Phage integrases for the construction and manipulation of transgenic mammals

Roger P Hollis1,2,6, Stephanie M Stoll1,3,6, Christopher R Sclimenti1,4, Jennifer Lin5, Yanru Chen-Tsai5 and Michele P Calos*1

Address: 1Department of Genetics, Stanford University School of Medicine, Stanford, CA 94305-5120, USA, 2Present address: Children’s Hospital of Los Angeles, Research Immunology/BMT, Mail Stop #62, 4650 Sunset Boulevard, Los Angeles, CA 90027, USA, 3Present address: University of California San Francisco, Department of Surgery, 513 Parnassus Ave, HSW 1622, Box 0522, San Francisco, CA 94122-0522, USA, 4Present address: Poetic Genetics LLC, 8638 Mitten Rd., Burlingame, CA 94010, USA, 5Stanford Transgenic Research Facility, Stanford University School of Medicine, Stanford, CA 94305-5120, USA and 6Joint first authors

Email: Roger P Hollis - RHollis@chla.usc.edu; Stephanie M Stoll - ststoll@itsa.ucsf.edu; Christopher R Sclimenti - csclimenti@poetgene.com; Jennifer Lin - jennifer.lin@stanford.edu; Yanru Chen-Tsai - ychen@stanford.edu; Michele P Calos* - calos@stanford.edu

* Corresponding author

Abstract

Phage integrases catalyze site-specific, unidirectional recombination between two short att recognition sites. Recombination results in integration when the att sites are present on two different DNA molecules and deletion or inversion when the att sites are on the same molecule. Here we demonstrate the ability of the φ31 integrase to integrate DNA into endogenous sequences in the mouse genome following microinjection of donor plasmid and integrase mRNA into mouse single-cell embryos. Transgenic early embryos and a mid-gestation mouse are reported. We also demonstrate the ability of the φ31, R4, and TP901-1 phage integrases to recombine two introduced att sites on the same chromosome in human cells, resulting in deletion of the intervening material. We compare the frequencies of mammalian chromosomal deletion catalyzed by these three integrases in different chromosomal locations. The results reviewed here introduce these bacteriophage integrases as tools for site-specific modification of the genome for the creation and manipulation of transgenic mammals.

Background

The ability to manipulate the mammalian genome in vivo is important in many areas of research. There is a growing need for site-specific recombinases that function in the mammalian environment to catalyze genomic integration and excision, for example in the construction of transgenic organisms and the study of gene function in vivo. The site-specific integrases of bacteriophages φC31, R4, and TP901-1 have been shown to function in human and mouse cells to catalyze excision on extrachromosomal vectors and chromosomal integration [1–4]. Phage φC31 integrase has been applied to achieve chromosomal integration in gene therapy studies [5–9]. These results suggested that these site-specific integrases would likely function in the mammalian embryonic environment and could have utility in the creation of transgenic animals and manipulation of the mammalian genome in vivo. The use of bacteriophage integrases in mammalian transgenesis is a new field. This article will discuss some of the potential applications of site-specific integrases in this area and the preliminary results our lab has obtained to date.
Site-Specific Integration for the Construction of Transgenic Mammals

Transgenic mammals are an essential component of biotechnology, having a wide variety of applications ranging from gene knockout studies to serving as living bioreactors. A common method of transgenic mammal production relies on microinjection of DNA directly into the pronucleus of a single-cell embryo, with genomic integration generally resulting by random integration. Alternatively, mouse embryonic stem (ES) cells can be modified by homologous recombination and placed into a developing embryo. The latter approach results in a mosaic animal that may or may not have germ cells modified. Random integration and homologous recombination both occur at low frequencies (10⁻⁴ and 10⁻⁶, respectively). Other methods have been developed for DNA delivery and maintenance in ES cells including mammalian artificial chromosomes [10,11], retroviruses [12–14], and the Sleeping Beauty transposon [15]. Nuclear transfer from more mature cells can also be used to introduce a new genome for a developing embryo, and this nuclear DNA can be first modified by any of the above methods [16,17].

The available technologies are hindered in that gene insertion is often random. In the resultant transgenic animals, the expression levels of the genes of interest are influenced by context factors that are difficult to control. For example, some influencing factors include the regulatory and bacterial elements in the insertion construct itself [18], the position of the insertion in the genome [19], and the number of integrated constructs [19,20]. Production of a desired clone often takes many rounds of screening. Following screening there may be additional problems caused by the transgene insertion resulting in sterility or pre-adolescent lethality, preventing preservation of the newly developed strain. Efficient, targeted single copy integrations would be helpful for the improvement of transgenic animal production. Site-specific integration systems could potentially be useful in this context. A bacteriophage integrase, such as that of phage φC31, can be used to integrate a donor plasmid into the mammalian genome, and integrants can be screened by a simple PCR reaction to identify integrations into known and desirable targets. A quantitative PCR approach can be used to ensure that only single integration events are selected. In the creation of multiple transgenic animal lines, candidate clones can be screened and selected so that the genomic integration event in each animal line is in the identical genomic location and orientation, thereby eliminating positional effect differences between the different transgenic lines. Any genomic location is likely to exert context effects that affect the expression of an integrated transgene, but the use of a site-specific integration system will help to ensure that these effects are not varied among different lines. To date, site-specific recombinases have not been used for the creation of transgenic mammals, though some, such as Cre, have been valuable for genomic manipulations in transgenic animals, as discussed below.

Phage Integrase Biology in the Mammalian Environment

Members of the serine family of site-specific recombinases [21], including integrases from phages φC31 [1,2,5,6,9], R4 [3], and TP901-1 [4], are good candidates for mammalian genomic modification. In nature, each integrase recognizes a sequence in the phage genome, the phage attachment site attP. Each integrase also recognizes a bacterial attachment site, attB, in the bacterial genome. Integrase of the serine family of recombinases are capable of recombing these short attP and attB attachment sites, whose minimal sizes are 35–50 bp. The products, two hybrid sites, attL and attR, are no longer substrates for recombination by the integrase, so the reactions are unidirectional [22]. While the integrase alone can only catalyze a recombination between an attP and an attB site, the result of the recombination can vary based on the location and orientation of these two att sites. When located on different DNA molecules, recombination between an attP and an attB results in an integration event. However, if the two sites are located on the same DNA molecule (plasmid or chromosome), two outcomes are possible: if the attP and the attB are in the same orientation on the DNA, the net result will be a deletion event; if the two att sites are in opposite orientations, the net result will be an inversion. Thus, by knowing the locations and orientations of compatible att sites, specific recombination events can be created and selected for.

We have demonstrated that the φC31 integrase is able to mediate efficient integration of attB-bearing plasmids into the human and mouse genomes, at inserted wild-type attP sites and at endogenous "pseudo" attP sites that have partial identity to attP [2]. We previously identified 26 pseudo attP sites in the human genome, and 57 pseudo attP sites in the mouse genome [2]. We have not detected the presence of pseudo attB sites in mammalian genomes, which should reduce the likelihood that endogenous genomic sequences are inadvertently deleted during an integration attempt. There may not be pseudo attBs that could be recognized and recombined with a pseudo attP by the integrase.

Site-Specific Chromosomal Deletions

In addition to chromosomal integration events, often times a chromosomal deletion is desired, such as in the generation of knockout transgenic animals. One way to accomplish deletion of a sequence of interest is through the use of site-specific recombinases, such as the Cre recombinase of coliphage P1 and its loxP recognition sites [23]. The deletion can be precisely confined to a specific
developmental time or tissue, based on control of recombinase expression. This outcome is most commonly achieved by creating a transgenic mouse that contains the gene of interest flanked byloxP sites, and then crossing that mouse with a Cre-expressing mouse [24]. The Cre gene may, for example, be under a tissue-specific promoter. Once crossed, the gene of interest will be deleted in all cells in which Cre is expressed, thereby generating a tissue-specific knockout [24–30]. The same approach can be used to activate a gene of interest, by inserting a stop cassette between the promoter and the gene of interest and flanking the stop cassette withloxP sites [31]. If the gene knockout is desired in the entire animal, the "floxed"-transgenic mouse oocytes can be transfected with a Cre-expression plasmid or Cre mRNA, so that expression is transient, but sufficient to delete the gene of interest at an early developmental stage, thereby producing an animal devoid of the gene of interest [32,33].

Alternatively, excision can be controlled in a temporalspecific manner by expressing Cre from an inducible promoter. A fusion protein of Cre recombinase and a mutated hormone-binding domain of the murine estrogen receptor showed temporal-specific activity in transgenic mice that was controlled by the presence of tamoxifen [34]. Temporal-specific recombination has also been accomplished by driving Cre expression from inducible promoters, such as a tetracycline-inducible promoter [35]. Simultaneous temporal- and tissue-specific Cre expression has been demonstrated using tissue-specific promoters to drive expression of the Cre ORF fused to a hormone receptor ligand-binding domain [29,36].

Some circumstances require more complicated deletion/insertion events, such as sequential gene inactivation and exon swapping [37]. Additionally, it may be desirable to remove extraneous sequences in an integrated vector, such as bacterial or mammalian drug selection markers [18]. The FLP recombinase from the 2µ plasmid of S. cerevisiae has been used in transgenic mice as an alternative to Cre [38,39] and in combination with Cre [40], although FLP functions at only ~10–25% efficiency compared to Cre in generating chromosomal deletions [41]. The Cre and FLP recombinases are the only two enzymes currently available for genomic manipulation experiments, making complicated experimental schemes difficult to perform [25,40]. The increasing desire to perform complex genomic manipulations creates a need for additional site-specific recombinase tools that function efficiently in the mammalian chromosomal environment.

Results

Phage φC31 Integrate Functions in the Mouse Embryonic Environment and Mediates Genomic Integration

To demonstrate that the φC31 integrase would be functional in the embryonic environment, single-cell embryos of the FVB/NafCBr mouse strain were injected with 3 ng/µl nucleic acid solutions directly into the pronucleus. Two groups of injected embryos were studied. One group received pCMVInt plasmid DNA [1] as a source of integrase (n = 31), and the other received φC31 integrase mRNA (n = 32). Both groups were co-injected with the intramolecular-integration assay vector pBCPB [1] carrying attB and attP sites. The embryos were frozen at the two-cell stage, total DNA was extracted [42], and the expected wild-type attL junction was PCR amplified from recombined plasmid. A product of correct size was observed in both groups of embryos. This product is specific for the recombination event and is not present unless φC31 integrase is expressed [1]. Embryos co-injected with integrase mRNA produced a band of correct size with just ~25% of the template that was required to detect the same band from the pCMVInt/pBCPB injections. This result suggested that integrase mRNA was more efficient in the embryonic environment than DNA, presumably due to the elimination of the requirement for transcription. This result was consistent with observations that injection of Cre mRNA into mouse oocytes provided efficient Cre activity in vivo [33]. The use of mRNA instead of DNA also precluded undesirable integration of the integrase gene by random integration.

Phage φC31 integrase can efficiently catalyze a site-specific cassette exchange reaction using introduced wild-type att sites in mouse ES cells [43] and can recognize and insert attB-containing plasmids into pseudo attP sites in the murine genome [2,6]. A hotspot for φC31 integrase-mediated integration was identified in hepatocytes on chromosome 2 (GenBank: AC 079573) and named mpsL1 (mouse pseudo attP site in liver) [6]. Integration at mpsL1 was also detected in mouse NIH3T3 cells [2], indicating that mpsL1 may be a preferred integration site in the mouse genome. A GFP vector, pEGFPB2, carrying an attB site was injected into one-cell embryos either with (n = 35) or without (n = 27) φC31 integrase mRNA. GFP expression was used to indicate visually the number of embryos with integrated DNA, because non-integrated DNA should be gradually lost during a 5 day period of rapid cell division to the multi-cell or blastocyst stage. By fluorescent microscopic analysis, the two groups had similar numbers of detectably GFP-positive embryos, 6/35 with integrase and 5/27 without integrase. Embryos were positive for GFP expression in all cells; however, the group that received the integrase mRNA contained embryos with an overall more intense green fluorescence.
DNA extracted from the embryos injected with pEGFPB2 with and without φC31 integrase mRNA was used as a template in a PCR assay designed to amplify the attL junction produced when φC31 integrase mediated site-specific integration at mpsL1. The group that received φC31 integrase mRNA produced a band of predicted size that was subsequently cloned and sequenced. The DNA sequence from the PCR product, shown in Figure 1A, demonstrated φC31 integrase-mediated site-specific integration at mpsL1 and proved that site-specific integration occurred within the embryonic genome. No integration into mpsL1 was detected in the GFP-positive embryos that did not receive integrase.

**Construction Of Site-Specific Integrant Transgenic Animals**

FVB/NacfBR single-cell embryos were injected with an attB-bearing donor plasmid, pBChFIXattB [6], with or without φC31 integrase mRNA. Injected embryos were implanted into surrogate mothers at ~20 embryos per animal. Pups were harvested mid-gestation at approximately 2 weeks post-implantation, and genomic DNA was isolated and screened for integrated donor plasmids by PCR. To determine whether an animal contained an integrated donor vector, PCR was employed to amplify a 406-bp internal hFIX fragment of the donor plasmid. To determine whether donor vector was integrated site-specifically by a φC31-mediated event, PCR primers were designed to distinguish an intact attB, signifying random integration, from an attL, signifying φC31-mediated integration. Animals whose DNA produced a 406-bp donor band but not an intact attB band were likely to have site-specifically disrupted the attB sequence, indicating an integrase-mediated event (Table 1).

A mid-gestation mouse found to have a plasmid inserted site-specifically in its genome was further studied by PCR and sequence analysis. Using a nested PCR rescue tech-

---

**Table 1: FVB/NacfBR transgenics obtained following injection with attBhFIX-containing vector with or without φC31 integrase mRNA.**

| ng/µl attB DNA/Integrase mRNA | Embryos injected | Embryos implanted | Pups harvested | hFIX positive | φC31-mediated |
|-------------------------------|------------------|-------------------|---------------|---------------|---------------|
| 3 DNA + 0 mRNA                | 68               | 48                | 16            | 0             | 0             |
| 3 DNA + 3 mRNA                | 144              | 102               | 39            | 3             | 1             |
| 3 DNA + 6 mRNA                | 415              | 284               | 66            | 1             | 0             |

**Figure 1**

Integration at mouse genomic sequences. (A) Sequence from an embryo-derived PCR clone displays the characteristic site-specific crossover between the attB arm, shown in gray, and the murine pseudo attP site, mpsL1, with the TT core (in bold) located at the point of crossover. A one-base mismatch (underlined) located downstream of the crossover event may be a polymorphism between the mouse FVB/NacfBR strain and the C57Bl/6 strain that was sequenced in the database. (B) Comparison of the sequences of embryo-derived pseudo attP sites mpsL1 and mpsT1 with the wild-type φC31 attP. The common TT core is shown in bold. Both of these embryo-derived pseudo attP sites are ~30% identical to the wild-type φC31 attP. The matching base pairs with a 40 base attP are shown in gray.
nique [2], it was determined that this animal contained the donor plasmid site-specifically integrated into a sequence on chromosome 4 termed mpsT1 (mouse pseudo attP site in Transgenic 1), shown in Figure 1B. This integration site is located in the last intron of a putative gene detected in a 16-day neonate (Genbank: AK052928). The integration event did not prevent embryogenesis in this experiment, indicating that an integration event at mpsT1 was unlikely to disrupt development and this location was potentially a suitable target site for future transgene integrations. Like the mpsL1 pseudo site, mpsT1 has limited identity to the wild type attP sequence (Figure 1B). Chromatin structure may influence which pseudo sites are used [6]. Repeatedly targeting an identical genomic locus such as mpsT1 could allow for the creation of multiple transgenic lines with transgenes integrated at the same, predictable genomic locus. Creation, by directed evolution [44], of integrases with tighter site-specificity may allow for such a possibility.

The use of an engineered mouse strain containing an inserted wild-type attP site could help resolve position effects by providing a method to preferentially target transgenes to a predetermined location. The use of an attP transgenic mouse strain might also increase the integration frequency, because the integrase might work more efficiently at a perfect attP site, compared to pseudo attP sites. The detection of site-specific integrants would also be simplified because the sequence of the target site was known. Therefore, similar experiments were performed on a strain of transgenic mice we prepared that contained a wild-type φC31 attB site. Injections into this attB-bearing mouse strain were performed with φC31 mRNA and the attB donor plasmid pBChFiXattB. Of 7 injected pups that carried the wild-type attP site, none were found to have an insertion of the attB donor plasmid anywhere in their genomes. Although the number of pups examined was low, this negative result may indicate a poor genomic location of the φC31 attP site in this transgenic strain, with respect to chromatin access for the φC31 integrase. The creation of a strain with a more accessible attP site might result in a higher percentage of φC31-mediated integrations at the attP site.

Chromosomal Deletions Catalyzed by Phage Integrases

The integrases of phages φC31, R4, and TP901-1 function efficiently in mammalian cells to complete precise deletions on extrachromosomal introduced DNA [1–4]. It was recently shown that the φC31 integrase could perform intrachromosomal deletions at a frequency ~10% that of Cre [41]. Addition of a nuclear localization signal (NLS) to φC31 integrase increased this frequency to ~80% of that of Cre [41]. The ability of R4 and TP901-1 integrases to catalyze chromosomal deletions was unknown. To facilitate their use as genomic tools, it would be useful to have a direct comparison of chromosomal deletion activity among these site-specific integrases. To determine the frequency of chromosomal deletions catalyzed by these integrases, we established human cell lines with an integrated copy of a vector containing the φC31, R4, and TP901-1 attB and attP recognition sites (Figure 2). Transfection of these cell lines with integrase expression plasmids, followed by quantitative PCR and Southern analysis, allowed us to calculate the chromosomal deletion frequency of each integrase. Because all three recognition site pairs were in the same vector, and thus in the same chromosomal context in each cell line, we were able to obtain a directly comparable measure of the chromosomal deletion frequency mediated by the φC31, R4, and TP901-1 integrases.

Mammalian expression vectors were created for each of the three integrases (Figure 2), in which the cytomegalovirus immediate early (CMV) promoter drove expression of the integrase, followed by an internal ribosome entry site (IRE), and the gene for green fluorescent protein (GFP). Cell lines were transfected with a plnt-l-hrGFP vector, and GFP-positive cells were sorted 72 h post-transfection, expanded, and genomic DNA was prepared. Separately for each sample, the total integrated vector copy number was determined by quantitative amplification of a region of the hygromycin resistance gene. The number of deletion products was determined by quantitative amplification of attR junctions. Deletion frequency was calculated as the ratio of attR junctions to total vector copies. The results are shown in Figure 3A.

Quantitative PCR analysis of TP901-1 integrase-catalyzed deletion was unsuccessful. The TP901-1 attR is highly A+T-rich, making Taqman primer/probe design difficult. Because we could not accurately determine TP901-1 deletion frequency by qPCR, Southern blot analysis was performed. This analysis was done for φC31 and R4 as well, in order to directly compare with the TP901-1 frequency and to independently verify the deletion frequency determined by qPCR. Genomic DNA from integrase-transfected cell lines was analyzed by Southern blot, and deletion frequency was quantified. Chromosomal recombination resulted in reduction of the genomic vector by deletion of 4.8 kb (φC31), 3.7 kb (R4), or 4.2 kb (TP901-1) of intervening vector sequence. Representative Southern blots of φC31-, R4-, and TP901-1-transfected cell line genomic DNA are shown in Figure 3B. Quantification of band intensity indicated that Southern analysis of chromosomal deletion frequency was in agreement with that calculated by qPCR analysis for the φC31 and R4-transfected cell lines.

The φC31 integrase consistently performed more efficiently than the R4 and TP901-1 integrases. In cell line...
293-3BP3, chromosomal deletion was catalyzed by the \( \Phi C31 \) integrase with an average frequency of 51%, by the R4 integrase at an average frequency of 37%, and by the TP901-1 integrase at a frequency of 39%. The R4 and TP901-1 integrases performed equally well, though at a lower frequency than the \( \Phi C31 \) integrase, in cell line 293-3BP6. The frequency of deletion catalyzed by the \( \Phi C31 \) and R4 integrases was lower in all other cell lines to varying degrees, while TP901-1 integrase showed no measurable activity in these other cell lines.

The chromosomal deletion frequency of the \( \Phi C31 \) integrase ranged from 51% in cell line 293-3BP3 to only 1.7% in the 293-3BP8 cell line. Similar ranges were observed for the R4 integrase. These variations in activity most likely represented a chromosomal position effect of the integrated vectors in different regions of the genome, which may affect the ability of the integrase to efficiently access the \( \text{att} \) sites. We have hypothesized that the \( \Phi C31 \) integrase preferentially integrates into pseudo \( \text{attP} \) sites that are in actively expressed regions; an open chromatin configuration may be required for efficient targeting and recombination by the integrase [6]. This hypothesis was supported by the apparent chromosomal position effects on excisive recombination observed here. The TP901-1 integrase appeared to be even more susceptible to chromosomal position effects, as evidenced by its ability to perform deletions at ~35% efficiency in two of the cell lines, the two with the highest deletion frequencies by \( \Phi C31 \) and R4 integrases, but its lack of deletion formation in the other cell lines analyzed.

Discussion
The experiments reported here suggest that the \( \Phi C31 \) integrase can be used to create transgenic mice carrying site-specific transgene insertions. The numbers of \( \Phi C31 \)-mediated transgenic animals generated were low, with approximately 2.6% of harvested pups representing \( \Phi C31 \) integrase-mediated transgenic animals, in the best experiment. Our sequence data demonstrated that the \( \Phi C31 \) integrase functioned in the embryonic nuclear environment to integrate plasmid DNA site-specifically into the genome. Additional experiments with greater numbers of animals will be required to optimize the procedure and determine the efficiency of such an approach for generating transgenic mammals. Such a tool would be beneficial because the genomic location of the integration site, either at an inserted wild-type \( \text{attP} \) site or at a known endogenous pseudo \( \text{attP} \) site, can be specifically detected by PCR amplification, potentially allowing the creation of multiple transgenic lines with different integrated transgenes in the same genomic context. This feature would permit creation and comparison of multiple transgenic animal strains without confounding chromosomal posi-
We also demonstrated that the integrases of bacteriophages ϕC31, R4, and TP901-1 performed efficiently in creating intrachromosomal deletions and should be of great value for manipulating mammalian genomes, both in vitro and in vivo. The maximum observed deletion frequencies presented here were in the range of 30–50%, but this range probably underestimates the activity of these integrases. The criteria used for sorting GFP-positive, integrase-transfected cells may have included untransfected cells. Following expansion of these cells, quantitative analysis of chromosomal deletion frequency would be depressed by the presence of cells that were never exposed to integrase. However, even at the observed frequencies of 30–50%, the ϕC31, R4, and TP901-1 integrases constitute efficient tools for the manipulation of mammalian chromosomes. In the experiments reported here, the ϕC31, R4, and TP901-1 integrases used were in their native prokaryotic form, with no addition of an NLS. A ϕC31-NLS integrase [41] may function significantly better, nearing 100%, in this chromosomal deletion assay, and addition of an NLS may have a similar positive effect on the efficiencies of the R4 and TP901-1 integrases. Such chromosomal deletion frequencies would make the ϕC31, R4, and TP901-1 integrases equivalent to Cre recombinase and valuable as molecular tools for mammalian chromosome engineering. Additional experiments will need to be conducted in order to determine the efficiency of these integrases in catalyzing chromosomal deletions in vivo. These serine family site-specific phage integrases clearly constitute new and valuable tools for genome engineering.

**Materials and Methods**

**Transgenic Animal Experiments**

**Vector construction**

The ϕC31 integrase gene was amplified to include the NdeI restriction site from pTA-Int [1] by using the primers "Native 5' NdeI (+)" (CGA CTA GTC ATA TGG ACA CGT ACG CGG GTG CT) and "3' BamHI ϕC31" (AGC CGG ATC CGG GTG TCT CGC TA). The pET11 vector (Novagen, Madison, WI) received the amplified ϕC31 integrase gene directionally inserted into NdeI and BamHI sites so that the T7 promoter drove ϕC31 integrase mRNA production (pET11ϕC31). The oligonucleotides BamHI-polyA-top (GAT CGC GCC AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA CGG) and BamHI-polyA-bottom (GAT CGG GTT TTT TTT TTT TIT TTT TIT TTT TTT TTT TTT CCG CGC) were annealed, digested with BamHI, and inserted into the BamHI site in pET11ϕC31 to create pET11ϕC31polyA. The assay construct pEGFPB2 is based on the vector pEGFP-N2 (Clontech, Palo Alto, CA). A lac promoter to drive GFP expression in bacteria and codons for the first 26 amino acids of β-galactosidase were cloned into the Nhel and HindIII sites within the multicloning site of pEGFP-N2, placing it in frame with eGFP to create...
pCMVlacGFP. The attB fragment was released from pTA-attB [1] with EcoRI, and the ends of the fragment were blunted. The blunt attB fragment was inserted into pCMVlacGFP vector blunted at the DraIII restriction site. This cloning of the attB with the T3 core pointing towards the GFP terminator sequence resulted in creation of the assay vector pEGFPB2.

Nucleic acid preparations for microinjection
The φC31 integrase mRNA was transcribed from the vector pET114φC31polyA using the mMessage mMachine (Ambion, Austin, TX). φC31 RNA and DNA were purified and diluted in microinjection TE (miTE) buffer (10 mM Tris and 0.1 mM EDTA, pH 7.4) to reach a final concentration of 3 ng/µl. For the injections 3 or 6 ng/µl solutions of each nucleic acid were used.

PCR screening for integration
Standard embryo DNA retrieval technique for PCR was employed [42]. The primers attBF2 (ATG TAG GTC ACG GTC TCG AAG C) and attP1+ (TGG CGG CCG TCT TAG AAC TA) were used to specifically amplify the wild type attl junction in integrase-reacted pBCPB. A 175 bp product will be amplified. For detection of mpsL1 integration in embryos, a primary PCR using the primers attBF3 (CGA AGC CGG GGT GCG) and mpsL1-R2 (GTA AAT GTT ATT GCG GCT CT) was purified and used as a template in a secondary PCR using attBF3 and the nested primer mpsL1-R1 (TGA GGA GGA GCC TTA GCA AC). The secondary PCR amplifies a nested product of 208 bp. For detection of the GFP transgene, the primers GFPfor (CTG GAC GCC GAC TAG AAT GTA CCG) and GFPrev (GGC GAC GGC GAC GTA AAC GGC CAC) and were used to specifically amplify a 453 bp segment of the GFP transgene. The hFIX positive cells were identified using the primers hFIXend (5'-CCGGGTACACCCGCAGAGTGTACCCACAAGCAGTACCACTC-3', followed by 5'-CCGGGTACACCCGCAGAGTGTACCCACAAGCAGTACCACTC-3') and will be referred to as pBB-TP(BP). This plasmid contains the attP fragment. For attP transgenic mice, the plasmid pTA-attP333 [4] carries the full length TP901-1 attP site. The φC31 attP site was isolated as a blunt SpeI fragment from pBCPB [1] and ligated into the blunted AvaI site of pTA-attP333, creating the plasmid pBI(-). The R4 attP site was isolated from the plasmid PCRCattP64 [3] by digestion with PvuII and SnaBI and blunt-ending with T4 polymerase. It was then ligated into the DraIII Pbf(-), creating the plasmid pTRI(+). pTRI(+) was digested with BglII and HindIII to liberate a 1.9 kb fragment containing the three attP sites. This triple-attP fragment was gel isolated and purified using the QIAexII gel extraction kit (Qiagen, Valencia, CA) and eluted with miTE (10 mM Tris-HCl, 0.1 mM EDTA, pH 7.4). DNA was prepared to a concentration of 25 ng/µl. To eliminate debris from the prep, the DNA was passed through a small 0.45 µm Millipore spin filter (Millipore, Billerica, MA). DNA was transferred to the Stanford Transgenic Research Facility for standard transgenic animal production by microinjection.

Chromosomal Deletion Experiments
Vector construction
The assay plasmid for detecting chromosomal deletions was generated as follows. Plasmid pBB-B304-P333 has been described [4] and will be referred to as pBB-TP(BP). This plasmid contains the full-length TP901-1 attP and attB recognition sites flanking the 3.5 kb lacZ gene. The 64 bp R4 attP site was generated by kinasing and annealing the oligonucleotides 5’-CCGGGCATGTTCCCAAAAGCGATATCACGTTAGAAGCATG GTACTCGTTGTGGTTACACTCTGGGGTGTAC-3’ and 5’-CCGGGTACACCCGACAATGTACCCACAAGCAGTACCACTC-3’ followed by ligation into the XhoI site of pBB-TP(BP), creating the plasmid pBB-TP(BP)-R4(P). Into the SalI site of this plasmid, the 224 bp 5’-CC31 attP site was cloned as an XbaI-SalI fragment from plasmid pBCPB+ [1], resulting in the plasmid pBB-TP(BP)-R4(P)-φC31(P). Into the XhoI site of this plasmid, the 313 bp 5’-CC31 attB site was cloned as an XhoI fragment released from pBCPB+, generating plasmid pBB-TP(BP)-R4(P)-φC31(BP). Into the BamHI site of this plasmid, the 295 bp R4 attB site was ligated as an XhoI fragment released from the plasmid pBB-C-R4PB [3], creating the plasmid pBB-TP(BP)-R4(P)-φC31(BP), generating the assay plasmid pBB-TP(BP)-R4(P)-φC31(BP)-Hyg, referred to as plasmid p3BP.

Each integrase was independently cloned into the expression plasmid pIREShrGFP-1a (Stratagene, La Jolla, CA).
The TP901-1 integrase was cloned into the EcoRI-XhoI sites of pIRES-hrGFP-1a as a 1.5 kb EcoRI-XhoI fragment from pCS-TPInt [4], generating the plasmid pTP-I-hrGFP. A 1.4 kb EcoRI fragment from pTA-sre [3], containing the R4 integrase, was cloned into the EcoRI site of pIRES-hrGFP-1a, creating the expression plasmid pR4-I-hrGFP. The φC31 integrase was cloned from pCSI [1] as a 1.9 kb blunted Spel-BamHI fragment and ligated into the Smal-BamHI sites of pIRES-hrGFP-1a, to generate the vector ϕC31-excised plasmid. The φC31 integrase was expressed in E. coli strain DH-Int, which expresses the ϕC31-excised plasmid. The attR junction was PCR amplified and TOPO-cloned into vector pCR2.1 (Invitrogen, Carlsbad, CA), creating the plasmid pTA-ϕex. Next, the R4 attR was generated by intramolecular integration reaction in 293 cells co-transfected with p3BP and pCMV-sre [3], creating p3BP-R4ex. The R4 attR junction was PCR amplified from p3BP-R4ex with primers that added Spel to each end of the attR. The PCR product was digested with Spel and ligated into the Spel site of pTA-ϕex, creating plasmid pTA-ϕex-R4ex. The TP901-1 attR junction was generated by transforming p3BP into DH-TPInts [4], generating plasmid p3BP-Tpex. The TP901-1 attR was then PCR amplified, with addition of BamHI sites to each end. The PCR product was digested with BamHI and ligated into the BamHI site of pTA-ϕex-R4ex, creating the plasmid pTA-ϕex-R4ex-Tpex. Next, a region of human Rad52 gene was PCR amplified from 293 genomic DNA with primers that added Clal sites to each end. The PCR product was digested with Clal and ligated into the Clal site of pTA-ϕex-R4ex-Tpex, creating the plasmid pTA-3attR-Rad52. Finally, a portion of the hygromycin gene was PCR amplified from p3BP, digested with XbaI, and ligated into the XbaI site of pTA-3attR-Rad52, generating the standard vector p3attR-Rad52-Hyg. This vector was linearized with XcmI, purified, and serially diluted for use in qPCR to generate a standard curve.

Cell lines
Plasmid p3BP was digested with XmnI to generate a linear molecule. Five µg of linearized p3BP plasmid was electroporated into 293 cells using a Bio-Rad Gene Pulser according to the manufacturer's recommendations. The cells were then allowed to recover in nonselective Dulbecco’s modified eagle medium, supplemented with 9% FBS and 1% penicillin-streptomycin. After 24 h, the cells were placed under selection using medium containing 200 µg/ml hygromycin B (Calbiochem, La Jolla, CA). After 14–21 d under selection, single, well-isolated colonies were picked and expanded. All cell lines were screened for the presence of intact triple-attP and triple-attB sites by PCR and by Southern analysis. Six 293-3BP cell lines were selected for further analysis.

Chromosomal deletion assay
293-3BP cell lines were transfected with 5 µg of plnt-I-hrGFP plasmid using Fugene 6 (Roche Applied Scientific, Indianapolis, IN), according to manufacturer’s protocol. At 72 h post-transfection, GFP-positive cells were sorted and collected using a MoFloPs FACS machine, thus sorting for cells that received integrase expression plasmid. Approximately 15 d after sorting and expansion, cells were harvested and genomic DNA was prepared using the Blood and Cell Culture DNA Maxi Kit (Qiagen, Valencia, CA).

Quantitative PCR analysis
Quantitative PCR on genomic DNA was performed using the following primer/probe combinations: TP901 attR: fwdTPEX-qPCR 5’-TGTATTCATCTGAGATAATGATATGAT-3’, revTPex-qPCR 5’-ATTAAAAATTCAGGAAAGACGTTT-3’, TPEX-attrR-probe 5’-CGAGTTTTATTTCTCATTATTCAATCAGGTTAAATGTC-3’; Φ C31 attR: fwdPhiEX-qPCR 5’-GGCCCTACGGTCTCCAGGT-3’, revPhiEX-qPCR 5’-CCAGATGCGGTGCGGTAGCT-3’, PhiEX-attrR-probe 5’-CTGCGGTAACCTTTGCCTCCTCC-3’, R4 attR: fwdR4ex-qPCR 5’-TCTCATGCTAGAAGGCCGCCC-3’, revR4ex-qPCR 5’-GGCTACACGGGACGGACC-3’, R4ex-attrR-probe 5’-CGATACCACTGAGCAGTCGAGAAGGC-3’.

The hygromycin resistance gene was also amplified as an internal control for integrated p3BP vector, using the following primer/probe combination: fwdHyg-qPCR 5’-CTCGGACGGCGAATCTC-3’, revHyg-qPCR 5’-CGACGTTTTTTACCTACCG-3’, Hyg-probe 5’-TCAGCTTCGATGAGAGGCG-3’. Each of these probes was 5’ labeled with the fluorophore 6-FAM and 3’ labeled with the quencher TAMRA.

In addition to the primers and probes described above, PCR reagents used were from the TaqMan PCR Core Reagent Kit (Applied Biosystems, Foster City, CA). Amplification was performed in an ABI 7700 machine and the results were analyzed using SDS v1.7 software (Applied Biosystems). Each sample and standard was analyzed in triplicate.

Southern analysis
Genomic DNA was digested overnight with HindIII. Twenty µg of digested genomic DNA was loaded onto a
0.6% TBE agarose gel and separated. The gel was depurinated in 0.25 M HCl, denatured in 0.5 M NaOH, neutralized in 0.5 M Tris- HCl (pH 7.0), and transferred to an S&S Nytran blotting membrane (Schleicher & Shuell, Keene, NH) in 20X SSC transfer buffer. The membrane was probed with a 1.6 kb fragment from the hygromycin resistance gene or with a 1.6 kb genomic fragment from 8p22 encompassing φC31 human pseudo attP site A (hpsA) [2], both labeled with 32P by random primer extension, using the Ready-To-Go DNA Labeling Kit (dCTP) (Amersham Pharmacia Biotech). Hybridization occurred at 65°C for ~20 h in Church buffer (1% BSA, 1 mM EDTA, 0.5 M NaHPO4 pH 7.2, 7% SDS). Membranes were washed in 2x - 0.2x SSC at 65°C. Membranes were exposed to a phosphor screen and analyzed using a Storm 840 PhosphorImager (Molecular Dynamics, Sunnyvale, CA). Southern blot quantitation was performed using Kodak 1D software.

Acknowledgments
This work was supported by NIH grants HL68112 and DK58187 to MPC and by the Stanford Transgenic Research Facility.

References
1. Groth AC, Olivares EC, Thyagarajan B and Calos MP: A phage integrase directs efficient site-specific integration in human cells. Proc Natl Acad Sci U S A 2000, 97:5995-6000.
2. Thyagarajan B, Olivares EC, Hollis RP, Ginsburg DS and Calos MP: Site-specific genomic integration in mammalian cells mediated by phage φC31 integrase. Mol Cell Biol 2001, 21:3926-3934.
3. Olivares EC, Hollis RP and Calos MP: Phage R4 integrase mediates site-specific integration in human cells. Gene 2001, 278:167-176.
4. Stoll SM, Ginsburg DS and Calos MP: Phage TP901-I site-specific integrase functions in human cells. J Bacteriol 2002, 184:3657-3663.
5. Olivares EC and Calos MP: Phage C31 Integrase-Mediated Site-Specific Integration for Gene Therapy. Gene Therapy, Regulation 2003 in press.
6. Olivares EC, Hollis RP, Chalberg TW, Messe L, Kay MA and Calos MP: Site-specific genomic integration produces therapeutic Factor IX levels in mice. Nat Biotechnol 2002, 20:1124-1128.
7. Ortiz-Urda S, Thyagarajan B, Keene DR, Lin Q, Calos MP and Khavari PA: Stable nonviral genetic correction of inherited human skin disease. Nat Med 2002, 8:166-170.
8. Ortiz-Urda S, Thyagarajan B, Keene DR, Lin Q, Calos MP and Khavari PA: PhIC31 integrase-mediated nonviral genetic correction of junctional epidermolysis bullosa. Hum Gene Ther 2003, 14:923-928.
9. Groth AC and Calos MP: Phage integrases: biology and applications. Journal of Molecular Biology 2003 in press.
10. Huxley C: Mammalian artificial chromosomes and chromosome transgenics. Trends Genet 1997, 13:345-347.
11. Wang B, Lazaris A, Lindenaubem M, Stewart S, Co D, Perez C, Drayer J and Karatagliches CN: Expression of a reporter gene after microinjection of mammalian artificial chromosomes into pronucli of bovine zygotes. Mol Reprod Dev 2001, 60:433-438.
12. Nagano M, Brinster CJ, Orwell FE, Ryu BY, Avarbock MR and Brinster RL: Transgenic mice produced by retroviral transduction of male germ-line stem cells. Proc Natl Acad Sci U S A 2001, 98:13090-13095.
13. Nagano M, Watson DJ, Ryu BY, Wolfe JH and Brinster RL: Lentiviral vector transduction of male germ line stem cells in mice. FEBBS Lett 2002, 524:111-115.
14. Lois C, Hong EJ, Pease S, Brown EJ and Baltimore D: Germline transmission and tissue-specific expression of transgenes delivered by lentiviral vectors. Science 2002, 295:868-872.
15. Dupuy AJ, Clark K, Carlson CM, Fritz S, Davidson AE, Markley KM, Finley K, Fletcher CF, Eekers SC, Hackett PB, Horn S and Largaespada DA: A transgenic germ-line vector system for transgenesis by transposition. Proc Natl Acad Sci U S A 2002, 99:4495-4499.
16. Wilmut I, Schnieke AE, McWhir J, Kind AJ and Campbell KH: Viable offspring derived from fetal and adult mammalian cells. Nature 1997, 385:810-813.
17. Rougeot WM 3rd, Hochvedinger K, Kyba M, Daley GQ and Jaenisch R: Correction of a genetic defect by nuclear transplantation and combined cell and gene therapy. Cell 2002, 109:17-27.
18. Nagy A, Moens C, Ivanyi E, Pawling J, Gertsenstein M, Hadjantonakis AK, Pinty M and Rossant J: Dissecting the role of N-myc in development using a single targeting vector to generate a series of alleles. Curr Biol 1998, 8:661-664.
19. Dobie K, Mehtali M, Mcleagahan M and Lathe R: Variegated gene expression in mice. Trends Genet 1997, 13:127-130.
20. Garrick D, Fiering S, Martin DI and Whitelaw E: Repeat-induced gene silencing in mammals. Trends Genet 1999, 15:267-273.
21. Smith MC and Thorpe HM: Diversity in the serine recombinases. Mol Microbiol 2002, 44:299-307.
22. Thorpe HM and Smith MC: In vitro site-specific integration of bacteriophage DNA catalyzed by a recombinase of the resolvase/ invertase family. Proc Natl Acad Sci U S A 1998, 95:5505-5510.
23. Sauer B and Henderson N: Cre-stimulated recombination at loxP-containing DNA sequences placed into the mammalian genome. Nucleic Acids Res 1989, 17:47-161.
24. Orban PC, Chui D and Marshall JD: Tissue- and site-specific DNA recombination in transgenic mice. Proc Natl Acad Sci U S A 1992, 89:6861-6865.
25. Gu H, Marth JD, Orban PC, Mossmann H and Rajewsky K: Deletion of a DNA polymerase beta gene segment in T cells using cell-specific gene targeting. Science 1994, 265:103-108.
26. Indra AK, Warot X, Brocard J, Bornert JM, Xiao JH, Chambon P and Metzger D: Temporally-controlled site-specific mutagenesis in the basal layer of the epidermis: comparison of the recombinase activity of the tamoxifen-inducible Cre-ER(T) and Cre-ER(T2) recombinases. Nucleic Acids Res 1999, 27:4324-4327.
27. Cao T, Longley MA, Wang XJ and Roop DR: An inducible mouse model for epidermolysis bullosa simplex: implications for gene therapy. J Cell Biol 2001, 152:651-656.
28. Peifer M, Metzger D and Chambon P: Temporally controlled targeted somatic mutagenesis in the mouse brain. Eur J Neurosci 2001, 14:1777-1783.
29. Kellendonk C, Tronche F, Casanova E, Anlag K, Opferk Ch and Schutz G: Inducible site-specific recombination in the brain. J Mol Biol 1999, 285:175-182.
30. Minamoto T, Gausain V, DeMayo F and Schneider MD: Inducible gene targeting in postnatal myocardium by cardiac-specific expression of a hormone-activated Cre fusion protein. Circ Res 2001, 88:587-592.
31. Lakso M, Sauer B, Moestinger BJ, Lee JE, Manning RW, Yu SH, Mulder KL and Westphal H: Targeted oncogene activation by site-specific recombinase in transgenic mice. Proc Natl Acad Sci U S A 1992, 89:6232-6236.
32. Araki K, Araki M, Miyazaki J and Vassalli P: Site-specific recombination of a transgene in fertilized eggs by transient expression of Cre recombinase. Proc Natl Acad Sci U S A 1995, 92:160-164.
33. de Wit T, Drabek D and Grosfeld F: Microinjection of cre recombinase RNA induces site-specific recombination of a transgene in mouse oocytes. Nucleic Acids Res 1998, 26:676-678.
34. Zhang Y, Riesterer C, Ayrall AM, Sablitzky F, Littlewood TD and Reth M: Inducible site-directed recombination in mouse embryonic stem cells. Nucleic Acids Res 1996, 24:543-548.
35. Meese AR, Nikitin AY, Avarbock MR, Temporal, spatial, and cell type-specific control of Cre-mediated DNA recombination in transgenic mice. Nat Biotechnol 1999, 17:1091-1096.
36. Yasioukhin V, Degenstein L, Wise B and Fuchs E: The magical touch: genome targeting in epidemical stem cells induced by tamoxifen application to mouse skin. Proc Natl Acad Sci U S A 1999, 96:8551-8556.
37. Tronche F, Casanova E, Turiault M, Sahly I and Kellendonk C: When reverse genetics meets physiology: the use of site-specific recombinases in mice. FEBS Lett 2002, 529:116-121.
38. Rodríguez CI, Buchholz F, Galloway J, Sequerra R, Kasper J, Ayala R, Stewart AF and Dymecki SM: High-efficiency deleter mice show that FLPe is an alternative to Cre-loxP. Nat Genet 2000, 25:139-140.

39. Schaft J, Ashery-Padan R, van der Hoeven F, Gruss P and Stewart AF: Efficient FLP recombination in mouse ES cells and oocytes. Genesis 2001, 31:6-10.

40. Meyers EN, Lewandoski M and Martin GR: An Fgf8 mutant allelic series generated by Cre- and Flp-mediated recombination. Nat Genet 1998, 18:136-141.

41. Andreas S, Schwenk F, Kuter-Luks B, Faust N and Kuhn R: Enhanced efficiency through nuclear localization signal fusion on phage PhiC31-integrase: activity comparison with Cre and FLPe recombinase in mammalian cells. Nucleic Acids Res 2002, 30:2299-2306.

42. Page RL, Canseco RS, Russell CG, Johnson JL, Velander WH and Gwazdauskas FC: Transgene detection during early murine embryonic development after pronuclear microinjection. Transgenic Res 1995, 4:12-17.

43. Belteki G, Gertsenstein M, Ow DW and Nagy A: Site-specific cassette exchange and germline transmission with mouse ES cells expressing phiC31 integrase. Nat Biotechnol 2003, 21:321-324.

44. Sclimenti CR, Thyagarajan B and Calos MP: Directed evolution of a recombinase for improved genomic integration at a native human sequence. Nucleic Acids Res 2001, 29:5044-5051.