Site-specific integrase-mediated transgenesis in mice via pronuclear injection

Bosiljka Tasic[a], Simon Hippenmeyer[a], Charlene Wang[b], Matthew Gamboa[b], Hui Zong[c], Yanru Chen-Tsai[b,1], and Liqun Luo[a,1]

[a]The Howard Hughes Medical Institute and Department of Biology, Stanford University, Stanford, CA 94305; [b]Transgenic Facility, Stanford Cancer Center, Stanford University School of Medicine, Stanford, CA 94305; and [c]Institute of Molecular Biology, University of Oregon, Eugene, OR 97403

Edited* by Matthew P. Scott, Stanford University/The Howard Hughes Medical Institute, Stanford, CA, and approved March 16, 2011 (received for review December 29, 2010)

Microinjection of recombinant DNA into zygotic pronuclei has been widely used for producing transgenic mice. However, with this method, the insertion site, integrity, and copy number of the transgene cannot be controlled. Here, we present an integrase-based approach to produce transgenic mice via pronuclear injection, whereby an intact single-copy transgene can be inserted into predetermined chromosomal loci with high efficiency (up to 40%), and faithfully transmitted through generations. We show that neighboring transgenic elements and bacterial DNA within the transgene cause profound silencing and expression variability of the transgenic marker. Removal of these undesirable elements leads to global high-level marker expression from transgenes driven by a ubiquitous promoter. We also obtained faithful marker expression from a tissue-specific promoter. The technique presented here will greatly facilitate murine transgenesis and precise structure/function dissection of mammalian gene function and regulation in vivo.

Production of transgenic mice via microinjection of DNA into zygotic pronuclei (1–3) has served mammalian genetics for 30 years. Although still the predominant method used to produce transgenic mice, it has several limitations: the insertion site, integrity, and copy number of the transgene cannot be controlled. Insertion of DNA into different chromosomal loci at random could disrupt the function of endogenous genes. Transgenes generated in this manner can be influenced by the local chromatin environment (i.e., position effect) that can lead to transgene silencing or ectopic expression (4–7). Moreover, transgenic DNA concatemerized into a large array is subject to repeat-induced gene silencing (8).

Single-copy transgenesis in mice can be achieved with retroviruses (9) and transposons (10, 11), but these approaches integrate transgenes throughout the genome. As a result, the transgenes are subjected to the local chromatin environment and can cause endogenous gene disruption, although the mutagenic properties of transposons can be desirable for particular applications (10). These problems can be overcome by targeting the transgene to a specific chromosomal locus via homologous recombination in embryonic stem (ES) cells (12, 13). However, this method is significantly more laborious and time-consuming, as it involves creation of modified ES cells and mouse chimeras, as well as eventual germline transmission of the transgene.

Integrase enzymes from a variety of sources have been used to catalyze integration of transgenes in heterologous systems (14, 15). Integrases catalyze irreversible recombination between appropriate *attB* and *attP* sites (14, 16). ϕC31 integrase from a *Streptomyces* phage has previously been used for transgene integration in flies (17–19). In mice, ϕC31 integrase has been used to catalyze integration of circular DNA into pseudo-*attP* sites in the genome for gene therapy (20) or low-efficiency transgenesis (21), for transgenesis in mouse ES cells (22), and for removal of undesirable transgene portions or reporter activation (23, 24).

Here, we describe an integrase-mediated method for site-specific transgenesis in mice via pronuclear microinjection, with integration efficiencies as high as 40%. We use ϕC31 integrase to catalyze recombination between one or two *attB* sites in a recombinant DNA with one or more tandem *attP* sites that we previously inserted into specific loci in the mouse genome (Fig. 1). We show that the plasmid bacterial backbone within the transgene and nearby transgenic elements dramatically decrease expression of our transgenes, and that the absence of these elements results in global, high-level transgene expression from a ubiquitous promoter. Finally, we show that a promoter for the murine transcription factor Hb9 integrated into one of the predetermined loci drives proper tissue-specific marker expression.

Results

Strategy and Proof of Principle for Integrase-Mediated Site-Specific Transgenesis. To generate embryos containing *attP* sites for ϕC31 integrase-mediated transgenesis, we used standard homologous recombination-based methods in mouse ES cells (12, 25). We inserted three shortened tandem ϕC31 integrase *attP* sites (*attPx3*) or a single "full-length" *attP* site (14) into two loci: the *Rosa26* locus on mouse chromosome 6 (26) and an intergenic *Hipp11* (*H11*) locus on mouse chromosome 11 (27) (Fig. 1, *Left*, and *SI Appendix*, Fig. S1). The *Rosa26* locus supports global expression of a single copy knock-in transgene driven by a combination of the CMV enhancer and the chicken β-actin promoter (*pCA*) (28, 29). Knock-in experiments confirmed that *H11* supports high-level global gene expression from the *pCA* promoter and a higher rate of mitotic (interchromosomal) recombination compared with *Rosa26* (27). The latter property suggested that *H11* might allow better access to ϕC31 integrase than *Rosa26*. The modified ES cells were used to produce chimeric mice, and mice with germline-transmitted alleles were used to establish mouse colonies homozygous for the knock-in cassettes.

For most experiments, we integrated transgenes into the *H11* locus because homozygous insertions into this locus are not predicted to disrupt any endogenous genes, and the resulting mice are completely healthy and fertile (27). We injected embryos homozygous for *attP* or *attPx3* at the *H11* locus (*H11P* or *H11P3*, respectively) with ϕC31 mRNA together with *attB-pCA-GFP*, a minicircle DNA that was generated by removal of the plasmid backbone to enable proper transgene expression (as detailed later). The *attB-pCA-GFP* minicircle contains a full-length *attB* site (14) and the sequence for a thermotolerant GFP (30, 31) driven by the ubiquitous pCA promoter (*SI Appendix*, *SI Materials and Methods*). We obtained integration of the transgene (Fig. 1, *Top Right*, and Table 1, rows 1–3; *SI Appendix*, Table S1, provides more details).

As an alternative to insertion of DNA at a single attP site, we also injected a plasmid in which *pCA-GFP* is flanked by two *attB*

Author contributions: B.T., H.Z., Y.C.-T., and L.L. designed research; B.T., S.H., C.W., and M.G. performed research; B.T. and L.L. analyzed data; S.H. contributed new reagents/analytic tools; and B.T. and L.L. wrote the paper.

Conflict of interest statement: Y.C.-T. is a consultant and co-founder of Applied StemCell, Inc.

*This Direct Submission article had a prearranged editor.

Freely available online through the PNAS Open Access Option.

[1]To whom correspondence may be addressed. E-mail: ychen@stanford.edu or lluo@stanford.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1019507108/-/DCSupplemental.

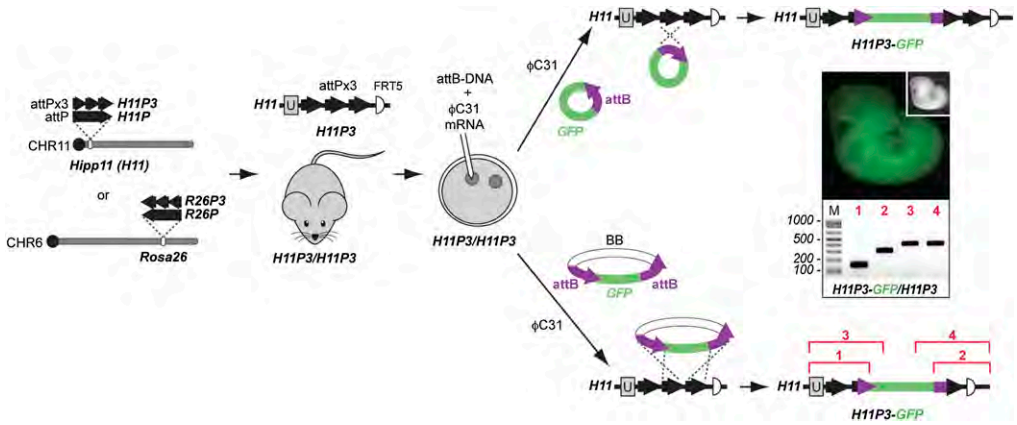

Fig. 1. Schematic summary for site-specific ϕC31 integrase-mediated transgenesis via pronuclear injection in mice. A single or three tandem *attP* sites were knocked into the *Hipp11* (*H11*) or *Rosa26* loci via homologous recombination in ES cells (*Left*; for details, see Fig. S1). Mice homozygous for one of the modified loci (*H11* is shown as an example) served as embryo donors. A mix of DNA and in vitro transcribed ϕC31 mRNA was injected into a single pronucleus of each zygote. The integration of plasmid bacterial backbone (BB) that decreases the transgene expression was avoided either by injecting a minicircle DNA with a single *attB* site (top branch; in this case, ϕC31 catalyzes a typical integration reaction), or by injecting plasmid DNA where the gene of interest (e.g., GFP) was flanked by two *attB* sites (bottom branch; in this case, ϕC31 catalyzes a recombinase-mediated cassette exchange reaction). *Right*, *Middle*: Green fluorescence of a representative transgenic F0 embryo and the corresponding PCR results that indicate site-specific insertion. The red numbers for the PCR results correspond to the numbers on the *H11P3-pCA-GFP* transgene scheme below; the same PCR tests can be used for the transgene above. The particular embryo shown was obtained by cassette exchange. *Inset*: Bright-field image of the same embryo.

sites (*pattB-pCA-GFP-attB*) into *H11P3* homozygous embryos. In principle, this approach should allow ϕC31 integrase to catalyze a recombinase-mediated cassette exchange reaction (Fig. 1, *Bottom Right*). We tested the integration of the circular *pattB-pCA-GFP-attB* plasmid into *H11P3* and obtained successful cassette exchange as confirmed by PCR (Fig. 1, *Bottom Right*, and Table 1, row 4). In these experiments, we never detected insertion of a full plasmid.

The integrated transgenes were properly transmitted from founders to progeny (*SI Appendix*, Table S2). Both strategies resulted in broad and high-level GFP expression in *pCA-GFP* transgenic mice (Fig. 1, *Right*, Fig. 2*D*, and Fig. S1*C*, *Bottom*). These data provide the proof of principle for our site-specific integrase-mediated transgenesis in mice. In the subsequent sections, we describe in more detail the optimization process that led to these results.

Transgenic Integrase Did Not Enable Site-Specific Integration. Our original knock-in cassettes contained a mammalian codon-optimized ϕC31 integrase (ϕC31o) (23) driven by a fragment of the mouse *VASA* promoter sufficient for germline expression (32), and a neomycin resistance gene, flanked by FRT5 sites (33) (*SI Appendix*, Fig. S1). This "*NVϕ* cassette" (i.e., *Neo-VASA-ϕC31o*) was designed to provide the integrase in embryos in situ. We injected mouse embryos homozygous for the *H11P3NVϕ* knock-in with a plasmid containing the *attB-pCA-GFP* transgene (*pattB-pCA-GFP*) but did not obtain any site-specific integration (0/32 F0s; Table 1, row 5) despite occasional random integrations. However, coinjecting *pattB-pCA-GFP* with ϕC31o mRNA into homozygous *H11P3NVϕ* embryos produced site-specific integrations (Table 1, row 6). Thus, the *VASA* promoter does not promote sufficient ϕC31o expression in situ to enable site-specific insertions.

Because the *NVϕ* cassette did not perform as expected we removed it from the *H11P3NVϕ* allele by FLP-mediated recombination to generate the *H11P3* allele (*SI Appendix*, Fig. S1, and *SI Appendix*, *SI Materials and Methods*). Similarly, the *H11P* allele was derived from *H11PNVϕ*. Coinjection of *pattB-pCA-GFP* and ϕC31o mRNA into homozygous *H11P* or *H11P3* embryos produced site-specific integrants (Table 1, rows 8–10). Thus, we coinjected integrase mRNA with DNA for all subsequent experiments.

Table 1. Efficiency of site-specific integration

Row	DNA*	DNA type	DNA size, kb	Strain	Background	F0 (n)	SS F0 (n)	Significance[¶]	SS, % (of F0)	R F0 (*n*)	R, % (of F0)
1	attB-pCA-GFP	Minicircle	~3	H11P	Mix	21	1	NS vs. row 2	4.8	1	4.8
2	attB-pCA-GFP	Minicircle	~3	H11P3	Mix	39	4		10.3	1	2.6
3	attB-pCA-GFP	Minicircle	~3	H11P3	FVB N4	15	6	$P < 0.05$ vs. row 2	40.0	3	20.0
4	attB-pCA-GFP-attB	Plasmid	~6	H11P3	FVB N4	38[‡]	6[‖]	—	15.8	1	2.6
5	attB-pCA-GFP, no RNA	Plasmid	~6	H11P3NVϕ	Mix	32[‡]	0	—	0.0	5	15.6
6	attB-pCA-GFP	Plasmid	~6	H11P3NVϕ	Mix	64[§]	10	NS vs. row 7	15.6	4	6.3
7	attB-pCA-GFP	Plasmid	~6	H11PNVϕ	Mix	30[‡]	2		6.7	0	0.0
8	attB-pCA-GFP-FRT5	Plasmid	~6	H11P	Mix	51	5	NS vs. row 9	9.8	3	5.9
9	attB-pCA-GFP-(FRT5)[†]	Plasmid	~6	H11P3	Mix	61	4		6.6	9	14.8
10	attB-pCA-GFP-FRT5	Plasmid	~6	H11P3	FVB N4	8	3	$P < 0.05$ vs. row 9	37.5	1	10.3
11	attB-pHB9-GFP-FRT5	Plasmid	~14	H11P3	FVB N4	66	2	—	3.0	2	3.0
12	attB-pCA-GFP	Plasmid	~6	R26P3NVϕ	Mix	22[‡]	2	—	9.1	2	9.1

Abbreviations: F0, embryos or animals obtained from injections; SS, site-specific integration; R, random integration; mix, mixed background of 129, C57BL/6 and DBA2; FVB N4, mice of the mixed background were outcrossed for 4 generations to the FVB strain and then intercrossed.

*All DNA was coinjected with ϕC31o mRNA, except for row 5.

[†]Both FRT and non-FRT versions of *attB-pCA-GFP* were used.

[‡]F0s were analyzed only as E10 or E11 embryos.

[§]F0s were analyzed either as E10 or E11 embryos or as live pups.

[‖]The six founders listed contained *pCA-GFP* without the bacterial backbone; five more founders with cassette exchange contained only the bacterial backbone. Therefore, the total number of founders with cassette exchange is 11 (29%).

[¶]Fisher's exact test. NS, not significant.

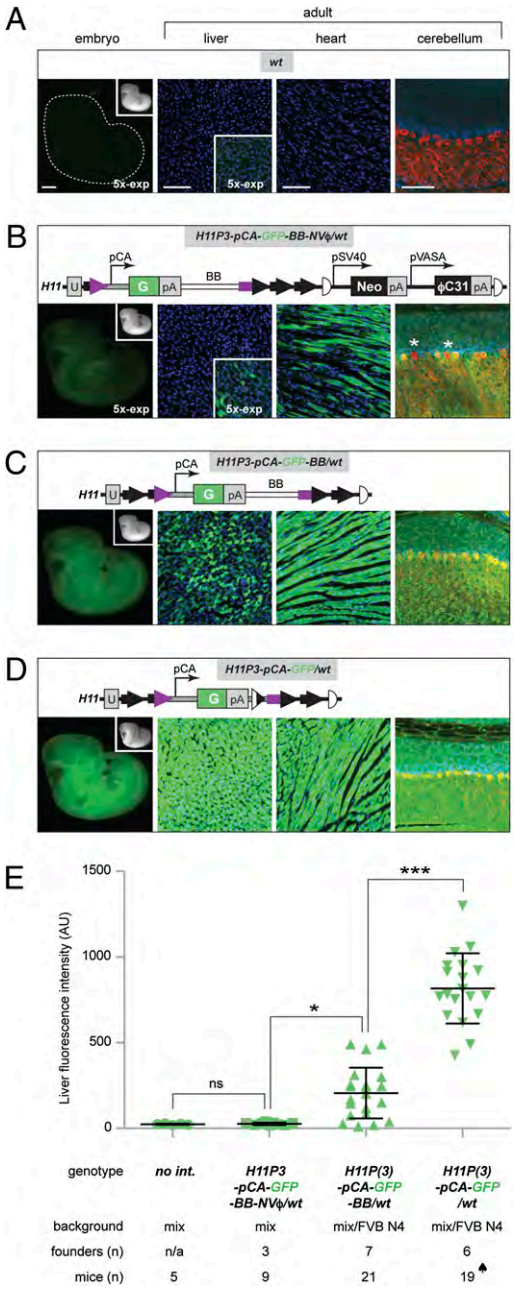

Fig. 2. GFP expression in animals carrying site-specific *pCA-GFP* transgenes introduced by ϕC31 integrase-mediated transgenesis. (*A–D*) Representative fluorescence images from embryonic day 11 embryos and adult livers, hearts, and cerebella of N1 or N2 transgenic animals corresponding to the genotypes and schematics of transgenes shown (*Upper*). Embryos or same tissues were imaged under identical conditions, except that "5×-exp" designates fivefold longer exposure time than for the rest of the images in the same column. Whole-mount embryos were imaged for GFP fluorescence; corresponding bright-field images of each embryo are also shown (*Insets*). The livers and hearts are represented by epifluorescence images of 10-μm sections stained only by DAPI (blue). The green signal is GFP fluorescence. The cerebella are represented by confocal images of sections stained by anti-GFP antibody (green), anti-calbindin (red) for Purkinje cells, and DAPI (blue). Two Purkinje cells labeled by asterisks appear negative for GFP. ϕC31 *attL* and *attR* are the product of recombination of an *attP* (black arrows) and *attB* site (purple) and are therefore half black and half purple. Half circles represent *FRT5* sites. Half white/half black triangle represent λ-integrase *attB* site created during minicircle production. pSV40, *SV40* promoter. pVASA, *VASA* promoter. U, unique sequence. pCA, CMV enhancer and β-actin promoter. G, GFP. pA, polyA signal. BB, plasmid bacterial backbone. (Scale bars: 1 mm for embryos, 100 μm for tissue sections.) (*E*) Average fluorescence in arbitrary units (AU) in the GFP channel for liver sections from animals of genotypes shown below (no int., no integration; represents *wt*, *H11P3NVϕ/wt*, *H11P3/wt*, and *H11P/wt* genotypes). Each dataset is represented by mean ± SD. The numbers of individual animals and founders analyzed for each genotype are listed below the genotypes. When samples from multiple founders were combined to obtain an average, each founder was represented by the same number of animals except in the case labeled by a spade. The fluorescence intensities differ significantly among the groups by one-way ANOVA [$F(3, 50) = 84.09$, $P < 0.0001$]. Tukey's post-hoc test was used for pair-wise comparisons (ns, not significant; $^*P < 0.05$ and $^{***}P < 0.001$).

Bacterial Backbone and the *NVϕ* Cassette Affect Proper Transgene Expression. Despite proper transmission of the site-specifically integrated transgenes produced from the *pattB-pCA-GFP* plasmid, the progeny of these transgenic founders exhibited a wide range of GFP expression levels as seen in whole-mount embryos (*SI Appendix*, Fig. S1*C*). Moreover, the GFP expression in the progeny was mosaic in several internal tissues, including the heart, brain, and particularly the liver (Fig. 2*B* and *SI Appendix*, Fig. S2). We first observed this variable and mosaic expression with *H11P3NVϕ* as the host. We suspected that the nearby germline-specific *VASA* promoter and/or other elements within the *NVϕ* cassette (e.g., the neomycin resistance gene) could affect the expression of *pCA-GFP* inserted at *H11*. Indeed, transgenic embryos containing *pCA-GFP* inserted at *H11P3*, which lack the *NVϕ* cassette, produced more uniform GFP expression (Fig. 2*C* and *SI Appendix*, Fig. S2). However, considerable variability of *pCA-GFP* expression still persisted especially in the livers of N1 or N2 animals derived from a number of transgenic founders (Fig. 2*C* and *SI Appendix*, Fig. S2).

It has been reported that plasmid bacterial backbone can decrease the expression of integrated and episomal transgenes (34–37). To test if this is the case in our system, we developed an in vitro method to produce minicircle DNA: circular DNA containing desired transgene elements, but devoid of the bacterial backbone (*SI Appendix*, Fig. S3 and *SI Appendix*, *SI Materials and Methods*). We injected the minicircle DNA into *H11P3* embryos to produce transgenic animals (*SI Appendix*, Fig. S1, and Table 1). For simplicity, we designate hereafter the transgenic alleles derived from integration of the entire *pCA-GFP* plasmid (which contains the bacterial backbone) as *pCA-GFP-BB* (Fig. 2*C*), and the mice derived from integration of the *pCA-GFP* minicircle as *pCA-GFP* (Fig. 2*D*). Transgenic animals derived from the *pCA-GFP* minicircle exhibited higher and more uniform expression in embryos and all adult tissues examined (Fig. 2*D* and *SI Appendix*, Fig. S2). The removal of bacterial backbone from the *pCA-GFP-BB* transgene by crossing to our newly generated *GFP-FLPo* transgenic mice (*SI Appendix*, *SI Materials and Methods*) also resulted in elevated transgene expression (*SI Appendix*, Fig. S4). These data demonstrate that *pCA-GFP* at the *H11* locus can express GFP ubiquitously in the absence of the *NVϕ* cassette and the bacterial backbone.

Because the greatest variability was observed in the liver (Fig. 2 and *SI Appendix*, Fig. S2), we used it as a model to determine the relative contributions of the *NVϕ* cassette and the bacterial backbone to GFP expression variability in these transgenic mice. In addition, to test for the possible differences between a single *attP* site and *attPx3*, we analyzed transgenic mice obtained by insertion into the *H11P* allele (Table 1, rows 1 and 8). Finally, to probe the effect of genetic background, we analyzed transgenes inserted into the *H11P3* allele in mice that had been outcrossed to the FVB inbred strain for four generations (Table 1, rows 3 and 10). We compared the total GFP fluorescence of liver sections from different transgenic animals under identical conditions. We observed no statistically significant differences in GFP fluorescence between transgenes that differed only in the number of *attP* copies or in the genetic background of the mouse strain used for transgenesis (*SI Appendix*, Fig. S5). Therefore, we grouped all data according to the presence of the *NVϕ* cassette and/or the bacterial backbone (Fig. 2*E*). We found that, in the

GENETICS

presence of both the *NVϕ* cassette and the bacterial backbone, GFP expression was detectable in the liver in a small number of cells and at a very low level (Fig. 2*B*), but total fluorescence was statistically indistinguishable from negative controls (Fig. 2*E*, second column vs. first column). In other organs analyzed (heart and brain), GFP expression was apparent but mosaic (Fig. 2*B* and *SI Appendix*, Fig. S2). When the *NVϕ* cassette was removed but the bacterial backbone was still present, average GFP fluorescence intensity became significantly higher (Fig. 2*E*, third column vs. second column). Finally, when the bacterial backbone was removed, average GFP fluorescence intensity became even higher (Fig. 2*E*, fourth column vs. third column). Thus, both the *NVϕ* cassette and the bacterial backbone significantly reduced transgene expression. The reduction of total fluorescence intensity could be caused by low-level of expression in every cell, absence of expression in a subset of cells, or a combination of the two. As is evident from Fig. 2 *A–D* and *SI Appendix*, Fig. S2, both mechanisms contributed to the reduced level of transgene expression in the presence of the *NVϕ* cassette and/or the bacterial backbone.

***H11* Can Support Tissue-Specific Expression.** To test if a tissue-specific promoter can provide appropriate expression using our transgenesis method, we integrated *pattB-pHb9-GFP-FRT5* into *H11P3* (Table 1, row 11). This plasmid contains an approximately 9-kb promoter fragment from the murine transcription factor *Hb9* gene that has been shown to be sufficient to direct appropriate tissue- and cell-specific expression in transgenic animals (38). We examined tissue-specific marker expression before and after removal of the bacterial backbone by using the *GFP-FLPo* transgene (*SI Appendix, SI Materials and Methods* and Fig. 3*A*). In agreement with the reported expression pattern (38), we observed GFP expression in motor neurons in the ventral spinal cord and the tail tip (Fig. 3*B*). In this case, the removal of bacterial backbone did not appear to affect the expression level of the transgene (Fig. 3*C*). Double-labeling with endogenous Hb9 protein confirmed the motor neuron-specific expression of the transgene (Fig. 3*C*). These experiments indicate that our integrase-based strategy can be used for faithful tissue-specific expression of transgenes.

Integration Efficiency. We compared the integration efficiency for *attP*-modified loci, expressed as the percentage of F0 animals with site-specific integrations obtained from the total number of F0s (Table 1; *SI Appendix*, Table S1, provides more details). Although in pooled data (*SI Appendix*, Table S3), *H11P3* (three copies of shortened *attP*) appeared somewhat more efficient than *H11P* (one copy of the full-length *attP*), the efficiencies of site-specific insertions into these two loci were statistically indistinguishable (*SI Appendix*, Table S3, compare rows 1 vs. 2; and Table 1, compare rows 1 vs. 2, 6 vs. 7, and 8 vs. 9). In contrast, outcrossing the *H11P3* mice to the FVB strain for four generations (FVB N4) significantly increased the integration efficiency to approximately 40% (Table 1, compare rows 2 vs. 3 and 9 vs. 10; and *SI Appendix*, Table S3, compare rows 2 vs. 3). This efficiency is comparable to or better than the efficiency of traditional transgenesis with random integration. Circular DNAs with sizes from 3 to 6 kb appeared to have similar efficiencies of integration (Table 1), but larger DNA (14 kb) showed decreased integration efficiency (~3%; Table 1, row 11). The efficiency of cassette exchange by using *H11P3* is approximately 30%, but because identical *attB* sites in the plasmid and identical *attP* sites in the genome were used, cassette exchange could result in either integration of the transgene of interest or the bacterial backbone. Therefore, only half of the cassette-exchange insertions (~16%) contained GFP and the other half contained the plasmid backbone (Table 1, row 4).

Although we used circular DNA for injections, we also observed insertions at locations other than our intended *attP* sites (Table 1). In 20 of 23 founders that transmitted their site-specific transgenes to the progeny, the site-specific integrants contained a single-copy transgene and did not contain a second random insertion as judged by PCR and quantitative PCR (*SI Appendix, SI Materials and Methods*). In rare cases, when site-specific and random integration occurred in the same transgenic founder, the two distinct transgene integrations could be readily segregated in the N1 progeny.

Site-Specific Integration into *Rosa26*. To test whether ϕC31-mediated integration is applicable to other genomic loci, we injected *pattB-pCA-GFP* into embryos homozygous for *R26P3NVϕ* (*attPx3+NVϕ* integrated into the *Rosa26* locus). We obtained site-specific integrants (Table 1, row 12). We have removed the *NVϕ* cassette from *R26P3NVϕ* by using Flpo, and have recently created homozygous *R26P3* mice to provide a second locus for integrase-mediated transgenesis.

Discussion

Here we describe an efficient method for producing transgenic mice containing an intact, single-copy transgene integrated into a predetermined locus via pronuclear injection. Our method is considerably simpler than transgenesis using homologous recombination in ES cells and offers many technical advantages compared with the current method of random integration of transgenes via pronuclear injection (1–3). Transgenes produced from our site-specific integration method are intact, have a defined copy number and chromosomal environment, and do not disrupt endogenous genes (at least at the *H11* locus). These properties will facilitate many transgenesis-based experiments and will increase their reliability and efficiency. For example, the relationships be-

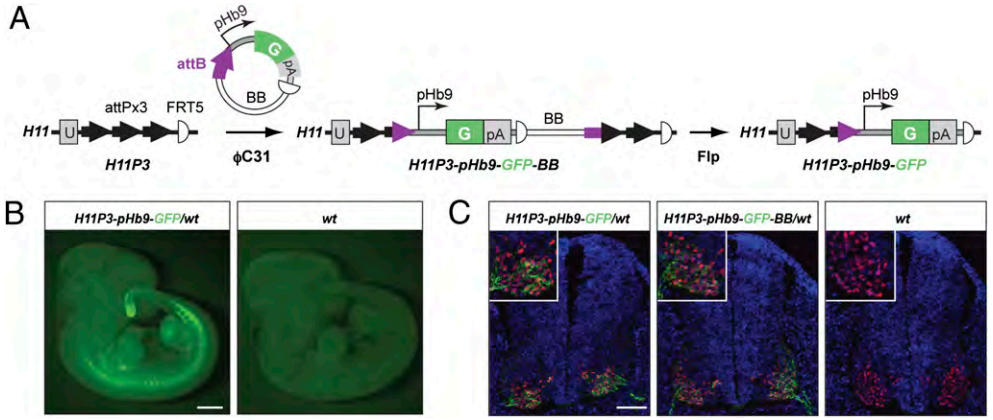

Fig. 3. GFP expression in embryos carrying a single copy of *pHb9-GFP* transgene site-specifically integrated in *H11*. (*A*) Schematic representation of the generation of *H11P3-pHb9-GFP* allele. After site-specific integration of the plasmid *pBT366* (*SI Appendix, SI Materials and Methods*), the bacterial backbone (BB) was removed by crossing to the *GFP-FLPo* transgenic line (*SI Appendix, SI Materials and Methods*). The embryos that inherited only the *Hb9* allele but not the *Flpo* transgene were tested for GFP expression. (*B*) GFP expression in a whole-mount representative embryonic day 11 embryo containing the *H11P3-pHb9-GFP* allele. A WT littermate is also shown (*Right*). (Scale bar, 1 mm.) (*C*) Immunofluorescence of sections from embryonic day 11 spinal cords at limb level with anti-GFP signal in green, anti-Hb9 signal in red, and DAPI in blue. *Insets*: Magnified bottom left portions of each image containing Hb9-positive nuclei. (Scale bar, 100 μm.)

tween amino acid sequences or domain structures of a protein and its in vivo biological functions can be more reliably compared if a series of transgenes encoding different variants of a protein are expressed at the same level. The regulatory elements that control gene expression can also be systematically dissected when reporter transgenes from the same integration site are compared. Subtle differences in levels or patterns of transgene expression that would be overwhelmed by positional effects and differences in copy numbers in randomly integrated transgenes are more appropriately compared by using site-specific integration of transgenes.

Recently, two other approaches for site-specific transgenesis in mice using pronuclear injection were reported (39–41). One approach relied on zinc-finger nucleases to create site-specific double stranded breaks that were repaired by injected recombinant DNA via homologous recombination (39, 41). The two reports using this approach achieved transgene integration with a frequency of approximately 2.5% (two events among 80 embryos) into the *Rosa26* locus (39), and approximately 5% (two of 40 for GFP insertion) into the *Mdr1a* locus (41). Both reports have yet to demonstrate germline transmission and proper adult expression of the transgenes, although proper germline transmission for the same technique in rats was reported (41). The other approach used the Cre recombinase to catalyze cassette exchange in the *Rosa26* locus and a tissue-specific locus, *H2-Tw3,* at an average frequency of approximately 4.3%, with proper expression and germline transmission of the transgenes (40). One drawback of the Cre-based method is that it cannot be used to generate Cre-activated transgenes (containing *loxP-STOP-loxP*), which are frequently used for reporting Cre activity and perturbing gene function in cells in which Cre is expressed. The integration frequencies of these studies and the present study cannot be easily compared, as the studies used different strains, loci, and constructs. However, our approach with the FVB strain and the *H11* locus consistently produces higher integration efficiencies with 3- to 6-kb plasmids than either of the other two approaches.

The present study also revealed that plasmid bacterial backbone and a nearby transgenic cassette (*NVϕ*) have profound effects on the expression reliability of the GFP transgenes driven by a ubiquitous promoter. We did not observe any obvious change in *HB9-GFP* transgene expression upon removal of the bacterial backbone; this observation could be the consequence of small numbers of animals compared, or a result of the possibility that the bacterial backbone could have different effects on different promoters. The effect of bacterial backbone has been reported for randomly integrated and episomal transgenes (34–37), and other native bacterial sequences like the *lacZ* gene or the neomycin resistance gene have been linked to variegation in transgenic animals (40, 42–44). Our experiments based on site-specific integration enabled us to systematically and quantitatively characterize these effects for single-copy chromosomally integrated transgenes.

We have generated mice that allow integration at two defined loci, the widely used *Rosa26* locus (26) and the new *Hipp11* locus (27), which support high-level ubiquitous expression of integrated transgenes. We describe three different approaches to create transgenes devoid of the bacterial backbone: (*i*) use of minicircle DNA for transgenesis, (*ii*) flanking the gene of interest with two *attB* sites in a plasmid to enable cassette exchange, (*iii*) removing the bacterial backbone from a transgene generated from a plasmid by crossing the transgenic mice to *GFP-Flpo* transgenic mice. Although currently only half of the cassette exchange events are desirable, the cassette exchange strategy removes the bacterial backbone without the need to produce minicircle DNA or to remove the plasmid backbone by subsequent crossing to *GFP-Flpo* mice. Therefore, the cassette exchange may be the approach of choice because of its combination of convenience and good integration efficiency. A future improvement could use two mutually noncompatible pairs of *attB* and *attP* sites to control for the direction of insertion. In addition, to expand the application of our method for producing transgenic mouse models, we are in the process of introducing the *attP* sites into the frequently used C57BL/6 genetic background. In summary, the present study facilitates murine transgenesis, highlights the requirements for gene expression reliability in mammals, and provides an efficient system for studies of gene expression and function in vivo.

Materials and Methods

Recombinant DNA. We used standard methods of recombinant DNA to construct all plasmids used in this study. Construction details are described in *SI Appendix, SI Materials and Methods*.

Gene Targeting in Mouse ES Cells. We used standard techniques to modify mouse ES cells (45). *SI Appendix, SI Materials and Methods*, provides more details.

Mouse Breeding and Maintenance. All experimental procedures were carried out in accordance with the Administrative Panel on Laboratory Animal Care protocol and the institutional guidelines by the Veterinary Service Center at Stanford University.

The F1 *attP* knock-in animals obtained from the cross of chimeras to B6D2F1/J females (stock no. 100006; Jackson Laboratories) were crossed to each other to establish homozygous knock-in mouse lines. These lines were maintained by intercrosses between homozygous animals. To outcross the mice to FVB (Charles River), we started from a homozygous transgenic male and bred him and his transgenic male progeny to FVB females for a total of four generations. During the outcrossing, we preferentially selected transgenic mice of white coat color. The fourth generation outcrossed mice (FVB N4) were crossed to each other to make homozygous males and females that were subsequently used to produce zygotes for microinjection. The FVB N4 homozygous line was subsequently maintained by homozygous crosses. For testing transgenic founders we crossed the founder (F0) animals to WT CD1 mice (Charles River) to generate the N1 generation. For the N2 and N3 generations, we continued crossing to CD1. We have generated homozygous mice from the founder E1 (*SI Appendix*, Table S2), and they were healthy and fertile.

Preparation of mRNA and DNA for Microinjection. Capped mRNAs for ϕC31o and Flpo were generated by using a mMESSAGEmMACHINE in vitro transcription kit (Ambion) according to the manufacturer's instructions from BamHI-digested *pBT317* (*SI Appendix, SI Materials and Methods*) and BssHII-digested *pFlpo* (23), respectively. The integrity of the RNA was assessed by electrophoresis on a 1% agarose gel. Before loading on the gel, the RNA was denatured by using the loading buffer provided in the Ambion kit according to the manufacturer's instructions.

Plasmid DNA was prepared using a modified Qiagen miniprep procedure and was subsequently extracted with phenol/chloroform (*SI Appendix, SI Materials and Methods*). The DNA was diluted to 6 ng/μL by sterile microinjection TE buffer (0.1 mM EDTA, 10 mM Tris, pH 7.5) and was kept at −80 °C until the injection. The DNA was tested to be RNase-free by incubation with an in vitro transcribed RNA at 37 °C for 1 h and then by analyzing the mix on a 1% agarose gel. Before loading on the gel, the RNA was denatured as described earlier. *SI Appendix, SI Materials and Methods* provides details on preparation of minicircle DNA.

Microinjection for Generation of Site-Specific Integrants. Microinjection was performed with an established setup at the Stanford Transgenic Facility. Superovulated homozygous *attP*-containing females were bred to corresponding males to generate homozygous *attP*-containing zygotes. A DNA/mRNA mix of interest was microinjected into a single pronucleus and cytoplasm of each zygote by using a continuous flow injection mode. The surviving zygotes were implanted into oviducts of pseudopregnant CD1 (Charles River) recipient mothers. All injection mixes contained 3 ng/μL DNA and 48 ng/μL of in vitro transcribed ϕC31o mRNA in microinjection TE buffer (0.1 mM EDTA, 10 mM Tris, pH 7.5). The injection mixes were prepared fresh before each injection by mixing equal volumes of 6 ng/μL DNA solution and 96 ng/μL mRNA solution.

PCR. To test F0 animals for site-specific and random insertions, we performed three PCRs: one for the 5′-end junction, one for the 3′-end junction, and one internal to the transgene. These PCRs cannot detect random insertions that occurred in mice with site-specific insertions. For that purpose, see PCR analysis of N1 generation below. For testing integration into *H11P* or *H11P3* alleles, we used *PCR1*, *PCR2*, and *PCR6* (*SI Appendix, SI Materials and Methods*). To test if a particular site-specific insertion into *H11P* or *H11P3* originated from appropriate cassette exchange or minicircle insertion, we used *PCR3* and *PCR4* or *PCR4′* (*SI Appendix, SI Materials and Methods*). These PCRs demonstrated that all injections of minicircle DNA produced only

GENETICS

minicircle insertions, suggesting that contamination of minicircle preps with full-length plasmid was negligible.

To test germline transmission of both site-specific and random insertions to N1 animals, we performed three PCRs on the N1 progeny: one for the 5′-end junction, one for the 3′-end junction, and one with internal primers. Correlation of 100% between the GFP-specific and site-specific integration PCRs on DNA from N1 animals suggested that the corresponding F0 founder most likely contained only a single site-specific insertion. This conclusion was reinforced by quantitative PCR (*SI Appendix, SI Materials and Methods*) for GFP to show that a selected number of N1 animals indeed had a single-copy transgene.

SI Appendix, SI Materials and Methods (46, 47), provides additional information on PCR procedures used in this study, and *SI Appendix*, Table S4 provides primer sequences.

Tissue Preparation and Immunohistochemistry. The procedures were performed essentially as described (29). *SI Appendix, SI Materials and Methods*, includes further details.

Quantification of GFP Fluorescence in Liver Sections. At least three individual images were taken from randomly chosen 10-μm sections for each liver by a camera connected to a fluorescence microscope (Nikon) with a 20× objective. The regions of interest were consistently chosen to contain minimal number of large blood vessels, so that the majority of every image would be covered by hepatocytes. All images were taken with the same exposure time (5 ms), same gain, and during two consecutive days of imaging. At this condition, even the samples with brightest fluorescence had no saturated pixels. Total fluorescence for each image was calculated by using ImageJ. Averaged total fluorescence from all images of the same liver was plotted on a graph (*SI Appendix*, Fig. S2). The fluorescence images shown in figures represent the same fields that were used for the measurements, but exposed four times longer for easier visualization.

Animal and Reagent Availability. Plasmids (containing attB sites or integrase cDNA) and *H11P3* and *R26P3* homozygous frozen embryos and mice will be distributed through Applied StemCell, Inc. (www.appliedstemcell.com), for prices comparable to those of other distributors (e.g., Addgene for plasmids, Jackson Labs for mice). Applied StemCell will also provide services for making customized, integrase-mediated site-specific transgenic mice.

ACKNOWLEDGMENTS. We thank Yanfeng Li, Jennifer Lin, Hong Zeng, Ying Jiang, and Carlota Manalac for technical support; Michele Calos for plasmids; Russell Fernald for real-time PCR machine; Silvia Arber for anti-Hb9 antibody; Dritan Agalliu for help with embryo dissections; and Kazunari Miyamichi and Tim Mosca for comments on the manuscript. This work is supported by National Institutes of Health Grant R01-NS050835. B.T. was a Damon Runyon Fellow and was supported by Damon Runyon Cancer Research Foundation Grant DRG-1819-04. S.H. was supported by postdoctoral fellowships from the European Molecular Biology Organization (ALTF 851-2005), Human Frontier Science Program Organization (LT00805/2006-L), and Swiss National Science Foundation (PA00P3_124160). L.L. is an investigator of The Howard Hughes Medical Institute.

1. Gordon JW, Scangos GA, Plotkin DJ, Barbosa JA, Ruddle FH (1980) Genetic transformation of mouse embryos by microinjection of purified DNA. *Proc Natl Acad Sci USA* 77:7380–7384.
2. Gordon JW, Ruddle FH (1981) Integration and stable germ line transmission of genes injected into mouse pronuclei. *Science* 214:1244–1246.
3. Brinster RL, et al. (1981) Somatic expression of herpes thymidine kinase in mice following injection of a fusion gene into eggs. *Cell* 27:223–231.
4. Milot E, et al. (1996) Heterochromatin effects on the frequency and duration of LCR-mediated gene transcription. *Cell* 87:105–114.
5. Pedram M, et al. (2006) Telomere position effect and silencing of transgenes near telomeres in the mouse. *Mol Cell Biol* 26:1865–1878.
6. Gao Q, et al. (2007) Telomeric transgenes are silenced in adult mouse tissues and embryo fibroblasts but are expressed in embryonic stem cells. *Stem Cells* 25:3085–3092.
7. Williams A, et al. (2008) Position effect variegation and imprinting of transgenes in lymphocytes. *Nucleic Acids Res* 36:2320–2329.
8. Garrick D, Fiering S, Martin DI, Whitelaw E (1998) Repeat-induced gene silencing in mammals. *Nat Genet* 18:56–59.
9. Lois C, Hong EJ, Pease S, Brown EJ, Baltimore D (2002) Germline transmission and tissue-specific expression of transgenes delivered by lentiviral vectors. *Science* 295:868–872.
10. Ding S, et al. (2005) Efficient transposition of the piggyBac (PB) transposon in mammalian cells and mice. *Cell* 122:473–483.
11. Mátés L, et al. (2009) Molecular evolution of a novel hyperactive Sleeping Beauty transposase enables robust stable gene transfer in vertebrates. *Nat Genet* 41:753–761.
12. Doetschman T, et al. (1987) Targetted correction of a mutant HPRT gene in mouse embryonic stem cells. *Nature* 330:576–578.
13. Thomas KR, Capecchi MR (1987) Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. *Cell* 51:503–512.
14. Groth AC, Olivares EC, Thyagarajan B, Calos MP (2000) A phage integrase directs efficient site-specific integration in human cells. *Proc Natl Acad Sci USA* 97:5995–6000.
15. Keravala A, et al. (2006) A diversity of serine phage integrases mediate site-specific recombination in mammalian cells. *Mol Genet Genomics* 276:135–146.
16. Thorpe HM, Smith MC (1998) In vitro site-specific integration of bacteriophage DNA catalyzed by a recombinase of the resolvase/invertase family. *Proc Natl Acad Sci USA* 95:5505–5510.
17. Groth AC, Fish M, Nusse R, Calos MP (2004) Construction of transgenic Drosophila by using the site-specific integrase from phage phiC31. *Genetics* 166:1775–1782.
18. Venken KJ, He Y, Hoskins RA, Bellen HJ (2006) P[acman]: A BAC transgenic platform for targeted insertion of large DNA fragments in D. melanogaster. *Science* 314:1747–1751.
19. Bischof J, Maeda RK, Hediger M, Karch F, Basler K (2007) An optimized transgenesis system for Drosophila using germ-line-specific phiC31 integrases. *Proc Natl Acad Sci USA* 104:3312–3317.
20. Olivares EC, et al. (2002) Site-specific genomic integration produces therapeutic Factor IX levels in mice. *Nat Biotechnol* 20:1124–1128.
21. Hollis RP, et al. (2003) Phage integrases for the construction and manipulation of transgenic mammals. *Reprod Biol Endocrinol* 1:79.
22. Belteki G, Gertsenstein M, Ow DW, Nagy A (2003) Site-specific cassette exchange and germline transmission with mouse ES cells expressing phiC31 integrase. *Nat Biotechnol* 21:321–324.
23. Raymond CS, Soriano P (2007) High-efficiency FLP and PhiC31 site-specific recombination in mammalian cells. *PLoS ONE* 2:e162.
24. Sangiorgi E, Shuhua Z, Capecchi MR (2008) In vivo evaluation of PhiC31 recombinase activity using a self-excision cassette. *Nucleic Acids Res* 36:e134.
25. Capecchi MR (1989) Altering the genome by homologous recombination. *Science* 244:1288–1292.
26. Soriano P (1999) Generalized lacZ expression with the ROSA26 Cre reporter strain. *Nat Genet* 21:70–71.
27. Hippenmeyer S, et al. (2010) Genetic mosaic dissection of Lis1 and Ndel1 in neuronal migration. *Neuron* 68:695–709.
28. Zong H, Espinosa JS, Su HH, Muzumdar MD, Luo L (2005) Mosaic analysis with double markers in mice. *Cell* 121:479–492.
29. Muzumdar MD, Tasic B, Miyamichi K, Li L, Luo L (2007) A global double-fluorescent Cre reporter mouse. *Genesis* 45:593–605.
30. Siemering KR, Golbik R, Sever R, Haseloff J (1996) Mutations that suppress the thermosensitivity of green fluorescent protein. *Curr Biol* 6:1653–1663.
31. Okada A, Lansford R, Weimann JM, Fraser SE, McConnell SK (1999) Imaging cells in the developing nervous system with retrovirus expressing modified green fluorescent protein. *Exp Neurol* 156:394–406.
32. Gallardo T, Shirley L, John GB, Castrillon DH (2007) Generation of a germ cell-specific mouse transgenic Cre line, Vasa-Cre. *Genesis* 45:413–417.
33. Seibler J, Bode J (1997) Double-reciprocal crossover mediated by FLP-recombinase: A concept and an assay. *Biochemistry* 36:1740–1747.
34. Townes TM, Lingrel JB, Chen HY, Brinster RL, Palmiter RD (1985) Erythroid-specific expression of human beta-globin genes in transgenic mice. *EMBO J* 4:1715–1723.
35. Chen ZY, He CY, Ehrhardt A, Kay MA (2003) Minicircle DNA vectors devoid of bacterial DNA result in persistent and high-level transgene expression in vivo. *Mol Ther* 8:495–500.
36. Chen ZY, He CY, Meuse L, Kay MA (2004) Silencing of episomal transgene expression by plasmid bacterial DNA elements in vivo. *Gene Ther* 11:856–864.
37. Suzuki M, Kasai K, Saeki Y (2006) Plasmid DNA sequences present in conventional herpes simplex virus amplicon vectors cause rapid transgene silencing by forming inactive chromatin. *J Virol* 80:3293–3300.
38. Arber S, et al. (1999) Requirement for the homeobox gene Hb9 in the consolidation of motor neuron identity. *Neuron* 23:659–674.
39. Meyer M, de Angelis MH, Wurst W, Kühn R (2010) Gene targeting by homologous recombination in mouse zygotes mediated by zinc-finger nucleases. *Proc Natl Acad Sci USA* 107:15022–15026.
40. Ohtsuka M, et al. (2010) Pronuclear injection-based mouse targeted transgenesis for reproducible and highly efficient transgene expression. *Nucleic Acids Res* 38:e198.
41. Cui X, et al. (2011) Targeted integration in rat and mouse embryos with zinc-finger nucleases. *Nat Biotechnol* 29:64–67.
42. Montoliu L, Chávez S, Vidal M (2000) Variegation associated with lacZ in transgenic animals: a warning note. *Transgenic Res* 9:237–239.
43. Cohen-Tannoudji M, Babinet C, Morello D (2000) lacZ and ubiquitously expressed genes: Should divorce be pronounced? *Transgenic Res* 9:233–235.
44. Fiering S, et al. (1995) 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* 9:2203–2213.
45. Nagy A, Rossant J, Nagy R, Abramow-Newerly W, Roder JC (1993) Derivation of completely cell culture-derived mice from early-passage embryonic stem cells. *Proc Natl Acad Sci USA* 90:8424–8428.
46. Zhao S, Fernald RD (2005) Comprehensive algorithm for quantitative real-time polymerase chain reaction. *J Comput Biol* 12:1047–1064.
47. Li L, et al. (2010) Visualizing the distribution of synapses from individual neurons in the mouse brain. *PLoS ONE* 5:e11503.

Supporting Information

SI Materials and Methods

Recombinant DNA construction. All PCR for DNA construction was done with Phusion DNA polymerase (Finnzymes, Finland). All DNA fragments that were amplified by PCR were completely sequenced after cloning. For primer sequences used in construction see Table S4.

pBT296 (*pBS-U-attP-FRT5-pSV40-Neo-pA-FRT5)*: Into a modified pBluescript, we subcloned:
1) a unique sequence "*U*" from the promoter of yeast *his3* gene:
(GGTGATAGGTGGCAAGTGGTATTCCGTAAGGATATC).
2) the single "full-length" *attP* site from *pTA-attP* (gift of Michelle Calos) (Ref. 1).
3) *FRT5* (GAAGTTCCTATTCCGAAGTTCCTATTCTTCAAAAGGTATAGGAACTTC) (Ref. 2, 3)-flanked neomycin resistance gene driven by an *SV40* promoter.

pBT298 *(pBS-U-attPx3-FRT5-pSV40-Neo-pA-FRT5):* Same as *pBT296*, but the "full-length" *attP* site was replaced by three sequential *attP* sites (70 bp each, sequence of a single site: CGGGAGTAGTGCCCCAACTGGGGTAACCTTTGAGTTCTCTCAGTTGGGGGCGTAGGG TCG), synthesized by Celtek Genes (Nashville, TN).

pBT305 *(pBS-U-attPx3-FRT5-pSV40-Neo-pA-PL-FRT5):* PL represents a polylinker: SbfI-HpaI-AatII. The plasmid was generated by inserting annealed oligos PR402 and PR403 into the XbaI site of *pBT298*, thereby destroying the XbaI sites on both ends of the PL.

pBT307 *(pBS-U-attPx3-FRT5-pSV40-Neo-pA-ϕC31o-pA-FRT5):* ϕC31o was amplified by PCR from *pPGKPhiC31obpA* (Addgene plasmid 13795) (Ref. 4) and subcloned into *pBT305*.

pBT308b *(pTOPO-pVasa*): A previously described fragment of the murine VASA promoter (5) was amplified by PCR from genomic DNA of the FVB strain and TOPO-cloned into *pTOPO* (Invitrogen).

pBT309a *(pBS-U-attPx3-FRT5-pSV40-Neo-pA-pVasa-ϕC31o-pA-FRT5)*: *pVasa* was subcloned from *pBT308b* into *pBT307*.

pBT310 *(pBS-U-attP-FRT5-pSV40-Neo-pA-pVasa-ϕC31o-pA-FRT5):* The NheI/AscI fragment from *pBT309a* was subcloned into the NheI/AscI-digested *pBT296*.

pBT311 *(pH11-U-attPx3-FRT5-pSV40-Neo-pA-pVasa-ϕC31o-pA-FRT5*): The SwaI/AscI fragment from *pBT309a* was subcloned into the PmeI/AscI-digested *pHipp11* (Ref. 6).

pBT312 *(pH11-U-attP-FRT5-pSV40-Neo-pA-pVasa-ϕC31o-pA-FRT5*): The SwaI/AscI fragment from *pBT310* was subcloned into the PmeI/AscI-digested *pHipp11* (Ref. 6).

pBT313 *(pR26-U-attPx3-FRT5-pSV40-Neo-pA-pVasa-ϕC31o-pA-FRT5*): The SwaI/AscI fragment from *pBT309a* was subcloned into the SwaI/AscI-digested *pROSA26* (Ref. 7).

pBT314 *(pR26-U-attP-FRT5-pSV40-Neo-pA-pVasa-ϕC31o-pA-FRT5*): The SwaI/AscI fragment from *pBT310* was subcloned into the SwaI/AscI-digested *pROSA26* (Ref. 7).

pBT316 *(pattB-pCA-GFP-pA)*: The SalI fragment from *pTA-attB* (1) containing the "full-length" *attB* site was subcloned into the SalI site of *pBT255* (*pCA-GFP4m-pA*).

pBT317 (*pET-ϕC31o-pA* – **for preparation of ϕC31o mRNA**): ϕC31o gene was amplified by PCR from *pPGKϕC31obpA* (Addgene plasmid 13795) (Ref. 4) using primers PR437 and PR438. The PCR was digested with BamHI and MseI and cloned into BamHI/NdeI-digested *pET11ϕC31pA* (8).

pBT340 *(pattB-pCA-GFP-pA-FRT5-pPGK-Flpo-pA)*. A fragment containing *FRT5-pPGK-Flpo-pA* was PCR amplified from *pPGKFLPobpA* (Addgene plasmid 13793) (Ref. 4) and cloned between NotI and AscI sites of *pBT316*.

pBT344 *(pattB-pCA-GFP-pA-FRT5* – **for cloning any DNA fragment to be integrated as a full plasmid using ϕC31; the bacterial backbone can be subsequently removed by crossing to *GFP-Flpo* mice***)*: The *pPGK-Flpo-pA* portion was removed from *pBT340*.

pBT346 *(pλ-attB-pCA-GFP-pA* – **for producing attB-containing minicircle *in vitro*)***:* I-SceI restriction site and the λ *attL1* were amplified from *pENTR-TopoD* using PR493 and PR494. The PCR product was digested with Acc65I and XhoI, and inserted into the Acc65I/XhoI-digested *pBT316* to generate a construction intermediate. Subsequently, the λ *attR1* site was amplified from *pENTR-TopoD* using PR495 and PR496 and cloned into the SacI/NotI-digested construction intermediate to generate *pBT346*.

pBT366 *(pattB-Hb9-GFP-pA-FRT5)*. The filled-in *Hb9-GFP* XhoI fragment from *pHB9-EGFP* (Addgene plasmid 16275) (Ref. 9) was subcloned into the PacI/AscI digested and filled-in *pBT344*.

pBT374 *(pattB-pCA-GFP-pA-attB^{Swa})*: The filled-in SalI fragment containing the *attB* site from *pBT316* was subcloned into the SwaI site of *pBT316*. This plasmid was used for initial cassette exchange tests, but for future use we recommend *pBT378* below, as it contains more convenient restriction sites.

pBT378 *(pattB-pCA-GFP-pA-attB* – **for ϕC31-mediated cassette exchange**): It contains more convenient restriction sites than *pBT374* that enable replacement of the *pCA-GFP-pA* insert with an insert of choice (ClaI, HindIII, PacI, PmeI, PstI between the first *attB* and *pCA*, and SwaI, AscI, SpeI and NotI between the *pA* and the second *attB*). It was created by subcloning the filled-in SalI fragment containing the *attB* site from *pBT316* into the BstXI-linearized and filled-in *pBT316*.

Preparation of plasmid DNA by a modified Qiagen mini-prep procedure. We started from 4 ml of DH5α bacterial culture grown in LB broth for not more than 10 h at 37ºC. We collected the bacteria in 2 ml tubes by spinning 2 ml of culture twice in the same tube. We doubled the recommended volumes of P1, P2 and N3 (Qiagen). After loading the samples onto the Qiagen

mini-prep columns, we washed the columns twice with buffer PB and then twice with buffer PE (Qiagen). The PB washes diminish but do not abolish RNase contamination. We eluted the DNA in 3-fold diluted EB (Qiagen). The plasmid DNA yield from a single prep was usually in the range of 7.5 to 25 μg. To obtain more DNA, several preps can be performed at the same time and pooled. We determined the concentration using the Nanodrop spectrophotometer (Thermo Scientific) and used at least 5-20 μg in 200 μl of solution for subsequent extractions to remove residual RNase (see below).

Phenol/chloroform extraction of plasmid DNA. At least 5 μg of DNA in 200 μl of solution were extracted twice with a phenol:chloroform (50:50) mix and twice with chloroform only. The DNA was mixed with 1/10 volume of 3M sodium-acetate pH 5.2, precipitated with 2.7 volumes of ethanol, and subsequently dissolved in sterile and RNase-free microinjection TE buffer (miTE; 0.1 mM EDTA, 10 mM Tris pH 7.5). The DNA was filtered through a sterile 0.2 μm filter (Millipore, Cat. No. SLGV004SL) and the concentration was determined using the Nanodrop spectrophotometer (Thermo Scientific).

Preparation of minicircle DNA. We started from 4 μg of *pBT346* plasmid DNA purified by the modified Qiagen mini-prep procedure above. The 200 μl-recombination reaction consisted of 40 μl of LR clonase II (Invitrogen, Cat. No. 11791-020) and 160 μl of the DNA diluted in miTE buffer. The reaction was incubated for 3 h at 25ºC in the PCR machine. The reaction was purified using the QIAquick PCR Purification kit (Qiagen) and the DNA was eluted in 35 μl of 3-fold diluted EB (Qiagen). The DNA was digested in a 50 μl reaction with 20U each of SacI and KpnI. The DNA was analyzed on 1% agarose gel (Figure S3) and the minicircle DNA was purified using the MinElute Gel Extraction kit (Qiagen). The DNA was eluted in 12 μl of 3-fold diluted EB (Qiagen), filtered through a sterile 0.2 μm filter (Millipore, Cat. No. SLGV004SL) and the concentration was determined using the Nanodrop spectrophotometer (Thermo Scientific). The overall yield of the DNA with this procedure is about 3% of the starting DNA. The DNA was diluted to 6 ng/μl in miTE buffer and stored at -80ºC before injection. We have noticed that this DNA is more difficult to microinject than plasmid DNA.

Preparation of mouse genomic DNA No. 1 – for genotyping by long-range PCR. Tissue samples from mouse pups (~5 mm of each tail tip) were collected in 1.5-ml tubes. Each tail was digested in 0.5 ml of lysis buffer (TrisHCl pH 8-8.5, 100 mM; EDTA pH 8, 5 mM; SDS, 0.2%; NaCl, 200 mM; proteinase K, 0.2 mg/ml) at 55ºC overnight. The digestion was centrifuged on the next day for 5 min. at ≥10,000 g, and 450 μl of the supernatant were transferred to a new tube. After adding 450 μl of 5M NaCl, the tubes were rocked for 5 min. at room temperature. The samples were centrifuged at ≥10,000 g for 10 min. 750 μl of the supernatant were transferred to a new tube and precipitated with 750 μl of isopropanol. The samples were centrifuged for 15 min at ≥10,000 g at room temperature. The pellet was washed with 500 μl of 70% ethanol and the tubes were air dried for 5-10 min. The pellet was dissolved in 200 μl of TE (1 mM EDTA, 10 mM Tris pH 7.5), and extracted twice with a phenol:chloroform (50:50) mix and twice with chloroform only. The DNA was precipitated with 1/10 volume of 3M sodium-acetate pH 5.2 and 2.7 volumes of ethanol, and dissolved in 200 μl of TE. 1 μl of this solution was used as template in long-range PCR (see below).

Preparation of mouse genomic DNA No. 2 – for genotyping by short-range PCR. Tissue samples from embryos or pups (~2 mm of each tail tip) were collected in 96-well plates, so that many subsequent steps could be done with a multi-channel pipet. The plate was sealed with the plastic cover (ThermalSeal, E&K Scientific, Cat. No. 100-THER-PLT) and briefly centrifuged before the next step to make sure that the tissue samples were on the bottom of the wells. Each tissue sample was lysed with 120 μl of 50 mM NaOH. The plate was sealed with a new plastic cover, incubated in PCR machine at 95ºC for 38 min., briefly centrifuged to collect possible condensation, and the cover was peeled away. At this moment, some gas may be released from the samples and cause droplets of lysate to come close to the rim of the wells. We collected any solution that was close to the rim of the wells by blotting it away carefully with a kimwipe. The lysates were neutralized with 30 μl of 1 M Tris (pH 7.5), tightly sealed with a new plastic seal, vortexed (using a flat head vortex), and briefly centrifuged. 1 μl of this prep was used for PCR (Materials and Methods).

Long-range genomic PCR. We used LA Taq (Takara Bio; Cat Nos. RR02AG and RR002M) and the following primers for *H11* 5' arm: PR374 and PR432; *H11* 3' arm: PR351 and PR422; *Rosa26* 5' arm: Rosa3 and PR432, and *Rosa26* 3' arm: PR351 and PR395. The complete PCR reactions had a volume of 20 μl, and contained 1 μl of genomic DNA that was prepared by the DNA preparation protocol No. 1 above. For *H11* 5'arm, we used the LA PCR buffer II and the following program: 94ºC, 3 min., 40 cycles of: [94ºC, 20 sec.; 60ºC, 30 sec.; 68ºC, 5 min. 30 sec.], 72ºC, 15 min. For *H11* 3' arm, we used the GC buffer I and the following program: 94ºC, 3 min., 40 cycles of: [94ºC, 30 sec.; 56ºC, 30 sec.; 72ºC, 3 min. 30 sec.], 72ºC, 15 min. For *Rosa26* 3' arm, we used the LA PCR buffer II and the following program: 94ºC, 3 min., 40 cycles of: [94ºC, 20 sec.; 58ºC, 30 sec.; 68ºC, 5 min.], 72ºC, 15 min. For *Rosa26* 5' arm, we used the GC buffer I and the following program: 94ºC, 3 min., 40 cycles of: [94ºC, 30 sec.; 60ºC, 30 sec.; 72ºC, 2 min.], 72ºC, 5 min.

Short-range genomic PCR. All short-range PCRs were performed in 20 μl reactions containing 1 μl of prepared DNA (see Preparation of mouse genomic DNA No. 2 above), using Taq polymerase (Qiagen), and the following program: 94ºC, 3 min.; 32 cycles of [94ºC, 20 sec., 60ºC, 25 sec., 72ºC, 45 sec.]; 72ºC, 5 min. Taq polymerase from Qiagen has proven more reliable than polymerases from other manufacturers with this particular DNA preparation. The products were analyzed on a 2% agarose gel.

Primer combinations and expected product sizes for the PCRs used in this study are:

PCR1 in Fig. 1 and Fig. S1; 5'-junction: PR425 and PR436. Expected sizes are: 147 bp, 217 bp, 287 bp, and 244 bp, for insertion into the first *attP*, second *attP*, third *attP*, or full-length *attP*, respectively.

PCR2 in Fig. 1 and Fig. S1; 3'-junction: PR522 and PR387. Expected sizes are: 371 bp, 301 bp, 231 bp, and 313 bp, for insertion into the first *attP*, second *attP*, third *attP*, or full-length *attP*, respectively.

PCR3 in Fig. 1; 5'-junction*:* PR425 and PR551. Expected sizes are 395 bp, 465 bp, 535 bp, and 492 bp, for insertion into the first *attP*, second *attP*, third *attP*, or full-length *attP*, respectively.

PCR4 in Fig. 1; 3'-junction: PR488 and PR387. In the case of cassette exchange with *pBT374*, expected sizes are 502 bp, 432 bp, and 362 bp, for insertion into the first *attP*, second *attP*, and third *attP*, respectively. For insertions of the minicircle derived from *pBT346*, expected

sized are: 463 bp, 393 bp, 323 bp, and 405 bp for insertion into the first *attP*, second *attP*, third *attP*, and full-length *attP*, respectively.

PCR4' in Fig. S1; 3'-junction**:* This PCR can be used instead of ***PCR4. It detects the same junction, but instead of PR488, it uses PR487. The expected sizes for minicircle insertions are: 544 bp, 474 bp, 404 bp, and 405 bp for insertion into the first *attP*, second *attP*, third *attP*, and full-length *attP*, respectively.

PCR5 in Fig. S1; 3'-junction: PR21 and PR387. Expected sizes are 498 bp, 428 bp, 358 bp, and 440 bp, for insertion into the first *attP*; second *attP*, third *attP*, or full-length *attP*, respectively. The products will be obtained only if the full plasmid is integrated.

PCR6 in Fig. S1; internal: FACS G5' and GFP2-Hermie. Expected size: 420 bp. This PCR amplifies a portion of the GFP cDNA.

PCR7+8 in Fig. S1: SH176, SH178 and PR432. The expected sizes are: 147 bp for any knockin or site-specifically integrated allele into *H11* that has the unique sequence "U" at 5' end (see plasmids above), 321 bp for wt, 726 bp for *H11P*, and 687 bp for *H11P3*.

PCR9 in Fig. S1; 3' junction: PR522 and PR428. Expected sizes are: 178 bp, 248 bp, 318 bp, and 260 bp for insertion into the first *attP*; second *attP*, third *attP*, or full-length *attP*, respectively.

To confirm that a particular insertion into *H11P(3)* detected by ***PCR1*** and ***PCR2*** originated from the plasmid, we performed an additional PCR for the 3' junction, ***PCR5*** (**Fig. S1**). The products were indeed obtained only when the full plasmid was integrated. This PCR was also used to test the cassette exchange founders that were positive for ***PCR1*** and ***PCR2*** and negative for ***PCR3*** and ***PCR4***. Indeed, all those founders were positive for ***PCR5***, thereby confirming that they contain only integration of the plasmid bacterial backbone. For detection of integration into the *H11PNVϕ* or *H11P3NVϕ* alleles we used ***PCR1*** and ***PCR6***, and instead of ***PCR2*** we used ***PCR9***. For detection of any knock-in or site-specifically integrated allele into *H11* we used ***PCR7+8***. For detection of the *H11P* or *H11P3* Flp-out alleles we used ***PCR8***.

For the majority of integrants that were analyzed in detail by sequencing of the recombinant junctions (22 out of 28 founders in **Table S2**), ϕC31 catalyzed precise recombination between *attP* and *attB*. In 6 cases, integration appeared to occur at two different *attP* sites, or it caused the deletion of one or more *attP* sites (for example, see **Fig. S1B**, top panel). These imprecise events occurred only when transgenesis was performed on embryos with three tandem *attP* sites.

For detection of site-specific integration into *R26P3NVϕ* we used ***PCR1***. For detection of any knock-in or site-specifically integrated allele into *Rosa26* we used primers: Rosa10, Rosa11, and PR432. Expected sizes are: 168 bp for any knockin or site-specifically integrated allele into *Rosa26* that has the unique sequence "U" at 5' end (see plasmids above), 330 bp for wt, and 696 bp for *R26P3*. For detection of the *R26P* or *R26P3* Flp-out alleles we used primers Rosa10 and Rosa11.

Quantitative PCR. Quantitative PCR was performed as previously described (46, 47). Each sample was tested in triplicate both for GFP and for the internal control. The primers used for GFP were: LL84 and LL85, and they generate a product of 187 bp. The internal control primers were: IMR0015 and IMR0016, and they generate a product of 200 bp.

Gene targeting in mouse ES cells. After electroporation of targeting constructs *pBT311*, *pBT312*, *pBT313*, and *pBT314* into mouse ES cells of 129 strain origin (10), individual G418-

resistant clones were evaluated for homologous recombination by long-range PCR (see above). The clones containing correctly recombined targeting vectors were used to generate mouse chimeras by injection into C57BL/6 blastocysts. The chimeras were crossed to B6D2 F1 females (F1 females from a cross between C57BL/6J and DBA2/J mice; Stock No. 100006, Jackson Lab). Agouti F1 progeny were genotyped for the presence of approproate knockin allele using the same long-range PCRs that were used for screening ES cells. Subsequent genotyping was performed with short-range PCR described above.

Microinjection for generation of site-specific integrants, additional notes. We have tested Qiagen maxi-prep DNA that was filtered through the 0.2 μm filter for injections. We noticed that although this DNA was RNase-free, it was more difficult to inject. Phenol/chloroform extractions followed by filtration as described above greatly facilitated the injection of this DNA.

We would not recommend the use of homozygous *attP*-containing F0 animals that do not contain site-specific integrations and were obtained from injections in subsequent injections, as they may contain random insertions or conversions of *attPx3* into *attPx2* or a single *attP*, which, although infrequent, have been observed. The use of these animals for subsequent injections is recommended only after proper control PCRs exclude the animals with undesirable events mentioned above.

To test for integrity of RNA after each injection, we analyzed the remaining DNA/RNA injection mix on 1% agarose gel (after incubation with the Ambion loading buffer as described for the analysis of *in vitro* transcribed RNA).

Microinjection of Flpo mRNA for generation of Flp-out alleles. To remove the *NV*ϕ cassette, which is flanked by *FRT5* sites (2, 3), we injected the capped *in vitro* transcribed Flpo mRNA obtained from *pFlpo* (Addgene plasmid 13792) (Ref. 4) at 48 ng/μl in miTE into the cytoplasm of *attP*-homozygous embryos. The average efficiency of Flp-out was 25.6% (32 out of 125) for *H11P(3)NV*ϕ and 7.4% (6 out of 81) for *R26P(3)NV*ϕ. None of the animals obtained were homozygous Flp-outs (n=206) based on the PCR that can detect both the Flp-out and non-Flp-out alleles. We mated the animals containing the same Flp-out allele to each other to create homozygous Flp-out mouse lines.

Generation of *GFP-Flpo* transgenic mice. *GFP-Flpo* mice were generated as random integrants from an experiment in which *pBT340* was co-injected with ϕC31o mRNA into *H11P3NV*ϕ homozygous embryos of mixed background in an attempt to achieve site-specific integration of this plasmid and subsequent removal of the bacterial backbone by Flpo-mediated self-excision. The site-specific integration was not successful (0/106 F0 founders screened). However, several random insertions were retained for further characterization, in order to select an efficient Flpo line that can be detected by ubiquitous GFP expression (the progeny can be screened for GFP fluorescence with a UV lamp). The activity of one of the *GFP-Flpo* lines was initially evaluated by analyzing the Flp-out frequency for the *H11P3NV*ϕ allele (removal of the *NV*ϕ cassette) to create the *H11P3* allele. As all F1 progeny from this founder, which was generated by a random insertion of *pBT340* into *H11P3NV*ϕ homozygous embryos, were heterozygous for the *H11* knock-in, we were able to establish the efficiency of Flp-out after a single cross to wt mice. We analyzed all F1 progeny that was negative for *GFP-Flpo*, and by the nature of the cross, heterozygous for the *H11* knock-in allele to detect *H11P3* and *H11P3NV*ϕ

alleles. Based on this experiment, the efficiency of Flp-out was 100%, as only *H11P3* and no *H11P3NVϕ* alleles could be detected among the F1 progeny.

Removal of bacterial backbone from site-specific transgenes by crossing to *GFP-Flpo* mice. The bacterial backbone can be removed from site-specific transgenes if the plasmid that was used to generate the transgene contains an *FRT5* site between the 3' end of the transgene and the plasmid bacterial backbone (e.g., *pBT344* or *pBT366*). This procedure requires two crosses: first one to create double heterozygous animals containing a site-specifically integrated allele and *GFP-Flpo*, and the second one to remove the *GFP-Flpo* transgene. For both *H11P3-pCA-GFP-BB* and *H11P3-pHb9-GFP-BB*, all progeny from the second cross that did not contain *GFP-Flpo*, but contained the site-specific integration allele, had the bacterial backbone removed (i.e., detected by genotyping as *H11P3-pCA-GFP/wt* and *H11P3-pHb9-GFP/wt*, respectively). Therefore, with this crossing scheme and our *GFP-Flpo* mice, the corresponding alleles without the bacterial backbone were generated at 100% efficiency.

We also examined the F1 progeny from the cross of *H11P3-pHb9-GFP-BB/wt* (male) to *GFP-Flpo* (female) for possible bacterial backbone excision by maternal contribution of the Flp recombinase. We did not detect any Flp-out in F1 animals from this cross that were positive only for the site-specifically inserted *Hb9* allele.

We also compared the efficiency of our *GFP-Flpo* line with *Rosa-Flpe* (Jackson Labs, Stock No. 003946) (Ref. 11). *Rosa-Flpe* generated Flp-out only in a small minority of F2 progeny following the two-cross scheme described above.

Tissue preparation and immunohistochemistry. Tissues were obtained from postnatal day 21 (± 2 days) mice that were transcardially perfused with 4% paraformaldehyde (PFA) in phosphate buffered saline (PBS). The tissues were post-fixed overnight in 4% PFA, washed once with PBS and cryoprotected in 30% sucrose overnight. The tissues were embedded in OCT (Tissue-Tek) and stored at -80ºC before cryosectioning. Tissue sectioning was performed using a Leica cryostat. The livers and hearts were sectioned coronally to obtain 10 μm-thick sections, washed in PBS, stained with DAPI and mounted in Fluorogel (Electron Microscopy Sciences, Cat. No. 17985-11). The livers and hearts were imaged with a fluorescence microscope (Nikon). The brains were sectioned sagittally to obtain 30 μm-thick sections, the sections were washed 3 times in PBS and incubated overnight at 4ºC with chicken anti-GFP antibody (Aves Labs) at 1:500 dilution and monoclonal mouse anti-calbindin antibody (Sigma) at 1:3000 dilution. Following incubation with fluorophore-conjugated secondary antibodies (Jackson ImmunoResearch) and DAPI, the slides were washed 3 times in PBS, mounted in Fluorogel, and imaged with a Zeiss confocal microscope.

E11 embryos were dissected in ice-cold PBS, fixed for 2 h at 4ºC with shaking, washed 3 times in ice-cold PBS, and cryoprotected in 30% sucrose overnight. Embryos were embedded into OCT coronally and sectioned at 12 μm thickness with a Leica cryostat. The sections were washed 3 times in PBS and incubated overnight at 4ºC with chicken anti-GFP antibody (Aves Labs) at 1:500 dilution and polyclonal rabbit anti-N-terminal Hb9 antibody (generous gift of S. Arber) (Refs. 9, 12) at 1:1000 dilution. Following incubation with fluorophore-conjugated secondary antibodies (Jackson ImmunoResearch) and DAPI, the slides were washed 3 times in PBS, mounted in Fluorogel, and imaged with a Zeiss confocal microscope.

References:

1. Groth AC, Olivares EC, Thyagarajan B, & Calos MP (2000) A phage integrase directs efficient site-specific integration in human cells. *Proc Natl Acad Sci U S A* 97(11):5995-6000.
2. Seibler J & Bode J (1997) Double-reciprocal crossover mediated by FLP-recombinase: a concept and an assay. *Biochemistry* 36(7):1740-1747.
3. Seibler J, Schubeler D, Fiering S, Groudine M, & Bode J (1998) DNA cassette exchange in ES cells mediated by Flp recombinase: an efficient strategy for repeated modification of tagged loci by marker-free constructs. *Biochemistry* 37(18):6229-6234.
4. Raymond CS & Soriano P (2007) High-efficiency FLP and PhiC31 site-specific recombination in mammalian cells. *PLoS ONE* 2(1):e162.
5. Gallardo T, Shirley L, John GB, & Castrillon DH (2007) Generation of a germ cell-specific mouse transgenic Cre line, Vasa-Cre. *Genesis* 45(6):413-417.
6. Hippenmeyer S*, et al.* (2010) Genetic mosaic dissection of Lis1 and Ndel1 in neuronal migration. *Neuron* 68(4):695-709.
7. Srinivas S*, et al.* (2001) Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus. *BMC Dev Biol* 1:4.
8. Hollis RP*, et al.* (2003) Phage integrases for the construction and manipulation of transgenic mammals. *Reprod Biol Endocrinol* 1:79.
9. Arber S*, et al.* (1999) Requirement for the homeobox gene Hb9 in the consolidation of motor neuron identity. *Neuron* 23(4):659-674.
10. Nagy A, Rossant J, Nagy R, Abramow-Newerly W, & Roder JC (1993) Derivation of completely cell culture-derived mice from early-passage embryonic stem cells. *Proc Natl Acad Sci U S A* 90(18):8424-8428.
11. Farley FW, Soriano P, Steffen LS, & Dymecki SM (2000) Widespread recombinase expression using FLPeR (flipper) mice. *Genesis* 28(3-4):106-110.
12. Thaler J*, et al.* (1999) Active suppression of interneuron programs within developing motor neurons revealed by analysis of homeodomain factor HB9. *Neuron* 23(4):675-687.

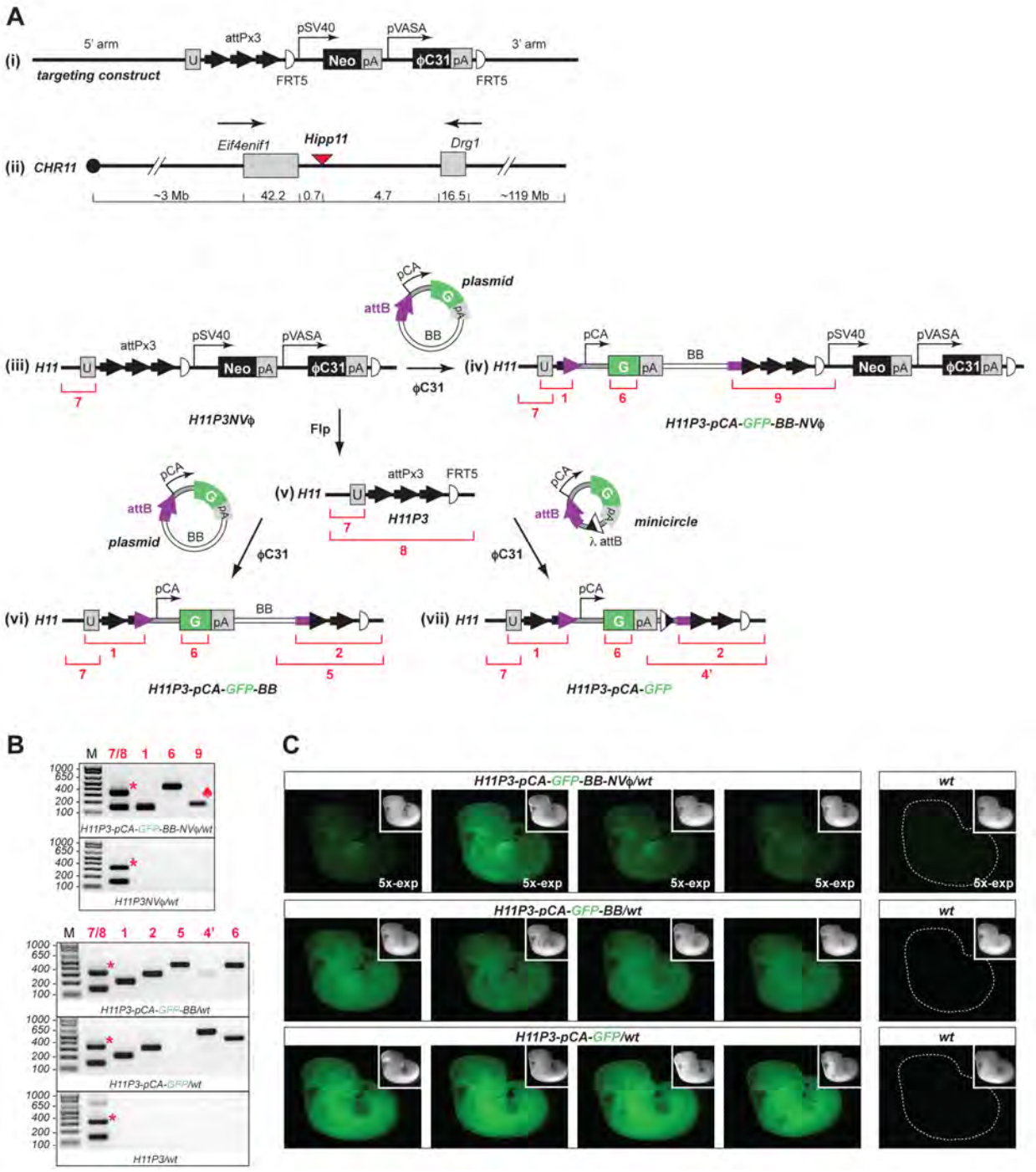

Figure S1. Site-specific transgenesis – proof of principle.
(***A***) Generation of *H11* knockin alleles containing ϕC31 *attP* sites. (**i**) A schematic of the recombinant DNA construct for introduction of three *attP* sites and the *NVϕ* cassette into *H11* via homologous recombination in mouse ES cells. Two versions, containing either a single "full-length" *attP* site or three tandem short *attP* sites, were generated. (**ii**) mouse chromosome 11 with the *Hipp11 (H11)* locus designated as a red triangle. The scale below is in kilobases, except where megabases (Mb) are indicated. (**iii**) The *H11P3NVϕ* allele resulting from homologous

recombination between (**i**) and (**ii**). (**iv**) *H11P3-pCA-GFP-BB-NVϕ* allele obtained by ϕC31-catalyzed site-specific insertion of the *pattB-pCA-GFP* plasmid *(pBT316)* into the first *attP* site. All three *attP* sites are suitable recipients for the transgene and the site used in any particular case can be determined by PCR. (**v**) *H11P3* locus that was generated from (**iii**) by Flpo mRNA injection into the cytoplasm of mouse embryos carrying (**iii**) (see SI Materials and Methods). The *H11* locus with a single *attP* site and the *Rosa26* locus with either a single or three *attP* sites were generated in the same manner. (**vi**) and (**vii**), two products obtained by ϕC31-catalyzed site-specific insertion of the *pattB-pCA-GFP* plasmid (*pBT316*, left) or the *attB-pCA-GFP* minicircle (generated from *pBT346*, right). The corresponding alleles are: *H11P3-pCA-GFP-BB* (left) and *H11P3-pCA-GFP* (right), respectively.

(***B***) PCR results on N1 animals confirming site-specific integrations. The DNA template for each PCR panel was obtained from a mouse of the genotype designated below each gel. The red numbers correspond to the PCR products designated on the schemes by red brackets and numbers in (*A*). The primer set #9 amplified a band smaller than expected due to the deletion of two *attP* sites during integration (spade). The wt *H11* locus is also amplified by primer set #8 to generate a 321 bp band (asterisk, see schematic (**v**)).

(***C***) GFP expression in N2 mouse embryos at embryonic day 11. Each row shows representative embryos from a single pregnancy with genotypes designated above. Images were obtained under identical conditions, except that "5x-exp" designates five-fold longer exposure time than for the rest of the images. Insets represent the corresponding bright field images of each embryo.

Abbreviations: pSV40, *SV40* promoter; pVASA, *VASA* promoter; U, unique sequence; FRT5, a mutant version of *FRT* that is compatible with itself but not with wt *FRT*; pCA, β-actin promoter and CMV enhancer; G, GFP; pA, *polyA* signal; BB, plasmid bacterial backbone; attB and attP, ϕC31 *attB* and *attP* sites.

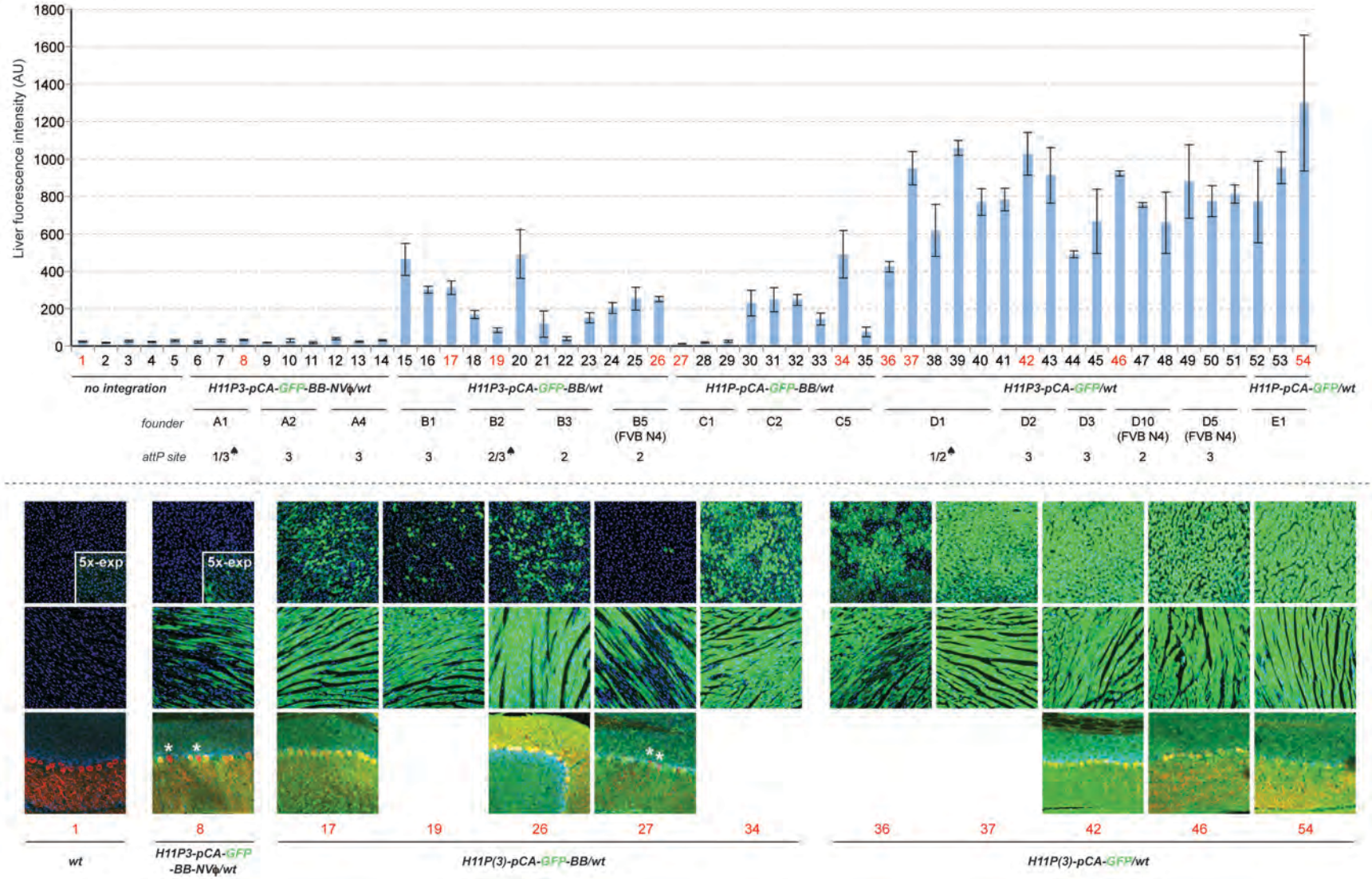

Figure S2. GFP expression in individual adult animals carrying *pCA-GFP* site-specific transgenes introduced by ϕC31 integrase-mediated transgenesis. Each column on the graph represents average fluorescence in arbitrary units (AU) in the GFP channel for liver sections from individual animals represented by numbers and genotypes below ("no integration" represents *wt*, *H11P3NVϕ/wt*,

H11P3/wt and *H11P/wt* genotypes). The animals are grouped according to the genotype and founder (designated further below). For each founder, one of the three sites from *H11P3* into which the site-specific integration occurred is indicated below. In cases labeled by spades, recombination appeared to occur at two different *attP* sites, or it resulted in deletion of some of the *attP* sites. For a subset of animals, representative images of liver and heart sections are shown below. The images show native GFP fluorescence in green and nuclei stained by DAPI in blue. Further below, corresponding confocal images of cerebellar sections are shown for a subset of animals. The sections are stained by anti-GFP antibody (green), anti-calbindin (red) for Purkinje cells, and DAPI (blue). The numbers below the images correspond to the numbers of individual animals below the chart. The Purkinje cells in animals #8 and #27 that appear calbindin(+) but GFP(-) are indicated by asterisks. GFP expression in the liver appears most sensitive to the presence of the bacterial backbone and the *NVϕ* cassette. 17 out of 19 transgenic animals (#36-54), which do not contain the *NVϕ* cassette and the bacterial backbone, show uniform GFP expression. The exceptions are animals #36 (the strongest variability observed in this set) and #38. The tissue section images for animals #1, 8, 17, and 42 also appear in Fig. 2.

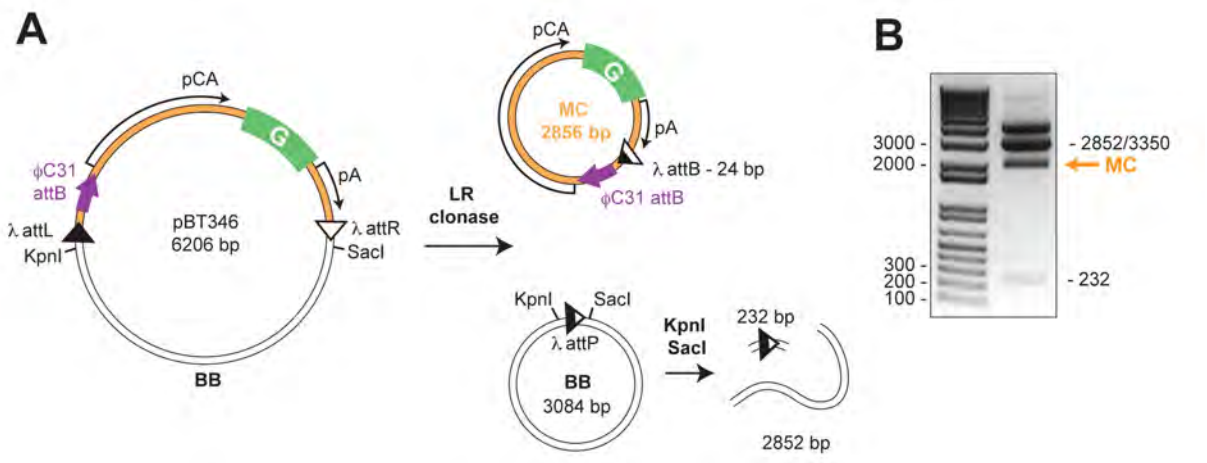

Figure S3. Generation of minicircle DNA with λ integrase and excisionase (LR clonase, Invitrogen) *in vitro*. (***A***) From left to right: The starting plasmid (*pBT346*) contains λ *attL* and *attR* sites, which recombine in the LR clonase-catalyzed reaction to generate two minicircles: one (*MC*) contains the ϕC31 *attB* site and *pCA-GFP*, and the other contains the plasmid bacterial backbone (*BB*). After recombination, the DNA is treated with appropriate restriction endonucleases to selectively digest the *BB* minicircle and the starting plasmid. (***B***) The recombined and digested DNA is run on 1% agarose gel. *MC* DNA (orange arrow) migrates faster than the linear *BB* or plasmid DNA and is purified from the gel for microinjection.

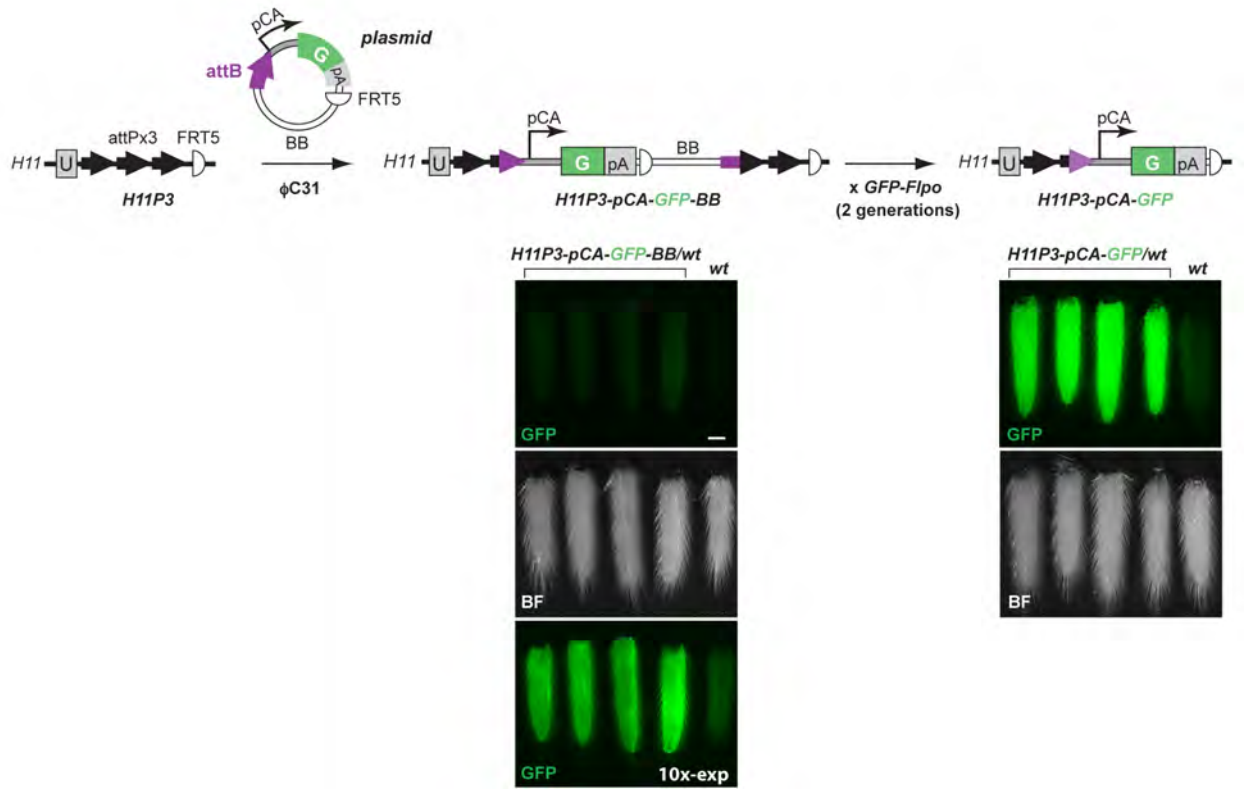

Figure S4. GFP expression from a site-specifically integrated *H11P3-pCA-GFP-BB* transgene increases upon the removal of the bacterial backbone. Top, schematic of the generation of the *H11P3-pCA-GFP-BB* transgene (from *pBT344*) and subsequent derivation the *H11P3-pCA-GFP* transgene from it, by crossing to the *GFP-Flpo* transgenic mouse to remove the bacterial backbone (SI Materials and Methods). Below, four tail tips for each genotype designated above and a tail tip of a wt littermate were imaged for GFP fluorescence using identical imaging conditions. Bright field (BF) images of the same tails are shown further below. 10x-exp, the same tails above were imaged for GFP fluorescence with 10-times longer exposure. Scale bar, 1 mm.

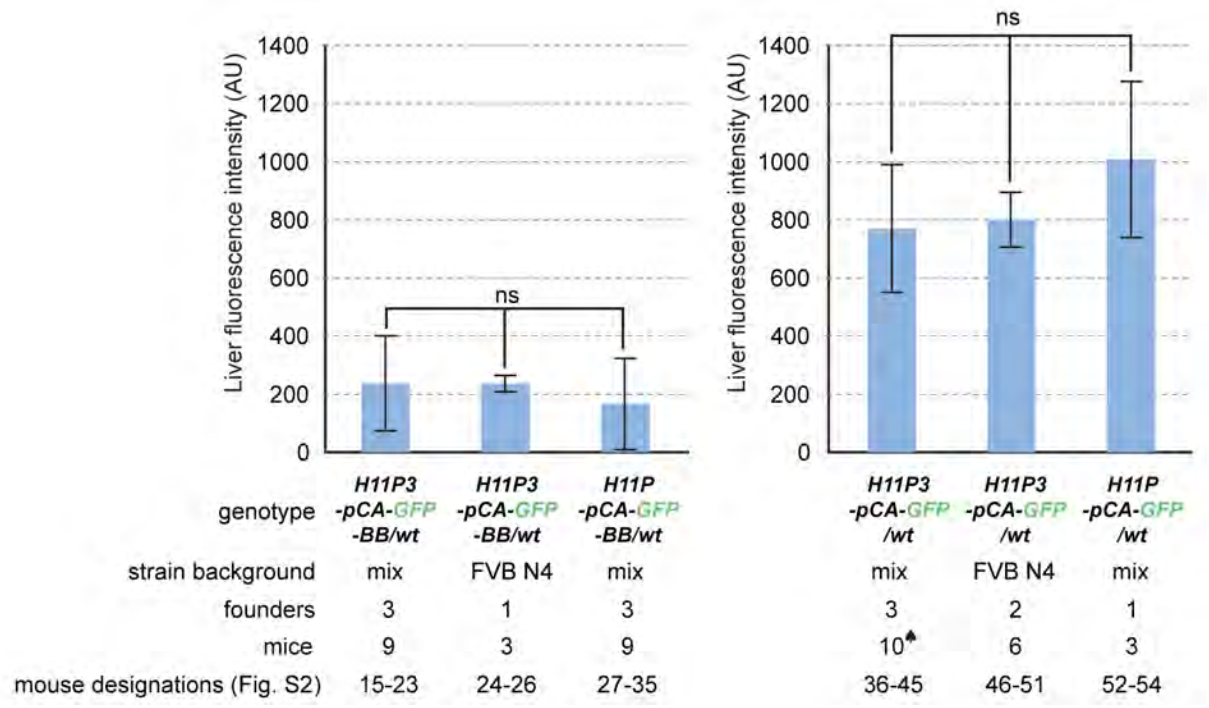

Figure S5. Average GFP fluorescence in livers does not differ between mice containing an insertion of the same transgene into one of the three *attP* sites from *H11P3* or insertion into a single *attP* site from *H11P* (compare 1st and 3rd columns in both graphs). The GFP fluorescence is also not affected by the genetic background of the injected embryos (compare 1st and 2nd columns in both graphs). Each dataset is represented by mean ± standard deviation. The numbers of individual animals and founders analyzed for each genotype are listed below the genotypes. When samples from multiple founders were combined to obtain an average, each founder was represented by the same number of animals except in the case labeled by a spade. Mouse designations are numbers used to represent each mouse in Fig. S2. Statistical comparisons were performed with one-way ANOVA.

Row	DNA[a]	DNA type	DNA size (kb)	Strain	Back-ground	Embryos injected (n)	Embryos implanted (n)	Implanted/ Injected (%)	F0 (n)	F0/ Implan-ted (%)	SS (n)	SS% (of injected)	SS% (of implan-ted)	**SS% (of F0)**	R (n)[h]	R% (of implan-ted)	R% (of F0)	Experi-ments (n)
1	***attB*-*pCA-GFP***	minicircle	~3	*H11P*	mix	141	115	82	21	18	1	0.7	0.9	**4.8**	1	0.9	4.8	4
2	***attB*-*pCA-GFP***	minicircle	~3	*H11P3*	mix	168	136	81	39	29	4	2.4	2.9	**10.3**	1	0.7	2.6	4
3	***attB*-*pCA-GFP***	minicircle	~3	*H11P3*	FVB N4	122	115	94	15	13	6	4.9	5.2	**40.0**	3	2.6	20.0	3
4	***attB*-*pCA-GFP*-*attB***	plasmid	~6	*H11P3*	FVB N4	128	119	93	38[c]	32	6[e]	15.8	5.0	**15.8**	1	0.9	2.6	1
5	***attB*-*pCA-GFP*, no RNA**	plasmid	~6	*H11P3NVϕ*	mix	160	78	49	32[c]	41	0	0.0	0.0	**0.0**	5	6.4	15.6	1
6	***attB*-*pCA-GFP***	plasmid	~6	*H11P3NVϕ*	mix	292	223	76	64[d]	29	10	3.4	4.5	**15.6**	4	1.8	6.3	3
7	***attB*-*pCA-GFP***	plasmid	~6	*H11PNVϕ*	mix	140	89	64	30[c]	34	2	1.4	2.2	**6.7**	0	0.0	0.0	2
8	***attB*-*pCA-GFP-FRT5***	plasmid	~6	*H11P*	mix	264	232	88	51	22	5[f]	1.9	2.2	**9.8**	3[g]	0.9	5.9	5
9	***attB*-*pCA-GFP-(FRT5)*[b]**	plasmid	~6	*H11P3*	mix	142	129	91	61	47	4[f]	2.8	3.1	**6.6**	9[g]	9.9	14.8	4
10	***attB*-*pCA-GFP-FRT5***	plasmid	~6	*H11P3*	FVB N4	50	43	86	8	19	3[f]	6.0	7.0	**37.5**	1[g]	0.0	10.3	1
11	***attB*-*pHB9-GFP-FRT5***	plasmid	~14	*H11P3*	FVB N4	305	267	88	66[d]	25	2[f]	0.7	0.7	**3.0**	2[g]	0.7	3.0	2
12	***attB*-*pCA-GFP***	plasmid	~6	*R26P3NVϕ*	mix	83	63	76	22[c]	35	2	2.4	3.2	**9.1**	2	3.2	9.1	1

Abbreviations: F0, embryos or animals obtained from injections; SS, site-specific integration; R, random integration; mix, mixed background of 129, C57BL/6 and DBA2; FVB N4, mice of the mixed background were outcrossed for 4 generations to the FVB strain and then intercrossed.

[a] All DNA was injected at the final concentration of 3 ng/μl. All DNA was coinjected with ϕC31o mRNA at the final concentration of 48 ng/μl, except for the experiment in row 5. [b] Both FRT and non-FRT versions of *pattB-pCA-GFP* (*pBT316 or pBT344*) were used. [c] F0s were analyzed only as E10 or E11 embryos. [d] F0s were analyzed either as E10 or E11 embryos or as live pups. [e] The 6 founders with cassette exchange listed here contain *pCA-GFP* without the bacterial backbone; 5 more founders with cassette exchange contained only the bacterial backbone. Therefore the total number of founders with cassette exchange is 11 (29%). [f] One F0 with site-specific insertion, also contains a random insertion. [g] One random insertion was found in a founder that also contains a site-specific insertion. [h] The number of random insertions may be somewhat underreported when F0 embryos were included in the analysis, as random insertions are not detectable by GFP-specific PCR if they occur in embryos that also contain site-specific insertions.

Table S1. Stepwise efficiency of site-specific integration

Transgene	Founder	Back-ground	Sex	Insertion into attP 1, 2 or 3	GT of SS (Y/N)[a]	N1 (n)	N1 SS (n)	N1 SS (%)[c]	N2/3 (n)	N2/3 SS (n)	N2/3[d] SS (%)	R (Y)	N1 R (n)	N1 R (%)
H11P3-pCA-GFP-BB-NVϕ	A1	mix	M	1/3[b]	Y	55	36	65.5	25	15	60.0		0	0
	A2	mix	M	3	Y	62	12	19.4***	16	10	62.5		0	0
	A3	mix	M	1/3[b]	Y	36	10	27.8*					0	0
	A4	mix	M	3	Y	63	29	46.0	14	4	28.6		0	0
H11P3-pCA-GFP-BB	B1	mix	M	3	Y	143	19	13.3***	48	20	41.7	Y	11	7.7
	B2	mix	M	2/3[b]	Y	29	11	37.9	85	36	42.4		0	0
	B3	mix	Fem	2	Y	44	9	20.5***	76	34	44.7		0	0
	B4	mix	Fem	2	Y	20	4	20.0*					0	0
	B5	FVB N4	M	2	Y	24	9	37.5					0	0
	B6	FVB N4	Fem	2	N									
	B7	FVB N4	Fem	2/3[b]	Y	8	3	37.5				Y	3	37.5
H11P-pCA-GFP-BB	C1	mix	Fem	n/a	Y	15	11	73.3					0	0
	C2	mix	M	n/a	Y	44	9	20.5***					0	0
	C3	mix	Fem	n/a	Y	9	4	44.4				Y	2	22.2
	C4	mix	Fem	n/a	N	1	0	0					0	0
	C5	mix	M	n/a	Y	67	14	20.9***	62	29	46.8		0	0
H11P3-pCA-GFP	D1	mix	Fem	1/2[b]	Y	20	8	40.0	25	15	60.0		0	0
	D2	mix	M	3	Y	33	10	30.3	12	4	33.3		0	0
	D3	mix	Fem	3	Y	23	5	21.7*					0	0
	D4	mix	M	3	N	59	0	0***					0	0
	D5	FVB N4	Fem	3	Y	32	6	18.8**					0	0
	D6	FVB N4	Fem	3	Y	41	10	24.4**					0	0
	D7	FVB N4	M	3	N									
	D8	FVB N4	Fem	3	Y	14	2	14.3*					0	0
	D9	FVB N4	M	1/3[b]	N									
	D10	FVB N4	Fem	2	Y	27	8	29.6	20	10	50.0		0	0
H11P-pCA-GFP	E1	mix	M	n/a	Y	39	11	28.2*					0	0
H11P-pHB9-GFP	F1	FVB N4	Fem	2	Y	45	9	20.0***					0	0

Abbreviations: SS, site-specific integration; R, random integration; GT, germline transmission; mix, mixed background of 129, C57BL/6 and DBA2; FVB N4, the mice of the mixed background were outcrossed for 4 generations to the FVB strain and then intercrossed; N1, progeny from the first generation where a founder was crossed to CD1 wt animal; N2 or N3, progeny from the second or third generation where N1 or N2 animals were crossed to CD1 wt animals, respectively.

[a] The founders that did not transmit were ether sterile (D7 and D9), died upon delivery (C4), or cannibalized the pups (B6). [b] Based on site-specific PCR, integration appeared to occur at two different *attP* sites, or it caused the deletion of one or more *attP* sites.

[c] Frequency of germline transmission for site-specific integrations from some founders was sub-Mendelian, suggesting mosaicism in the founders (Fisher's exact test, *, p<0.05; **, p<0.01; ***, p<0.001). [d] Frequency of transmission for site-specific integrations for subsequent generations (N2 and/or N3) is statistically indistinguishable from Mendelian transmission (Fisher's exact test).

Table S2. Complete list of transgenic founders (n=28) and their germline transmission efficiency.

Row	Pooled rows from Table 1[a]	Strain	Background	F0 (n)[a]	SS F0 (n)[a]	Significant?[b]	**SS% (from F0)**
1	1, 7, 8	*H11P(NVϕ)*	mix	102	8	} ns	**7.8**
2	2, 6, 9	*H11P3(NVϕ)*	mix	164	18		**11.0**
3	3, 10	*H11P3*	FVB N4	23	9	} p=0.0015	**39.1**

Abbreviations: SS, site-specific integration; mix, mixed background of 129, C57BL/6 and DBA2; FVB N4, mice of the mixed background were outcrossed for 4 generations to the FVB strain and then intercrossed.
[a] Pooled results from Table 1 for the same strain and background (regardless of the presence of the *NVϕ* cassette) used for injections of 3-6 kb DNA. The pooling did not include Table 1 entry 4 (as the mechanism of integration may be different); entry 11 (as the DNA used is dramatically larger); entry 5 (as no RNA was coinjected); and entry 12 (as it used the *Rosa26* locus). [b] Statistical significance was evaluated using Fisher's exact test. ns, not-significant; α=0.05.

Table S3. Comparisons of integration efficiencies: single vs. triple *attP* alleles, and mixed vs. FVB background.

Name	Sequence
21	ctgcaaggcgattaagttgg
351	aataaGCTAGCctcgagGATATCctgtgccttctagttgccag
374	atgtgaggcaggagatgagagaggaatgactggtcac
387	gtgggactgctttttccaga
395	gttgagggcaatctgggaaggt
402	ctagCCTGCAGGaattaaGTTAACaattaaGACGTC
403	ctagGACGTCttaattGTTAACttaattCCTGCAGG
422	ccatttttagtacccctctacactcctcc
425	ggtgataggtggcaagtggtattc
428	ccgaaaagtgccacctgaataat
432	GATATCCTTACGGAATACCACTTGCCACCTATCACC
436	atcaactaccgccacctcgac
437	AACCAACCttaaCCGCCACCATGGATACCTAC
438	AATAggatccTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTctcgagTCACACTTTCCGCTTTTTCTTAGG
487	tccccctgaacctgaaacat
488	gcaatagcatcacaaatttcacaa
493	aaagaGGTACCagttacgctagggataacagggtaatatagCAAATAATGATTTTATTTTGACTGATAG
494	aaataCTCGAGagcctGCTTTTTTGTACAAAGTTG
495	aagaaGCGGCCGCacaagtttgtacaaaaaagcTGAACG
496	AAGAAgagctcCATAGTGACTGGATATGTTGTGTTTTA
522	CGATGTAGGTCACGGTCTCG
551	GGCTATGAACTAATGACCCCGTA
FACS G5'	CTTCAAGTCCGCCATGCCCGA
GFP2-Hermie	TCCAGCAGGACCATGTGATCGC
IMR0015	CAAATGTTGCTTGTCTGGTG
IMR0016	GTCAGTCGAGTGCACAGTTT
LL84	AAGTCGTGCTGCTTCATGTG
LL85	ACGTAAACGGCCACAAGTT
Rosa10	CTCTGCTGCCTCCTGGCTTCT
Rosa11	cgaggcggatcacaagcaata
Rosa3	ccactgaccgcacggggattc
SH176	tggaggaggacaaactggtcac
SH178	ttccctttctgcttcatcttgc

Table S4. List of primers used in this study.