminal of human AML sequence and replacing the HSB5 with a GFP sequence by homologous recombination. The GFP version construct was mostly used for *in vitro* validation experiments and the SB transposon version was used to generate knockin mice and for tumour watch experiment. S – *Stu*I restriction sites. The double head arrows below the schematic graph indicate the DNA fragment length in southern blot after *Stu*I digestion.

Characterization work (see results) showed that the *AML1* cDNA transcript is precisely fused with *Tel*, resulting in expression of *Tel-AML1* transcript from the *Tel* locus *in vivo* at levels approximating endogenous expression. By immunoprecipitation using an anti-Flag tag antibody, the fusion protein can be readily detected by western blotting. As part of the strategy to identify the 'secondary hits', the *Sleeping Beauty* transposase is also expressed from *Tel* locus and expression of the transposase results *'*jumping*' in vivo* in the experimental mouse after crossing of the *Tel-AML1* mouse with the *T2/Onc* transposon mouse.

The goal for *in vivo* study with this mouse model is to derive of disease mimicking human ALL and to identify cooperative mutations associated with *Tel-AML1*. The *Tel-AML1*$^{+/-}$ knockin mice were crossed with the *T2/Onc* transposon mice to generate experimental mice for a tumour watch study (**Figure 4-3**). After tumour generation in these mice the tumour types were analyzed by histological methods and FACS for disease identification. In the end DNA could be extracted from these tumours to identify cooperative mutations by splinkerette PCR (**Figure 4-3**).

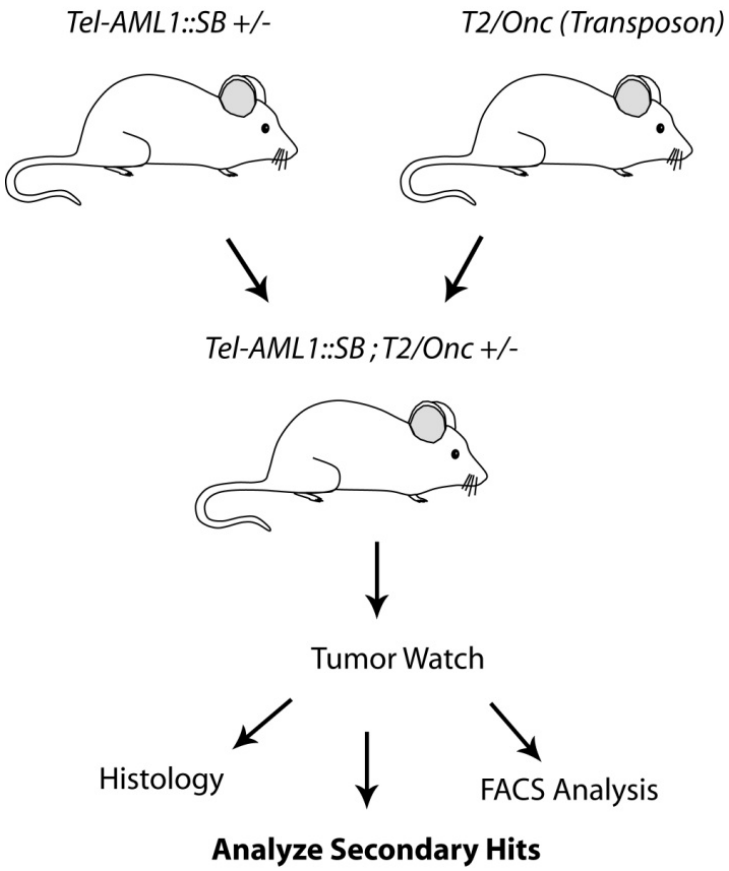


Figure 4-3. Crossing strategy for tumour watch and subsequent characterization studies in the *Tel-AML1* mouse model.

The tumour watch was initiated by crossing *Tel-AML1* knockin mouse with the *T2/Onc* transposon mouse to activate transposition and derive leukaemia formation in the experimental mice. The tumours generated were subjected to histology and FACS analysis to identify the tumour types. These tumours will subsequently been analyzed to identify secondary mutations using splinkerette PCR.

4.3 Materials and Methods

4.3.1 Targeting construct generation

To generate the *Tel-AML1* targeting construct (**Figure 4-2**), a genomic fragment of *Tel* was cloned into pBlueScript SK+ vector by gap repair from a 129S7 BAC clone (bMQ-66F22) using two homologous arms (PCR primers for 5' arm: FWD: 5'-GACAAAGTAGATGG CACCAGTGCAGTG-3', REV: 5'-GATGAGTGGTCAGGGGGGCAAAGAAGGAAAAA AAACCTTACAGAAA-3' ; 3' arm: FWD: 5'-GGTTGAAGGGCAGAGCTCTAGTGTCAA TTTG-3', REV: 5'-GGCTGGAGGGCAAACCAGGTACCATTACAGCAC TAGAAACCA GAGA-3') of 497 bp each. The AML1-SB-Puro cassette and the Lox66 and Lox71 sites were synthesized by GENEART. This cassette was inserted as a *Hpa*I fragment into a *Stu*I site within the *Tel* genomic fragment. This targeting construct was subsequently modified by molecular cloning to add a Flag tag to the end of *AML1* sequence: First a Flag tag sequence was introduced at the C-terminus of AML1 with primers: FWD: 5'-GCTCGCCGCCGCGCATCCT-3'; REV: 5'-GGCCTTAATTAATCACTTGTCGTCATCGTCCT-3'. The PCR product was then digested overnight with restriction enzyme *Pac*I. A subsequent double digestion of the targeting construct was performed overnight with *Afe*I/*Pac*I, and products were ligated at 4°C for 16 hours. The resultant plasmid was transfected into *E. Coli*, and colonies picked for PCR screening. Positive clones were verified by sequencing.

To generate a GFP version construct of the *Tel-AML1* I a *ccdB* negative selection marker was cloned by PCR and used to exchange the transposase sequence by recombineering. Two *Sbf*I restriction sites were introduced on the two ends of the *ccdB* cassette. A GFP sequence was cloned by PCR with *Sbf*I restriction ends and was used to exchange the ccdB fragment by *Sbf*I digestion and ligation. All constructs were sequenced in full to ensure that PCR had not introduced any mutations. PCR primers:

CCDB: 5'-CGGGGACGTGGTTTTCCTTTGAAAAACACGATGATAATATGGCCACAA CCCCTGCAGGGCATTAGGCACCCCAGGCT TTACAC-3' (FWD) and 5'-TAGATGCA TGCTCGA GCGGCCGCCAGTGTGATGGATATCTGCAGAGAATTCCTGCAGGTGCA GACTGGC TGTGTATAAGGGAG-3' (REV)

GFP: 5'-GTTTTCCTTTGAAAAACACGATGATAATATGGCCACAACCCCTGCAGG ATGGTGAGCAAGGGCGAGGAGCTGT-3' (FWD) and 5'-CTCGAGCGGCCGCCAGTG

TGAT GGATATCTGCAGAGAATTCCTGCAGGTTACTTGTACAGCTCGTCCATGC CG-3' (REV).

4.3.2 ES cell transfection and selection

For transfection of the targeting construct into embryonic stem cells, wild type E14 cells (Strain Name: 129P2 Ola) were fed in M-15 medium till 70-80 % confluent. To harvest cells, one culture dish of cells (~3×10^7 cells) was washed twice with phosphate buffered saline (PBS) and then treated with 3 ml Trypsin (1×) at 37°C for 7 minutes. The trypsin reaction was stopped by adding 10 ml M-15 medium and dissociated cells were spun down at 400×g for 4 min. The supernatant was removed and cells were washed two times with PBS. After the second wash, cell pellet was dissolved in PBS to give a final concentration of 1.4×10^7 cells/ml. Plasmid DNA was prepared using a JETstar Plasmid Purification Maxi Kit (Cat. No. 220020) and linearized with *Pvu*I restriction enzyme prior to electroporation. 0.9 ml of the cell suspension was added into a 0.4 cm BIO-RAD Gene Pulser Cuvette. 40 µg plasmid DNA was mixed well with cells in the cuvette. Electroporation was performed at 800 V, 25 mF with a BIO-RAD Gene Pulser II Electroporator. Cells were rested for 5 minutes after electroporation and were plated on 10 cm culture dish with confluent feeder cells. Cells were cultured in M-15 medium for two days, and then 3 µg/ml puromycin added to the medium to allow selection over a further 10 days of growth. Colonies were picked on day 12 and correctly targeted clones were screened for by southern blot analysis.

4.3.3 Generation of pMSCV expression constructs

The pMSCV-GFP expression vector was constructed by inserting an IRES2-EGFP sequence into the *Bgl*II restriction site of pMSCVneo (Clontech) downstream of the 5' LTR. The human *TEL-AML1* sequence was amplified from a GFP-*TEL-AML1* construct described previously (151) using the following primers (5' primer: 5'-GGCCGAATTCATGTCTGA GACTCCTGCTCA-3'; 3' primer: 5'- AAGATCTTCACTTGTCGTCATCGTCCTTGTAGT CCCGCGGGTAGGGCCTCCACACGGCCT-3'), thus incorporating a Flag tag at the C-Terminus of *TEL-AML1*. These constructs were sequenced in full to ensure that PCR had not introduced any mutations. The two mouse alternative initiation transcripts of *Tel-AML1* sequences (M1 and M43) were amplified from cDNA of targeted ES cells and were cloned into the pMSCV-human Tel-AML1-GFP construct using *EcoR*I and *Bcl*I restriction sites to insert the mouse *Tel-AML1* sequence into the vector (**Figure 4-5 A**).

4.3.4 Immunocytochemistry

The immunostaining using anti-Flag M2 antibody (Sigma, F1804) was performed following manufacturer's instructions (Sigma). Briefly, human 293T embryonic kidney (H293T) cells were seeded one day before transfection on cover slips in 10 cm dishes. The next day 10 µg plasmid DNA for each expression construct were transfected into H293T cells by calcium phosphate transfection using a ViraPack transfection reagent (Stratagene). Two days after transfection, cover slips were fixed in 6 well plates with 4 % paraformaldehyde in PBS (Sigma) with 4 % sucrose (BDH). Cells were then permeabilized with 0.25 % Triton X-100 (Sigma). Blocking was performed using 10 % bovine serum albumin (BSA) in PBS for 30 min at 37 ˚C. Coverslips were then incubated with ANTI-FLAG M2 antibody (1:1000 dilution) in 3 % BSA/PBS for 2 hours at 37 ˚C in a humidified chamber. After three washes in PBS, cells were incubated with Alexa Fluor 568 anti-mouse secondary antibody (Molecular Probes) for 45 min at 37 ˚C and mounted in VECTASHIELD (Vector) with DAPI. Slides were dried at room temperature and analyzed by fluorescent microscopy.

4.3.5 Immunoprecipitation of Flag Tagged Proteins

The Dynabeads protein G (Invitrogen, 0.5 ml beads) were first washed three times (buffer: 24.5mM Citric Acid, 51.7 mM Dibasic Sodium phosphate (Na_2HPO4) dehydrate, pH = 5.5). One microgram of anti-Flag M2 antibody (Sigma, F1804) was incubated with the beads in 20 µl of bead wash buffer for 40 min at room temperature. After incubation beads were washed three times with beads wash buffer with 0.1 % Tween-20 (Sigma). To prepare a cell lysate, cells were treated with protein lysis buffer (50 mM Tris pH 8.0, 450 mM NaCl, 0.2 % Nonidet P-40 (Igepal), 1 mM DTT, 1 mM EDTA, 1X Protease inhibitor (Roche) for 15 min on ice, then collected by centrfiguation at maximum speed using a desktop centrifuge for 15 min at 4 ˚C. Cell lysate was collected and incubated with the antibody conjugated beads for 1 hour at 4 ˚C with gentle shaking. The beads were collected after incubation and washed three times with protein wash buffer (same formula as protein lysis buffer except using 150 mM NaCl and 0.1 % NP-40 concentration). For Western blotting, 30 µl loading buffer were added to the beads after pull down. The beads were then boiled for 10 min at 95 ˚C and supernatants were loaded directly on SDS page gels (5 %, Bio-rad). Western blotting was using anti-Flag M2 antibody (Sigma, F1804) and performed following manufacturer's instructions from Sigma.

4.3.6 RNA isolation and cDNA preparation

The total RNA from cell culture or mouse tissue was isolated using TRIZOL reagent (Invitrogen, Cat. No. 15596-018) and manufacturer's standard protocol. Briefly, tissue samples collected from mice were crushed through a 70 □m nylon cell strainer (BD Falcon, Cat. No. 352350) in PBS to create a single-cell suspension. The cells were then pelleted (1,500 rpm for 5 min at 4 ˚C) and resuspended in 1 ml TRIZOL reagent. Cell pellet was lysed in TRIZOL for 5 minutes at room temperature and 0.2 ml chloroform added. After vigorously shaking the tube was stood at room temperature for approximately 5 minutes and then spun down by centrifuge at 12,000 g for 15 minutes at 4 ˚C. After centrifugation, the upper aqueous phase was carefully transferred into a new tube. The RNA was precipitated by adding 0.6 ml isopropyl alcohol and centrifuged at 12,000 g for 10 minutes at 4 ˚C. The RNA pellet was washed by 75 % ethanol and allowed to dry at room temperature. RNA was dissolved in 100 µl H_2O and stored at -20 ˚C for quantitative PCR. The first strand total cDNA was reverse transcribed using a SuperScript First-Strand RT-PCR kit from Invitrogen (Cat. No. 11904-018) following the Random Hexamers first-stand synthesis protocol from the manufacturer.

4.3.7 Quantitative PCR

The quantitative PCR was performed using an ABsolute™ QPCR ROX Mix kit (Cat. No. AB-4138/B) according to previously described procedure with slight modifications (152). The probe was ordered from MWG Operon which was labelled with FAM at 5' end and TAMRA at 3' end. The probe was first diluted in ddH_2O to 4 pmol/µl and then master mix was prepared using following formula per reaction: QPCR Mix kit 12.5 µl; Forward/Reverse Primer (100 nM) 0.25 µl each; Probe (4 pmol/µl) 0.5 µl; ddH_2O 6.5 µl. The reactions were set up in a standard 96 well plate. 20 µl master mix was added into each well with 5 µl reverse transcribed total cDNA. Two parallel reactions were setup for each cDNA sample, and 5 µl ddH_2O was mixed with 20 µl master mix as negative control. The plate was sealed using a MicroAmp Optical Adhesive Film (ABgene, 4311971) and PCR reaction was performed on the ABI PRISM 7900HT sequence detection system (Applied Biosystems). Thermal cycling was initiated with an incubation step at 50 °C for 2 min, followed by a first denaturation step at 95 °C for 15 minutes, and continued with 40 cycles of 95 °C for 15 seconds, 60 °C for 1 minute. The fluorescent signal representing the transcripts expression level was normalized with the β-Actin control. Probe and primer sequences were as follows:

Tel-AML1 Probe: 5'-AGCACGCCATGCCCATTGGG-3';
FWD Primer: 5'-CTTGAACCACATCATGGTCTCTATG-3';
REV Primer: 5'-TCGTGCTGGCATCTGCTATT-3'.
Tel Probe: 5'- CACGCCATGCCCATTGGGAGAA-3';
FWD Primer: 5'-TCTCTATGTCCCCACCGGAAG-3';
REV Primer: 5'-CATAATCCCAAAGCAGTCTACAGTCT-3'.
β-Actin Probe: 5'-TTTGAGACCTTCAACACCCCAGCCA-3';
FWD Primer: 5'-CGTGAAAAGATGACCCAGATCA-3';
REV Primer: 5'-CACAGCCTGGATGGCTACGT-3'.

4.4 Results

4.4.1 Generating the *Tel-AML1* knockin mouse model and characterizing the targeted ES cells

A *Tel-AML1* knockin targeting construct was first generated by conventional DNA cloning and recombineering in mouse ES cells. In the targeting construct, the human *AML1* cDNA was inserted into intron 6 of mouse *Tel* allele (Ensembl transcript ID: ENSMUST00000111963) so that a *Tel-AML1* fusion would be generated by splicing of the *Tel* transcript to *AML1* from the endogenous mouse *Tel* locus (**Figure 4-2**). The hyperactive *Sleeping Beauty* transposase variant HSB5 (provided by Steven Yant, Stanford) was cloned behind the *AML1* cDNA after an internal ribosomal entry site (IRES). This design results in the production of a bicistronic messenger composed of the *Tel-AML1* fusion and the *Sleeping Beauty* transposase (**Figure 4-2**). After introducing this construct into ES cells, targeting was analysed by southern blot analysis, with successfully targeted constructs having a 9.7 kb band (**Figure 4-4 A**). Using RT-PCR (Figure 4-2, primers P1 + P2) the expression of fusion transcript could be detected as two splicing variants (**Figure 4-4, B**), representing the alternative splicing forms M1 and M43 from *Tel* exon 2 (**Figure 4-4, D**). Sequencing results subsequently confirmed a precise in-frame fusion between *Tel* and *AML1* (**Figure 4-4, C**). The targeted ES cells were injected and the allele had been successfully transmitted through the germ line.

4.4.2 Detection of the Tel-AML1 fusion protein by immunoprecipitation

Although the *Tel-AML1* fusion transcript was easily detected from the knockin ES cells, at the protein level the fusion protein could not be detected by western blotting in cell lysate using either the anti-TEL or anti-AML1 antibody (data not shown). This may be due to very low level expression of the protein from the endogenous *Tel* locus, or alternatively that the fusion protein is not stable in ES cells.

The fact that the *Tel-AML1* fusion transcript but not the protein could be easily detected, a more sensitive detection method was required. Since only the presence of the oncogenic fusion protein itself would be the direct validation for the mouse model the original *Tel-AML1* targeting construct was modified by recombination to add a Flag tag sequence after the *AML1* cDNA (**Figure 4-2 A**) so that a highly-specific anti-Flag antibody could be used to detect the fusion protein. After targeting of this construct into ES cells and concentration of the fusion protein from cell lysate by immunoprecipitation (see Materials and Methods), the Flag tagged Tel-AML1 was successfully detected by western blotting. Two splicing isoforms were identified at ~100 and ~110 KDa (**Figure 4-2 E**) in two targeted cell lines (BM2 and BM3), representing the translation started from M1 and M43 alternative start codons, respectively (153).

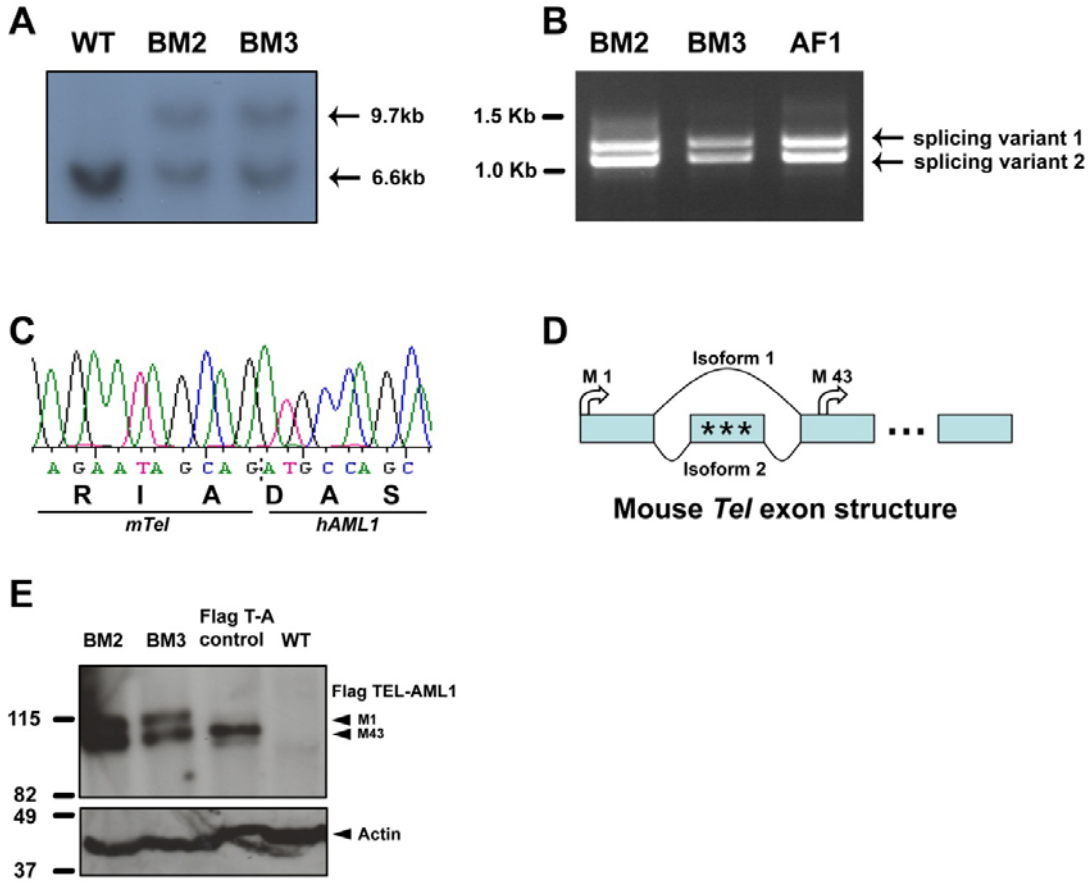


Figure 4-4. Characterization of *Tel-AML1* targeted ES cells

(A) Southern blotting detected a targeted band at 9.7 kb for two targeted ES cell clones (BM2 and BM3), in addition to the wild type band at 6.6 kb. (B) RT-PCR detected two splicing variants in three targeted ES cell clones. The clone AF1 is the actual injected clone used to derive the F1 mice. (C) Sequencing trace for the RT-PCR product showing in-frame fusion between *Tel* and *AML1* transcripts. (D) Schematic diagram of mouse Tel exon structure. Two alternative start codons (M1 and M43) are indicated. Three stop codons on exon 2 are represented with *'*'*. (E) Western blot showing two isoforms of FLAG tagged TEL-AML1 fusion proteins (M1 and M43) detected in two targeted ES cell clones after immunoprecipitation. The Flag T-A positive control is the *in vitro* expressed TEL-AML1-FLAG protein, and WT is the E14 wild type ES cells which served as a negative control.

4.4.3 *In vitro* characterization of the mouse TEL – human AML1 fusion protein

Previous studies have suggested that expression of the human TEL-AML1 is oncogenic in both mouse and zebrafish (151,154,155), however no study has ever characterized the mouse TEL-human AML1 fusion protein which is expressed in our mouse model. To prove that the mouse-human fusion is biologically identical to human TEL-AML1, localisation studies of the mouse and human TEL-AML1 fusions were performed, since localization study could be a good indication to protein's biological function. The full length mouse version transcripts, including both the M1 and M43 splicing variants, were cloned from targeted ES cells and inserted into the pMSCV-GFP expression vector (**Figure 4-5, A**). The human version of TEL-AML1 was generated as a positive control and an empty GFP vector was used as a negative control. All the fusion proteins contained a Flag tag at the C terminus and could be detected using an anti-FLAG antibody on a Western blot after transfection into human embryonic kidney 293T cells (**Figure 4-5, B**). The localization of these fusion proteins was examined using immunocytochemistry. Both the human TEL-AML1 and two isoforms of mouse-human TEL-AML1 were localized in the cytoplasm in 293T cells (**Figure 4-5, C**). This localization for human TEL-AML1 has been suggested from previous reports (156). In addition, no obvious localization difference between the M1 and M43 isoforms can be seen (**Figure 4-5, C**). Although this localization study is not a directly proof of the function, it demonstrates that the fusion transcript expressed in the knockin mouse can give rise to a stable protein which shares the same localization pattern with the human fusion protein in 293T cells.

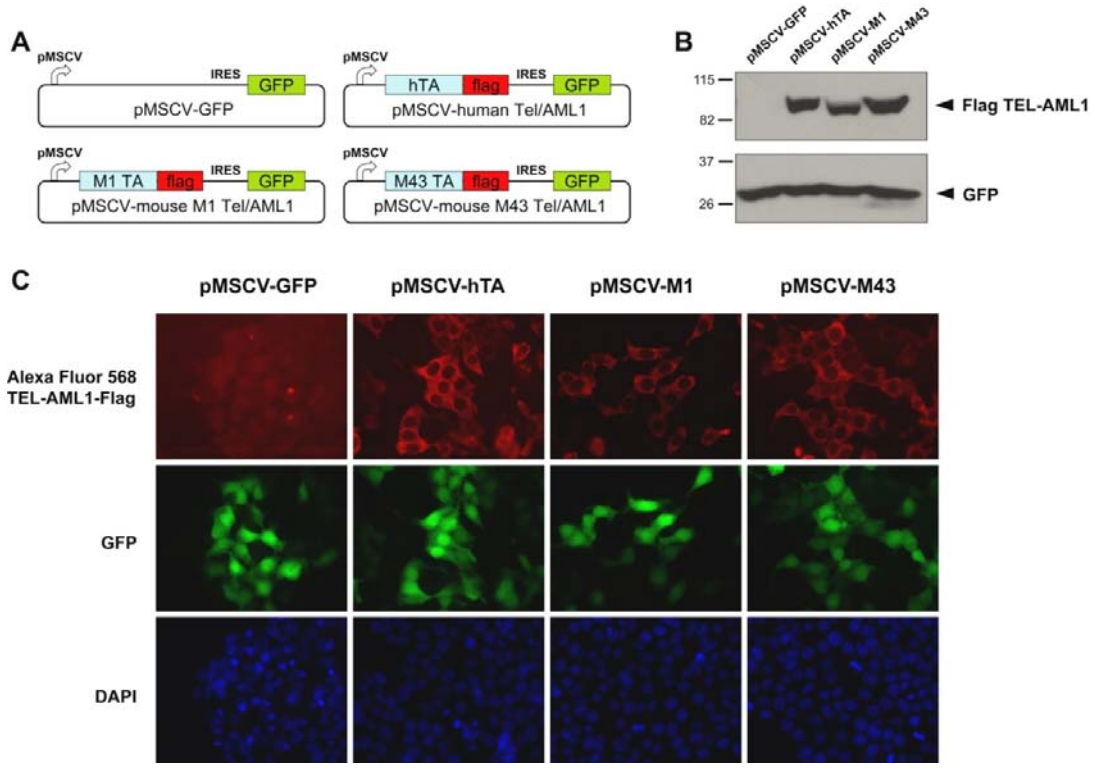


Figure 4-5. Characterization of the mouse-human TEL-AML1 protein by *in vitro* studies

(A) Generation of the pMSCV overexpression constructs for expressing TEL-AML1 fusion proteins. The mouse Tel-AML1 was expressed as two isoforms (M1 and M43). Human TEL-AML1 was used as a positive control and the original pMSCV-GFP construct was used as a negative control. (B) Western blot to verify expression of the fusion proteins in 293T cells using anti-FLAG antibody. (C) Immunostaining in 293T cells transfected with empty vector, human TEL-AML1 or two isoforms of mouse Tel-AML1 (M1 and M43).

4.4.4 Analysis of the *Sleeping Beauty* Transposon system in the knockin mouse

In our *Tel-AML1* mouse model, a transposon mediated insertional mutagenesis system was designed to incorporate with TEL-AML1 expression as a combined strategy to identify the cooperative mutations associated with formation of the *TEL-AML1*. To achieve this, a hyperactive variant of the *Sleeping Beauty* transposase (HSB5) was introduced into the *Tel-AML1* knockin mouse (**Figure 4-2**). *In vitro* tests were first performed to characterize the transposon system before *in vivo* application. By western blotting, the SB transposase protein expression could be detected in targeted ES cells (**Figure 4-6, A**). To determine if the transposon could be mobilized *in vivo*, a splicing PCR was performed on mouse tail DNA using primers flanking the SB transposon. A 225 bp 'jumping' band was detected on the electrophoresis gel with the *Tel-AML1::SB* $^{+/-}$ mouse crossed with the *T2/Onc* transposon mouse (88) but not the *T2/Onc* transposon mouse alone (**Figure 4-6, B**). These results indicated that the transposon system is active in our *Tel-AML1* knockin mouse.

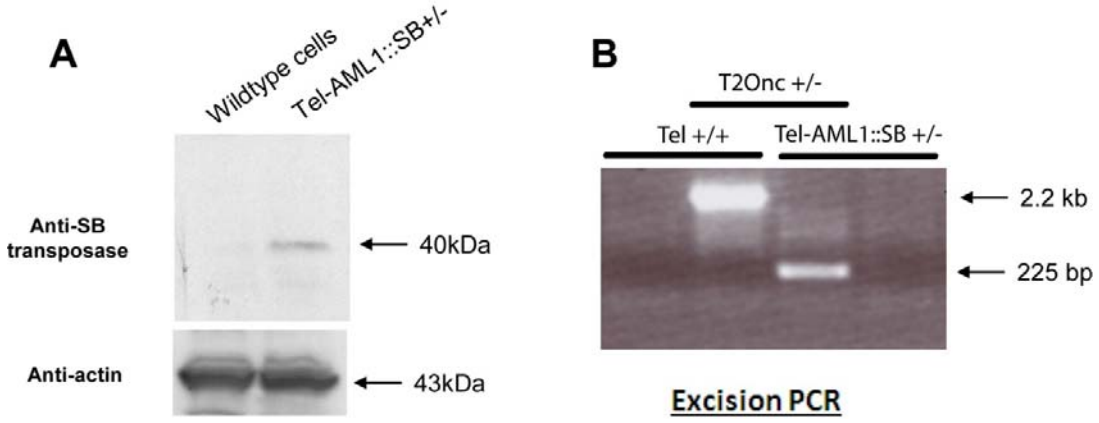


Figure 4-6. Characterization of the transposon system in *Tel-AML1* mouse model

(A) Western blotting on *Sleeping Beauty* transposase in the targeted *Tel-AML1* ES cells showing a band at 40 kDa representing the transposase protein. (B) Excision PCR for detecting transposon 'jumping' in the knockin mice using primer pairs surrounding the transposon sequence that has been described (88). A 225 bp band could be amplified in the knockin mouse crossed with *T2/Onc* transposon mouse indicating the transposon has been mobilized *in vivo* (lane 3). The *T2/Onc* transposon mouse along only generated a 2.2 kb 'unjumped' band (lane 2). The knockin mouse DNA and wild type mouse DNA was served as negative control and generated no band (lane 1 and 4).

4.4.5 Validation of the *Tel-AML1* expression level by real-time qPCR

The *in vitro* validation experiment in targeted ES cells detected the expression of the correct fusion transcripts from endogenous *Tel* locus, but was not a quantitative analysis. A real-time qPCR experiment was carried out to quantify and compare the expression level of the *Tel-AML1* transcript with the wild type *Tel* expressed from the alternate allele (**Figure 4-7 A**). The experiment was performed on hematopoetic tissues (bone marrow, spleen, thymus) as they are mostly relevant to the disease in question. To detect the fusion transcript expression two-week old mice were sacrificed and tissues were taken to prepare RNA and cDNA. The cDNA sample were first validated by Reverse Transcriptase PCR (RT-PCR) for the presence of fusion transcript in three hematopoetic tissues in the knockin mice (**Figure 4-7 B**). The real-time qPCR was performed using an ABsolute™ QPCR ROX Mix kit following the protocol described (see Materials and Methods). The expression level of *Tel-AML1* and *Tel* cDNA were normalized with β-Actin expression level and compared. From the results the *Tel-AML1* is expressed at a lower but within a comparable range (less than one fold difference in expression level) with the *Tel* transcript expressed from the other allele in the knockin mice (**Figure 4-7 C**). This result indicates that the knockin allele strategy was working *in vivo*, resulting in expression of comparable amount of *Tel-AML1*.

The expression level of *Tel* transcript in five knockin mice and five wild type mice was compared. Although *Tel* was expressed from only one allele in the knockin mice rather than from both alleles in the wild type mice, the expression level was similar indicating that there might be a compensation mechanism for *Tel* expression in the knockin mice (**Figure 4-7 D**).

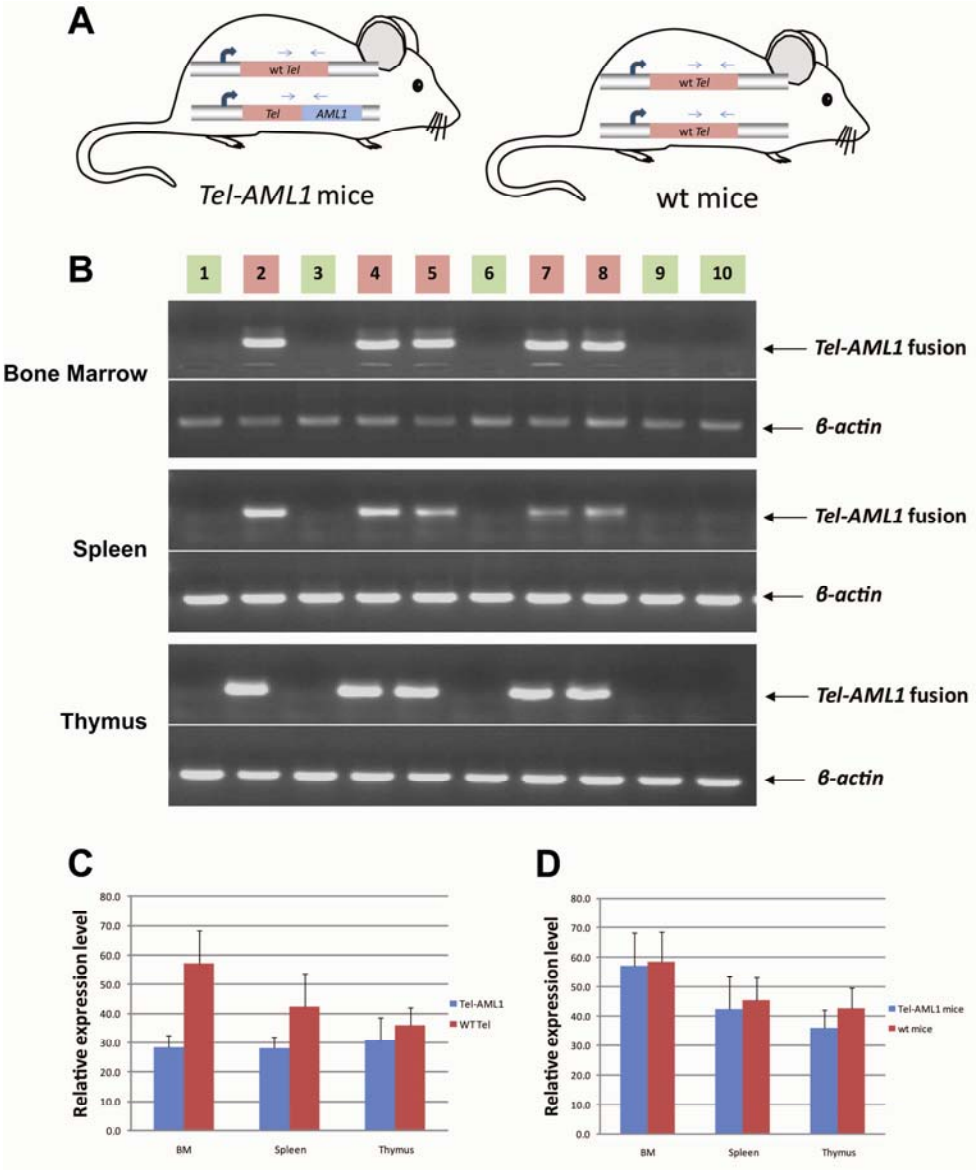


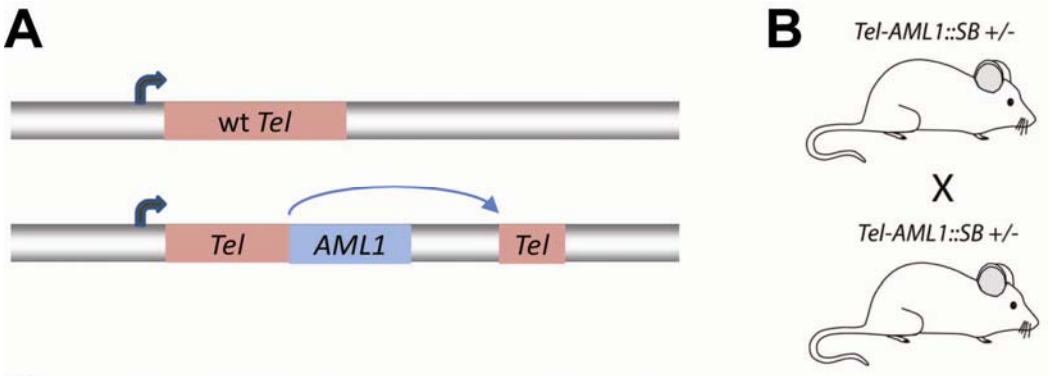
Figure 4-7. *In vivo* validation of the *Tel-AML1* expression strategy and real-time qPCR

(A) Genotypes of the *Tel-AML1* mice and wild type mice. The arrows represent the primer pairs used for RT-PCR analysis; (B) Detection of expression of *Tel-AML1* fusion transcripts by RT-PCR in *Tel-AML1* knockin mice (n = 5) and wild type mice (n = 5). The numbers in red box represent the *Tel-AML1* mouse number and in green box represent the wt mouse number; (C) Comparison of the relative expression level of the *Tel-AML1* transcript and wild type *Tel* transcript in *Tel-AML1* knockin mice (n = 5); (D) Compare the relative expression level of wild type *Tel* transcript in *Tel-AML1* mice (n = 5) and wild type mice (n = 5). The relative expression level was normalized to β-Actin.

4.4.6 Analysis of cryptic splicing in *Tel-AML1* knockin mice

One of the potential problems in designing a knockin experiment is cryptic splicing over the knockin allele i.e. part or full length of the knockin transcript could not be properly spliced. In *the Tel-AML1* knockin, the human *AML1* sequence was inserted into the *Tel* intron sequence between exon 6 and 7, leaving the 3' *Tel* sequence intact. Theoretically the splicing could jump over the *AML1* sequence and splice into the endogenous 3' *Tel* splicing site to generate wild type *Tel* transcript from the knockin allele (**Figure 4-8 A**). To assess cryptic splicing in the knockin mice, two heterozygous mice were crossed to generate homozygous embryos (**Figure 4-8 B**). According to Mendelian principles of inheritance this cross would generate progenies with three genotypes for *Tel-AML1* (+/+, +/-,-/-) at a rate of 1:2:1. However, genotyping for 95 live born mice resulted in no live born homozygous *Tel-AML1* mice (**Figure 4-8 C**), indicating an embryo lethal phenotype associated with the *Tel-AML1*$^{-/-}$ embryo.

To confirm the embryo lethality of *Tel-AML1*$^{-/-}$ mice, embryos were examined 10.5 days into pregnancy. The double knockin embryo, if no cryptic splicing taking place, should be identical to homozygous *Tel* knockout embryos which are embryonic lethal at E10.5 day (157). 27 embryos were harvested at E10.5 day and a ratio of 8:13:6 was obtained from three genotypes, close to the Mendelian ratio of 1:2:1 (**Figure 4-8 C**). From the embryo imaging all the wild type and heterozygous embryos have quite typical E10.5 day embryo morphology but the homozygous *Tel-AML1* embryos were smaller than the wildtypes and heterozygotes and some were under degradation indicating the embryo lethality (**Figure 4-8 D**). These results confirmed that the *Tel-AML1* knockin allele should be non-permissive for cryptic splicing.



C ***Tel-AML1* Intercross Results**

	$Tel^{+/+}$	$Tel^{+/AML1}$	$Tel^{AML1/AML1}$
Genotype at E10.5d	8	13	6
Genotype at live born	45	50	0

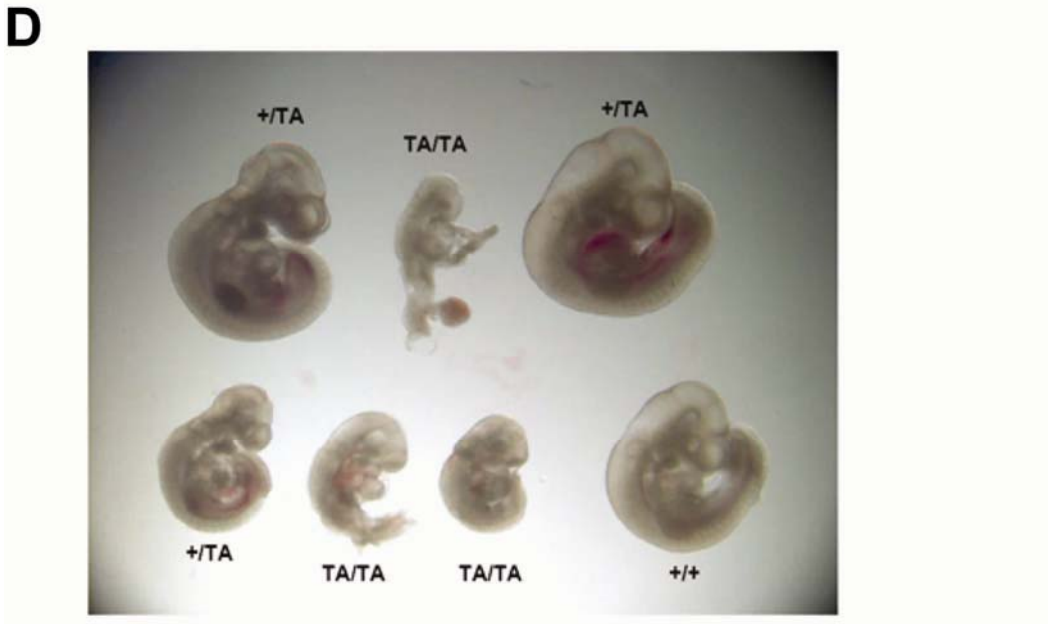


Figure 4-8. Investigation of cryptic splicing in *Tel-AML1* knockin mice by intercrossing

(A) Cartoon of the wild type and knockin alleles in the *Tel-AML1* knockin Embryo and heterozygous mice. Potentially, the splicing could jump over the AML1 knockin sequence and splice into the endogenous 3' Tel sequence. (B) The strategy of the intercrossing experiment. (C) Results of the intercrossing experiment showing the number of embryo (mouse) obtained at two different stages (Embryo day 10.5 and live born). (D) Morphology of the embryo at E10.5 day with genotypes indicated.

4.4.7 Tumour watch study in *Tel-AML1* knockin mice

To investigate the oncogenic property of *Tel-AML1* for acute lymphoblastic leukaemia (ALL) in the knockin mouse model, the *Tel-AML1*$^{+/-}$ mice were crossed with T2Onc$^{+/-}$ transposon mice to generate transposon 'jumping' mice (*Tel-AML1*$^{+/-}$; T2Onc$^{+/-}$ or *TAOnc*, *n* = 90) (**Figure 4-3**) or non-jumping mice (*Tel-AML1*$^{+/-}$; T2Onc$^{+/+}$ or *TA*, *n* = 54). The progenies of other two cohorts without *Tel-AML1* knockin (*Tel-AML1*$^{+/+}$; T2Onc$^{+/-}$ or *Onc* (n = 50), *Tel-AML1*$^{+/+}$; T2Onc$^{+/+}$ or *WT*, *n* = 22)) from the same cross were served as control groups. The experiment and control cohort mice were raised in an animal facility receiving standard care and routine check on a daily basis. Once the phenotype of sickness or cancer was identified, mice were sacrificed for histological and haematological studies for disease identification. In the 'jumping ' cohort mice (*TA*$^{+/-}$,*T2Onc*$^{+/-}$) tumour progression started as early as 25 weeks and the population also declined the quickest among the four cohorts as depicted in a Kaplan-Meier curve (**Figure 4-9 A**). All the 'non-jumping' cohorts have very similar curve patterns with each other except the *Tel-AML1* knockin cohort had slightly higher accelerated death rate than the other two control cohorts. This result indicates that the *Tel-AML1* knockin and *Sleeping Beauty* transposon system co-ordinately accelerated the tumour progressing.

The tumours isolated from these mice were subsequently identified by Dr. Brian Huntly - a haematologist from Cambridge Addenbrooke's Hospital following standard Bethesda criteria for lymphoid and non-lymphoid murine malignancies. Control mice in the *Onc* and *WT* cohorts occasionally presented with acute leukemias (10/72 mice, 14 %), however the incidence was significantly increased in the cohorts expressing *Tel-AML1* (73/90 mice, 83 % for the *TAOnc* cohort and 13/54, 24 % for *TA*) (**Figure 4-9 B**). Among the *Tel-AML1* knockin cohorts, a significant proportion of *TAOnc* mice (15/73, 21 %) developed B cell progenitor (BCP)-ALL, as did a slightly smaller proportion of the TA cohort (2/16 cases, 13 %). Importantly, BCP-ALL disease was only seen in mice carrying the *TEL-AML1* allele. FACS plots from the bone marrow of a representative mouse demonstrate only background Gr-1/Mac1 myeloid cells, with the majority of cells having a B220^{+}/CD19^{-}/sIg^{-} phenotype, in keeping with BCP-ALL (**Figure 4-9 C**). The histology study showed the presence of lymphoblasts in the peripheral blood, as is the infiltration of the spleen, bone marrow and liver, with effacement of the normal cellular architecture and replacement by nucleolated blasts (**Figure 4-9 D**). These results showed the *Tel-AML1* knockin mouse model could be able to derive BCP-ALL, i.e. the modelling strategy was successful. For the other types of

leukaemia identified from *TAOnc* mice, acute myeloid leukemia (AML) is predominant (34/75 cases, 44 %), with T-cell ALL also seen (14/75, 19 %) (**Figure 4-9 B**).

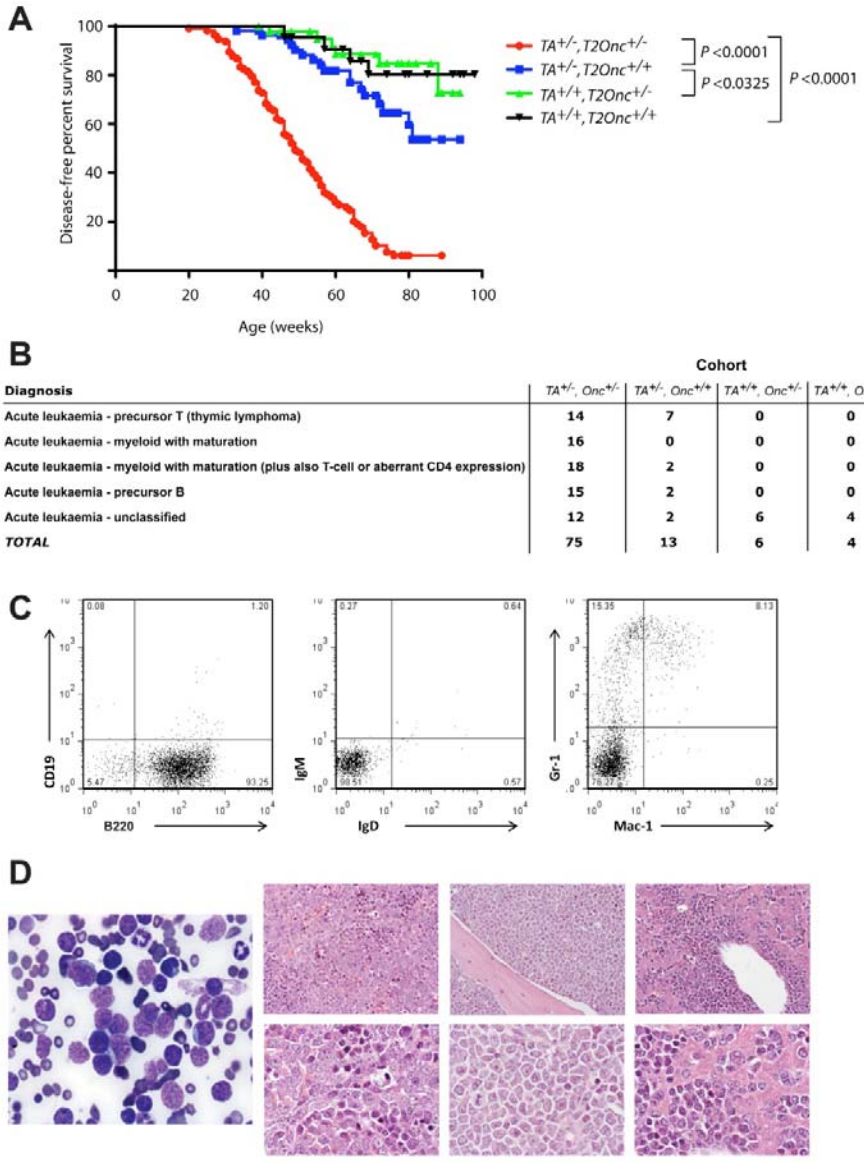


Figure 4-9. Tumour progression and hiostological analysis in *Tel-AML1* knockin mice

(A) Kaplan-Meier curves showing the tumour latency of 'jumping' Tel+/AML1; T2Onc+/- mice and 'non-jumping' control mice. (B) Classification of the malignancies developed by mice according to the Bethesda criteria for lymphoid and non-lymphoid murine malignancies (158,159). (C) FACS analysis from the bone marrow of a representative *Tel-AML1*$^{+/-}$ mice where transposon mutagenesis resulted in development of B cell precursor acute lymphoblastic leukemia; only background Gr-1/Mac1 myeloid cells and a majority of B220+/CD19-/sIg- cells are detected, in keeping with B-ALL. (D) Peripheral blood, spleen, bone marrow and liver from a representative mouse with B-ALL, showing phenotypes recapitulating features of the human disease.

4.4.8 An alternative mouse model of TEL-AML1

As an alternative to the knockin mouse model, I also generated a *'*conditional*'* *TEL-AML1* mouse model where a GFP tagged TEL-AML1 is expressed from the Rosa26 locus and also bicistronically the HSB5 *Sleeping Beauty* transposase (**Figure 4-10**). This mouse model contains a human version *TEL-AML1* fusion flanked by two inverted *LoxP* sites, which can flip the *GFP-TEL-AML1* fusion into the 'active' orientation following expression of Cre recombinase. In this construct the GFP TEL-AML1 fusion protein is expressed from the CAGGS promoter and as such is an overexpression model. Because TEL-AML1 is tagged with GFP this model may prove useful in downstream validation experiments. ES cells carrying this construct have also been injected and the chimeras are being bred for germ line transmission. However, as the knockin mouse successfully reproduced the characteristics of human ALL, I did not pursue my study on this backup mouse model after it was made.

Figure 4-10. The targeting construct of Rosa 26 GFP-TEL-AML1 mouse model.

The human version *TEL-AML1* with a N-terminus GFP tag is inserted as a reverse orientation into Rosa 26 targeting vector downstream of CAGGS promoter, together with a hyperactive transposase coding sequence (HSB5) following an IRES site. The TEL-AML1-HSB5 cassette is flanked by two inverted LoxP sites to allow *in vivo* switching on using Cre recombinase. This targeting construct is targeted into Rosa 26 locus by two homologous arms of 4.0 and 4.2 kb, respectively.

4.5 Discussion

4.5.1 Advantages of knockin mouse model for characterizing human *TEL-AML1* oncogenic translocation

As a result of this research, a *Tel-AML1* mouse model was generated to study the *in vivo* consequence(s) of this oncogenic translocation. The *TEL-AML1* is one of the most frequent translocations identified in paediatric cancer and occurring in 25 % of cases of childhood ALL. Although many studies indicated *TEL-AML1* might cooperate with secondary mutations to cause cancer, the detailed molecular mechanism required to induce leukaemogenesis is still remains unclear. In addition, previous transgenic *TEL-AML1* mouse models have all been unsuccessful, creating an opportunity to generate a knockin mouse model that better recapitulates this disease. Therefore, our mouse models bring the opportunity to help understanding the molecular mechanism of a common human cancer.

Compare to mouse models been generated by other strategies, this mouse model was generated through direct knockin of a downstream fusion partner into the original translocation locus have several advantages: 1) The fusion transcript is expressed at the endogenous level. The oncoprotein expression level has already been implicated to be critical for cancer development in KRAS and other mouse models (160), therefore appropriate expression levels will hopefully result the development of leukaemia in the mouse model; 2) The knockin allele also contains a transposase coding sequence. By one cross with the transposon mouse, secondary mutation(s) can be generated to cooperate in the cancer development; linker-based PCR technology could allow quickly identify these *'*secondary hits*'* in mouse tumours; 3) The expression of fusion protein and transposase are temporally and spatially regulated in the same way as in endogenous disease induction, therefore the cancer should develop at the right lineage and right stage.

4.5.2 Expression level of oncogenic fusion protein

Difficulty was experienced in identifying expression of the Tel-AML1 fusion protein in the mouse model and therefore analysis of a large amount of cell lysate by pull down assay was required to identify the fusion protein. This is probably due to low endogenous expression level, and ES cells may not be a good host for expression of the oncogenic fusion protein. Another possibility for low expression level is due to cryptic splicing over the knockin allele, however the *Tel-AML* transcript was able to be detected by real time qPCR, indicating

comparable expression to the endogenous *Tel*. In addition, the homozygous *Tel-AML1* were embryonic lethal, suggesting the cryptic splicing might not count for the low expression of the fusion protein. Therefore a functional correlation may exist between the oncogenic property of the fusion protein and low expression level. This is also indicated by previous publications that oncogenic fusion proteins expressed at a much lower level compared with their wild type fusion partners (153,161,162). It has also been shown that the low expression level(s) might be due to a low rate of synthesis but not a decreased half-life (161). One possible mechanism is that the fusion protein is expressed at low level in the predisposition stage of cancer formation but could be up-regulated by cooperative mutations during malignancy. In this mouse model the fusion protein incorporates a Flag tag, allowing the future testing of this hypothesis in mouse tumours.

4.5.3 The choice of mouse or human AML1

The mouse model presented here expressed a *Tel-AML1* fusion between mouse *Tel* and human *AML1*. It was decided to knockin the human *AML1* cDNA because other mouse models using human fusion protein developed a pre-leukaemic phenotype with an expanded pro-B cell population, consistent with the phenotype of *TEL-AML1* associated cALL in human. Furthermore, the zebrafish model of *TEL-AML1* developed leukaemia, indicating the human *TEL-AML1* could induce leukaemia formation in other model organisms. Both TEL and AML1 protein sequences between mouse and human are highly similar, sharing 88 % and 95 % sequence identity in the TEL and AML1 sequence respectively. Therefore it would be reasonable to assume both the mouse and human version TEL-AML1 fusion protein would have very similar oncogenic properties. However, as part of the fusion protein that has already been characterized, human *AML1* was used sequence in the mouse model.

4.5.4 The *Tel-AML1* knockin mice as a model for human ALL

The *Tel-AML1* knockin mice have derived certain numbers of B cell leukaemia under transposon-mediated mutagenesis, suggesting this knockin strategy could be successfully used for modelling human cALL by *TEL-AML1* translocation, a disease which could not be modelled by previously studies in mice. The previous *TEL-AML* mouse models, however, have based on retrovirus-mediated transplantation or transgenic method for ectopically delivering and expressing *TEL-AML1*. One hypothesis to explain the lack of success of these *TEL-AML* mouse models is that the oncogenic fusion protein needs to be expressed at a

physiological level to induce the right type of disease. In our mouse model, the *Tel-AML1* fusion is expressed from the *Tel* locus at physiological level, thus recapitulating the *TEL-AML1* oncogenic fusion expressed from the human patient. In addition, from the tumour watch studies it seems that SB transposon-mediated mutagenesis could greatly accelerate the disease in the knockin mouse, suggesting the requirement of secondary mutations to derive B cell leukaemia. The next step of this study would be to isolate insertion sites from the appropriate tumour type and identify these cooperative mutations. In turn candidate genes as 'secondary hits' could then be validated using this *Tel-AML1* mice for understanding the pathogenesis of this disease and investigate treatment strategies for use in human patients.

In contrast to human *TEL-AML1* patients where the disease is predominantly derived from pre-B cells, the acute leukaemia derived from the *Tel-AML* mice can be originated from other cellular types such as myeloid and precursor T cells. The reason for this different disease spectrum could be well caused by the differences between mouse and human in hematopoietic system. For instance, the common cell surface markers for undifferentiated hematopoietic stem cells in mouse are $CD34^{low/-}CD38^{+}$ lin^{-} but for human are $CD34^{+}CD38^{low/-}$ lin^{-}. In addition, the cell signalling pathways are also quite different between mouse and human in hematopoietic system.