Establishing a genetically engineered mouse ES cell line expressing an inducible Xist transgene along chromosome 19

Tan, Christina En Hui

2018

Tan, C. E. H. (2018). Establishing a genetically engineered mouse ES cell line expressing an inducible Xist transgene along chromosome 19. Master's thesis, Nanyang Technological University, Singapore.

https://hdl.handle.net/10356/89676

https://doi.org/10.32657/10220/46341

Establishing a genetically engineered mouse ES cell line expressing an inducible Xist transgene along chromosome 19

CHRISTINA TAN EN HUI

SCHOOL OF BIOLOGICAL SCIENCES

A thesis submitted ot the Nanyang Technological University in partial fulfilment of the requirement for the degree of Master of Science (SBS)

2018

Acknowledgements

To my supervisors, Asst Prof Zhang Li-Feng and Asst Prof Francesc Xavier Roca Castella, and past and current lab members, Dr Hong Ru, Ms Lai Lan-Tian, Mr Norbert Ha, Ms Mehnaz Huq, Ms Siti Nadirah Bte Ismail, Ms Zheng Yashu, for their guidance and support throughout my research.

Table of Contents

List of figures

List of abbreviations

µg	Microgram
ºC	Degree Celsius
bp	Base pair
Cas9	CRISPR associated protein 9
CRISPR	Clustered regularly interspaced short palindromic repeats
Dox	Doxycycline
DSBs	Double-strand breaks
ES cell	Embryonic stem cell
FISH	Fluorescence in situ hybridization
gDNA	Genomic DNA
HDR	Homology-directed repair
lncRNA	Long non-coding RNA
NHEJ	Non-homologous end-joining
rtTA	Reverse tetracycline transactivator
sgRNA	Single guide RNA
TRE	Tetracycline response element
XCI	X chromosome inactivation
Xi	Inactive X chromosome
Xic	X chromosome inactivation center
Xist	X-inactive specific transcript
Xm	Maternal X chromosome
Xp	Paternal X chromosome

Abstract

Xist (X-inactive specific transcript) RNA plays an important role in X chromosome inactivation (XCI), a dosage compensation mechanism mammalian females have evolved to transcriptionally silence one of two copies of X chromosomes they carry. Xist RNA remains in the nucleus, and during XCI onset, it is transcribed only from the future inactive X, coating the chromosome from which it is being produced. This results in the exclusion of RNA polymerase II and accumulation of epigenetic modifications, and eventually, silencing of the selected inactive X. Using RNA fluorescence in situ hybridisation (FISH), Xist RNA coating can be visualised as “clouds” within the cells. How Xist RNA propagates along the chromosome has been widely researched upon, yet the working mechanisms of Xist RNA still remain largely elusive.

This project aims to address the question of whether Xist RNA is robust enough to recognise the boundaries of the chromosome from which it is being expressed from. This was achieved through establishing a genetically engineered cell line expressing an inducible Xist transgene along chromosome 19. As the smallest autosome in mice, occupying a much smaller area compared to the relatively large X, the stringency of the Xist RNA would be put to test, to recognise the smaller chromosome boundary of 19. We chose the Ainv15 cell line for this purpose as it contained a reverse tetracycline-controlled transactivator (rtTA) gene integrated into a constitutive locus, one of the components required to create a tractable, inducible system. A gene targeting site was inserted onto chromosome 19 via homology-directed repair (HDR) using the CRISPR-Cas9 genome editing method, and Cre-loxP mediated integration facilitated the insertion of the Xist transgene at the targeting site. The cell line established was verified by PCR genotyping, observation of colony fluorescence, and RNA FISH analysis. Further research with this model created can then be used to elucidate Xist cloud dynamics.

1. Introduction

1.1 Dosage compensation

In mammals, sex determination is dependent on the sex chromosomes, namely, X and Y, which evolved from a common ancestral pair of autosomes. Individuals carrying two X develop as females (homogametic sex), while those carrying an X and a Y develop as males (heterogametic sex). Over the course of evolution, the Y chromosome has degenerated dramatically, losing most of its content such that it presently contains only a few protein-coding genes. In contrast, the X chromosome, having conserved a relatively high gene density, contains approximately 1,000 genes (Graves, 2006). Given that embryonic lethality or severe developmental defects are inevitable consequences of changes in autosomal copy number, it is highly plausible that such a drastic imbalance in X chromosome copy number between the sexes would lead to similar detrimental effects. Yet, this twofold difference in X chromosome copy number between females and males is part of normal development, and the phenotypes of individuals who inherit abnormal copy numbers of X, for example as in XXY males and XXX or XO females, are fairly mild compared to those with autosomal aneuploidies. The explanation for this remarkable behaviour is this: while the XY sex-determination system evolved to distinguish between the sexes, dosage compensation mechanisms have co-evolved to offset potential damaging consequences associated with disparities in hundreds of X-linked genes (Pontier & Gribnau, 2011).

Although Mary Lyon was the first to put forth the hypothesis in 1961 that dosage compensation is attained through the inactivation of one of two X chromosomes in each cell of mammalian females, earlier studies had reported findings that hinted at this phenomenon. The first of which was made by Barr and Bertram in 1949, when they observed the presence of a darkly stained aggregate of DNA in neuronal nuclei

of female but not male cats (Barr & Bertram, 1949). Just over a decade later, Ohno et al. deduced using female rat liver cells, that this structure represents the condensed X, while the other appears euchromatic like autosomes (Ohno et al., 1959). Subsequently, drawing extensively from the works of those before her, in addition to her independent studies of the mosaic appearance of female mice heterozygous for an X-linked fur colour gene, Lyon proposed that it is the random inactivation of either one of the X chromosomes in each cell, early on in embryonic development of female mice that accounts for this observation (Lyon, 1961). X chromosome inactivation (XCI), the chromosome-wide silencing process of one X in females, therefore ensures that levels of X-linked gene expression are equivalent between mammalian males and females throughout development and adult life.

Although XCI occurs in all mammalian species, there are diverse techniques for achieving it. Marsupials display imprinted inactivation of the paternal X (Xp) in all tissues (Sharman, 1971). In contrast, eutherians are commonly understood to exhibit random inactivation in somatic tissues, where the maternal (Xm) or Xp have an equal chance of being silenced (Chow & Heard, 2009). Interestingly, imprinted inactivation of Xp is observed in the extraembryonic tissues of some eutherian mammals. However, even among eutherian mammals, recent studies have highlighted significant differences in the strategies employed for XCI initiation during early development (Furlan & Rougeulle, 2016; Okamoto et al., 2011; X. Wang et al., 2014; Waters & Waters, 2015). Regardless, one mutual characteristic across all mammals is the importance of long noncoding RNAs' (lncRNA) participation in XCI regulation. As the murine model has generally been preferred for XCI research, and as is the case in this project which uses murine embryonic stem (ES) cells, subsequent elaboration will be made with respect to this model.

1.2 XCI: the process

In the early mouse embryo, XCI occurs in two rounds. At approximately the 4 to 8 cell stage, the Xp is exclusively inactivated in all cells of the developing embryo (imprinted XCI). As it develops into a blastocyst, this imprinted XCI is reversed, resulting in reactivation of the Xp. This, however, occurs only in the inner cell mass (ICM); imprinted XCI persists in the extra-embryonic tissues. The second round of XCI is initiated shortly after implantation, around embryonic day 5.5, and this time, both Xp and Xm have equal chance to be inactivated (random XCI) (Pontier & Gribnau, 2011). Random XCI can be broken down into three phases: initiation, establishment, and maintenance (Maduro et al., 2016). In the first phase, XCI activators and inhibitors regulate the process by which the cell ensures that only one X is active per diploid genome. This leads to the monoallelic upregulation of the X-linked long noncoding RNA gene, Xist (X-inactive specific transcript). Elements regulating Xist transcription are all located within the X inactivation center (Xic), an X-linked control locus. In the next phase, Xist RNA associates in cis with the X from which it is being expressed, coating the entire presumptive inactive X (Xi). This induces a cascade of chromatin modifications as well as spatial reorganisation of the chromosome, ultimately contributing to the stable repression of nearly all genes on this chromosome. Once established, the Xi is clonally propagated to all daughter cells, and maintained through cell divisions; only before meiosis is it reversed in primordial germ cells (Ohhata & Wutz, 2013).

1.3 X-inactive specific transcript (Xist)

Xist was first described in the early 90s and plays a vital role in the XCI process. The full length mouse Xist gene is approximately 23 kb in length, comprising of 7 exons flanked by 6 introns. The Xist transcript is spliced and polyadenylated in the nucleus, similar to most other mRNAs, resulting in its functional form measuring 17 kb in length. Rather than being exported to the cytoplasm, Xist remains in the nucleus, and upon XCI onset, is transcribed solely from the future Xi, coating it to form typical "clouds" when visualised using RNA fluorescence in situ hybridisation (FISH) or live-cell imaging (Borsani et al., 1991; Brockdorff et al., 1991; Brown et al., 1991). The earliest event following Xi coating is RNA polymerase II exclusion from the Xist compartment, and then the steady amassing of a plethora of epigenetic marks (Pontier & Gribnau, 2011). Initial gene silencing is established with the recruitment of active genes into this compartment devoid of transcription factors, and subsequently epigenetic modifications that contribute to long-term silencing such as H3K27me3 enrichment, a hallmark of facultative heterochromatin, mediated by the Polycomb repressive complex 2 (PRC2) which is recruited by Xist to the Xi (Chadwick & Willard, 2004; Chaumeil et al., 2006; Marks et al., 2009). Importance of this lncRNA in the initiation of XCI was demonstrated not long after its discovery, when deletion of the gene was found to result in XCI failure (Penny et al., 1996) as well as lethality in female knockout mice (Marahrens et al., 1997). Although these deletion experiments in mice have shown that this lncRNA is necessary and sufficient for initiating XCI, it appears that Xist is largely dispensable with regards to maintaining the inactive state in differentiated cells, and is unable to initiate gene silencing when upregulated in this context (Brown & Willard, 1994).

A number of highly repetitive blocks lie within Xist and are highly conserved at the sequence level, in the otherwise poorly conserved RNA. Referred to as repeat A – E, the A-repeat has been shown to be the most highly conserved among them, evidenced by deletion experiments (Wutz et al., 2002). Other sequences seem to act in combination to facilitate the Xist localisation to the Xi, but they seem to serve redundant functions and no single sequence has been found to be necessary (Gendrel & Heard, 2014).

RNA FISH or live-cell imaging methods have demonstrated that the localisation of Xist RNA is confined to the Xi territory in the nucleus. Such findings indicate that precise mechanisms limit Xist spreading and binding to the chromosome from which it is expressed. However, Xist RNA localisation appears to be highly dynamic, with properties changing during the cell cycle and differing from species to species (Ng et al., 2011). The fact that Xist can spread over autosomes and cause repression of certain autosomal genes, demonstrates that the phenomenon of Xist RNA binding and XCI are not exclusively restricted to sequences on X (Heard et al., 1999; Herzing et al., 1997; Lee & Jaenisch, 1997; Lyon, 1998; Tang et al., 2010). How Xist RNA propagates along the chromosome is an intriguing question, and specifically, whether Xist RNA coating is precisely confined in its host chromosome's territory has yet to be explored.

1.4 Genome editing and CRISPR-Cas9

Primarily the hallmark of an adaptive immune system in bacteria, Class 2 Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) also form the basis for CRISPR-Cas9 (CRISPR-associated protein 9) genome editing technology. CRISPRs were first reported in 1993, but only two decades later were they successfully harnessed for genome engineering in eukaryotic cells (Cong et al., 2013). Before the

CRISPR-Cas9 system, editing approaches such as zinc finger nucleases (ZFN) or transcription-activator-like effector nucleases (TALENs) required a unique nuclease pair to be designed and generated for each genomic target, which proved to be tedious and cumbersome. In comparison, the CRISPR-Cas9 system composed of just two components: the Cas9 endonuclease and a guide RNA (gRNA), which itself consists of a scaffold sequence for Cas-binding and a user-defined approximately 20 nucleotide spacer that delineates the genomic target to be modified. Thus, the genomic target of the endonuclease can be changed easily just by changing the gRNA target sequence, making this system simple and highly adaptable. The system works by inducing double-strand breaks (DSBs) at specific sites recognised by the gRNA designed. In mammalian cells, these DSBs can be repaired either by the canonical non-homologous end-joining pathway (NHEJ) or the HDR pathway. While the most active repair mechanism, NHEJ however frequently generates small nucleotide insertions or deletions which often lead to frameshift mutations, and these are useful for the construction of knockout alleles. On the other hand, the latter pathway though much less efficient, can be leveraged to introduce precise genetic modifications by recombination with exogenously introduced repair templates (H. Wang et al., 2016). Apart from containing the desired edit, the repair template generated also needs to have additional homologous sequences immediately upstream and downstream of the target, and the lengths of these homologous sequences are dependent on the size of the modification being introduced.

2. Objective

This study aims to first establish a genetically engineered cell line with a gene targeting site on chromosome 19. Subsequently, an inducible full-length Xist transgene will be inserted into this engineered gene targeting site. In contrast to the X chromosome, chromosome 19 being the smallest autosome in mice, occupies a relatively smaller area. As such, challenging the ability of the Xist RNA to recognise and effect specifically on the smaller boundary of chromosome 19, from which it is being expressed. This transgenic murine ES cell line established would then provide the induced Xist RNA clouds covering chromosome 19, to be analysed in subsequent studies for the purposes of elucidating Xist cloud dynamics, specifically the question of the Xist RNA's robustness in recognising only the boundaries of the chromosome from which it is being expressed from.

2.1 Experiment outline and workflow

The murine ES cell line, Ainv15, was chosen for modification in this project (Kyba et al., 2002). The premise for selecting this cell line is the reverse tetracycline-controlled transactivator (rtTA) gene that was integrated into the constitutive ROSA26 locus on chromosome 6 of this cell line. As one of the two components required in a tetracycline-controlled transcriptional activation system known as Tet-On, the rtTA gene allows for the inducible expression of the transgene integrated at the loxP site in subsequent steps. However, it is also noted that because puromycin and neomycin selection had previously been used in the establishing of the Ainv15 cell line, the two resistance genes conferring resistance to these antibiotics cannot be used for selection purposes in this project. Figure 1 illustrates the above described modifications found in the Ainv15 cell line.

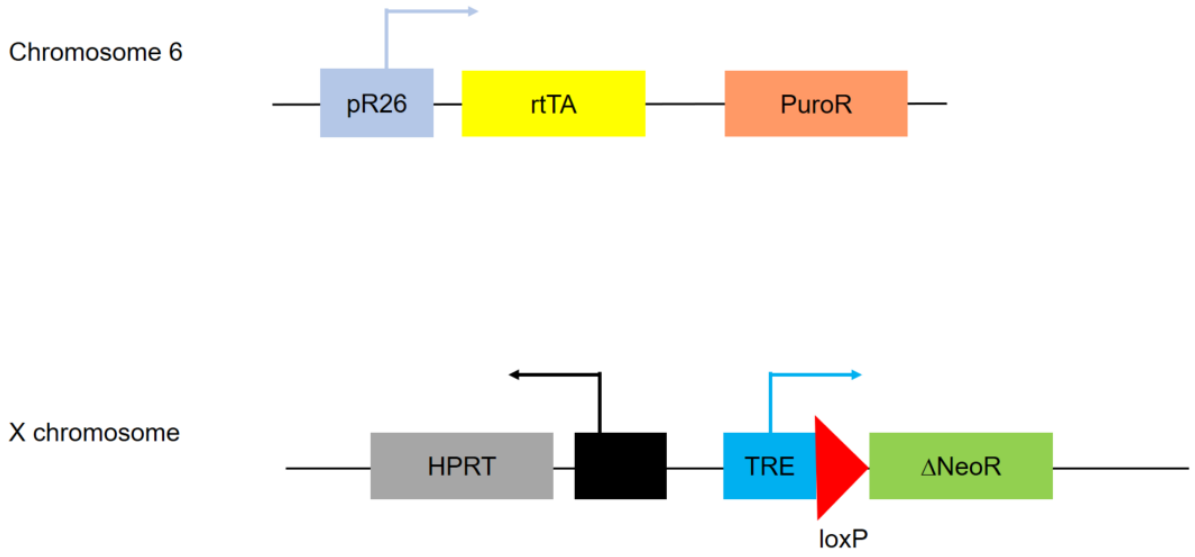


Figure 1: Schematic drawing of the Ainv15 cell line.

The rtTA is integrated into the constitutive ROSA26 locus on chromosome 6 while a tetracycline response element (TRE), loxP site and truncated neomycin resistance gene are integrated into the HPRT locus on the X chromosome (Adapted from Kyba et al., 2002).

In selecting suitable transgene insertion sites on chromosome 19, care was taken to select a locus that is constitutively active, with a similar epigenetic neighbourhood as ROSA26. This was to ensure that the gene of interest would not be silenced after insertion at the chosen site. The Cd6 gene was selected based on findings from a previously published paper which identified this locus as a transgene integration site that showed widespread reporter expression (Ichise et al., 2014). The Cd6 gene codes for CD6, a cell-surface receptor expressed on the majority of T-cells as well as a subset of B-cells in both humans and mice (Robinson et al., 1995). The observations that CD6 binds in *cis* to the TCR/CD3 complex (Gimferrer et al., 2004) and CD5 (Gimferrer et al., 2003) on T-cells, and also binds in *trans* to ALCAM

(CD166) (Bowen et al., 1995), 3A11 protein (Saifullah et al., 2004), and microbial components (Sarrias et al., 2007), to mediate cellular adhesion, suggests that CD6 plays multiple roles in lymphocyte function. As described in their paper, Ichise et al. had inserted their gene-of-interest into the intronic region between exon 1 and 2 of the Cd6 locus. As this intronic region is large, the primers used for PCR genotyping in the paper were used to approximate the location for homologous arm design in the repair template.

The CRISPR-Cas9 system was the genome engineering tool of choice given its simplicity and efficiency, requiring only a simple redesign of the gRNA to change target specificity to the desired locus on chromosome 19. As the modification to be inserted is large, a conventional double-stranded DNA targeting construct with homologous arms flanking the insertion sequence was used as a repair template. A total of 800 bp homology to the target locus on chromosome 19 was used. The modification to be inserted at the Cd6 locus consists of a TRE, a loxP site, and a hygromycin-resistant gene without the ATG start codon (ΔHygro). TRE working in combination with rtTA would complete the Tet-On system, allowing the control of Xist gene expression. The loxP site is required for the insertion of the Xist gene using a lox-targeting plasmid, iXist-3.1, previously generated by colleague Chelliah Rosi, via Cre recombination in subsequent steps. The successful targeting of the iXist-3.1 plasmid into the homing site on chromosome 19 would restore resistance to the antibiotic hygromycin, and thereby facilitate efficient isolation of transgenic cells.

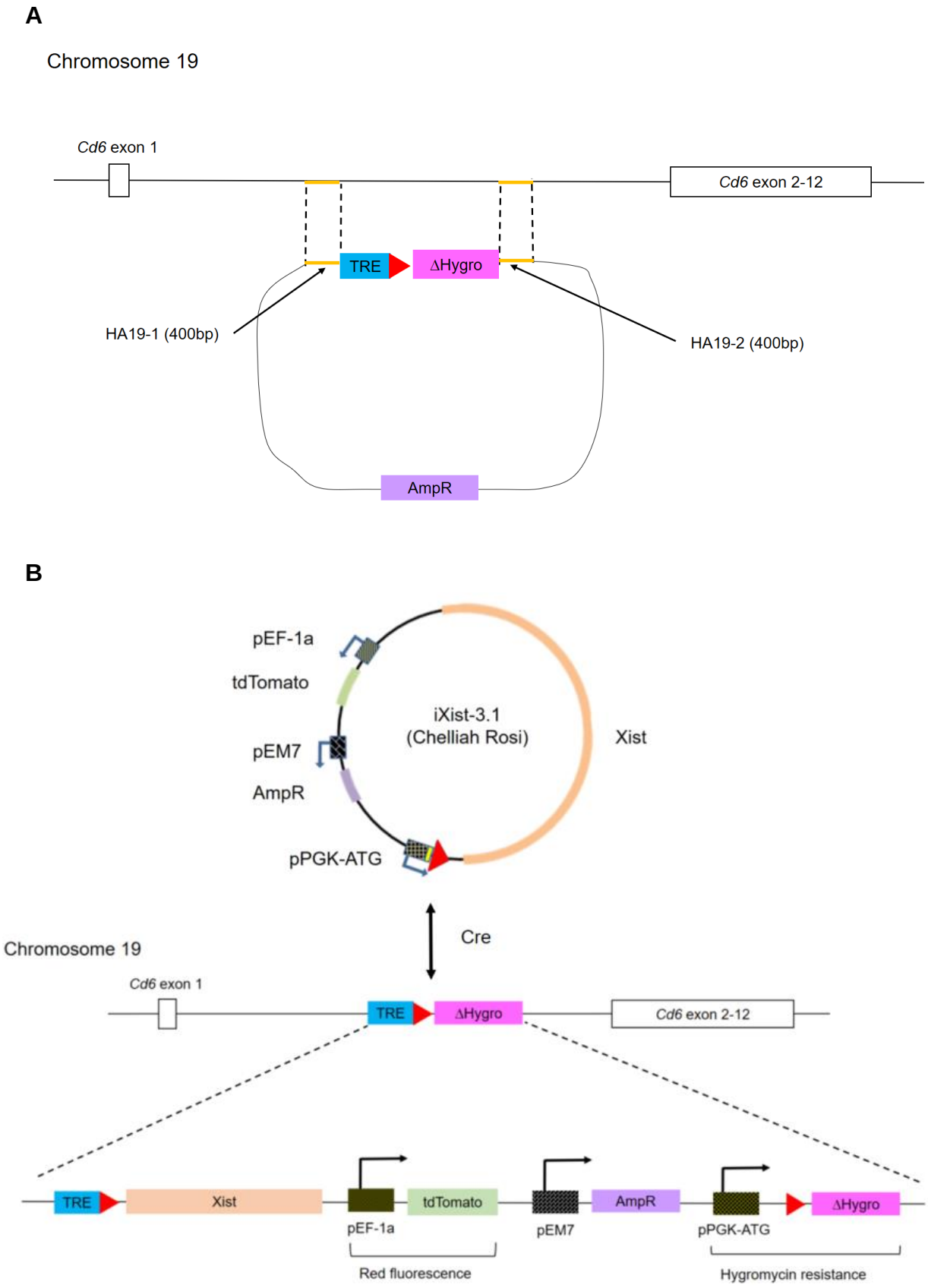


Figure 2: Schematic drawing of modification to be inserted at Cd6 locus.

(**A**) The CRISPR-Cas9 plasmids (not pictured) and the repair template are co-transfected into Ainv15 cells. Once the site-specific DSBs are generated by the CRISPR-Cas9 complexes, the repair template plasmid, which has homology to the sequences flanking the DSB, can serve as template for HDR, thereby inserting the cassette of TRE, loxP and truncated hygromycin resistance gene at the desired locus.

(**B**) The modified cells from the previous step are co-transfected with the pSALK CRE plasmid and the iXist-3.1 plasmid. As the iXist-3.1 plasmid contains a pPGK-ATG cassette, Cre-mediated recombination of the vector into the homing site on the Cd6 locus of chromosome 19 provides the ATG start codon required for the translation of the truncated hygromycin resistance gene, and therefore restoring hygromycin resistance to the transgenic cells. This process would also result in colonies that fluoresce red under a fluorescent microscope, as a result of the constitutive tdTomato gene.

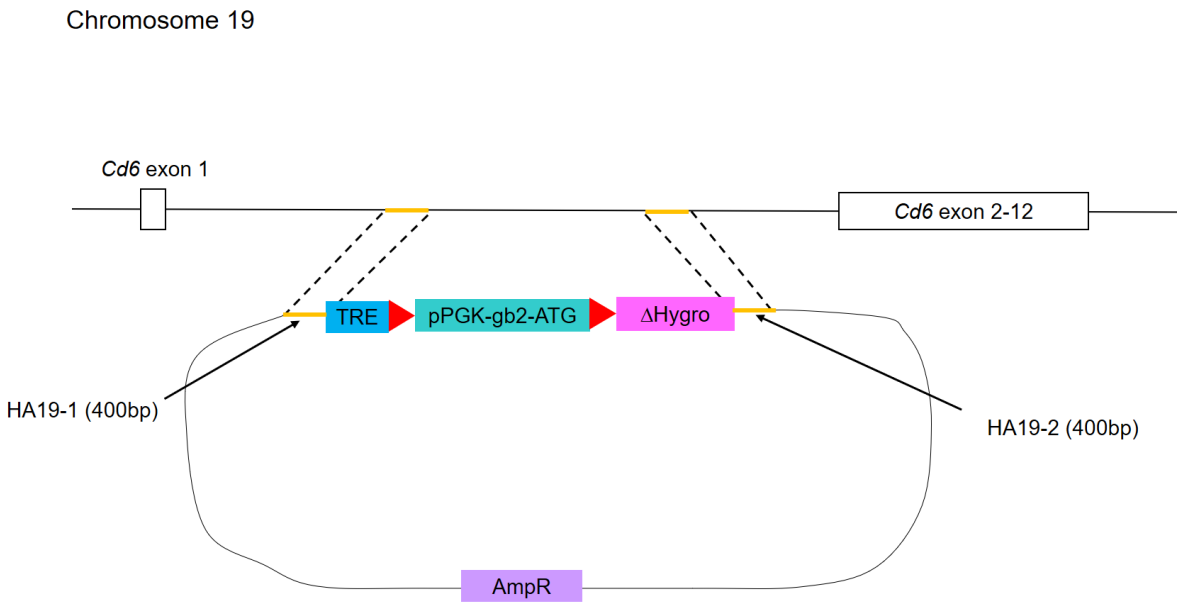
Once generated, the CRISPR-Cas9 and repair template plasmids would then be co-transfected into the Ainv15 cells. As no selection was introduced in the insertion of the desired modification, it was assumed that if a sufficiently large number of single colonies were picked from the pool of transfected cells, at least a handful of positive hits could be obtained. However, despite picking 96 single colonies, unfortunately none had the desired genotype following genomic DNA (gDNA) extraction and genotyping by a three-primer PCR system. This is likely due to the fact that the NHEJ repair pathway for DSBs that result from the Cas9-mediated cleavage is known to be more active than the HDR pathway. Hence, the chances of the modification being inserted are highly unlikely. As such, it was decided that the repair

template should be modified to introduce positive selection with hygromycin so as to facilitate more efficient isolation of the desired transgenic cells.

The revised repair template would introduce an additional loxP site and a PGK-gb2 fragment with the ATG start codon just after the TRE promoter on the original repair template (Figure 3A). The PGK-gb2 fragment consists of the eukaryotic PGK promoter and the prokaryotic gb2 promoter, thereby conferring hygromycin resistance to both transformed bacteria (on top of the existing ampicillin resistance), and the transfected ES cells subsequently. Once ES cells have been successfully isolated based on their hygromycin resistance, an additional step to excise the PGK-gb2-ATG cassette via Cre recombination is required to render the clones hygromycin sensitive again (Figure 3B).

Following this, clones that have been ascertained by PCR genotyping to have the PGK-gb2-ATG cassette excised would then be transfected with the iXist-3.1 and pSALK CRE plasmids as described earlier in Figure 2.

A



B

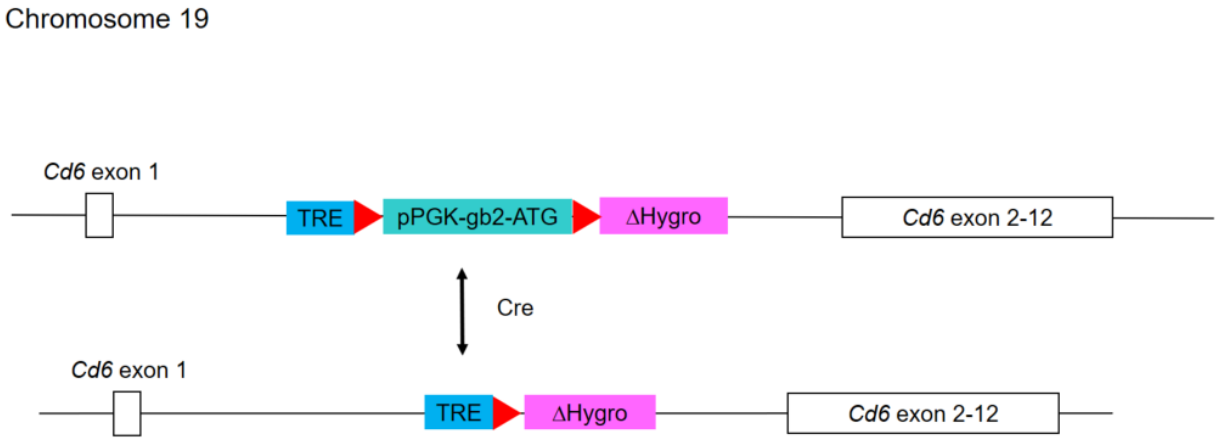


Figure 3: Schematic drawing of the revised repair template.

(**A**) The revised repair template utilises the original repair template as a “backbone”, and the loxP-PGK-gb2-ATG cassette is to be inserted between the TRE and the hygromycin gene.

(**B**) Following the generation of this new repair template to insert the desired modification into the Cd6 locus, an additional Cre recombination step is needed to excise the PGK-gb2-ATG cassette. Excision of the cassette would remove the ATG start codon required for translation of the truncated hygromycin resistance gene, and therefore restoring hygromycin sensitivity to the ES cells, allowing positive selection with hygromycin in the subsequent step.

3. Materials and Methods

3.1 Materials

3.1.1 List of culture media used

LB medium	10 g/L Tryptone (BD Biosciences; 211705), 5 g/L Yeast Extract (BD Biosciences; 212750), 5 g/L NaCl (Affymetrix; J21618), heat sterilization, and stored at 4 °C.
LB agar	40 g/L Difco LB Agar Miller (BD Bioscience; 244510), heat sterilization and stored at 4C.
2i/LIF medium	1 x Dulbecco's Modified Eagle Medium (Gibco; 12800-082), 15 % Fetal Bovine Serum (Hyclone; SH30071.03), 1 % Penicillin/Streptomycin (Gibco; 15140), 1 % sodium bicarbonate, 25 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (Gibco; 15630), 1 x nonessential amino acid (Gibco; 11140), 0.0008 % 2-mercaptoethanol (Sigma; M6250), 1,000 U/mL leukemia inhibitory factor (LIF; Merck; ESG1107), 3 µM CHIR99021 (Miltenyi Biotec; 130-095-555) and 1 µM PD0325901 (Miltenyi Biotec; 130-095-557), filter sterilization, stored at 4 °C and protected from light.

3.1.2 List of primers/oligos used for generating repair template and CRISPR-Cas9 plasmids

Name	Purpose	Sequence
HA19-1_F	Amplifying the left homologous arm from mouse gDNA.	5'- GTGTTGATCTTGAAGCACGCAGAACAG -3'
HA19-1_R		5'- TTACTAGGGACAGGATTGGTAAGAA GGGAGTTCTAAGATCCCAGACTAGC -3'
TRE_F	Amplifying TRE promoter from plasmid TRE-Tight-EGFP-backward donor.	5'- ACCAATCCTGTCCCTAGTAAAGCTTAGTACTG -3'
TRE_R		5'- ATGCAACTTCGTATAATGTATGCTAT ACGAAGTTATGGCGATCTGACGGTTCACTAAACG -3'
Hygro_F	Amplifying hygromycin resistance gene from plasmid pSUPER-hygro-sv40; primer was designed to include the loxP sequence such that it would be included in the repair template plasmid following Gibson Assembly.	5'- ATAACTTCGTATAGCATACATTATACGAA GTTGCATAAAAAGCCTGAACTCACCGCGAC -3'
Hygro-HA19-2_R		5'- CACTTTTCCTGTGTGTATTACAGACATG ATAAGATACATTGATGAGTTTGGACAAACC -3'

HA19-2_F	Amplifying the right homologous arm from mouse gDNA.	5’- TAATACACACAGGAAAAGTGATAAAGCAGCCTG -3’
HA19-2_R		5’- AGCTAGAGGTCGACGGTATACGGACA CTCTTTCTCTTTTTCCATTCCTTCA -3’
Ampbb_F	Amplifying vector backbone containing ampicillin resistance gene from plasmid pSUPER-hygro-sv40.	5’- TATACCGTCGACCTCTAGCTAGAGCTTG -3’
Ampbb-HA19-1_R		5’- GCGTGCTTCAAGATCAACACTAAAGCCT GGGGTGCCTAATGAGT -3’
PGK_F	Amplifying PGK-gb2 promoter from the loxP-PGK-gb2-hygro-loxP template (Genebridges; A011); primer was designed to include the ATG start codon and loxP sequence such that it would be included in the repair template plasmid following Gibson Assembly.	5’- CTACCGGGTAGGGGAGGC -3’
PGK-ATG-loxP_R		5’- TTCAGGCTTTTTATGCAACTTCGTATAAT GTATGCTATACGAAGTTATCATGGTTTAG TTCCTCACCTTGTCGTATTATACTATGC -3’
TRE-PGK_R	Combined with Hygro_F to amplify the vector backbone for the revised repair template plasmid from the original repair template plasmid generated.	5’- GCGCCTCCCCTACCCGGTAGATGCAAC TTCGTATAATGTATGCTATACGAAGTTATG GCGATCTGACGGTTCACTAAACG -3’

19.1_oligo 1	For cloning of the guide sequence into the sgRNA scaffold for CRISPR-Cas9 plasmid 1.	5’- CACCGCAATCAGATCTCACCCATTG -3’
19.1_oligo 2		5’- AAACCAATGGGTGAGATCTGATTGC -3’
19.2_oligo 1	For cloning of the guide sequence into the sgRNA scaffold for CRISPR-Cas9 plasmid 2.	5’- CACCGAGGGATTCAACTCTCAATAC -3’
19.2_oligo 2		5’- AAACAGGGATTCAACTCTCAATACC -3’

3.1.3 List of primers used for PCR genotyping

Name	Purpose	Sequence
19com1_F	Validating the insertion on the left homologous arm. An unmodified allele would result in the pairing of 19com1_F and 19wt1_R to generate a 522 bp band. A modified allele would result in the pairing of 19com1_F and mut1_R to generate an 823 bp band.	5'- TAGCCTCTAATGGCGGAGGC -3'
19wt1_R		5'- CCAAGCCAATCAGATCTCACCCATTG -3'
mut1_R		5'- CTATACGAAGTTATGGCGATCTGACGGTTC -3'
19com2_R	Validating the insertion on the right homologous arm. An unmodified allele would result in the pairing of 19com2_R and 19wt2_F to generate a 483 bp band. A modified allele would result in the pairing of 19com2_R and mut2_F to generate a 625 bp band.	5'- AAAAGATTAGACAGCACTGAACTGCATA GTAAACAA -3'
19wt2_F		5'- GGAAAAAGCTAAATAAGCACATTTTT CAATGATATTATTTCAGAAGA -3'
mut2_F		5'- CGCCGGCTGGATGATCCT -3'
cre_F	Validating the excision of the PGK-gb2-ATG cassette. An unmodified allele would result in the generation of a longer band of 1064 bp. A modified allele would result in a smaller 438 bp band.	5'- GCTAGTCTGGGATCTTAGAACTCCCTTCTT -3'
cre_R		5' – GTCGCGGTGAGTTCAGGCTTTTT -3'

NH-E1-0_R	Validating the left arm insertion of plasmid iXist-3.1. When paired with cre_F, a modified allele would generate a 533 bp band. An unmodified allele would not result in any band.	5’- CCCCCAAAGCTCCTTAGATGG -3’
loxcheck_F	Validating the right arm insertion of plasmid iXist-3.1. A modified allele would result in the pairing of loxcheck_F and loxcheck_R to generate a 508 bp band. An unmodified allele would not result in any band.	5’- GGCCACCATGGTGTCGATAAC -3’
loxcheck_R		5’- CGCGCATATGAAATCACGCC -3’

All primers were synthesised by Integrated DNA Technologies (USA) with standard 5’ and 3’ modification and standard desalting. Due to the length of these primers, polyacrylamide gel electrophoresis (PAGE) purification was selected for PGK-ATG-loxP_R and TRE-PGK_R to increase the purity of the yield.

3.1.4 List of plasmids used or generated

Name (Source)	Description
pSUPER-hygro-sv40	A plasmid containing hygromycin and ampicillin resistance gene. Used as template for PCR amplification of the hygromycin resistance gene and vector backbone.
TRE-Tight-EGFP-backward donor	A plasmid containing TRE promoter and gene encoding the enhanced green fluorescent protein. Used as template for PCR amplification of the TRE promoter.
pX330 (Addgene; #42230)	A CRISPR-Cas9 system plasmid carrying humanized SpCas9 gene and chimeric guide RNA expression cassette.
iXist-3.1 (R. Chelliah)	A plasmid containing full length Xist gene, loxP site, ATG start codon, and tdTomato reporter.
loxP-PGK-gb2-hygro-loxP (Genebridges; A011)	A plasmid encoding the hygromycin resistance gene which combines a prokaryotic promoter (gb2) with a eukaryotic promoter (PGK). Used as template for PCR amplification of the PGK-gb2 promoter cassette.
pSALK Cre	A Cre recombinase expression plasmid, to be used for Cre-loxP mediated excision and integration.

3.2 Methods

3.2.1 PCR and gel extraction

PCR amplification of the respective fragments were carried out using Herculase II Fusion DNA Polymerase (Agilent Technologies; 600675) in Applied Biosystems Veriti 96-Well Thermal Cycler. The annealing temperature used was -2°C the T_m of the primer pairs used in the amplification reaction of each fragment.

The amplified fragments were separated on 1% agarose gels for gel extraction using the QIAquick Gel Extraction Kit (QIAGEN; 28704). Extracted products were then purified using the QIAquick PCR Purification Kit (QIAGEN; 28104).

3.2.2 Gibson Assembly, transformation into bacteria and miniprep

Gibson Assembly (NEB; E2611) was performed to join the four DNA insert fragments with the vector backbone. In the case of generating the revised versions of the plasmid, the loxP-PGk-gb2-ATG cassette as a single fragment was ligated into a vector backbone that was amplified from the original version of the plasmid. The amount of fragment backbone was optimized at 50 ng, and 2-3 fold of excess insert fragments in pmols were used. The mixture was incubated at 50°C for 60 min. NEB 5-alpha Competent E.coli (NEB; C2987) were chemically transformed with 2 µl of the master mix/fragment mixture: heat shock at 42°C for 30 sec, snap chilled on ice for 2 min, 500 µl of room temperature LB media added to mixture and incubated at 37°C for 60 min, shaking vigorously (250 rpm). Following this, varying dilutions of the cell suspension were spread onto LB agar plates supplemented with 100 µg/ml ampicillin and incubated at 37°C overnight. The next morning, surviving colonies were picked and inoculated into tubes containing 5ml LB media supplemented with 100 µg/ml ampicillin. After shaking incubation of at least 16 hours, 500 µl of the bacterial culture

was mixed with equal portion of 80% glycerol to create a bacteria glycerol stock to be kept at -80°C, before the bacterial cells were harvested by centrifugation and the constructed plasmid was isolated using the QIAGEN Plasmid Mini Kit (QIAGEN; 12123).

3.2.3 Restriction enzyme digest validation, sequencing, and midiprep

1 µg of the repair template plasmid generated was digested with 1 µl of PstI in 1 µl of NEBuffer 3.1 (NEB; B7203S), topped up to 20 µl with ddH_2O. Digested products were separated on a 1% agarose gel, and the sample was sent for further sequencing by AITbiotech when the expected band pattern was observed. The samples which had passed both the restriction enzyme digest and sequencing validation were retrieved from their respective bacteria glycerol stocks and inoculated in 5 ml LB medium supplemented with 100 µg/ml ampicillin. After shaking incubation of at least 8 hours, 50 µl of the starter culture was aliquoted into a conical flask containing 100 ml LB media supplemented with 100 µg/ml ampicillin, and then placed in a shaking incubator overnight. The overnight bacterial culture was then harvested by centrifugation at 6000 x g for 15 min at 4°C and the constructed plasmid was isolated using the QIAGEN Plasmid Midi Kit (QIAGEN; 12143). The amount of plasmid obtained was quantified with Thermo Scientific NanoDrop spectrophotometer before restriction enzyme digest and sequencing validation once more to confirm the desired plasmid sequence.

3.2.4 CRISPR-Cas9 plasmid design

The CRISPR-Cas9 plasmid constructs were generated to target the intronic region between exon 1 and 2 of the Cd6 locus. Single gRNAs were generated using the online CRISPR design software provided by the Zhang lab (http://crispr.mit.edu:8079/). The tool identifies the most suitable gRNA target

sequences and also checks for off-target binding simply from the sequence input by the user. Two pairs of oligos were synthesized from choosing the top two most suitable gRNA target sequences identified by the tool. Two such plasmids were generated so as to increase the efficiency of DSBs induced at the target locus following transfection. The oligos were annealed respectively, phosphorylated by 5 U T4 PNK (NEB; M0201S), and then ligated by Quick Ligase (NEB; M2200S) into BbsI-linearized pX330 (Cong et al., 2013). The ligation reaction was treated with Plasmid-Safe^{TM} ATP-Dependent DNase (Lucigen; E3101K) to prevent unwanted recombination products, before the entire reaction mixture was transformed into competent *E. coli* cells in a similar fashion as described previously. The two CRISPR-Cas9 plasmids subsequently isolated by miniprep and sequence verified using the U6 promoter primer were named 19.1 and 19.2.

3.2.5 ES cell line and culture

Ainv15 ES cells were cultured on a feeder layer of Drug Resistant 4 Mouse Embryonic Fibroblasts (DRF-MEF) (Applied StemCell; ASF-1002) in 2i/LIF medium at 37°C and 5% CO_2. Medium was refreshed daily and the cells were passaged before they reached 80% confluency.

3.2.6 Transfection with CRISPR-Cas9 and repair template plasmids, hygromycin selection, and ES single colony picking

A total of 4 µg of plasmid DNA was transfected into 8 x 10^5 Ainv15 ES cells in suspension using Lipofectamine2000 (Invitrogen; 11668019): 0.5 µg of each CRISPR-Cas9 plasmid, and 3 µg of the repair template plasmid. 24 hours post transfection, selection was initiated using medium supplemented with 300 µg/ml hygromycin (ThermoFisher; 10687010). Following at least 14 days of hygromycin selection, the surviving ES colonies were trypsinised into single cells. Cells were

counted and 5 x 10^3 were seeded on a feeder layer of DRF-MEF in a 6 cm tissue culture dish. Single ES colonies took between 4-6 days to grow to a sufficiently large size for picking into a 96 well plate, and colonies that had overlapped with neighbouring ones were avoided to achieve clonal populations.

3.2.7 gDNA extraction and PCR genotyping (1)

Once the picked ES colonies had reached 80-90% confluence, each was split, with one half being sustained in culture for gDNA extraction while the other half was frozen down to be retrieved later should that particular clone pass the PCR genotyping. Once the cells set aside to culture for gDNA extraction were confluent, they were washed with 1 x PBS, before 100 µl of lysis buffer [0.1M NaCl, 0.05M Tris (1st BASE; BIO-1400), 0.015M EDTA (Affymetrix; 6381-92-6), and 1% SDS (Hoefer; 03-500-509) adjusted to pH 8] freshly supplemented with 500 µg/ml Proteinase K (Promega; V3021) was added to each well and the plate was incubated at 55°C overnight. The next morning, 100 µl of isopropanol was added to each well and the plate was shaken at 250 rpm on a horizontal shaker at room temperature for 4 hours. Following centrifugation at 3,000 g at 4°C for 5 min, the DNA precipitated in each well was washed twice with 75% ethanol. The plate was air-dried at room temperature for 1 hour and 75 µl TE buffer was added to the DNA pellets. The plate was left at 4°C overnight to allow the pellets to dissolve. The next morning, the gDNA in each well was quantified using Thermo Scientific NanoDrop spectrophotometer before PCR genotyping.

A three-primer system was used to genotype the clones as described in 3.1.3 List of primers used for PCR genotyping. The first three-primer set consists of 19com1_F, 19wt1-R, and mut1_R. The second three-primer set consists of 19com2_R, 19wt2_F, and mut2_F. Colonies that showed the desired band pattern, and therefore

genotype, were retrieved from frozen culture and expanded. A more thorough protocol [detailed in the following paragraph] to isolate pure gDNA from the retrieved and expanded clones was performed, to ascertain the genotypes of the clones again.

Cells were lysed with 700 µl lysis buffer freshly supplemented with 500 µg/ml Proteinase K, and an equal volume of UltraPure^{TM} Phenol:Chloroform:Isoamyl Alcohol (ThermoFisher; 15593031) was added. The tube was vortexed strongly until a homogenous mixture was obtained. The mixture was then centrifuged at 14,000 rpm for 5 min and 300 µl of the upper aqueous layer was carefully transferred into fresh Eppendorf tubes; the bottom layer of phenol chloroform as well as the middle layer of protein shreds were avoided. gDNA was precipitated in absolute ethanol and stored at -20°C overnight, and centrifuged at 12.5 k/min at 4°C for 30 min the next day. The DNA pellets were washed with 70% ethanol twice and then dissolved in water before quantification with Thermo Scientific NanoDrop spectrophotometer, to be used for PCR genotyping.

3.2.8 Cre-loxP mediated excision and ES single colony picking

1 µg of pSALK Cre plasmid was transfected into 8 x 10^5 cells of the selected clones in suspension using Lipofectamine2000 (Invitrogen; 11668019). The transfected cells were trypsinised 48 hours later into single cells and 5 x 10^3 cells were seeded on a feeder layer of DRF-MEF in a 6 cm tissue culture dish. Single ES colonies took between 4-6 days to grow to a sufficiently large size for picking in a 96 well plate, and colonies that had overlapped with neighbouring ones were avoided to achieve clonal populations. As no positive selection was used in this stage, more colonies were picked to increase the odds of selecting a colony that had the PGK-gb2-ATG cassette excised.

3.2.9 gDNA extraction and PCR genotyping (2)

gDNA extraction was performed as described above. A two primer system was used to genotype the clones as described in 3.1.3 List of primers used for PCR genotyping. The two primers in the system were cre_F and cre_R. Colonies that showed the desired band pattern, and therefore genotype, were retrieved from frozen culture and expanded. A more thorough protocol, as described earlier, to isolate pure gDNA from the retrieved and expanded clones was performed, to ascertain the genotypes of the clones again.

3.2.10 Cre-loxP mediated transgene insertion, hygromycin selection, ES colony picking

A total of 11 µg of plasmid DNA was transfected into 1.6 x 10^6 cells of the selected clones in suspension using Lipofectamine2000 (Invitrogen; 11668019): 2.5 µg of pSALK CRE plasmid, and 8.5 µg of iXist-3.1 plasmid. 24 hours post transfection, selection was initiated using medium supplemented with 300 µg/ml hygromycin (ThermoFisher; 10687010). Following at least 14 days of hygromycin selection, the surviving ES colonies were observed under a fluorescence microscope. Colonies that were fluorescing red were marked and picked into a 24 well plate for expansion.

3.2.11 gDNA extraction and PCR genotyping (3)

The more thorough gDNA extraction protocol was performed as described earlier. Two primer pairs were used to genotype the clones as described in 3.1.3 List of primers used for PCR genotyping. The primers in the first pair were cre_F and NH-E1-0_R. The primers in the second pair were loxcheck_F and loxcheck_R. Colonies that showed the desired band pattern and therefore genotype, were retrieved from frozen culture and expanded, to be induced with doxycycline (Dox).

3.2.12 Dox induction and slide preparation

The retrieved cell lines were grown in 6 well plates pre-treated with 0.2% gelatin (Sigma-Aldrich; G2500) and treated with 1 µg/mL of Dox (Clontech; 631311). Following 48 hours of Dox induction, the ES cells were harvested by trypsinisation and resuspended in a concentration of 8 x 10^5 cells per ml in 1 x PBS. The cells suspended were then cytospun onto Superfrost/Plus microscope slides (Fisher Scientific; Cat# 12-550-15) using cytofunnels in a Shandon Cytospin 4 centrifuge (Fisher Scientific) at 800 rpm for 10 min. The slides were air dried and rinsed in ice cold 1 x PBS for 5 min, before being fixed in 4% paraformaldehyde at room temperature for 10 min. After fixation, the slides were stored in 70% ethanol at 4ºC until use.

3.2.13 RNA FISH

Mouse Xist RNA was detected with Sx9 probe, a PI DNA construct containing a 40 kb genomic fragment covering the mouse Xic region. The nucleotide analog used in the probe labelling was Cy3-dUTP (Amersham; Cat# PA53022) and the probe was prepared using nick translation kit (Roche; Cat# 10976776001). The probe was denatured at 75ºC for 10 min and pre-annealed at 42ºC for 10 min before being applied onto slides for hybridization.

Slides were dehydrated through 80%, 90%, and 100% ethanol sequentially for 2 min each and air dried for 3 min. Denatured probe was added onto slides and hybridization was carried out for 3 h at 42ºC in a dark and humidified environment. After hybridization, slides were washed three times in 50% formamide, 2xSSC at 45ºC for 5 min each, followed by three washes in 2xSSC at 45ºC for 5 min each. The washing steps were done in a shaker-waterbath with slight agitation. Slides were rinsed in 1 x PBS plus 0.2% Tween-20 for 3 mins before being mounted with Dapi-containing mounting medium with antifade (Vector Laboratories Inc.; Cat#H1200).

Fluorescence images were collected on Eclipse Ti microscope (Nikon) using digital camera Clara Series model C01 (Andor) and NIS Elements AR imaging software (Nikon).

4. Results

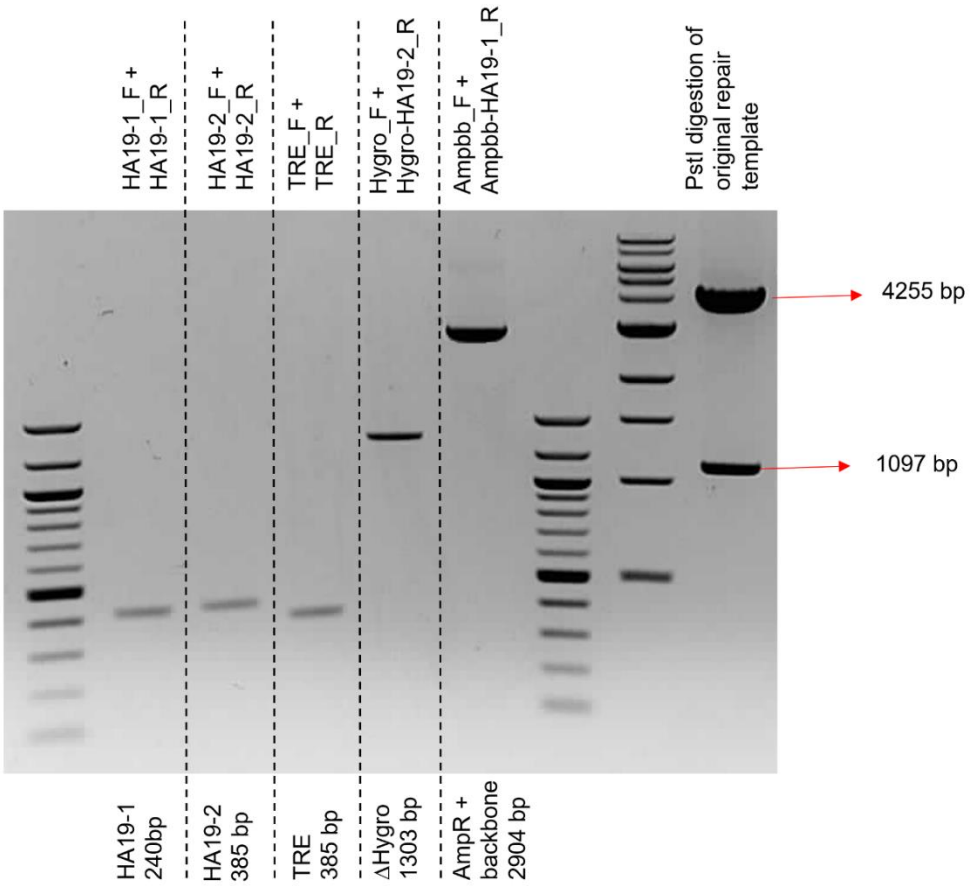
4.1 Construction of the repair template plasmids

As discussed in Section 2.1 Experiment outline and workflow, two repair template plasmids were designed: an original template with the insertion sequence carrying a TRE, loxP site, and ΔHygro gene, flanked by a pair of homologous arms, and the revised template which inserted a loxP-PGK-gb2-ATG cassette between the TRE and loxP site of the original template.

After the individual fragments required for the construction of the plasmids were PCR amplified, the products were extracted, purified, and verified by gel electrophoresis. The band sizes were as expected for each fragment as shown in Figure 4A & B. After Gibson Assembly, the repair templates were validated through restriction enzyme digestion. The band patterns observed were as expected (Figure 4A & B), following which, the bacteria glycerol stocks from which the correct repair template plasmids were retrieved for culture and bulk plasmid production.

The harvested repair template plasmids produced were subsequently co-transfected with the CRISPR-Cas9 plasmids into Ainv15 ES cells.

A


HA19-1_F + HA19-1_R
HA19-2_F + HA19-2_R
TRE_F + TRE_R
Hygro_F + Hygro-HA19-2_R
Ampbb_F + Ampbb-HA19-1_R
PstI digestion of original repair template
4255 bp
1097 bp
HA19-1 240bp
HA19-2 385 bp
TRE 385 bp
ΔHygro 1303 bp
AmpR + backbone 2904 bp

B

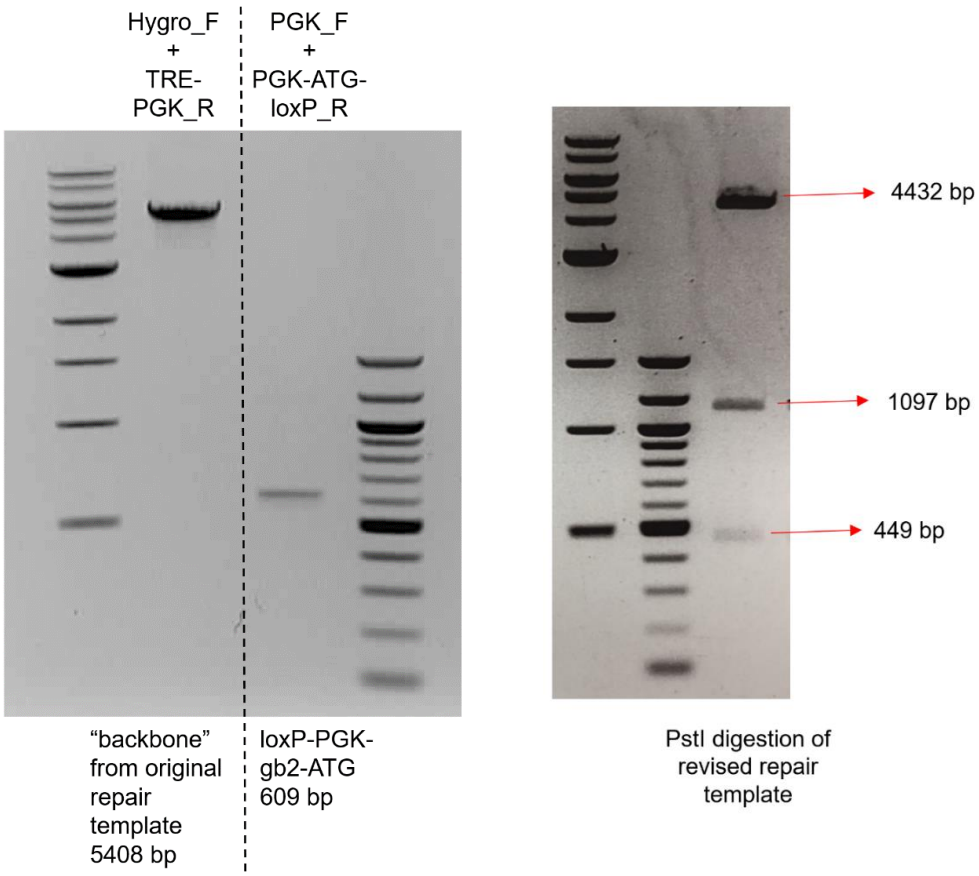

Hygro_F + TRE-PGK_R
PGK_F + PGK-ATG-loxP_R
4432 bp
1097 bp
449 bp
"backbone" from original repair template 5408 bp
loxP-PGK-gb2-ATG 609 bp
PstI digestion of revised repair template

Figure 4: Generation of the repair template plasmids.

(**A**) For each of the PCR amplified fragments, the band size following gel electrophoresis was as expected; sizes of each fragment are stated within the figure. The band pattern observed following PstI digestion of the original repair template was as expected as well, with two fragments generated at 4255bp and 1097 bp. The names of the primers used for amplifying the fragments are listed above the gel picture, while the respective name of the fragment generated and its size is listed below the gel picture.

(**B**) The band size of the vector backbone amplified from the original repair template plasmid and the PGK-gb2-ATG cassette, following gel electrophoresis was as expected. The band pattern observed following PstI digestion of the revised repair template was as expected, with three fragments generated at 4432 bp, 1097 bp, and 449 bp. The names of the primers used for amplifying the fragments are listed above the gel picture, while the respective name of the fragment generated and its size is listed below the gel picture.

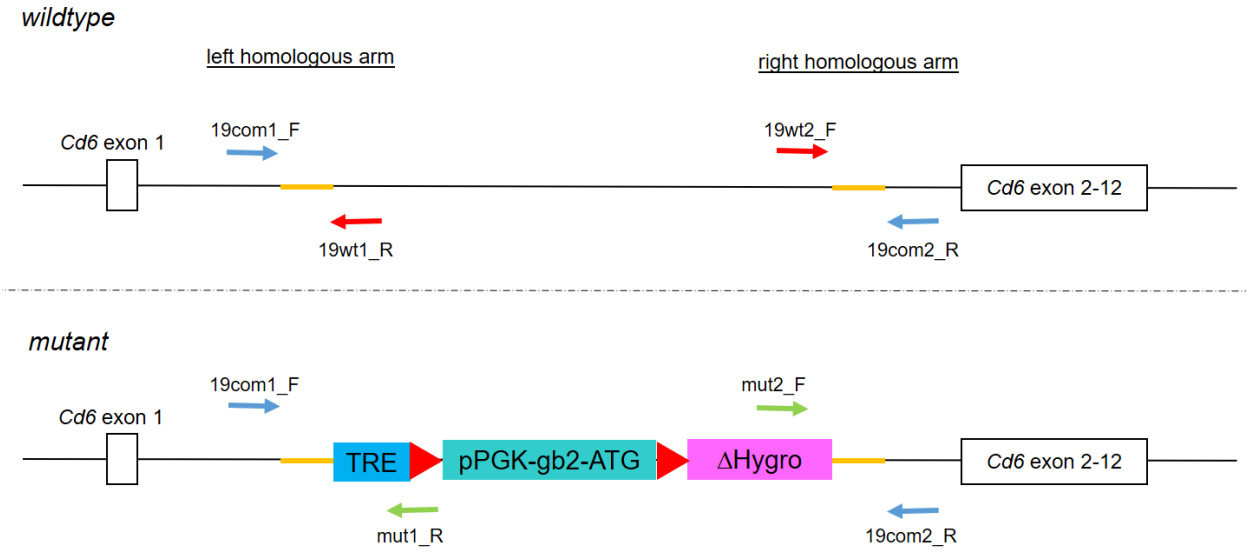
4.2 Genotyping results after transfection with CRISPR-Cas9 and repair template plasmids

As mentioned in Section 2.1 Experiment outline and workflow, the original repair template plasmid did not contain any positive selection, and therefore in an attempt to obtain at least a handful of positive hits, a large number of individual colonies were picked from the pool of cells following transfection with CRISPR-Cas9 and the original repair template plasmids. 96 individual colonies were picked, but unfortunately following PCR genotyping with the gDNA extracted from each of the 96 clones, none were identified to have the desired genotype. Hence these clones were discarded, and the revised repair template plasmid was designed (and produced as described in the Section 4.1 Construction of the repair template plasmids) to introduce positive selection with hygromycin.

For the transfection involving the revised repair template plasmid, 24 individual colonies were picked after surviving ES colonies were trypsinised into single cells and seeded following 14 days of hygromycin selection. PCR genotyping was performed using the gDNA extracted from each of the 24 clones, and 4 clones were identified to have the desired genotype.

After retrieving these 4 clones from frozen culture, clone E3 was selected to proceed with subsequent steps, given its optimal growth rate and clear genotyping pattern. Given that both the wildtype and mutant band can be observed, we can conclude that one chromosome 19 had been modified to contain the desired insertion, while the other remained unmodified (Figure 5).

A



B

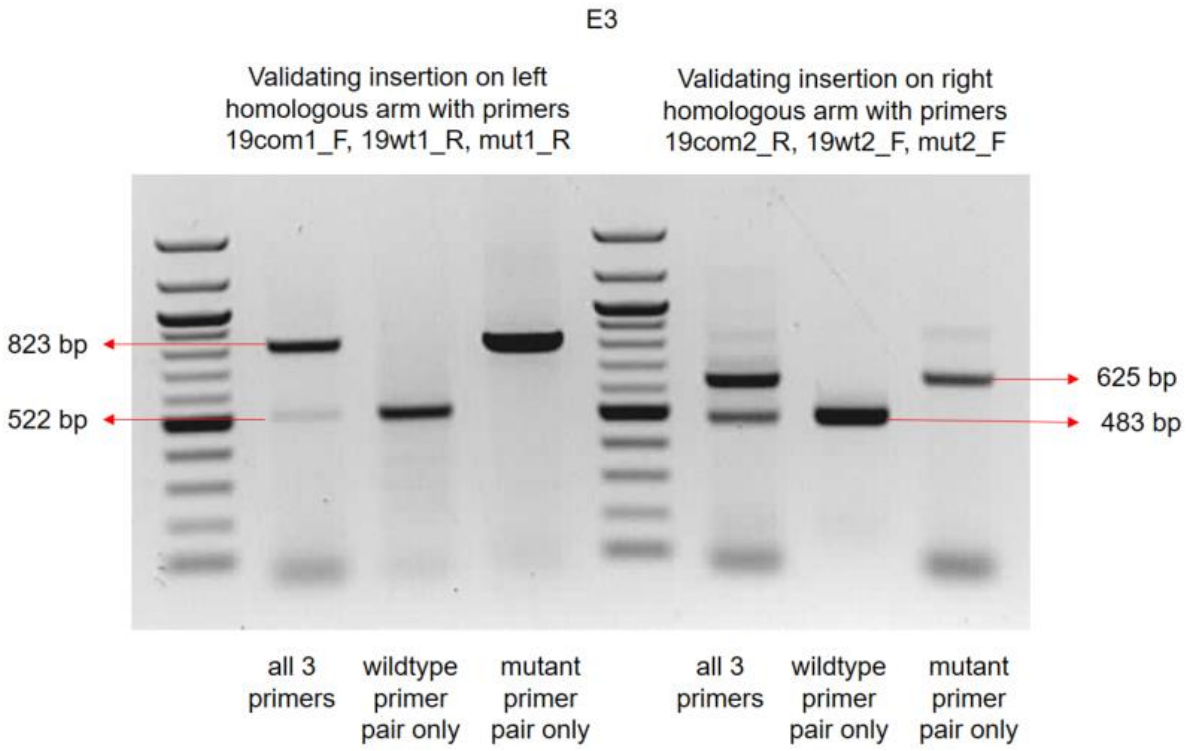


Figure 5: Genotyping primer design and genotyping results of clone E3 after transfection with CRISPR-Cas9 and repair template plasmids.

(**A**) Schematic drawing of the genotyping primer design for validation of the insertion on the left and right homologous arms. For the left homologous arm, primer 19com1_F can pair with either 19wt1_R or mut1_R to generate bands of 522 bp and 823 bp respectively. Similarly, for the right homologous arm, primer 19com2_R can pair with either 19wt2_F or mut2_F to generate bands of 483 bp and 625 bp respectively. The bands produced would indicate the genotype of the cell from which the gDNA was extracted from.

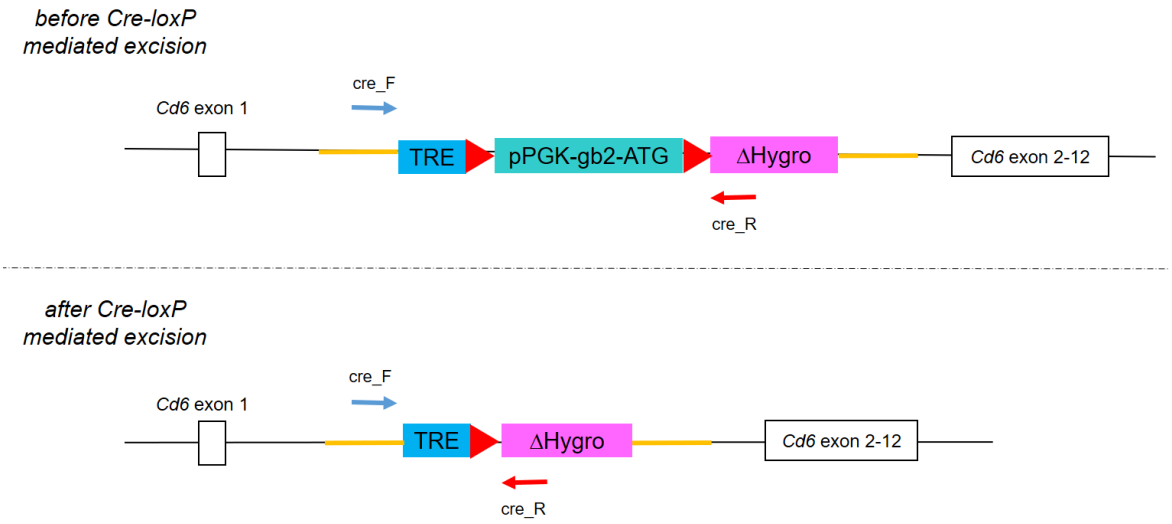
(**B**) When all 3 primers were included in the genotyping PCR, two bands each were observed in the validation of the left and right homologous arms. For the left homologous arm, the wildtype and mutant bands of 522 bp and 823 bp were observed respectively. For the right homologous arm, the wildtype and mutant bands of 483 bp and 625 bp were observed respectively. When separated into their respective primer pairs (either wildtype or mutant pair only), the same products were observed as well.

4.3 Pooled gDNA PCR genotyping results after Cre-loxP mediated excision

Following expansion, clone E3 was transfected with the pSALK Cre plasmid and a pooled gDNA was extracted from the cells, before trypsinisation for single colony picking. As expected, the genotyping PCR using pooled gDNA before Cre-loxP mediated excision generated a larger band, corresponding to the fact that the PGK-gb2-ATG cassette was still present. A smaller band was observed when pooled gDNA after Cre-loxP mediated excision was used, indicating successful excision of the PGK-gb2-ATG cassette.

With this confirmation that the desired ES cells could be found within the pool of transfected cells, the single cells were seeded and grown to sufficiently large sizes for single colony picking.

A



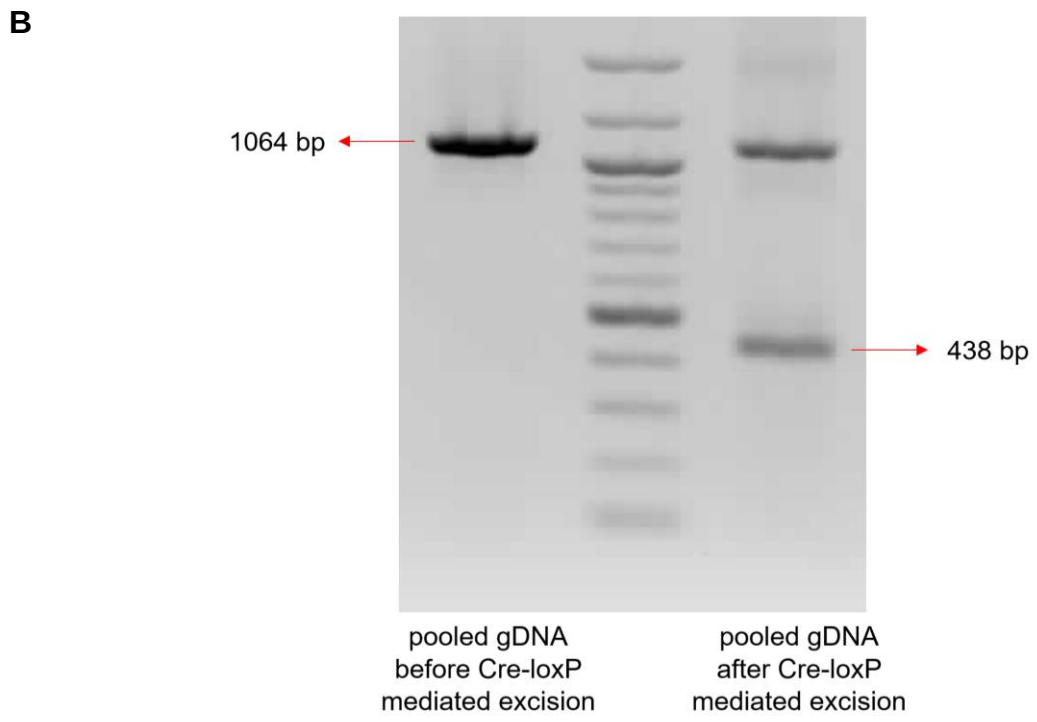


Figure 6: Genotyping primer design and genotyping results of pooled gDNA after pSALK Cre transfection.

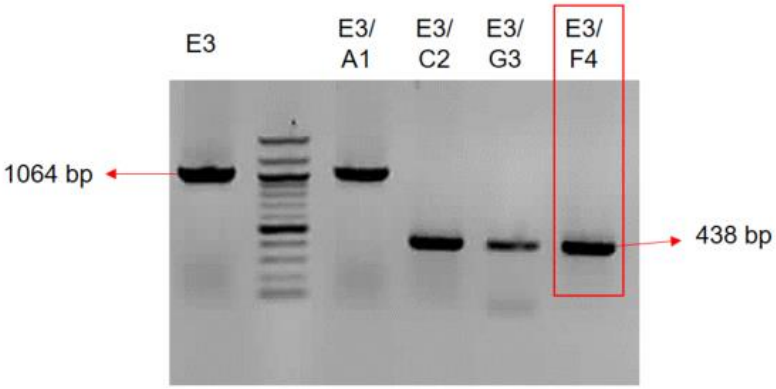
(**A**) Schematic drawing of genotyping primer design for validation of Cre-loxP mediated excision of the PGK-gb2-ATG cassette. Primers cre_F and cre_R can pair on both alleles, before and after Cre-loxP mediated excision. However, when the cassette has been excised as on the allele depicted after Cre-loxP mediated excision, the band produced would be smaller than before. Thereby allowing a differentiation between the two genotypes using the same pair of primers.

(**B**) Results from after Cre-loxP mediated excision showed a smaller band of 438 bp in addition to the larger band of 1064 bp that was also observed in the results from before Cre-loxP mediated excision. This provides evidence that the PGK-gb2-ATG cassette had been successfully excised in a subset of the ES cells.

4.4 Individual clones PCR genotyping results after Cre-loxP mediated excision

4 individual colonies were picked following trypsinisation and seeding of single cells. PCR genotyping was performed using the gDNA extracted from each of the 4 clones, and 3 of the 4 were identified to have the PGK-gb2-ATG cassette excised (Figure 7A). After retrieving the 3 clones from frozen culture, clone E3/F4 was selected to proceed with the subsequent and final step given its optimal growth rate and clear genotyping pattern (Figure 7B).

A



B

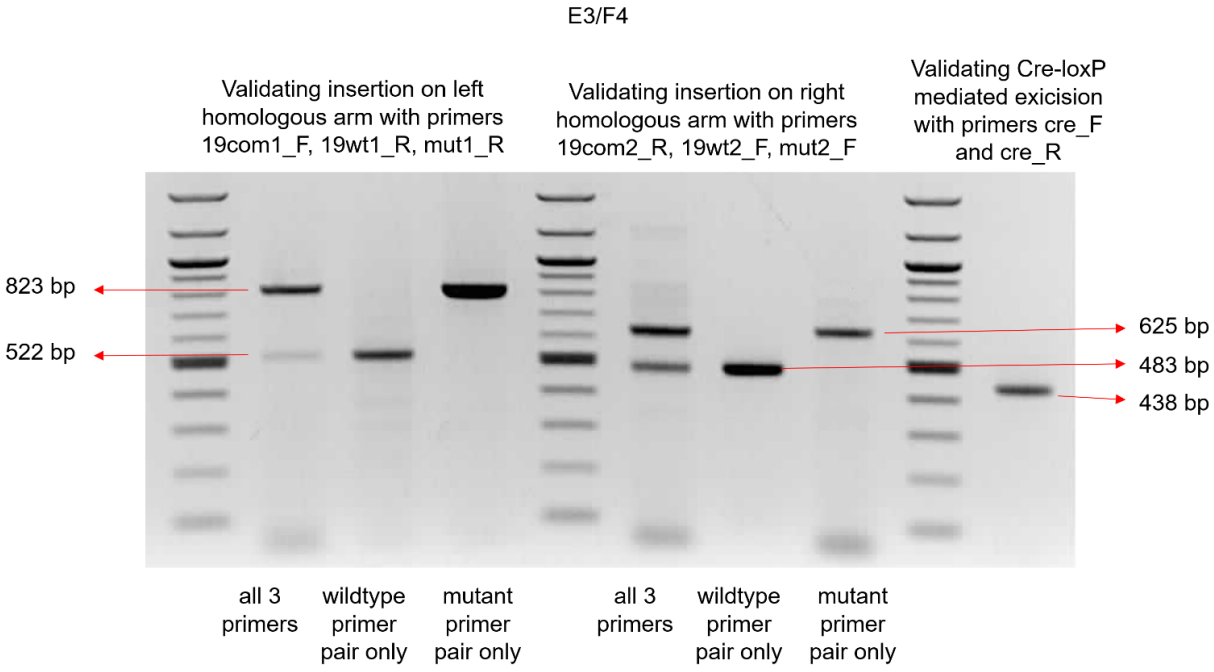


Figure 7: Genotyping results of individual clones after pSALK Cre transfection.

(**A**) Clone E3 produced a band of 1064 bp as expected, and 3 of the 4 clones picked after pSALK Cre transfection produced the smaller band of 438 bp, indicating that the PGK-gb2-ATG cassette had been excised. Primers cre_F and cre_R were used for this genotyping.

(**B**) Genotyping using primers for validating the insertion of the initial modification was repeated using gDNA extracted from clone E3/F4, to ensure that the TRE-loxP-ΔHygro modification was indeed still present.

4.5 Fluorescence of ES colonies after Cre-loxP mediated transgene insertion

CloneE3/F4 was transfected with the pSALK Cre and iXist-3.1 plasmids, and following 14 days of hygromycin selection, the surviving ES colonies observed under a fluorescence microscope appeared red. This fluorescence was as a result of the tdTomato red reporter, indicating that the transgene had been successfully integrated.

Following this, single colonies were to be picked and further validated by PCR genotyping.

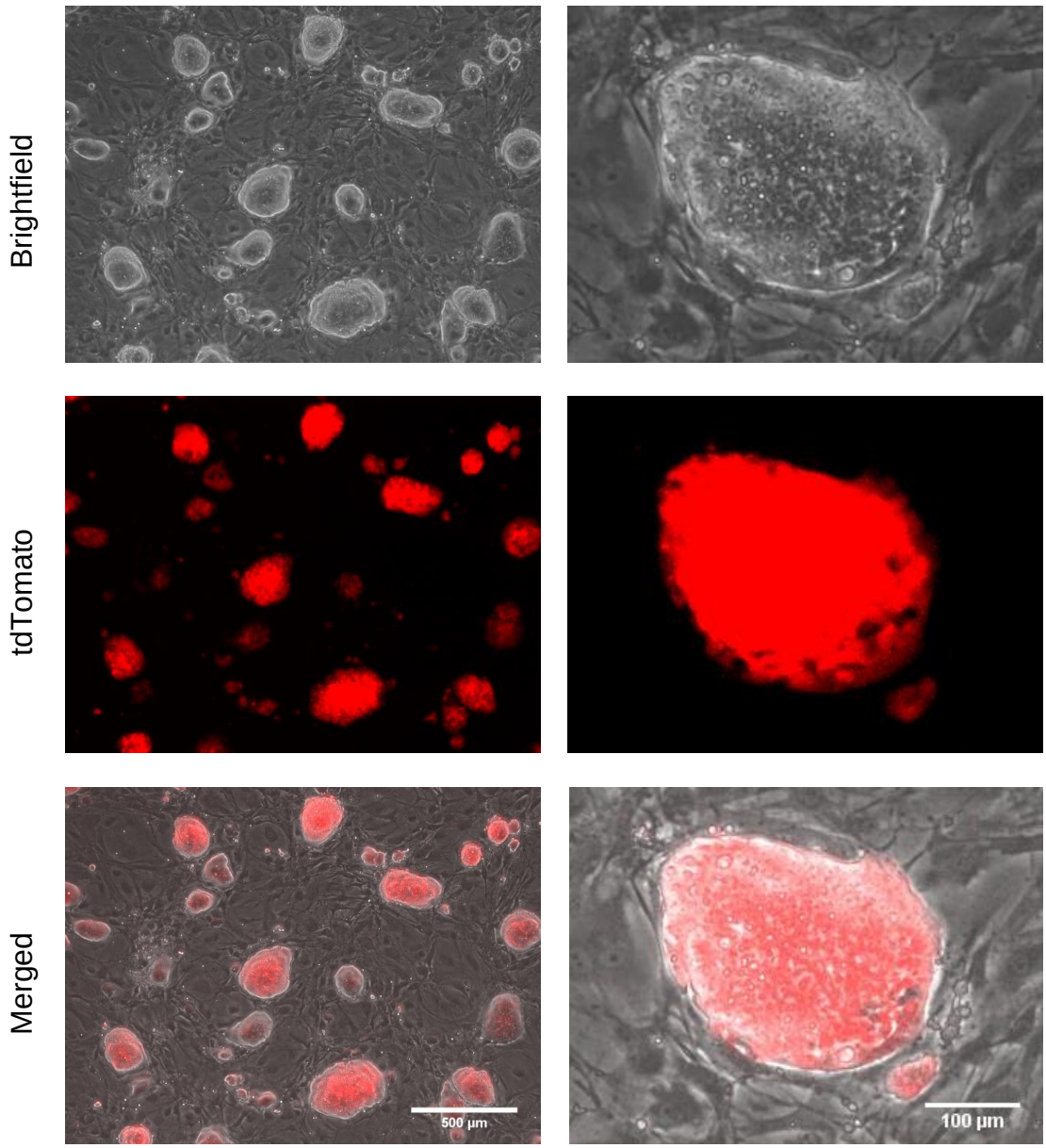


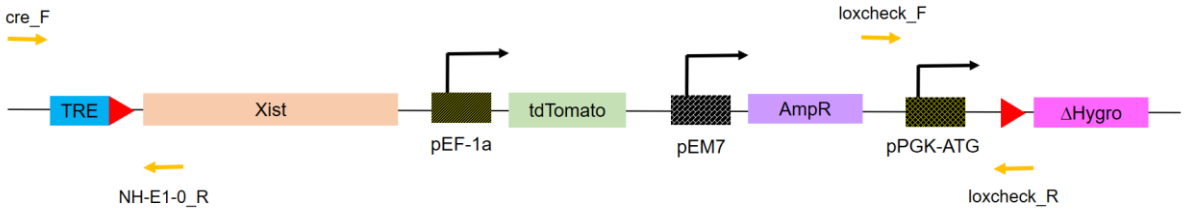
Figure 8: Detection of tdTomato fluorescence in transfected ES colonies.

Images of the colonies captured in brightfield and fluorescence overlapped when merged, signifying the successful integration of the transgene.

4.6 PCR genotyping results after Cre-loxP mediated transgene insertion

6 individual colonies were picked following trypsinisation and seeding of single cells. PCR genotyping was performed using the gDNA extracted from each of the 6 colonies, and all produced the expected band pattern (Figure 9). After retrieving the clones from frozen culture, clones E3/F4/A1 and E3/F4/A2 were selected for expansion and subsequent RNA FISH, as the final validation of transgene insertion.

A



B

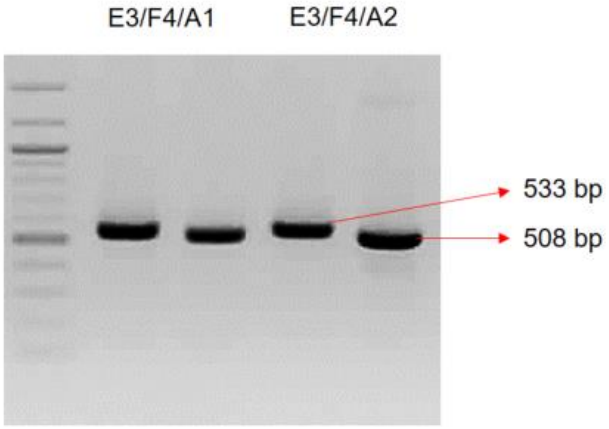


Figure 9: Genotyping results of 2 of 6 individual clones after transfection with pSALK Cre and iXist-3.1 plasmids.

(A) Schematic drawing of genotyping primer design for validation of Xist transgene insertion at the loxP site on chromosome 19. The primer pairs cre_F and NH-E1-0_R, and loxcheck_F and loxcheck_R, validated left and right of the transgene insertion respectively. Given that primers NH-E1-0_R, loxcheck_F and loxcheck_R were specific for the transgene sequence, and the primers would only pair after transgene insertion, bands would only be produced upon successful integration. In a scenario where the transgene had not been integrated, no bands would appear.

(B) Successful transgene insertion was further evidenced by the observed bands from validation of left and right of the transgene insertion.

4.7 RNA FISH results

The visualisation of the Xist cloud using the RNA FISH technique would clearly demonstrate that the transgene had been successfully integrated, and could be expressed upon Dox induction.

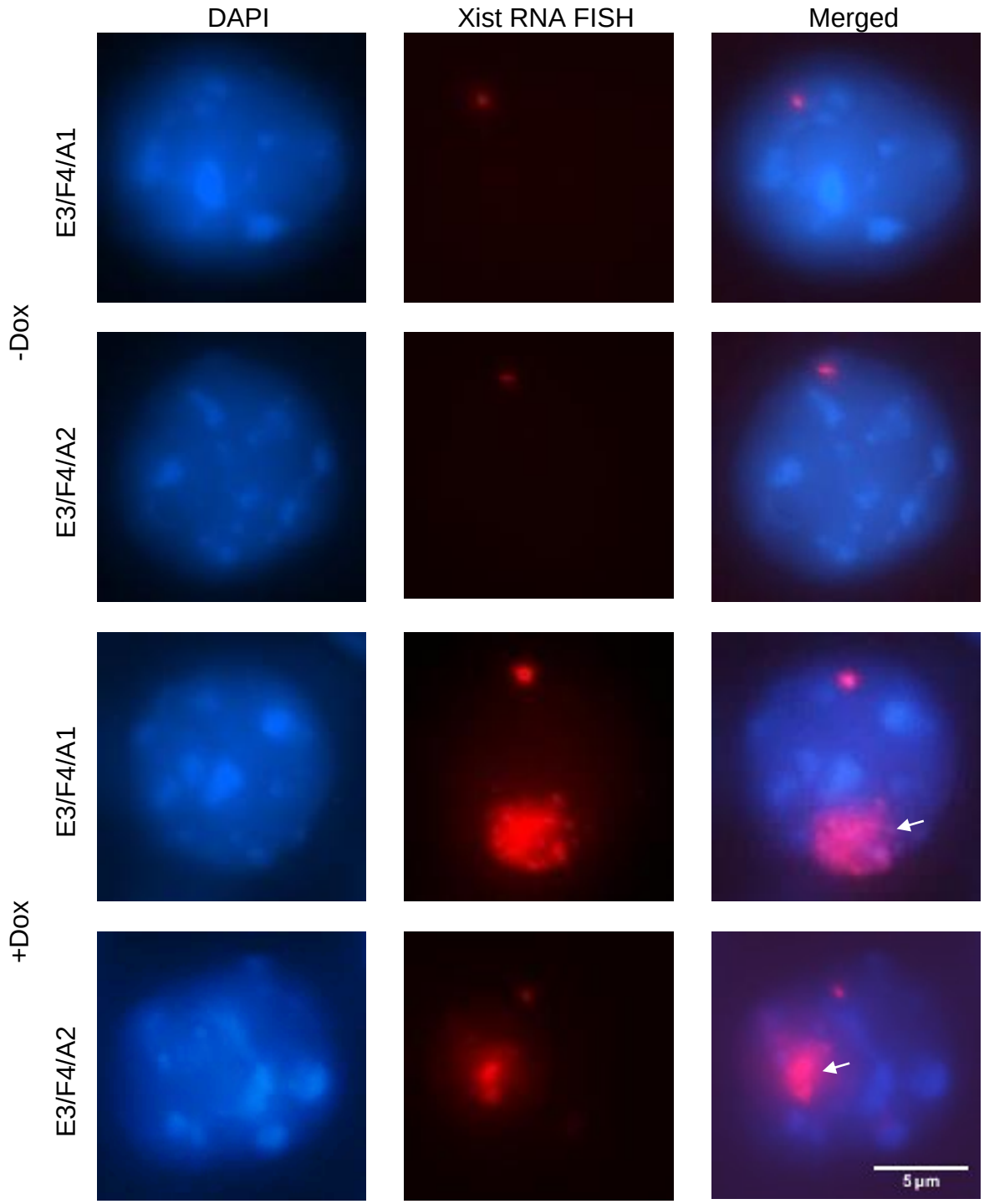
Clones E3/F4/A1 and E3/F4/A2 were induced with Dox for 48 hours, and were harvested to be cytospun onto slides. RNA FISH was performed on the prepared slides and fluorescence images were collected. Slides prepared from cells that were not induced with Dox were made as well to serve as a control.

Two Xist signals were expected to be observed in each cell: a pinpoint signal corresponding to the native Xist gene on the X chromosome, and a “cloud” signal arising from the coating of chromosome 19 with the induced transgenic Xist RNA transcripts.

That Xist clouds could be visualised in individual cells on slides prepared from colonies that had been treated with Dox, demonstrating that the transgene could be expressed. On the other hand, the lack of such an observation on the control slides confirmed that the transgene was induced as a result of Dox treatment. (Figure 10A)

In a field of 30 cells captured, Xist clouds can be visualised for a number of cells (Figure 10B). However, it is noted that not all the cells display the cloud as predicted.

A


DAPI
Xist RNA FISH
Merged
-Dox
+Dox
E3/F4/A1
E3/F4/A2
E3/F4/A1
E3/F4/A2
5 µm

E3/F4/A1

B

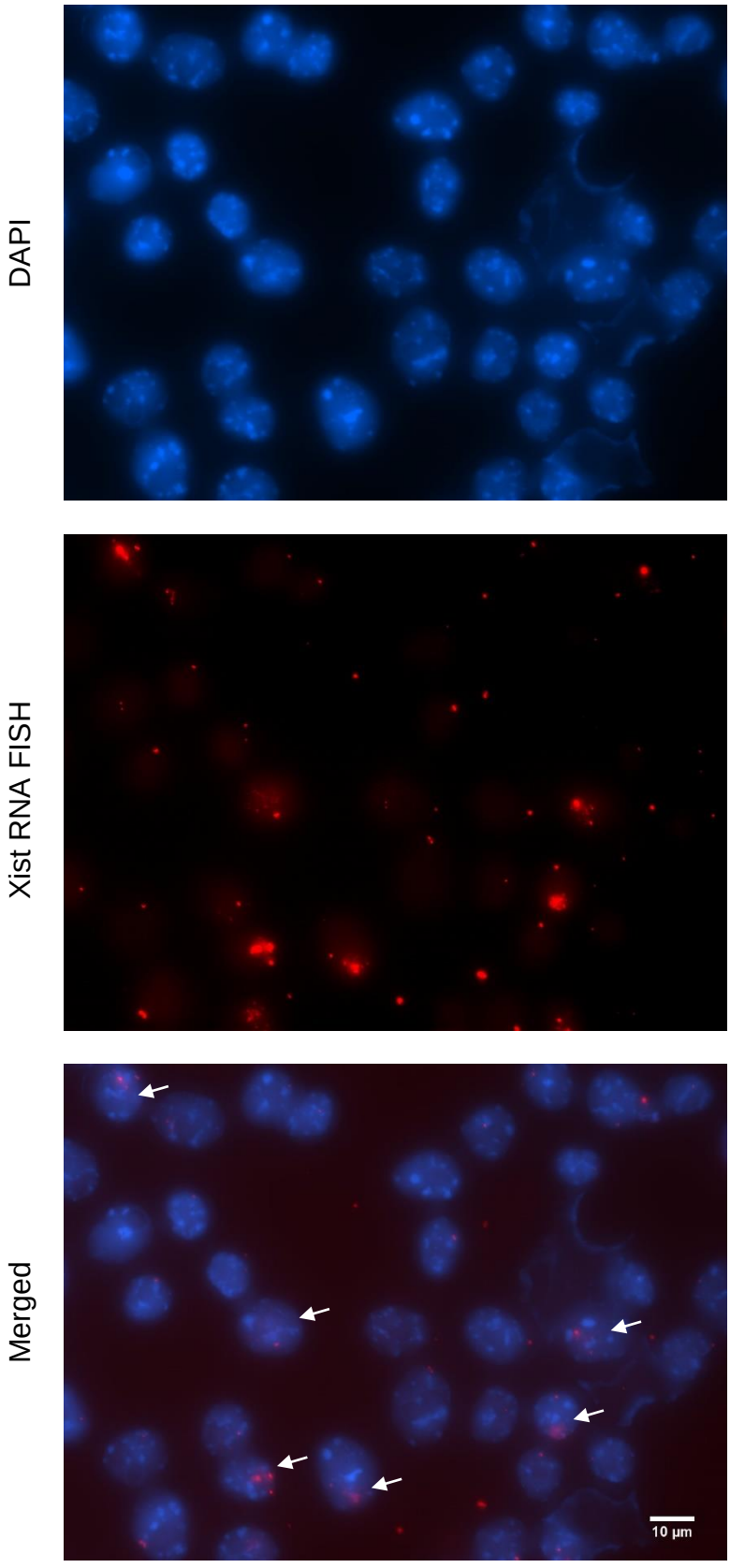


Figure 10: Xist RNA FISH validation.

(**A**) Pinpoint signals can be observed in individual cells on slides prepared from colonies that were not treated with Dox, while Xist clouds can be visualised in addition, in cells from slides prepared from colonies supplemented with Dox. The Xist cloud is highlighted in the respective cells with a white arrow.

(**B**) Xist RNA FISH results from a sample field of 30 cells. Cells in which the Xist cloud can be visualised are highlighted with a white arrow.

5. Discussion and Future Work

Two cell lines were established, namely clones E3/F4/A1 and E3/F4/A2, from the experiments detailed above. Having passed PCR genotyping, displayed tdTomato fluorescence, and visualised Xist clouds upon Dox induction, the objective stated in Section 2 has been achieved.

However it is noted that the Xist cloud was not visualised in all the cells as seen in Figure 10B. In the field of 30 cells sampled, only 6 cells were observed to have the Xist cloud, corresponding to 20% of cells showing the Xist cloud. This was despite the fact that the cells had been treated with Dox for 48 hours, which would be more than sufficient time for the Xist cloud to be induced.

One possibility for this could be that the cell lines established were not clones. However, this explanation is odd given that after hygromycin selection, surviving colonies were trypsinised into single cells before individual colonies were picked from the single cells that had been seeded. Therefore, the colonies that had grown and were picked from this procedure should have theoretically been clones.

Another possibility is the “position effect” on the inserted transgene. Although the chosen gene targeting site is the best possible on chromosome 19 given the supporting literature, it is however not the most commonly used in genetic engineering. A sub-optimal gene target site may result in the “position effect”, where the inserted transgene is accidentally silenced by an epigenetic mechanism in a mosaic and uncontrollable fashion. Nonetheless, subsequent work could attempt at subcloning the above established cell lines and repeat RNA FISH to achieve at least 80-90% of cells displaying the Xist cloud.

Another interesting observation is that the Xist cloud observed from RNA FISH is actually rather big (Figure 10A). Given that Xist is being induced from chromosome 19, the smallest chromosome in the murine ES cell, it was expected that the cloud visualised would occupy a smaller area, correlating to the size of chromosome 19. However, this initial observation is insufficient to ascertain the true size of the Xist cloud induced. Additional experiments, such as RNA-DNA FISH, should be done to examine how the Xist cloud overlaps with the territory of chromosome 19.

On a separate note, a similar attempt at inserting an inducible Xist transgene had earlier been explored (Jiang et al., 2013). In this paper, the authors had the idea of translating the mechanism of dosage compensation by the Xist RNA to correct gene imbalance in trisomic 21 cells. Using ZFNs, another genome editing tool, the authors inserted an inducible Xist transgene into the DYRK1A locus on chromosome 21, in induced pluripotent stem cells derived from a male Down's syndrome patient. Among a number of successfully genome-edited clones, Xist transgene expression was induced and a localised Xist RNA territory over one of three chromosome 21 copies was observed in more than 85% of the cells. While highly successful in establishing that a large Xist RNA cloud was observed over the DYRK1A locus following Dox treatment, further details such as measurements of the Xist RNA cloud size in comparison to chromosome 21 territory were not found in the paper. Although there are a number of differences between the systems produced in the Jiang et al. paper and in this project, such as the starting cell lines, genome editing method, and Xist transgene constructs used, both are nonetheless useful and can be studied in parallel to address the question this project posed on the robustness of the Xist RNA in recognising small chromosome boundaries, and to further investigate Xist dynamics and chromosome silencing on small autosomes.

6. Conclusion

We established two genetically engineered cell lines with a gene targeting site at chromosome 19. Using Cre-loxP mediated transgene insertion, the full length Xist gene was inserted onto chromosome 19, Xist gene expression was induced with Dox and Xist clouds were visualised in the cells following RNA FISH. Further experiments could look at subcloning to achieve a higher percentage of cells showing the Xist cloud, and performing RNA-DNA FISH. These cell lines can then be used to address the question of whether the Xist cloud is tight enough to recognise the chromosome boundaries from which it is expressed from, and reveal more insight into the dynamics of the Xist cloud.

7. References

Barr, M. L., & Bertram, E. G. (1949). A morphological distinction between neurones of the male and female, and the behaviour of the nucleolar satellite during accelerated nucleoprotein synthesis. *Nature, 163*(4148), 676.

Borsani, G., Tonlorenzi, R., Simmler, M. C., Dandolo, L., Arnaud, D., Capra, V., . . . Ballabio, A. (1991). Characterization of a murine gene expressed from the inactive X chromosome. *Nature, 351*(6324), 325-329. doi:10.1038/351325a0

Bowen, M. A., Patel, D. D., Li, X., Modrell, B., Malacko, A. R., Wang, W. C., . . . et al. (1995). Cloning, mapping, and characterization of activated leukocyte-cell adhesion molecule (ALCAM), a CD6 ligand. *J Exp Med, 181*(6), 2213-2220.

Brockdorff, N., Ashworth, A., Kay, G. F., Cooper, P., Smith, S., McCabe, V. M., . . . Rastan, S. (1991). Conservation of position and exclusive expression of mouse Xist from the inactive X chromosome. *Nature, 351*(6324), 329-331. doi:10.1038/351329a0

Brown, C. J., Ballabio, A., Rupert, J. L., Lafreniere, R. G., Grompe, M., Tonlorenzi, R., & Willard, H. F. (1991). A Gene from the Region of the Human X-Inactivation Center Is Expressed Exclusively from the Inactive X-Chromosome. *Nature, 349*(6304), 38-44. doi:DOI 10.1038/349038a0

Brown, C. J., & Willard, H. F. (1994). The human X-inactivation centre is not required for maintenance of X-chromosome inactivation. *Nature, 368*(6467), 154-156. doi:10.1038/368154a0

Chadwick, B. P., & Willard, H. F. (2004). Multiple spatially distinct types of facultative heterochromatin on the human inactive X chromosome. *Proc Natl Acad Sci U S A, 101*(50), 17450-17455. doi:10.1073/pnas.0408021101

Chaumeil, J., Le Baccon, P., Wutz, A., & Heard, E. (2006). A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced. *Genes Dev, 20*(16), 2223-2237. doi:10.1101/gad.380906

Chow, J., & Heard, E. (2009). X inactivation and the complexities of silencing a sex chromosome. *Curr Opin Cell Biol, 21*(3), 359-366. doi:10.1016/j.ceb.2009.04.012

Cong, L., Ran, F. A., Cox, D., Lin, S. L., Barretto, R., Habib, N., . . . Zhang, F. (2013). Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science, 339*(6121), 819-823. doi:10.1126/science.1231143

Furlan, G., & Rougeulle, C. (2016). Function and evolution of the long noncoding RNA circuitry orchestrating X-chromosome inactivation in mammals. *Wiley Interdiscip Rev RNA, 7*(5), 702-722. doi:10.1002/wrna.1359

Gendrel, A. V., & Heard, E. (2014). Noncoding RNAs and epigenetic mechanisms during X-chromosome inactivation. *Annu Rev Cell Dev Biol, 30*, 561-580. doi:10.1146/annurev-cellbio-101512-122415

Gimferrer, I., Calvo, M., Mittelbrunn, M., Farnos, M., Sarrias, M. R., Enrich, C., . . . Lozano, F. (2004). Relevance of CD6-mediated interactions in T cell activation and proliferation. *J Immunol, 173*(4), 2262-2270.

Gimferrer, I., Farnos, M., Calvo, M., Mittelbrunn, M., Enrich, C., Sanchez-Madrid, F., . . . Lozano, F. (2003). The accessory molecules CD5 and CD6 associate on the membrane of lymphoid T cells. *J Biol Chem, 278*(10), 8564-8571. doi:10.1074/jbc.M209591200

Graves, J. A. (2006). Sex chromosome specialization and degeneration in mammals. *Cell, 124*(5), 901-914. doi:10.1016/j.cell.2006.02.024

Heard, E., Mongelard, F., Arnaud, D., & Avner, P. (1999). Xist yeast artificial chromosome transgenes function as X-inactivation centers only in multicopy arrays and not as single copies. *Mol Cell Biol, 19*(4), 3156-3166.

Herzing, L. B., Romer, J. T., Horn, J. M., & Ashworth, A. (1997). Xist has properties of the X-chromosome inactivation centre. *Nature, 386*(6622), 272-275. doi:10.1038/386272a0

Ichise, H., Ichise, T., Sasanuma, H., & Yoshida, N. (2014). The Cd6 gene as a permissive locus for targeted transgenesis in the mouse. *Genesis, 52*(5), 440-450. doi:10.1002/dvg.22779

Jiang, J., Jing, Y., Cost, G. J., Chiang, J. C., Kolpa, H. J., Cotton, A. M., . . . Lawrence, J. B. (2013). Translating dosage compensation to trisomy 21. *Nature, 500*(7462), 296-300. doi:10.1038/nature12394

Kyba, M., Perlingeiro, R. C., & Daley, G. Q. (2002). HoxB4 confers definitive lymphoid-myeloid engraftment potential on embryonic stem cell and yolk sac hematopoietic progenitors. *Cell, 109*(1), 29-37.

Lee, J. T., & Jaenisch, R. (1997). Long-range cis effects of ectopic X-inactivation centres on a mouse autosome. *Nature, 386*(6622), 275-279. doi:10.1038/386275a0

Lyon, M. F. (1961). Gene action in the X-chromosome of the mouse (Mus musculus L.). *Nature, 190*, 372-373.

Lyon, M. F. (1998). X-chromosome inactivation spreads itself: effects in autosomes. *Am J Hum Genet, 63*(1), 17-19. doi:10.1086/301940

Maduro, C., de Hoon, B., & Gribnau, J. (2016). Fitting the Puzzle Pieces: the Bigger Picture of XCI. *Trends Biochem Sci, 41*(2), 138-147. doi:10.1016/j.tibs.2015.12.003

Marahrens, Y., Panning, B., Dausman, J., Strauss, W., & Jaenisch, R. (1997). Xist-deficient mice are defective in dosage compensation but not spermatogenesis. *Genes Dev, 11*(2), 156-166.

Marks, H., Chow, J. C., Denissov, S., Francoijs, K. J., Brockdorff, N., Heard, E., & Stunnenberg, H. G. (2009). High-resolution analysis of epigenetic changes associated with X inactivation. *Genome Res, 19*(8), 1361-1373. doi:10.1101/gr.092643.109

Ng, K., Daigle, N., Bancaud, A., Ohhata, T., Humphreys, P., Walker, R., . . . Wutz, A. (2011). A system for imaging the regulatory noncoding Xist RNA in living mouse embryonic stem cells. *Mol Biol Cell, 22*(14), 2634-2645. doi:10.1091/mbc.E11-02-0146

Ohhata, T., & Wutz, A. (2013). Reactivation of the inactive X chromosome in development and reprogramming. *Cell Mol Life Sci, 70*(14), 2443-2461. doi:10.1007/s00018-012-1174-3

Ohno, S., Kaplan, W. D., & Kinosita, R. (1959). Formation of the sex chromatin by a single X-chromosome in liver cells of Rattus norvegicus. *Exp Cell Res, 18*, 415-418.

Okamoto, I., Patrat, C., Thepot, D., Peynot, N., Fauque, P., Daniel, N., . . . Heard, E. (2011). Eutherian mammals use diverse strategies to initiate X-chromosome inactivation during development. *Nature, 472*(7343), 370-374. doi:10.1038/nature09872

Penny, G. D., Kay, G. F., Sheardown, S. A., Rastan, S., & Brockdorff, N. (1996). Requirement for Xist in X chromosome inactivation. *Nature, 379*(6561), 131-137. doi:10.1038/379131a0

Pontier, D. B., & Gribnau, J. (2011). Xist regulation and function explored. *Hum Genet, 130*(2), 223-236. doi:10.1007/s00439-011-1008-7

Robinson, W. H., Prohaska, S. S., Santoro, J. C., Robinson, H. L., & Parnes, J. R. (1995). Identification of a mouse protein homologous to the human CD6 T cell surface protein and sequence of the corresponding cDNA. *J Immunol, 155*(10), 4739-4748.

Saifullah, M. K., Fox, D. A., Sarkar, S., Abidi, S. M., Endres, J., Piktel, J., . . . Singer, N. G. (2004). Expression and characterization of a novel CD6 ligand in cells derived from joint and epithelial tissues. *J Immunol, 173*(10), 6125-6133.

Sarrias, M. R., Farnos, M., Mota, R., Sanchez-Barbero, F., Ibanez, A., Gimferrer, I., . . . Lozano, F. (2007). CD6 binds to pathogen-associated molecular patterns and protects from LPS-induced septic shock. *Proc Natl Acad Sci U S A, 104*(28), 11724-11729. doi:10.1073/pnas.0702815104

Sharman, G. B. (1971). Late DNA replication in the paternally derived X chromosome of female kangaroos. *Nature, 230*(5291), 231-232.

Tang, Y. A., Huntley, D., Montana, G., Cerase, A., Nesterova, T. B., & Brockdorff, N. (2010). Efficiency of Xist-mediated silencing on autosomes is linked to chromosomal domain organisation. *Epigenetics Chromatin, 3*(1), 10. doi:10.1186/1756-8935-3-10

Wang, H., La Russa, M., & Qi, L. S. (2016). CRISPR/Cas9 in Genome Editing and Beyond. *Annu Rev Biochem, 85*, 227-264. doi:10.1146/annurev-biochem-060815-014607

Wang, X., Douglas, K. C., Vandeberg, J. L., Clark, A. G., & Samollow, P. B. (2014). Chromosome-wide profiling of X-chromosome inactivation and epigenetic

states in fetal brain and placenta of the opossum, Monodelphis domestica. *Genome Res, 24*(1), 70-83. doi:10.1101/gr.161919.113

Waters, S. A., & Waters, P. D. (2015). Imprinted X chromosome inactivation: evolution of mechanisms in distantly related mammals. *AIMS Genetics, 2*(2), 110-126. doi:10.3934/genet.2015.2.110

Wutz, A., Rasmussen, T. P., & Jaenisch, R. (2002). Chromosomal silencing and localization are mediated by different domains of Xist RNA. *Nat Genet, 30*(2), 167-174. doi:10.1038/ng820