

Section3: Integration based on Site-specific recombination

Cre/loxP, Flp/FRT systems and pluripotent stem cell lines

Candice GT Tahimic¹, Kenji Sakurai², Kazuhiro Aiba^{3*}, Norio Nakatsuji^{2,3*}

¹Children's Hospital & Research Center Oakland, Oakland, California, USA

²Institute for Frontier Medical Sciences, Kyoto University, Kyoto, Japan

³Institute for Integrated Cell-Material Sciences (iCeMS), Kyoto University, Kyoto, Japan

* Corresponding authors:

Kazuhiro Aiba: kaiba@icems.kyoto-u.ac.jp

Norio Nakatsuji: nakatsu@icems.kyoto-u.ac.jp

E-mail addresses:

Candice GT Tahimic: ctahimic@chori.org

Kenji Sakurai: kenji.sakurai@kyowa-kirin.co.jp

Kazuhiro Aiba: kaiba@icems.kyoto-u.ac.jp

Norio Nakatsuji: nakatsu@icems.kyoto-u.ac.jp

Mailing address:

Institute for Integrated Cell-Material Sciences (iCeMS), Kyoto University

Yoshida-Ushinomiya-cho, Sakyo-ku, Kyoto 606-8501, Japan

Abstract

Cre and FLP recombinases are the most widely used site-specific recombinases in genome engineering. Both are members of the tyrosine class of recombinases and catalyze the reversible, site-specific recombination between two identical sequences of 34 bp length in the absence of accessory factors. The substrate sequences for Cre and Flp recombinases are called loxP and FRT sites respectively. Cre recombinase was discovered in the *E. coli* bacteriophage P1 where it plays a crucial role in the life cycle of P1 while Flp recombinase was originally derived from the 2 μ circle of *Saccharomyces cerevisiae* and catalyzes recombination between inverted repeats within the 2 μ plasmid. This chapter shall provide a brief historical perspective on the discovery and early development of the Cre/loxP and FRT/FLP systems citing key studies that paved the way for the application of these site-specific recombination technologies to the engineering of mammalian genomes. Also included are discussions on the mechanisms of Cre/loxP and FRT/FLP systems application of these site-specific recombinase technologies to the introduction of transgenes in human pluripotent stem cells. Certain studies using mouse pluripotent stem cells will also be discussed in order to highlight the possibility of adopting the same strategy in their human counterparts. Lastly, future prospects for these two site-specific recombinases will be presented.

Keywords (5 words or short phrases): Cre/loxP system, Flp/FRT system, pluripotent stem cells, site-specific recombination, site-specific transgene integration

3.1. A historical perspective on Cre/loxP and Flp/FRT systems

3.1.1 Cre/loxP

In 1981, Sternberg and Hamilton reported a high frequency of recABC-independent site-specific recombination in the *E. coli* bacteriophage P1 (Sternberg and Hamilton 1981). In this study, they identified two crucial components for recombination found in phage DNA, a site which they called loxP (for locus of crossing over (x), P1), and a trans-acting gene called Cre (cyclization recombination). In

another report that year, Sternberg and colleagues showed that P1 phage can integrate its DNA at low frequencies into *E. coli* chromosome (Sternberg et al. 1981). The site on the bacterial chromosome into which the P1 plasmid integrates was called loxB. They also showed that integration between loxP and loxB sites generated two hybrid sites which they called LoxL and LoxR. From 1982, major strides towards application of Cre/loxP technology to the engineering of mammalian genomes were made, starting from that year's publication by Hoess and colleagues that reported the nucleotide sequences of all four sites, loxP, loxB, loxL and loxR (Hoess et al. 1982). In their study, they showed that loxP has a perfect 13 bp inverted repeat sequence flanking each side of an asymmetric 8 bp spacer region (Figure 3.1.a). The next year, Abremski and colleagues defined the requirements for loxP/loxP recombinations in terms of co-factors and substrates (Abremski et al. 1983). They reported that no external energy cofactors are required for recombination and that *in vitro* recombination between loxP sites can occur whether the molecule containing them is a supercoiled, nicked-circle or linear DNA. Another major stride towards Cre/loxP-based engineering of mammalian genomes was the 1987 publication of Sauer demonstrating that Cre was functional in the more structurally complex genome of the eukaryote, *Saccharomyces cerevisiae* (Sauer 1987). In this study, Sauer showed that Cre recombinase can catalyze recombination between two loxP sites in the genome of *S. cerevisiae*. This remarkable property of Cre recombinase may be fully appreciated if one considers that in order to conduct more complex transactions on a eukaryotic chromosome, a prokaryotic protein must not only recognize DNA but must also be capable of managing synapsis, DNA cleavage, strand exchange, and religation (Sauer 1987). The year after, Sauer and Henderson further demonstrated that the Cre/loxP system was also functional in mammalian cells (Sauer and Henderson 1988). In their study, they showed that Cre recombinase can mediate the excision of DNA sequences flanking two loxP sites in the genome of the murine cell line C-127. In 1990, they reported the first targeted insertion of exogenous DNA into both yeast and mammalian cells. They demonstrated that Cre-mediated recombination can direct the site-specific integration of a loxP-containing circular DNA into a chromosomal loxP site and that the resulting integrants are predominantly simple single-copy insertions. In 1993, the use of the Cre/loxP system to manipulate the genome of mouse ES cells (ESCs)

was reported by Gu and colleagues (Gu et al. 1993). In their study, the Cre/loxP system was used to study Ig isotype switch recombination in a murine system. They described an approach wherein gene targeting was used to introduce two loxP sites flanking the J_H-E_μ sequences of the IgH locus of mouse ESCs. Next, transient expression of Cre recombinase via an expression vector was performed, resulting in the excision of the J_H-E_μ sequences. ESC clones carrying the IgH deletion were then introduced into blastocysts, thereby generating IgH mutant mice. Their study was a significant step towards the widespread use of the Cre/loxP system in the creation of null phenotypes in mouse. The significant achievements of their study were the demonstration that combined use of the Cre/loxP system and conventional gene targeting can efficiently generate deletions of defined length in mouse ESCs and that the combinatorial method allowed germline transmission.

The first report for the application of the Cre/loxP system in human ESCs came in 2007 (Irion et al. 2007) where it was used to introduce a fluorescent reporter gene in the newly discovered human homolog of the ROSA26 locus. The recombination strategies used in this study will be further discussed in one section of this chapter.

3.1.2. Flp/FRT

In 1978, Beggs reported the high recombination frequency in the 2μ plasmid of *S. cerevisiae* (Beggs 1978) and observed that plasmid sequences recombined intramolecularly during propagation in this yeast. Two years after, Broach and Hicks reported that the recombination proficiency of the 2μ plasmid was linked to an inverted repeat of 599 bp in length and a gene product which they called Flp, in reference to the ability of the gene product to promote recombination between two 2μ plasmids (Broach and Hicks 1980). By 1982, Broach and colleagues delineated the site for recombination to 65 bp within the inverted repeat and also showed that Flp was sufficient to induce intermolecular recombination between two inverted repeats (Broach et al. 1982). The next year, Vetter and colleagues demonstrated that the Flp gene product can catalyze *in vitro* recombination between two inverted repeats and that this process required the presence of cations. They also reported that Flp works efficiently on supercoiled,

relaxed circular, or linear DNA and noted that FLP possessed certain similarities with Cre recombinase (Vetter et al. 1983). In 1985, Andrews and colleagues, using an in vitro DNA protection experiment, defined the Flp binding site to an 8-bp spacer flanked on both sides by a 13 bp inverted repeat (Andrews et al. 1985) (Figure 3.1.b). The year after, McLeod and colleagues coined the word “FRT” to refer to the Flp recognition sequence. They also determined the actual recombination crossover point *in vivo* and confirmed that the essential sequences for the recombination event are limited to the 8-base-pair core sequence and the two 13-base-pair repeated units immediately flanking it (McLeod et al. 1986). Three years after, it was shown in *S. cerevisiae* that FLP can catalyze the recombination between FRT sites separated by a distance of up to 13.3 kb (Rank et al. 1989). In 1992, Dang and Perrimon described the use of the Flp/FRT system to create *Drosophila* embryonic mosaics (Dang and Perrimon 1992). In their work, they placed Flp under the control of a heat shock protein promoter and used this inducible expression system to promote mitotic exchange between homologous chromosomes containing FRT sequences. The next year, Fiering and colleagues demonstrated the utility of Flp-mediated site-specific recombination to study complex cis-regulatory DNA elements within their normal chromosomal context. In their study, Flp/FRT system was used to delete a selectable marker that they inserted into the locus control region (LCR) of the murine beta globin locus of a mouse erythroleukemia cell somatic hybrid (Fiering et al. 1993). As a follow-up publication, Fiering and colleagues showed that the system can also be applied to mouse ESCs. They also highlighted that the introduction of a selection marker to facilitate selection of correctly targeted clones may interfere in the analysis of complex regulatory elements. Hence, they suggested that the selectable marker be removed to avoid influencing the phenotype of the mutation (Fiering et al. 1995). For this purpose, the addition of FRT sites on both sides of the transgene was shown to be useful. The application of the Flp/FRT system in human ESCs was first reported by Patsch and colleagues (Patsch et al. 2010). Their study described the use of a cell-permeant Flp protein that can mediate enhanced site-specific recombination in human ESCs. In their system, they also used Cre and Flp as switches to regulate the expression of lacZ and the caudal-related homeobox protein CDX2 transgenes in human ESCs.

3.2. Mechanisms of Cre/loxP and Flp/FRT recombination

Cre and Flp belong to the tyrosine class of site-specific recombinases (SSRs). Tyrosine recombinases mediate site-specific recombination through nucleophilic attack on the phosphodiester backbone of DNA by a tyrosine hydroxyl group resulting in a covalent protein-DNA intermediate (Grindley et al., 2006, Anastassiadis et al. 2009). Cre is a 38.5 kDa protein product of the bacteriophage P1 cre gene that plays critical roles in the propagation and infectivity of P1 phage. Firstly, it provides an alternative mechanism for the circularization of the P1 genome after infection (Segev and Cohen 1981, Hochman et al. 1983) and secondly, it ensures proper partitioning of the P1 plasmid during bacterial cell division through resolution of plasmid dimers (Austin et al. 1981, Sauer 1988). Flp is a 48.6 kDa protein originally found in the 2 μ plasmid of *S. cerevisiae* and functions in the amplification of the plasmid in yeast. Both Cre and Flp exist as homotetramers in the presence of DNA (Guo et al. 1997, Chen et al. 2000).

Cre recognizes a 34 bp sequence called loxP, composed of an 8 bp non-palindromic spacer region flanked on each side by 13 bp inverted repeats. On the other hand, the full-length substrate of Flp is a 48 bp FRT site consisting of two inverted 13 bp repeats around an 8 bp spacer, and an additional 13 bp direct repeat. However, Flp can also recognize a truncated 34 bp version of FRT which lacks the third 13bp repeat. This 34 bp FRT sequence appears to be a more favorable substrate for the excision reaction. It has been suggested that the use of full length FRT sites may increase occurrence of the integration over the excision reaction (Baer and Bode 2001). For both loxP and FRT, the nonpalindromic nature of the spacer sequence provides directionality to recombination.

Cre and Flp both catalyze reciprocal conservative DNA recombination between recognition sequences. The process involves the following steps, (1) sequence recognition and pairing of recombination sequences, (2) cleavage, (3) strand exchange, (4) ligation of DNA and (5) release of the recombinant DNA products (Figure 3. 2). The recombinase initially binds to each of the recombination sites. Cleavage of the DNA strand is induced when the hydroxyl group of the tyrosine residue attacks a

phosphate group in the spacer region. This conserves the energy of the phosphodiester bond thereby allowing reversibility of the reaction without requiring a high-energy cofactor. Intermediate products are generated, namely free hydroxyl groups at the 5' ends of the DNA and covalent DNA-protein phosphotyrosines link at the 3' ends. The next step is an exchange reaction where the free 5' ends attack the 3' phosphotyrosines of the opposing DNA substrates to generate a Holliday junction. The attack of a nucleophilic tyrosine on a phosphate group is repeated for the intact strand and the new 5' ends migrate and attack the 3' phosphotyrosine linkage of the other DNA duplex. This leads to the resolution of the Holliday junction (Hoess and Abremski 1985, Grindley 2006).

Despite sharing a mechanism common to tyrosine recombinases, some differences exist between Flp and Cre. Flp does not display a significant preference for one or the other end of the spacer for initial cleavage. In contrast, Cre preferentially cleaves at the GC end of the spacer rather than at the AT end (Gosch et al. 2005, Grindley 2006). Kinetic studies have also shown that FLP has a lower binding affinity for FRT than Cre has for loxP. Cre dimers bind the two halves of their target sites cooperatively at approximately threefold higher cooperativity than FLP (Ringrose et al. 1998, Patsch 2010) and recombination efficiency of Flp is 10-fold lower than that of Cre (Ringrose et al. 1998). Moreover, FLP is known to be less thermostable than Cre with its activity significantly reduced at 37°C (Buchholz et al. 1996). This is partly due to the difference in optimal growth temperatures between the original hosts *S. cerevisiae* and *E. coli*. While *E. coli* grows well in 37°C, *S. cerevisiae* prefers temperatures lower than 37°C. In an attempt to circumvent the lower activity of Flp, molecular evolution was used to generate FLPe, a modified version of FLP exhibiting improved performance at 37°C. Despite the five-fold increase in activity of FLPe compared to the wild type version, FLPe has been shown to possess only 25% the activity of Cre (Buchholz et al. 1998).

Depending on the position and orientation of the two loxP sites relative to each other, Cre activity will result in different post-recombination outcomes (Figure 3.3). Recombination between two loxP sites of similar orientation located within the same molecule will result in the excision of DNA sequences between them. On the other hand, Cre activity will result in the inversion of DNA sequences between a

pair of loxP sites of opposite orientation. In scenarios wherein one loxP site is situated in two separate DNA molecules, Cre-mediated recombination will result in the integration of DNA sequences. Cre activity can also induce translocations between two chromosome segments that each carry a loxP site (Wu et al. 2007).

3.3. Mutant lox and FRT sites

Studies using mutant lox and FRT sites have revealed the major requirements for efficient recombination. The altered specificities of these mutant recombinase recognition sites have been exploited to create more sophisticated recombination strategies and to restrict the recombination reaction to one direction.

Homology within the central spacer sequence is necessary for efficient recombination between two loxP sites. It has been demonstrated that a single nucleotide substitution within the 8 bp spacer unit renders a loxP site incapable of recombining with the wild type sequence. Interestingly, this mutant loxP site can efficiently recombine with itself (Hoess et al. 1986). Examples of modified loxP sequences carrying mutations within the non-palindromic spacer region include lox511 (Hoess et al. 1986), lox2272 and lox5171 (Lee and Saito 1998).

Mutant loxP sites have also generated through the introduction of nucleotide substitutions in the left or right inverted repeat elements and are referred to as LE or RE lox. LE/RE lox recombination favors integration over excision (Araki et al. 2010). This is because recombination between LE and RE sites results in a wild type loxP and a double mutant lox site that is an inefficient substrate of Cre recombinase. Due to its unidirectionality, the LE/RE lox system has been incorporated in Cre-mediated recombination strategies to minimize the possibility of transgene excision after the initial integration reaction.

Homology within the central spacer sequence is also necessary for efficient recombination of FRT sites (Senecoff and Cox, 1986, Bode et al. 2000). However, the spacer sequence itself is not so important provided that homology is maintained between the two target spacer sequences (Senecoff and Cox, 1986, Lee and Saito 1998).

Table 3.1 shows some of the currently available mutant loxP and FRT. As will be further discussed in the succeeding sections, these additional sequences have increased the precision of both Cre and FRT-based site-specific recombination and spurred the development of more sophisticated genetic manipulation strategies.

3.4. Pluripotent stem cells, an overview

In 1981, Evans and Kaufman reported the derivation of murine embryonic stem cells (ESCs) from the inner cell mass of murine preimplantation blastocyst embryos (Evans and Kaufman 1981, Martin 1981). ESCs have two characteristic properties, self renewal which pertains to their ability to expand indefinitely, and pluripotency defined as the capacity to differentiate into derivatives of all three germ layers. More than a decade after, the derivation of ESCs from *Rhesus* macaque (Thomson et al., 1995), and the common marmoset *Callithrix jacchus* (Thomson et al. 1996) were reported. Human ESC lines were established in 1998 (Thomson et al. 1998) and a variety human and non-human primate ESC lines were generated by other groups thereafter (see Nakatsuji and Suemori 2002, Guhr et al. 2006 for reviews). Adult cell-derived pluripotent stem cells (PSCs) also became available with the discovery that differentiated cells can revert to an embryonic state via forced expression of transcription factors such as Oct3/4, Sox2, Nanog and c-Myc (Takahashi and Yamanaka 2006, Takahashi et al. 2007, Okita et al. 2007). These cells are referred to as induced pluripotent stem cells (iPSCs) and a number of mouse and human-derived lines are currently available.

Unlike mouse ESCs and iPSCs, their human counterparts cannot be routinely passaged as single cells. This greatly impedes the creation of human ES and iPS reporter cell lines as genetic modification is usually followed by clonal selection. Moreover, human ESCs and iPSCs exhibit lower gene targeting efficiencies of approximately 10^{-6} compared to values of 10^{-2} or 10^{-3} observed in murine counterparts. A proposed explanation for such differences is that hESCs are actually more similar in properties to mouse epiblast stem cells. Both cell types exhibit dependence on bFGF supplementation, flattened colony morphology, low tolerance for single cell disruption, and slower growth rate compared to mouse ESCs

(Brons et al. 2007, Tesar et al. 2007). A study has strengthened this theory by demonstrating that human iPSCs that mimic morphological, molecular, and functional properties of murine ESCs can be generated by performing the reprogramming process in the presence of leukemia inhibitory factor (LIF) (Buecker et al. 2010). These mESC-like human iPSCs, referred to as hLR5 iPSCs, have been shown to possess gene targeting rates similar to that found in mESCs. However, it should be noted that this particular cell line exists in a metastable state, meaning that LIF withdrawal and transcriptional inactivation of reprogramming factors will result in conversion into typical hESC behavior. Nevertheless, the study is a clear demonstration that low gene targeting rates is not necessarily an intrinsic property of human PSCs, but rather, is a consequence of an incomplete understanding of factors required for the establishment and maintenance of pluripotency. Further improvements in methodologies to derive PSCs should result in cell lines that are more amenable to genetic manipulation and downstream applications.

An ever-expanding list of PSC lines is currently available to researchers. Naturally, the choice of cell line will be dictated by the objectives of the study and other practical considerations. In studies that use a single hESC line to integrate various versions of the transgene at the same locus, the Cre/loxP and Flp/FRT systems may provide the added benefit of more robust recombination rates. For example, loxP or FRT sites can be introduced via conventional gene targeting and subsequent transgene integrations can be performed via the relatively more efficient Cre or Flp systems. This strategy will be further discussed in the next section.

3.5. Basic strategies for site-specific integration of transgenes via Cre/loxP and Flp/FRT systems

Although literature search will reveal numerous approaches to site-specific integration of transgenes using Cre/loxP and Flp/FRT systems, these strategies are combinations of three basic approaches which will be categorized as follows: (1) integration, (2) excision, (3) inversion, and (4) recombination-mediated cassette exchange (RMCE). For brevity, the Cre/loxP system will be used as an example to describe these approaches in detail. In the simple integration approach (Figure 3.4.a), gene targeting is initially performed to introduce a single loxP site. Correctly targeted clones are then subjected

to the next round of genetic manipulation involving co-transfection of a Cre expression vector containing a loxP site and the gene of interest. Cre activity results in the intermolecular recombination between the two loxP sites and the subsequent integration of the transgene into the chromosome of the host cell.

The excision approach (Figure 3.4.b) makes use excision capabilities of Cre. The initial gene targeting step involves the use of a vector containing two loxP sites that flank a gene of interest (GOI). Recombinant clones are then co-transfected with Cre expression plasmid and a plasmid carrying a single loxP site and another GOI. Intermolecular recombination between the two loxP sites in the chromosomal DNA will result in the excision of the sequences between these two sites, leaving one chromosomal loxP site.

The inversion approach (Figure 3.4.c) begins with a gene targeting step to introduce a GOI flanked by a pair of anti-parallel loxP sites. Cre recombinase recombines the two loxP sites and in the process inverts the transgene. In most applications, Cre-mediated inversion brings about reconstitution of transgene expression. In a straightforward application of this approach, no transgene is co-transfected with the Cre-expressing vector as the transgene is already present in the post-targeted allele.

RMCE (Figure 3.4.d) is a more sophisticated approach that involves an initial targeting step to introduce a GOI flanked by two mismatched (heterotypic) lox sites. Recombinant clones are then co-transfected with Cre expression vector and a plasmid carrying two heterotypic lox sites similar to that used in the targeting vector. A transgene is usually present between the heterotypic lox sites of the plasmid vector. Cre activity will result in pairing of the corresponding heterotypic sites in the chromosomal and plasmid DNA. The result is an equivalent exchange of sequences flanked by the heterotypic lox sites. This approach minimizes undesirable post-recombination outcomes since only the matching lox sites can effectively recombine.

3.6 Delivery of recombinases into PSCs

Conventionally, Cre or Flp-expressing plasmids are introduced into PSCs via electroporation. In an ideal situation, the transient burst of recombinase expression results in the desired recombination event

and the recombinase-expressing plasmid is subsequently lost as PSCs undergo cell division. Nevertheless, as in any gene delivery attempt involving DNA, there is always a risk of plasmid integration into the host genome. An alternative delivery method such as protein transduction can be employed to eliminate this risk. This method makes use of short peptides called cell penetrating peptides (CPPs) or protein transduction domains to carry macromolecules into cells (Gump and Dowdy 2007). Peitz and colleagues have developed a cell-permeable Cre protein called HTN-Cre (HTNC) (Peitz et al. 2002). HTNC is composed of an N-terminal cell permeation domain derived from the human immunodeficiency virus (HIV) transactivator of transcription (TAT) protein fused to a nuclear localization sequence (NLS). They showed that high recombination efficiencies can be achieved in fibroblasts and mouse ESCs using this method. As mentioned earlier, a cell-permeable version of the Flp protein has been generated by Patsch and colleagues and has been successfully used to perform genome manipulations in both mouse and human ESCs (Patsch et al. 2010).

3.7. Applications of Cre/loxP and Flp/FRT systems in the introduction of transgenes to PSCs

At the time of writing, there is less extensive history on the genetic modification of human ES for the creation of iPS reporter cell lines and disease models compared to efforts in mouse. This may be attributed to the relatively recent discovery of human ESCs and to technical challenges in their manipulation. Nevertheless, a number of elegant Cre/loxP and Flp/FRT-based strategies for transgene integrations have been reported and will be the focus of this section. Wherever necessary, studies using mouse PSCs will be described.

3.7.1. Creation of reporter cell lines

In its simplest sense, a reporter gene should allow easy identification of a cell population or event of interest. However, in more demanding applications such as the study of developmental genes and cell lineage specification, it is critical that reporter genes faithfully mimic the spatiotemporal patterns and expression levels of their physiological counterparts. Positional effects are always a concern in expression

data obtained via the random integration of transgenes. For example, random integration at hypermethylated regions of the genome may result in the transgene silencing. The opposite, but equally unwanted outcome is a loss of spatial and temporal control of gene expression when the transgene randomly inserts into an area of constitutive expression.

The site specificity of Cre and Flp recombination systems make them superior to random integration-based methods for the creation of reporter cell lines. However, it should be noted that the risk of undesirable integrations cannot be completely eliminated even with site-specific integration approaches. Moreover, some studies have reported the presence of pseudo loxP sites in mammalian genomes (Thyagarajan et al. 2000, Schmidt et al. 2000) that may potentially allow illegitimate Cre and Flp-mediated recombination. Although most of the reported pseudo sites recombine with legitimate loxP sequences at a lower efficiency compared to legitimate loxP pairings, a determination of transgene copy number must be performed for recombinant cell lines if the requirement of the study is single copy expression of a transgene.

3.7.2. Turning on transgenes via recombinase-mediated inversions

An early proof-of-principle on the use of the Cre/lox system to generate a reporter hESC line was performed by Irion and colleagues (Irion et al. 2007). This study made use of a combination of Cre-mediated excision and inversion reactions to reconstitute expression of a transgene (Figure 3.5). First, gene targeting introduced a viral splice acceptor followed by a loxP-flanked promoterless neomycin-resistance gene (neomycin^r). The neomycin gene was followed by an inverted tandem RFP variant, tdRFP which in turn was flanked by mutant lox2272 sites. Cre-mediated recombination may first induce inversion of the intervening DNA at either the loxP or the lox2272 sites. This will cause inversion of the promoterless tdRFP, with its ATG codon now at the same orientation as the endogenous promoter. Further Cre-mediated activity will then excise neomycin^r located between the two loxP or lox2272 sites, finally leaving two mismatched lox sites which are incapable of further recombination. This is indeed an elegant way of minimizing the potential for unwanted post-Cre recombination outcomes. However, a

word of caution is given to those who intend to eliminate a potentially recombinogenic lox site flanked by matching lox variants as one study has reported that the distance between the compatible lox sites should be at least 82 base pairs to ensure efficient recombination (Hoess et al. 1985).

The report by Irion and colleagues also demonstrated how heterotypic lox sites could be strategically positioned to limit the unwanted excision of the GOI, in this case, tdRFP (Irion et al. 2007). The other findings of this study mirror what has been established in the murine, that is, transgenes integrated into the ROSA26 locus retain their expression throughout differentiation into lineages of the three germ layers and does not adversely affect the pluripotency of the recombinant cell lines (Zambrowicz et al. 1997, Soriano 1999). Hence, the human homolog of ROSA26 is considered a candidate “safe-harbor” site for the introduction of transgenes into human PSCs.

3.7.3. Cre/loxP, Flp/FRT and chemically inducible transgene expression systems

The Cre and Flp systems may also be used to circumvent bottlenecks in the study of gene function. For example, expression of certain mutant forms of a gene may result in early lethality, making it impossible to generate viable cells past a certain developmental stage. This can be circumvented by placing the transgene under the control of inducible systems such as the tetracycline (Gossen and Bujard 1992) or tamoxifen (Littlewood et al. 1995) systems. We have shown proof of principle for this approach in one of our publications (Sakurai et al. 2010) where we performed gene targeting at the HPRT locus of a human female ESC line and subsequently introduced the tetracycline inducible elements via Cre-mediated recombination. (Figure 3.6). We first generated a mutant HPRT allele a loxP site 5' to a neomycin expression cassette. Another loxP site was located 5' to a promoterless hygromycin resistance gene lacking the ATG codon, with each end of the targeting construct flanked by an insulator sequence. After gene targeting, we performed co-transfection of a Cre-expression vector with a plasmid carrying an EF1 α promoter, a Kozak sequence, an ATG codon, a loxP site and an EGFP transgene under the control of tetracycline responsive element (TRE). Also in this plasmid is the “Tet-on” tetracycline transactivator, rtTA under the control of CAGGS promoter (fusion of cytomegalovirus immediate early enhancer /

chicken beta-actin promoter). Subsequent Cre-mediated recombination results in the excision of the neomycin expression cassette and the introduction of the plasmid carrying the EGFP transgene and the promoterless hygromycin. This also results in repositioning of the promoterless hygromycin gene 3' to the EF1 α promoter and the start codon, thereby reconstituting hygromycin expression. In our hands, all of the hygromycin-resistant clones we picked underwent the desired recombination event. Hence, this strategy can be employed to significantly minimize the number of clones to be picked and screened. However, in our hands, we observed that the response to doxycycline induction varied from clone to clone. Hence, we recommend testing a number of clones to verify levels of transgene induction as they go through continuous passage and directed differentiation.

3.7.4. Recombination-mediated cassette exchange (RMCE)

The RMCE approach involves an exchange between a transgene flanked by mismatched lox sites and a corresponding plasmid carrying a transgenes flanked by the same types of mismatched lox sites found in the chromosomal locus. Di Domenico and colleagues replaced a chromosomal EF1 α promoter-driven neomycin resistance transgene with a human Oct4 promoter-driven EGFP transgene using loxP and lox2272 sites (Di Domenico et al. 2008) (Figure 3.7). Clones that underwent the desired recombination event become resistant to blasticidin through Cre-mediated reconstitution of this selection marker.

It is important that the choice of heterotypic lox sites take into consideration the potential for illegitimate recombination between non-target lox sites. Certain mismatched lox sites have a greater tendency to illegitimately recombine more than others. In fact, one study has demonstrated cassette replacement via recombination between loxP and lox511 sites (Soukharev et al. 1999). Siegel and colleagues reported the rates of illegitimate recombination between certain mutant lox sites (Siegel et al. 2001).

3.7.5. Site-specific integration of BACs

The use of bacterial artificial chromosomes (BACs) in the creation of reporter cell lines is expected to confer faithful transgene expression compared to the use of shorter promoter constructs. This is due to ability of BACs to house longer DNA sequences that retain full length gene regulatory elements. Placantonakis and colleagues provided a proof of principle for the use of BACs in the creation of hESC reporter lines (Placantonakis et al. 2009). In their study, BACs carrying an EGFP transgene under the control of full-length regulatory elements of neuron-specific genes were introduced to hESCs via random integration. Using this approach, they were able to isolate clones that maintained proper spatio-temporal expression of EGFP as differentiation progressed towards the neuronal lineage. It will be worthwhile to examine whether introduction of BACs via site-specific integration into “safe harbor” sites will increase the efficiency of generating high fidelity human reporter cell lines.

However, a number of concerns have been raised in recombinase-mediated integration of BACs. Due to their larger size, BACs may be less efficiently introduced into chromosomal regions via Cre. The use of mammalian codon-optimized Fip (Raymond and Soriano 2007) and Cre sequences (Shimshek et al. 2002) or the introduction of cell-permeant versions (Nolden et al. 2006, Patsch et al. 2010) may potentially help offset the decreased efficiency of these recombinases on large constructs such as BACs. BACs may also contain too many unwanted prokaryotic genes or regulatory elements. It is therefore necessary to carefully consider intermediate steps in BAC construction to generate final constructs that are devoid of unwanted sequences or perhaps, employ a strategy to eliminate these sequences via a secondary recombination event in the PSC line.

To date, a number of human BAC libraries have been generated and are available to researchers (BACPAC Resources, Invitrogen; see references for websites). Specifically, BAC libraries have been created for the H1 and H9 hESC lines (BACPAC Resources; see reference for website). These two libraries were generated in a transposon-free *E. coli* host to minimize the possibility of sequence disruption caused by random insertions of prokaryotic transposon elements.

3.8. Technological improvements and future outlook

Site-directed integration of transgenes is usually preceded by a gene targeting step. Hence, Cre and Flp-mediated recombination is expected to gain more widespread application in the creation of human PSC reporter lines as better gene targeting methods and more robust human PSC lines are developed. Consequently, there will be a need to expand the current list of well-characterized loci available for site-specific integration. HPRT locus has been extensively studied and has been favored in many genetic manipulation attempts for both mouse and human PSCs due to the low risk of transgene silencing after differentiation. However, it is also known that null mutations in this gene result in Lesch-Nyhan syndrome. Hence, the utility of HPRT as an integration site may be limited to cell types that are not substantially affected by HPRT depletion.

At present, a number of candidate “safe-harbor” sites exist such as the adeno-associated virus integration site (AAVS) in chromosome 19 and the previously discussed ROSA26 locus. One group has also shown that integration of an EGFP transgene between the thrombopoietin (*TPO* or *THPO*) and solute carrier 25 (*SC25*) genes of human chromosome 12 does not disrupt any known gene and can facilitate reliable reporter gene expression in derivatives of the three germ layers (Costa et al. 2005). Another study has suggested that the region between IGFBP7 and LOC255130 in human chromosome 4 and that flanked by CLDN3 and CLDN4 may serve as potential integration sites that have a low risk of transgene silencing after differentiation (Du et al. 2009). However, further studies in human cells are needed to confirm that no refractory phenotypes will result from disruptions of these regions.

Any technology is always at risk of being overtaken by newer methodologies. As expected, the discovery of the PhiC31 recombination system and the development of zinc finger nuclease-assisted genetic manipulations, both discussed in other chapters of this book, have provided alternative approaches to site-directed transgene integration. Nevertheless, the relative simplicity, accessibility to researchers, the low cost and the availability of a number of pre-existing lox and FRT-carrying PSC lines will continue to make the Cre and Flp systems a practical method of choice for the site-directed integration of transgenes.

References

Abremski K, Hoess R, Sternberg N. Studies on the properties of P1 site-specific recombination: evidence for topologically unlinked products following recombination. *Cell*. 1983 Apr;32(4):1301-11. PubMed PMID: 6220808.

Albert H, Dale EC, Lee E, Ow DW. Site-specific integration of DNA into wild-type and mutant lox sites placed in the plant genome. *Plant J*. 1995 Apr;7(4):649-59. PubMed PMID: 7742860.

Anastassiadis K, Fu J, Patsch C, Hu S, Weidlich S, Duerschke K, Buchholz F, Edenhofer F, Stewart AF. Dre recombinase, like Cre, is a highly efficient site-specific recombinase in *E. coli*, mammalian cells and mice. *Dis Model Mech*. 2009 Sep-Oct;2(9-10):508-15. Epub 2009 Aug 19. PubMed PMID: 19692579.

Andrews BJ, Proteau GA, Beatty LG, Sadowski PD. The FLP recombinase of the 2 micron circle DNA of yeast: interaction with its target sequences. *Cell*. 1985 Apr;40(4):795-803. PubMed PMID: 3879971.

Araki K, Okada Y, Araki M, Yamamura K. Comparative analysis of right element mutant lox sites on recombination efficiency in embryonic stem cells. *BMC Biotechnol*. 2010 Mar 31;10:29. PubMed PMID: 20356367; PubMed Central PMCID: PMC2865440.

Austin S, Ziese M, Sternberg N. A novel role for site-specific recombination in maintenance of bacterial replicons. *Cell*. 1981 Sep;25(3):729-36. PubMed PMID: 7026049.

BACPAC Resources website. <http://bacpac.chori.org>

Baer A, Bode J. Coping with kinetic and thermodynamic barriers: RMCE, an efficient strategy for the targeted integration of transgenes. *Curr Opin Biotechnol*. 2001 Oct;12(5):473-80. Review. PubMed PMID: 11604323.

Beggs JD. Transformation of yeast by a replicating hybrid plasmid. *Nature*. 1978 Sep 14;275(5676):104-9. PubMed PMID: 357984.

Bode J, Schlake T, Iber M, Schübeler D, Seibler J, Snezhkov E, Nikolaev L. The transgeneticist's toolbox: novel methods for the targeted modification of eukaryotic genomes. *Biol Chem*. 2000 Sep-Oct;381(9-10):801-13. Review. PubMed PMID: 11076013.

Broach JR, Hicks JB. Replication and recombination functions associated with the yeast plasmid, 2 mu circle. *Cell*. 1980 Sep;21(2):501-8. PubMed PMID: 7407923.

Broach JR, Guarascio VR, Jayaram M. Recombination within the yeast plasmid 2mu circle is site-specific. *Cell*. 1982 May;29(1):227-34. PubMed PMID: 6286142.

Brons IG, Smithers LE, Trotter MW, Rugg-Gunn P, Sun B, Chuva de Sousa Lopes SM, Howlett SK, Clarkson A, Ahrlund-Richter L, Pedersen RA, Vallier L. Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature*. 2007 Jul 12;448(7150):191-5. Epub 2007 Jun 27. PubMed PMID: 17597762.

Buchholz F, Angrand PO, Stewart AF. Improved properties of FLP recombinase evolved by cycling mutagenesis. *Nat Biotechnol*. 1998 Jul;16(7):657-62. PubMed PMID: 9661200.

Buecker C, Chen HH, Polo JM, Daheron L, Bu L, Barakat TS, Okwieka P, Porter A, Gribnau J, Hochedlinger K, Geijsen N. A murine ESC-like state facilitates transgenesis and homologous

recombination in human pluripotent stem cells. *Cell Stem Cell*. 2010 Jun 4;6(6):535-46. PubMed PMID: 20569691.

Chen Y, Narendra U, Iype LE, Cox MM, Rice PA. Crystal structure of a Flp recombinase-Holliday junction complex: assembly of an active oligomer by helix swapping. *Mol Cell*. 2000 Oct;6(4):885-97. PubMed PMID: 11090626.

Costa M, Dottori M, Ng E, Hawes SM, Sourris K, Jamshidi P, Pera MF, Elefanty AG, Stanley EG. The hESC line Envy expresses high levels of GFP in all differentiated progeny. *Nat Methods*. 2005 Apr;2(4):259-60. Epub 2005 Mar 23. PubMed PMID: 15782217.

Dang DT, Perrimon N. Use of a yeast site-specific recombinase to generate embryonic mosaics in *Drosophila*. *Dev Genet*. 1992;13(5):367-75. PubMed PMID: 1292893.

Di Domenico AI, Christodoulou I, Pells SC, McWhir J, Thomson AJ. Sequential genetic modification of the hprt locus in human ESCs combining gene targeting and recombinase-mediated cassette exchange. *Cloning Stem Cells*. 2008 Jun;10(2):217-30. PubMed PMID: 18386992.

Du ZW, Hu BY, Ayala M, Sauer B, Zhang SC. Cre recombination-mediated cassette exchange for building versatile transgenic human embryonic stem cells lines. *Stem Cells*. 2009 May;27(5):1032-41. PubMed PMID: 19415769; PubMed Central PMCID: PMC2801346

Evans MJ, Kaufman MH. Establishment in culture of pluripotential cells from mouse embryos. *Nature*. 1981 Jul 9;292(5819):154-6. PubMed PMID: 7242681.

Fiering S, Kim CG, Epner EM, Groudine M. An "in-out" strategy using gene targeting and FLP recombinase for the functional dissection of complex DNA regulatory elements: analysis of the beta-globin locus control region. *Proc Natl Acad Sci U S A*. 1993 Sep 15;90(18):8469-73. PubMed PMID: 8378321; PubMed Central PMCID: PMC47378.

Fiering S, Epner E, Robinson K, Zhuang Y, Telling A, Hu M, Martin DI, Enver T, Ley TJ, Groudine M. Targeted deletion of 5'HS2 of the murine beta-globin LCR reveals that it is not essential for proper regulation of the beta-globin locus. *Genes Dev*. 1995 Sep 15;9(18):2203-13. PubMed PMID: 7557375.

Ghosh K, Lau CK, Gupta K, Van Duyne GD. Preferential synapsis of loxP sites drives ordered strand exchange in Cre-loxP site-specific recombination. *Nat Chem Biol*. 2005 Oct;1(5):275-82. Epub 2005 Sep 11. PubMed PMID: 16408057.

Gossen M, Bujard H. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc Natl Acad Sci U S A*. 1992 Jun 15;89(12):5547-51. PubMed PMID: 1319065; PubMed Central PMCID: PMC49329.

Grindley ND, Whiteson KL, Rice PA. Mechanisms of site-specific recombination. *Annu Rev Biochem*. 2006;75:567-605. Review. PubMed PMID: 16756503.

Gu H, Zou YR, Rajewsky K. Independent control of immunoglobulin switch recombination at individual switch regions evidenced through Cre-loxP-mediated gene targeting. *Cell*. 1993 Jun 18;73(6):1155-64. PubMed PMID: 8513499.

Guhr A, Kurtz A, Friedgen K, Löser P. Current state of human embryonic stem cell research: an overview of cell lines and their use in experimental work. *Stem Cells*. 2006 Oct;24(10):2187-91. Epub 2006 Jun 15. Review. PubMed PMID: 16778154.

Guo F, Gopaul DN, van Duyn GD. Structure of Cre recombinase complexed with DNA in a site-specific recombination synapse. *Nature*. 1997 Sep 4;389(6646):40-6. PubMed PMID: 9288963.

Gump JM, Dowdy SF. TAT transduction: the molecular mechanism and therapeutic prospects. *Trends Mol Med*. 2007 Oct;13(10):443-8. Review. PubMed PMID: 17913584.

Hochman L, Segev N, Sternberg N, Cohen G. Site-specific recombinational circularization of bacteriophage P1 DNA. *Virology*. 1983 Nov;131(1):11-7. PubMed PMID: 6228057.

Hoess RH, Ziese M, Sternberg N. P1 site-specific recombination: nucleotide sequence of the recombining sites. *Proc Natl Acad Sci U S A*. 1982 Jun;79(11):3398-402. PubMed PMID: 6954485; PubMed Central PMCID: PMC346427.

Hoess RH, Abremski K. Mechanism of strand cleavage and exchange in the Cre-lox site-specific recombination system. *J Mol Biol*. 1985 Feb 5;181(3):351-62. PubMed PMID: 3856690.

Hoess R, Wierzbicki A, Abremski K. Formation of small circular DNA molecules via an in vitro site-specific recombination system. *Gene*. 1985;40(2-3):325-9. PubMed PMID: 3007297.

Hoess RH, Wierzbicki A, Abremski K. The role of the loxP spacer region in P1 site-specific recombination. *Nucleic Acids Res*. 1986 Mar 11;14(5):2287-300. PubMed PMID: 3457367; PubMed Central PMCID: PMC339658.

Invitrogen website. <http://www.invitrogen.com>

Irion S, Luche H, Gadue P, Fehling HJ, Kennedy M, Keller G. Identification and targeting of the ROSA26 locus in human embryonic stem cells. *Nat Biotechnol*. 2007 Dec;25(12):1477-82. Epub 2007 Nov 25. PubMed PMID: 18037879.

Lee G, Saito I. Role of nucleotide sequences of loxP spacer region in Cre-mediated recombination. *Gene*. 1998 Aug 17;216(1):55-65. PubMed PMID: 9714735.

Littlewood TD, Hancock DC, Danielian PS, Parker MG, Evan GI. A modified oestrogen receptor ligand-binding domain as an improved switch for the regulation of heterologous proteins. *Nucleic Acids Res*. 1995 May 25;23(10):1686-90. PubMed PMID: 7784172; PubMed Central PMCID: PMC306922.

Martin GR. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc Natl Acad Sci U S A*. 1981 Dec;78(12):7634-8. PubMed PMID: 6950406; PubMed Central PMCID: PMC349323.

McLeod M, Craft S, Broach JR. Identification of the crossover site during FLP-mediated recombination in the *Saccharomyces cerevisiae* plasmid 2 microns circle. *Mol Cell Biol*. 1986 Oct;6(10):3357-67. PubMed PMID: 3540590; PubMed Central PMCID: PMC367081.

Nakatsuji N, Suemori H. Embryonic stem cell lines of nonhuman primates. *ScientificWorldJournal*. 2002 Jun 26;2:1762-73. Review. PubMed PMID: 12806169.

Nolden L, Edenhofer F, Haupt S, Koch P, Wunderlich FT, Siemen H, Brüstle O. Site-specific recombination in human embryonic stem cells induced by cell-permeant Cre recombinase. *Nat Methods*. 2006 Jun;3(6):461-7. PubMed PMID: 16721380.

Okita K, Ichisaka T, Yamanaka S. Generation of germline-competent induced pluripotent stem cells. *Nature*. 2007 Jul 19;448(7151):313-7. Epub 2007 Jun 6. PubMed PMID: 17554338.

Patsch C, Peitz M, Otte DM, Kessler D, Jungverdorben J, Wunderlich FT, Brüstle O, Zimmer A, Edenhofer F. Engineering cell-permeant FLP recombinase for tightly controlled inducible and reversible overexpression in embryonic stem cells. *Stem Cells*. 2010 May;28(5):894-902. PubMed PMID: 20333748.

Peitz M, Pfannkuche K, Rajewsky K, Edenhofer F. Ability of the hydrophobic FGF and basic TAT peptides to promote cellular uptake of recombinant Cre recombinase: a tool for efficient genetic engineering of mammalian genomes. *Proc Natl Acad Sci U S A*. 2002 Apr 2;99(7):4489-94. Epub 2002 Mar 19. PubMed PMID: 11904364; PubMed Central PMCID: PMC123675.

Placantonakis DG, Tomishima MJ, Lafaille F, Desbordes SC, Jia F, Socci ND, Viale A, Lee H, Harrison N, Tabar V, Studer L. BAC transgenesis in human embryonic stem cells as a novel tool to define the human neural lineage. *Stem Cells*. 2009 Mar;27(3):521-32. PubMed PMID: 19074416.

Rank GH, Arndt GM, Xiao W. FLP-FRT mediated intrachromosomal recombination on a tandemly duplicated YEp integrant at the ILV2 locus of chromosome XIII in *Saccharomyces cerevisiae*. *Curr Genet*. 1989 Feb;15(2):107-12. PubMed PMID:2663188.

Raymond CS, Soriano P. High-efficiency FLP and PhiC31 site-specific recombination in mammalian cells. *PLoS One*. 2007 Jan 17;2(1):e162. PubMed PMID: 17225864; PubMed Central PMCID: PMC1764711.

Ringrose L, Lounnas V, Ehrlich L, Buchholz F, Wade R, Stewart AF. Comparative kinetic analysis of FLP and cre recombinases: mathematical models for DNA binding and recombination. *J Mol Biol*. 1998 Nov 27;284(2):363-84. PubMed PMID: 9813124.

Sakurai K, Shimoji M, Tahimic CG, Aiba K, Kawase E, Hasegawa K, Amagai Y, Suemori H, Nakatsuji N. Efficient integration of transgenes into a defined locus in human embryonic stem cells. *Nucleic Acids Res*. 2010 Apr;38(7):e96. Epub 2010 Jan 13. PubMed PMID: 20071742; PubMed Central PMCID: PMC2853137.

Sauer B. Functional expression of the cre-lox site-specific recombination system in the yeast *Saccharomyces cerevisiae*. *Mol Cell Biol*. 1987 Jun;7(6):2087-96. PubMed PMID: 3037344; PubMed Central PMCID: PMC365329.

Sauer B, Henderson N. Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. *Proc Natl Acad Sci U S A*. 1988 Jul;85(14):5166-70. PubMed PMID: 2839833; PubMed Central PMCID: PMC281709.

Schmidt EE, Taylor DS, Prigge JR, Barnett S, Capecchi MR. Illegitimate Cre-dependent chromosome rearrangements in transgenic mouse spermatids. *Proc Natl Acad Sci U S A*. 2000 Dec 5;97(25):13702-7. PubMed PMID: 11087830; PubMed Central PMCID: PMC17639.

Segev N, Cohen G. Control of circularization of bacteriophage P1 DNA in Escherichia coli. *Virology*. 1981 Oct 30;114(2):333-42. PubMed PMID: 7027600.

Senecoff JF, Cox MM. Directionality in FLP protein-promoted site-specific recombination is mediated by DNA-DNA pairing. *J Biol Chem*. 1986 Jun 5;261(16):7380-6. PubMed PMID: 3711092.

Schlake T, Bode J. Use of mutated FLP recognition target (FRT) sites for the exchange of expression cassettes at defined chromosomal loci. *Biochemistry*. 1994 Nov 1;33(43):12746-51. PubMed PMID: 7947678.

Shimshek DR, Kim J, Hübner MR, Spergel DJ, Buchholz F, Casanova E, Stewart AF, Seeburg PH, Sprengel R. Codon-improved Cre recombinase (iCre) expression in the mouse. *Genesis*. 2002 Jan;32(1):19-26. PubMed PMID: 11835670.

Siegel RW, Jain R, Bradbury A. Using an in vivo phagemid system to identify non-compatible loxP sequences. *FEBS Lett*. 2001 Jun 15;499(1-2):147-53. Corrected and republished in: *FEBS Lett*. 2001 Sep 21;505(3):467-73. PubMed PMID: 11418130.

Soriano P. Generalized lacZ expression with the ROSA26 Cre reporter strain. *Nat Genet*. 1999 Jan;21(1):70-1. PubMed PMID: 9916792.

Soukharev S, Miller JL, Sauer B. Segmental genomic replacement in embryonic stem cells by double lox targeting. *Nucleic Acids Res*. 1999 Sep 15;27(18):e21. PubMed PMID: 10471751; PubMed Central PMCID: PMC148613.

Sternberg N, Hamilton D, Hoess R. Bacteriophage P1 site-specific recombination. II. Recombination between loxP and the bacterial chromosome. *J Mol Biol*. 1981 Aug 25;150(4):487-507. PubMed PMID: 6276558.

Sternberg N, Hamilton D. Bacteriophage P1 site-specific recombination. I. Recombination between loxP sites. *J Mol Biol*. 1981 Aug 25;150(4):467-86. PubMed PMID: 6276557.

Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*. 2006 Aug 25;126(4):663-76. Epub 2006 Aug 10. PubMed PMID: 16904174.

Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell*. 2007 Nov 30;131(5):861-72. PubMed PMID: 18035408.

Tesar PJ, Chenoweth JG, Brook FA, Davies TJ, Evans EP, Mack DL, Gardner RL, McKay RD. New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature*. 2007 Jul 12;448(7150):196-9. Epub 2007 Jun 27. PubMed PMID: 17597760.

Thomson JA, Kalishman J, Golos TG, Durning M, Harris CP, Becker RA, Hearn JP. Isolation of a primate embryonic stem cell line. *Proc Natl Acad Sci U S A*. 1995 Aug 15;92(17):7844-8. PubMed PMID: 7544005; PubMed Central PMCID: PMC41242.

Thomson JA, Kalishman J, Golos TG, Durning M, Harris CP, Hearn JP. Pluripotent cell lines derived from common marmoset (*Callithrix jacchus*) blastocysts. *Biol Reprod*. 1996 Aug;55(2):254-9. PubMed PMID: 8828827.

Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS, Jones JM. Embryonic stem cell lines derived from human blastocysts. *Science*. 1998 Nov 6;282(5391):1145-7. Erratum in: *Science* 1998 Dec 4;282(5395):1827. PubMed PMID: 9804556.

Thyagarajan B, Guimarães MJ, Groth AC, Calos MP. Mammalian genomes contain active recombinase recognition sites. *Gene*. 2000 Feb 22;244(1-2):47-54. PubMed PMID: 10689186.

Vetter D, Andrews BJ, Roberts-Beatty L, Sadowski PD. Site-specific recombination of yeast 2-micron DNA in vitro. *Proc Natl Acad Sci U S A*. 1983 Dec;80(23):7284-8. PubMed PMID: 6316354; PubMed Central PMCID: PMC390039.

Wu S, Ying G, Wu Q, Capecchi MR. Toward simpler and faster genome-wide mutagenesis in mice. *Nat Genet*. 2007 Jul;39(7):922-30. Epub 2007 Jun 17. PubMed PMID: 17572674.

Zambrowicz BP, Imamoto A, Fiering S, Herzenberg LA, Kerr WG, Soriano P. Disruption of overlapping transcripts in the ROSA beta geo 26 gene trap strain leads to widespread expression of beta-galactosidase in mouse embryos and hematopoietic cells. *Proc Natl Acad Sci U S A*. 1997 Apr 15;94(8):3789-94. PubMed PMID: 9108056; PubMed Central PMCID: PMC20519.

Table 3.1. List of selected mutant lox and FRT sequences.

Name	Sequence	References
wild type loxP	ATAACTTCGTATA GCATACAT TATACGAAGTTAT	Hoess et al. 1982
lox2272 [†]	ATAACTTCGTATA GgATACtT TATACGAAGTTAT	Lee and Saito 1998
lox5171 ^{††}	ATAACTTCGTATA GtAcACAT TATACGAAGTTAT	Lee and Saito 1998
lox511 ^a	ATAACTTCGTATA GtAtACAT TATACGAAGTTAT	Hoess et al. 1986
lox71 ^b	taccgTTCGTATA GCATACAT TATACGAAGTTAT	Albert et al. 1995
lox66 ^b	ATAACTTCGTATA GCATACAT TATACGAAcggta	Albert et al. 1995
wild type FRT	GAAGTTCCTATAC TTTCTAGA GAATAGGAACTTC	Broach et al. 1982, Andrews et al. 1985
F3 mutant ^{††}	GAAGTTCCTATAC TTcaaAta GAATAGGAACTTC	Schlake and Bode 1994
F5 mutant ^{††}	GAAGTTCCTATAC TTcaaAag GAATAGGAACTTC	Schlake and Bode 1994

The wild type sequences are provided as a reference. Small caps indicate differences with corresponding wild type sequence.

[†]Recombines only with a mutant site having the same sequence

^amay have significant cross-reaction with loxP

^blox71 and lox66 recombine to generate incompatible wild type loxP and lox78 sites.

Figure labels

Figure 3.1. LoxP (a) and FLP (b) sequences showing positions of inverted repeats and nonpalindromic spacer region.

Figure 3.2. Diagram of tyrosine recombinase-mediated DNA recombination. Two DNA duplexes (gray and black lines) are bound by four recombinase molecules (gray rounded rectangles). The black star represents DNA cleavage. One strand of each DNA duplex is cleaved (gray line), exchanged and then ligated to form a Holliday structure. Isomerization allows cleavage, exchange and ligation of the other strands (black lines). Finally, the recombined products are released.

Figure 3.3 Types of recombination activities mediated by Cre or FLP recombinase. The three basic outcomes are (a) integration, (b) excision and (c) inversion.

Figure 3.4. Basic strategies for site-specific integration of transgenes. (a) integration, (b) excision, (c) inversion, and (d) recombination-mediated cassette exchange (RMCE).

Figure 3.5. Cre recombinase-mediated excision-inversion reaction to reconstitute expression of a fluorescent transgene. Diagram shows targeting vector between exon 1 and 2 of the human homolog of the ROSA26 locus. Lox P: filled triangles, lox2272: white triangles, SA: splice acceptor. Two scenarios are presented. Cre-mediated activity catalyzes inversion of sequences between two loxP (left panel) or two lox2272 (right panel) sites. As a second Cre-mediated event, two lox2272 (left panel) or two loxP (right panel) sites recombine resulting in excision of the neomycin resistance marker. In both scenarios, expression of the tdRFP transgene is reconstituted.

Figure 3.6. Use of Cre recombinase for site-specific introduction of EGFP transgene under the control of tetracycline regulatory elements. Figure shows targeted HPRT locus and the integration cassette. Cre activity will result in excision of the PGK promoter-driven neomycin resistance marker and the integration of a cassette carrying elements of the tetracycline inducible system. Integration of the cassette also results in reconstitution of hygromycin expression.

Figure 3.7. RMCE approach to generate human ESC reporter line. LoxP (black triangles) and lox2272 (white triangles) in the targeted HPRT locus pair with their corresponding partners in the exchange cassette to facilitate transgene swapping. As the end result of RMCE, a human Oct4 (hOct4) promoter-driven EGFP transgene is introduced and expression of blastocidin (bsd) resistance marker is reconstituted. EF1a prom: elongation factor 1 alpha promoter, neo: neomycin resistance marker.

Figure 3.1.

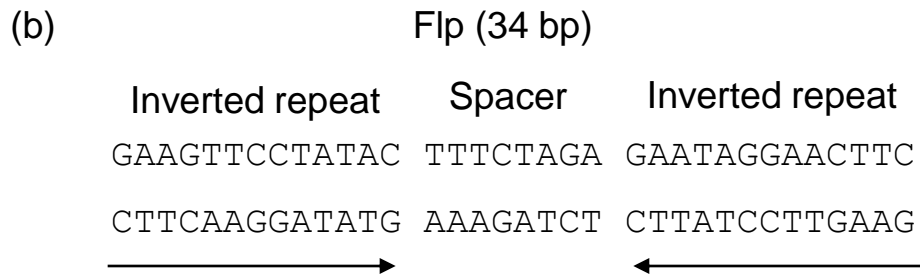
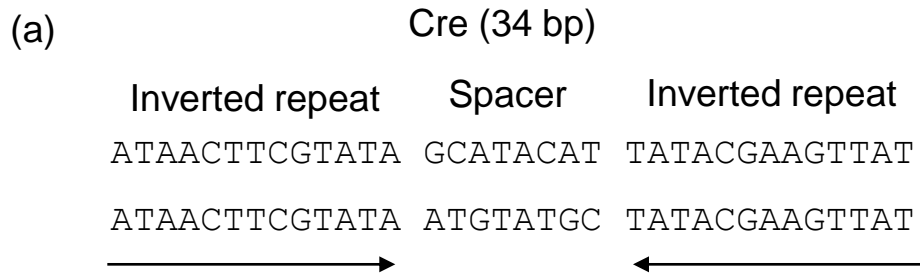
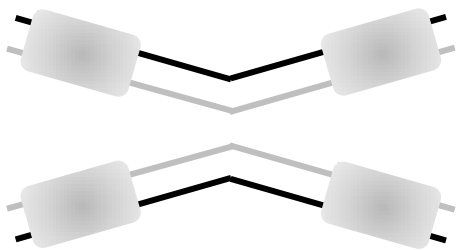
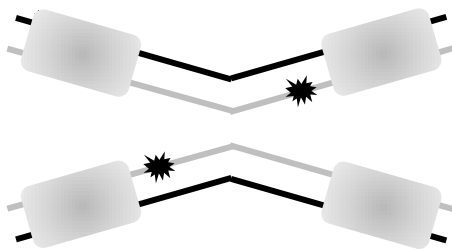


Figure 3.2.

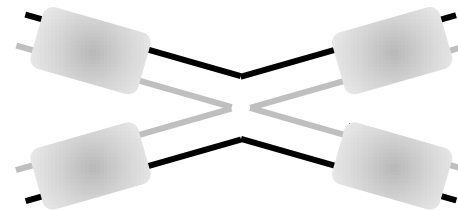
Sequence recognition and pairing



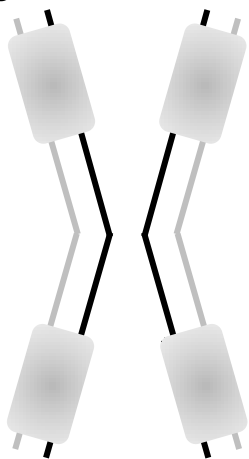
Cleavage



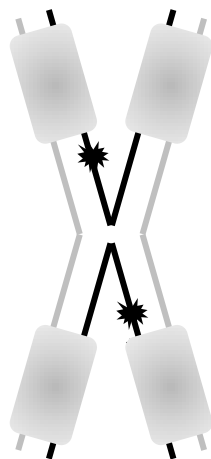
Strand exchange and ligation



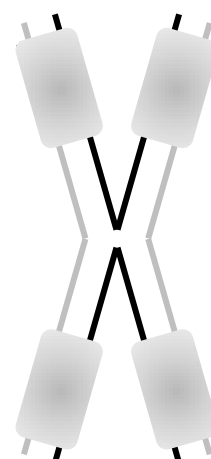
Ligation and release



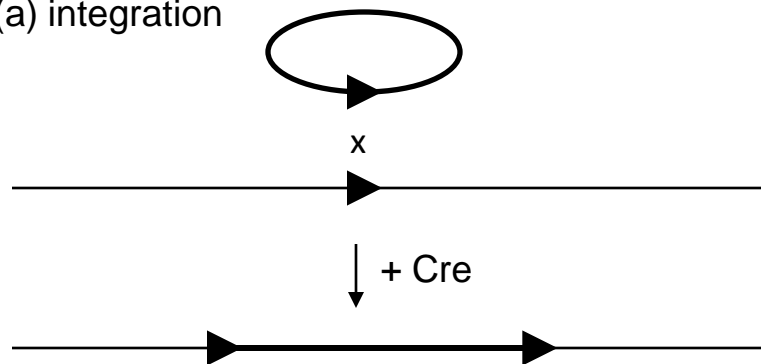
Cleavage



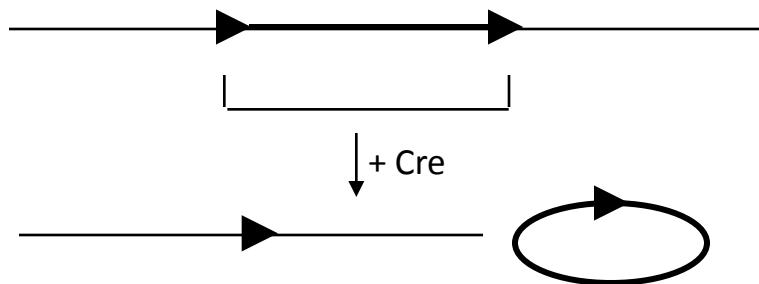
Isomerization



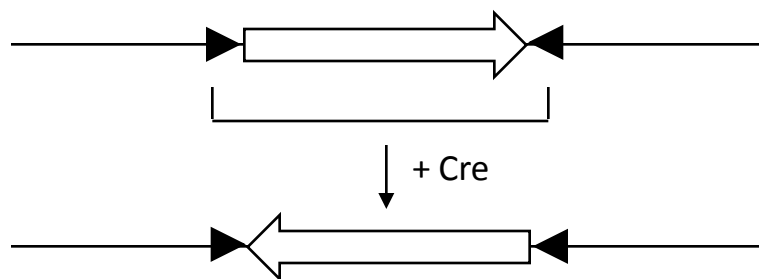
(a) integration



(b) excision



(c) inversion



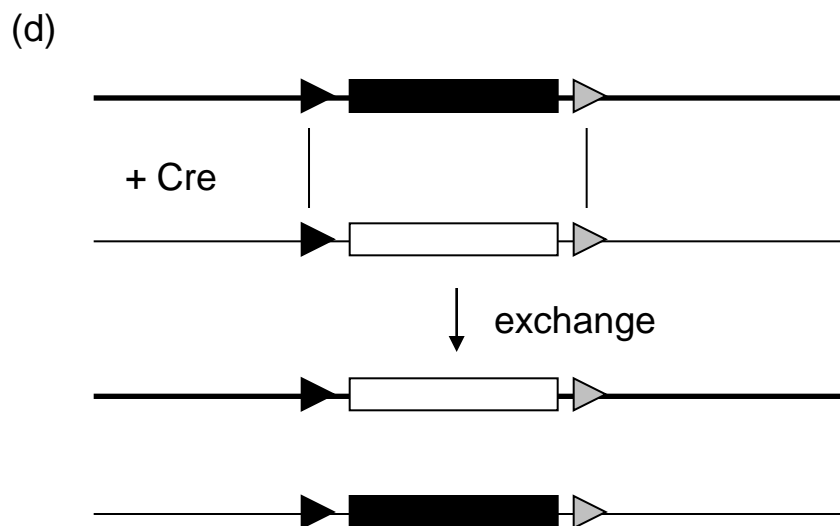
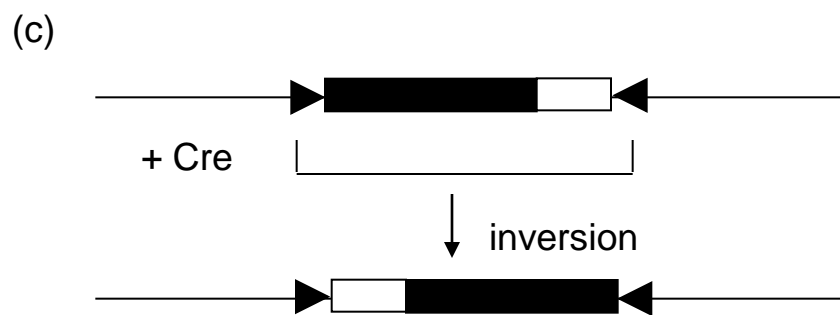
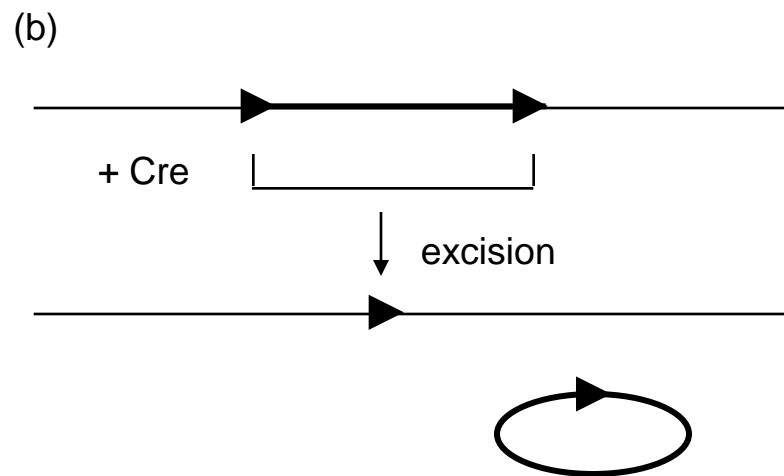
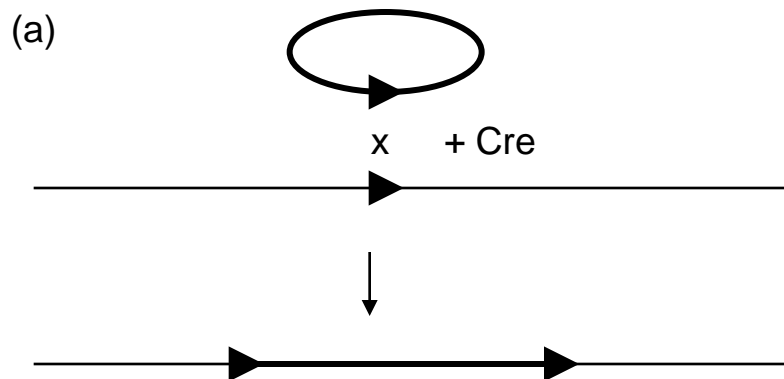


Figure 3.5.

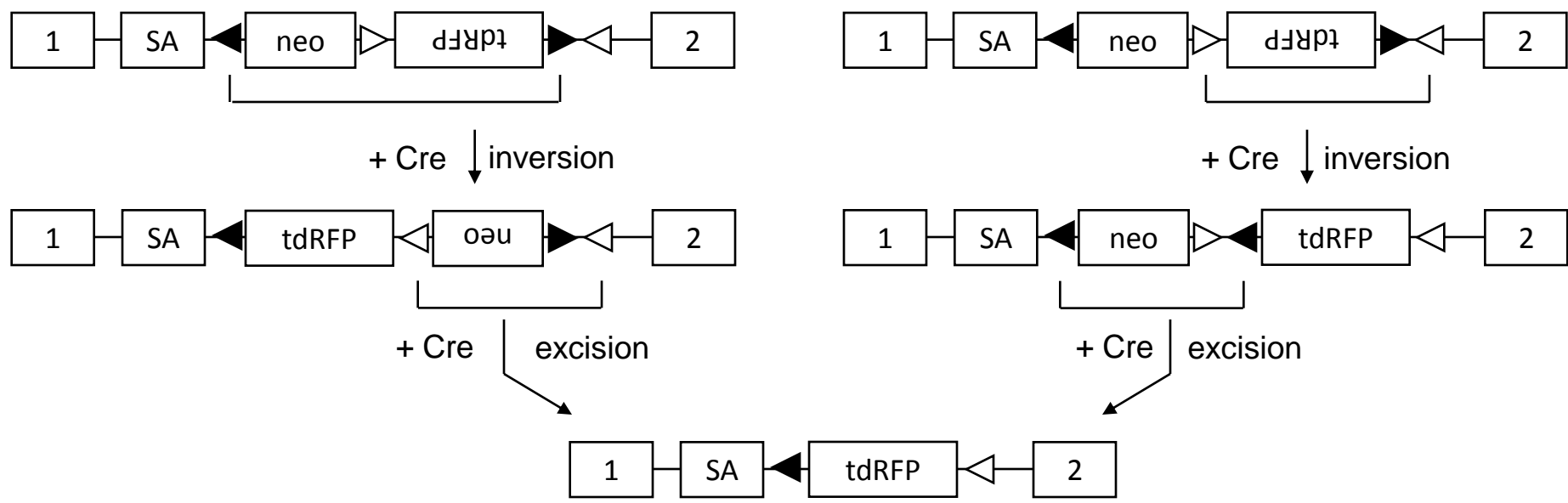


Figure 3.6.

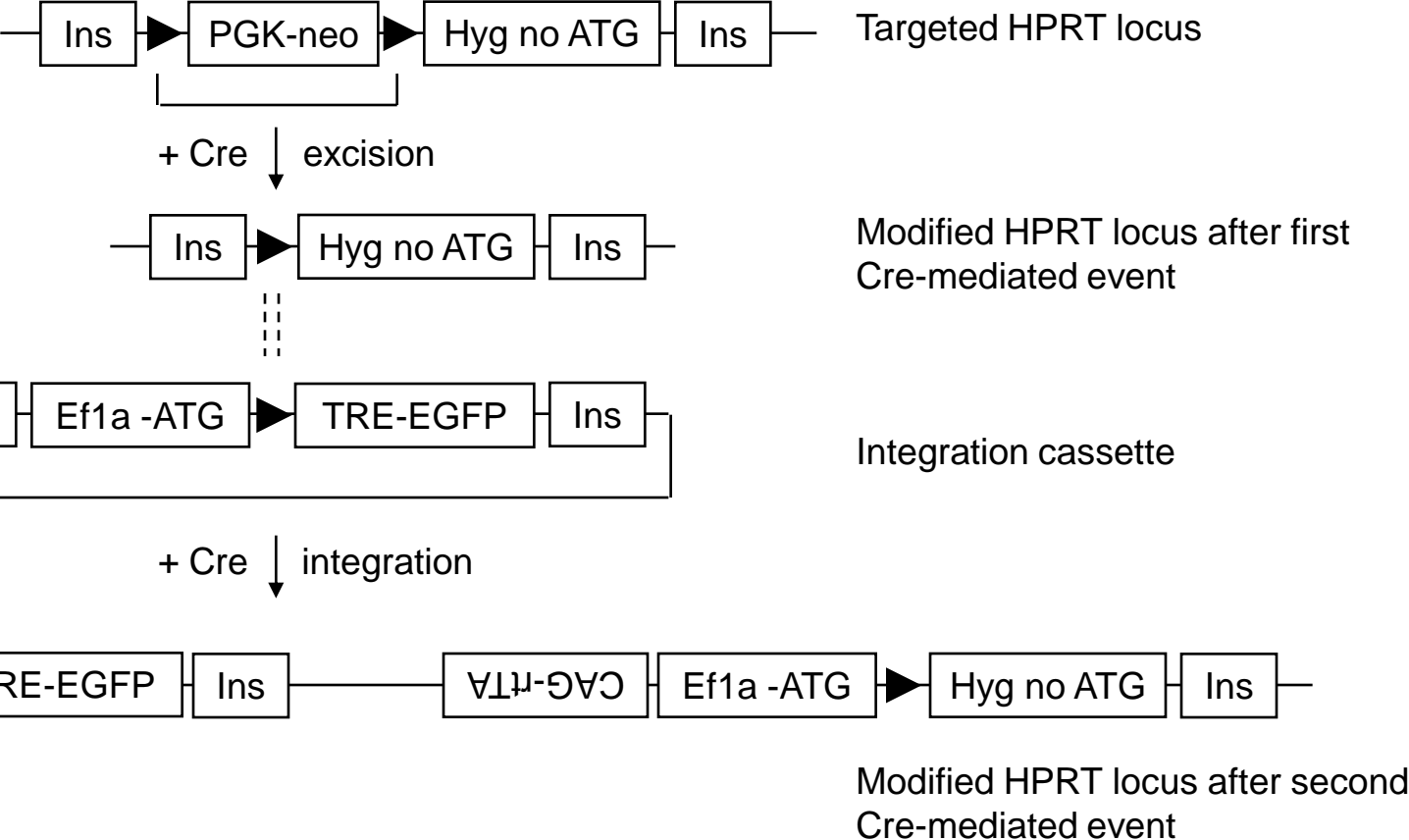


Figure 3.7.

