www.sciencemag.org/cgi/content/full/1134426/DC1

# Supporting Online Material for

## P[acman]: A BAC Transgenic Platform for Targeted Insertion of Large DNA Fragments in *Drosophila melanogaster*

Koen J. T. Venken, Yuchun He, Roger A. Hoskins, Hugo J. Bellen*

*To whom correspondence should be addressed. E-mail: hbellen@bcm.edu

Published 30 November 2006 on *Science* Express
DOI: 10.1126/science.1134426

**This PDF file includes:**

**SUPPORTING ONLINE MATERIAL**

**FOR**

# P[acman]: A BAC Transgenic Platform for targeted insertions of 100 kb DNA fragments in *Drosophila*

Koen J.T. Venken[1], Yuchun He[2,3], Roger A. Hoskins[5], Hugo J. Bellen[1,2,3,4,#]

[1] Program in Developmental Biology

[2] Department of Molecular and Human Genetics

[3] Howard Hughes Medical Institute

[4] Department of Neuroscience

Baylor College of Medicine

Houston, TX 77030

[5] Department of Genome Biology

Lawrence Berkeley National Laboratory

Berkeley, CA 94720-3200

# corresponding author; email: hbellen@bcm.edu

## SUPPLEMENTARY MATERIALS AND METHODS

### Bacterial transformation, colony PCR, primers, and enzymes

Bacteria were grown on Luria-Bertani (LB) broth. The following antibiotic concentrations were used: kanamycin (15 μg/ml: K15 or 30 μg/ml: K30), ampicillin (50 μg/ml: A50 or 100 μg/ml: A100), chloramphenicol (12.5 μg/ml: C12.5), tetracycline (10 μg/ml: T10), and streptomycin (60 μg/ml: S60). Electroporations were performed in 1 mm cuvettes using a Biorad Gene Pulser set at 1.8 kV, 200 Ohm and 25 μFD. Chemical transformation was carried out with chemocompetent cells according to the supplier's instructions (Invitrogen). SOC medium was used in recovery steps after electroporation and chemical transformation. Colony PCR screening was performed using the following parameters: a denaturation cycle (94°C for 10'), 35 amplification cycles (94°C, 30 seconds; 52°C, 30 seconds, 72°C, 1 minute) and a post-amplification cycle (72°, 10'). Primers were obtained from Operon. Restriction endonucleases, T4 DNA ligase, Klenow DNA polymerase and Calf Intestinal Alkaline Phosphatase (CIAP) were obtained from New England Biolabs. Taq polymerase and Pfu polymerase were obtained from QIAGEN and Stratagene, respectively. Gel and PCR purifications were carried out using the QIAquick Gel Extraction and the PCR Purification Kit, respectively (QIAGEN). Miniprep plasmid DNA was prepared using the QIAprep Miniprep Kit (QIAGEN).

### Recombineering and plasmid copy-number induction

Recombineering was performed as described previously (*1-3*) with minor modifications, using the DY380 strain (*3*) or mini-lambda (*2*). Specific details for each recombineering experiment are described in the relevant sections below. In general, single colonies of a bacterial

host with recombineering functions and containing a P[acman] or BAC plasmid were grown overnight at 30°C in 5 ml of LB containing appropriate antibiotics. A 300 µl aliquot of overnight culture was diluted 1:50 with LB containing appropriate antibiotics and grown further at 30°C to an $OD_{600}$ between 0.4 and 0.6. Uninduced control samples were processed in parallel. The induced sample was incubated for 15'at 42°C in a shaking water bath. The uninduced control was kept at 30°C for 15'. After 15', both samples were chilled in an ice-water-bath for 5'. Bacteria were pelleted and washed in 25 ml autoclaved milliQ water, twice, and pelleted again. Bacteria were resuspended in 1 ml 10% glycerol, pelleted again. Most of the 10% glycerol was removed, and the pellet was resuspended in the remaining liquid (about 120 µl) and subdivided into two aliquots of 60 µl cells, sufficient for two transformations. Cells were transformed with 2 µl of purified PCR product (for plasmid construction steps) or linearized plasmid (for gap-repair) using electroporation. Potential recombinants were selected on LB plates containing appropriate antibiotics at 30°C. True recombinants were identified by colony PCR. Plasmid DNA was isolated and transformed into EPI300 copy-up bacteria (Epicentre). Plasmid copy induction was performed according to the supplier's instructions (Epicentre).

**Chloramphenicol resistant P[acman] plasmid construction**

The *P* element backbone was assembled in pUni/V5-His-TOPO, which is maintained in the PIR1 strain using K30 (Invitrogen). Primer sequences used for P[acman] plasmid construction are listed in Table S3. First, an engineered MCS was subcloned in pBluescript-KS and PCR amplified using primers MCS-Cloning-F and MCS-Cloning-R, which incorporate sites for standard sequencing primers (T7, M13-F and T3, M13-R) flanking the MCS. During this process, two additional small MCSs were incorporated for downstream cloning steps and

excision of the final assembled *P* element backbone. The resulting PCR fragment was TOPO subcloned into pUni/V5-His-TOPO, resulting in plasmid MCS-2. Second, the *5'P* transposase recognition site was amplified from pP{CaSpeR-4} using primers P5-F and P5-R and subcloned as a *NsiI-SpeI* fragment in MCS-2, resulting in MCS-P5-3. Third, the *3'P* transposase recognition site was PCR amplified from pP{CaSpeR-4} using primers P3-F and P3-R and subcloned as a *NheI-AgeI* fragment into MCS-P5-3, resulting in MCS-P5-P3-1. Finally, an *EcoRI-NsiI white*$^{+mc}$ (*w*$^{+mc}$) mini-*white*$^{+}$ fragment, obtained from pP{CaSpeR-4} was subcloned into MCS-P5-P3-1, resulting in the final *P* element backbone MCS-P5-P3-w$^{+}$-7. The entire *P* element was released from MCS-P5-P3-w$^{+}$-7 as a *BamHI* fragment, blunt-ended with Klenow polymerase, and subcloned into a Klenow-blunt-ended and CIAP dephosphorylated 6.4 kb *SalI* fragment from the vector pJW360 (a gift from Waclaw Szybalski) (*4*). This *SalI* fragment contains the pBELOBAC11 vector backbone (*5*) and *oriV* (*4*). The ligation product was electroporated into the EPI300 strain (Epicentre), and selected on LB (C12.5), resulting in the chloramphenicol resistant P[acman], P[acman]-Cl$^{R}$. Plasmid copy number of single colonies was induced as described above. Ligations were performed using T4 DNA ligase. Each cloning step was verified by restriction fingerprinting and DNA sequencing.

**Modification of chloramphenicol resistant BAC's and P[acman] with an ampicillin resistance marker**

P[acman]-Cl$^{R}$ (pBeloBAC11 backbone) (*5*) or BAC clones BACR17E13 and BACR02B03 (pBACe3.6 backbone) (*6*) were transformed into the DY380 strain and grown on LB plates (C12.5/T10). The chloramphenicol marker in both BAC backbones was replaced by an ampicillin marker through recombineering to obtain the ampicillin modified P[acman],

P[acman]-Ap$^R$, or BACs (BAC-Ap$^R$). The ampicillin marker was obtained by PCR from pBluescript-KS using primers Amp-Chl-swap-F-pBS and Amp-Chl-swap-R-pBS (Table S3). The resulting PCR product was gel purified, digested with *DpnI* and PCR purified. Single DY380 colonies (induced versus uninduced for recombineering) containing P[acman]-Cl$^R$ or BAC were electroporated with 2 μl of purified PCR product. Potential recombinants were selected (A100) at 30°C. Plasmid DNA was isolated, electroporated into EPI300 cells, and plasmid copy number of single colonies was induced as described above. Correct recombination was confirmed with restriction fingerprinting (P[acman]-Ap$^R$ and BAC-Ap$^R$) and DNA sequencing (P[acman]-Ap$^R$).

**_attB_-P[acman] plasmid construction**

To obtain *attB* plasmids, a ϕC31 *attB* recognition site was incorporated through two-step recombineering in the DY380 strain using the positive-negative selectable marker pRpsl-Neo (GeneBridges) (*7*). P[acman]-Cm$^R$ and -Ap$^R$ vectors were electroporated in the DY380 strain and grown on LB plates (C12.5/T10 and A100/T10, respectively). In a first step, the rpsL-Neo marker was engineered in front of the *3'P* transposase site using recombineering. The rpsL-Neo cassette was PCR amplified from pRpsL-Neo using primers P[acman]-Neo-SalI-R and P3-Neo-R (Table S3). The resulting PCR product was gel purified, digested with *DpnI* and PCR purified. Single DY380 colonies (induced versus uninduced for recombineering) containing P[acman]-Cm$^R$ and -Ap$^R$ were electroporated with 2 μl of purified PCR product. Potential recombinants were selected on LB plates (C12.5/T10/K15 and A100/T10/K15, respectively) at 30°C. Correct recombination events were identified through colony PCR screening using primers pRpsL-Neo-screen-F and P3-R (Table S3) and tested for streptomycin sensitivity (C12.5/T10/K15/S60 and

A100/T10/K15/S60, respectively) and kanamycin resistance (C12.5/T10/K15 and A100/T10/K15 respectively). Only colonies that were kanamycin resistant, streptomycin sensitive and positive for the colony PCR assay were processed further.

In a second step, the inserted rpsL-Neo marker was replaced by a fragment containing the ϕC31 *attB* site through recombineering to obtain *attB* modified P[acman]-$Cm^R$ and -$Ap^R$ plasmids. The *attB* recognition site was obtained by PCR from pTA-attB (*8*) (gift from Michelle Calos) using primers P[acman]-attB-SalI-R and P3-attB-R (Table S3). The resulting PCR product was gel purified, digested with *DpnI* and PCR purified. Single DY380 colonies (induced versus uninduced for recombineering) containing rpsL-Neo modified P[acman]-$Cm^R$ and -$Ap^R$ were electroporated with 2 μl of purified PCR product. Potential recombinants were selected on LB plates (C12.5/T10/S60 and A100/T10/S60, respectively) at 30°C. True recombinants were identified through colony PCR using primers attB-screen-F and P3-R (Table S3) and by testing for kanamycin sensitivity (C12.5/T10/K15 and A100/T10/K15, respectively) and streptomycin resistance (C12.5/T10/S60 and A100/T10/S60, respectively). Only colonies that were kanamycin sensitive, streptomycin resistant and positive by colony PCR were processed further. Plasmids were electroporated into EPI300 cells, and plasmid copy number of single colonies was induced as described above. Correct integration of the *attB* site, resulting in *attB*-P[acman]-$Cm^R$ and -$Ap^R$ was confirmed by restriction digestion and DNA sequencing.

**Construction and integration of *piggyBac-yellow$^+$-attP* docking sites**

A 5.2 kb *SalI* fragment containing the intronless mini-*yellow$^+$* fragment (y+mDint25.2(S,S)) was obtained from pEPgy2 (*9*) and cloned into the *XhoI* site of the minimal *piggyBac* transposon pXL-BacII (*10*) (gift from Malcolm Fraser), resulting in *piggyBac-yellow$^+$*-

*3* (Element 1, *yellow* in the sense orientation) and *piggyBac-yellow*$^{+}$*-9* (Element 2, *yellow* in the antisense oreintation). A 168 bp fragment encompassing the *attP* site was obtained by PCR from pTA-attP (*8*) (gift from Michelle Calos) using primers attP-shorter-NotI-F and attP-XmaI-R (Table S3) and cloned as a *NotI-XmaI* into *NotI-XmaI* digested *piggyBac-yellow*$^{+}$*-3* and *piggyBac-yellow*$^{+}$*-9*, resulting in *piggyBac-yellow*$^{+}$*-attP-3B* and *piggyBac-yellow*$^{+}$*-attP-9A,* respectively. Both plasmids were transformed into the $y^{-}$ $w^{-}$ fly strain with two different co-injected helper plasmids (*11, 12*) (gifts from DGRC and Alfred Handler, respectively). Since pXL-BacII transposed inefficiently, some insertions were remobilized using two stably integrated *piggyBac* transposase sources $y^{1}$ $w^{1118}$*; CyO, P{FRT(w*$^{+}$*)Hsp70-PBac\T}2/wg*$^{Sp\text{-}1}$ (*11*) and *y* $w^{122}$*;* $wg^{Sp\text{-}1}$*/CyO, P{w*$^{+}$*,* $\alpha$*-1-tub-PBac\T}; TM2/TM6b* (gift from Arzu Celik). All insertions were mapped to a chromosome and balanced. Only homozygous viable docking sites were kept. Genomic DNA of the corresponding strains was isolated, and the presence of the *attP* site was confirmed using PCR with primers attP-F and attP-R (Table S3). The exact locations of insertions were mapped as previously described (*9*). The detailed protocol is available at http://flypush.imgen.bcm.tmc.edu/pscreen/. Genomic DNA was restriction digested with Sau3A I or Hpa II. Flanking sequences were amplified by inverse PCR with primers HY.5.F and HY.5.R for the 5' piggyBac end and HY.3.F and HY.3.R for the 3' piggyBac end. PCR products were directly sequenced with primers HY.5.R and HY.3.F, respectively. Sequence traces were basecalled, trimmed and aligned to the *Drosophila* genome sequence (Release 3) as previously described (*9*). Insertion sites were analyzed using the genome sequence annotations in the Flybase Genome Browser (http://www.flybase.org). Primers for inverse PCR and sequencing are listed in Table S3).

**Recombineering mediated gap-repair**

The protocol describes the retrieval of fragments from BAC (C12.5) or P1 (K30) clones into P[acman]-$Ap^R$ or *attB*-P[acman]-$Ap^R$. The protocol is easily adapted for the retrieval of fragments from BAC-$Ap^R$ (A50) or P1 (K30) clones into P[acman]-$Cm^R$ or *attB*- P[acman]-$Cm^R$ (C12.5). Other antibiotic combinations were required during certain steps. The protocol has also been used to retrieve fragments from high-copy plasmids. Alternatively, for fragments too small to be retrieved by gap repair, PCR amplification and conventional cloning was used. Primer sequences used for cloning of all genomic fragments are listed in Table S4.

BACs containing a gene of interest were identified using either the Release 3 or the Release 4.3 annotated *Drosophila* genome sequence in the Flybase Genome Browser (http://www.flybase.org) and obtained from BacPac Resources (http://bacpac.chori.org/). P1 clones were identified using BLASTN (*13*) or the BDGP ArmView browser (http://www.fruitfly.org/seq_tools/displays/ArmView.html). The region of interest was identified, and four primer sets were designed for each DNA fragment to be gap-repaired (Fig.S1). Primer sets 1 and 2, incorporating appropriate restriction enzyme cloning sites, are designed to PCR amplify two homology arms of 500 bp each, left arm (LA) and right arm (RA). Alternatively, primers were designed for hybrid PCR amplification, which seals two initially obtained PCR fragments together during a secondary PCR. Primer set 3 (5'-Check-R and 3'-Check-F) is used to screen for correct recombinants after gap-repair, in combination with the vector specific primers MCS-F and MCS-R. Primer set 4 (LA-Seq-F and RA-Seq-R) is used to sequence across the two junctions to confirm correct retrieval of the desired fragment into P[acman]. The homology arms were PCR amplified and cloned as *AscI-BamHI* (LA) and *BamHI-PacI* (RA) fragments in *AscI-PacI* digested and PCR purified P[acman] in a three component ligation

reaction. Alternatively, the LA-RA hybrid PCR fragment was digested with *AscI-PacI*, PCR purified and ligated into *AscI-PacI* digested P[acman] in a two component ligation. Ligation products were electroporated into EPI300 cells. Correct ligation was verified by restriction digestion and DNA sequencing, following plasmid copy number induction and plasmid DNA preparation. A 1 μg sample of purified plasmid was linearized with *BamHI* and PCR purified. BAC-containing DY380 or mini-lambda/BAC-containing DH10B colonies were electroporated with 2 μl of linearized P[acman]. Potential recombinants were selected on LB plates (A100) at 30°C. Correct recombination at both homology arms was identified by colony PCR using MCS-F and 5'-Check-R (LA) and MCS-R and 3'-Check-F (RA) (For MCS-F and –R, see Table S3). After recombineering, gap-repaired P[acman] construct DNA was isolated and electroporated into EPI300 cells. Presence of the correct recombination event was verified by colony PCR using the conditions described above. Plasmid copy number was induced, plasmid DNA was isolated, and correct recombination was confirmed by restriction fingerprinting and DNA sequencing using primer set 4 (LA-Seq-F and RA-Seq-R).

**Serial gap-repair**

For serial gap repair, a similar strategy was followed, but seven primer sets were used (Fig. S4). Primer sets 1, 2 and 5 are used to PCR amplify three homology arms of about 500 bp each: left arm 1 (LA1), right arm 1 (RA1) and right arm 2 (RA2). LA1, RA1 and RA2 are located at the left end, the region of overlap of the two BACs, and right end of the desired DNA fragment, respectively. Primers for LA1 and RA1 were used for hybrid PCR amplification which seals the two initially obtained PCR fragments together during a secondary PCR. The hybrid LA1/RA1 and RA2 PCR products were cloned as *MluI-EcoRI* and *EcoRI-PacI* fragments,

respectively, into *AscI/PacI* digested and PCR purified P[acman]. During this procedure, LA1 and RA1, and RA1 and RA2 are separated by restriction site for efficient linearization with *BamHI* and *AscI* respectively. Two additional primer sets are used at each step to screen for correct recombination events at both junctions after gap repair and for DNA sequencing across the junctions to confirm correct retrieval of the desired fragment. Primer set 3 (LA2-5'-Check-R and LA2-3'-Check-F) is used to screen for correct recombinants after gap repair following step 1, in combination with the vector-specific primers MCS-F and MCS-R. Primer set 4 (LA1-Seq-F and RA1-Seq-R) is used to sequence across the junction to confirm correct retrieval of the desired fragment. Primer set 6 (5'-Check-R and 3'-Check-F) is used to screen for correct recombinants after gap repair following step 2, in combination with primers specific for the first gap-repaired fragment (LA2-3'-Check-F) and vector (MCS-R), respectively. Primer set 7 (RA1-Seq-F and RA2-Seq-R) is used to sequence across the junction to confirm correct retrieval in step 2. For the serial gap repair of *ten-m*, small modifications were made in step 2 (Fig.3). First, 1 μg of the purified plasmid obtained in step 1 was linearized with *AscI* and PCR purified. Second, after gap repair, colony PCR screening was performed only at the right end. After plasmid isolation from the recombineering background and electroporation into EPI300 cells, colony PCR screening was performed for all three junctions: the left end (MCS-F and LA2-5'-Check-R), the region in the BAC overlap (LA2-3'-Check-F and 5'-Check-R) and the right end (3'-Check-F and MCS-R).

**_P_ element transformation**

P[acman]-$Cm^R$ was tested for *Drosophila P* element transformation. Embryos obtained from a cross between *yw* virgins and *yw*; *Ki Δ2-3* males were injected with 300-500 ng/μl DNA

in 10 mM Tris-HCl (pH8.0). A similar procedure was performed for small gap-repaired fragments (<20 kb) in P[acman]. For fragments larger tan 20 kb, a DNA concentration of 500-1,000 ng/μl was used.

**ϕC31-mediated transformation**

Supercoiled DNA for injection was isolated using a CsCl banding method followed by dialysis in micro-injection buffer (10 mM Tris pH7.5, 0.1mM EDTA, 10 mM NaCl in DEPC treated water) or the Large Construct Kit (QIAGEN) followed by resuspension in micro-injection buffer. ϕC31 mRNA encoding integrase was prepared using the mMESSAGE mMACHINE T7 Transcription Kit (Ambion) as described (*14*) with the following modifications. Before and after linearization with *BamHI*, pET11phiC31-polyA plasmid (gift from Michelle Calos) was PCR purified and resuspended in nuclease-free water (Ambion). After LiCl precipitation, mRNA was resuspended in micro-injection buffer. mRNA was mixed with plasmid to obtain a final injection mixture of 6 μl. Injection mixture was tested for mRNA integrity and plasmid conformation using non-denaturing gel electrophoresis before and after injection and stored at -80°C. Typically, 100 embryos for injection were obtained from fly stocks containing docking sites, *yw; piggyBac-yellow*$^{+}$*-attP*. Embryos were collected, dechorionated in 50% bleach for 2 min. and desiccated for 4 min. We typically obtained 25-50 fertile adults which were crossed to *yw* flies. The transformation efficiency is defined as the percentage of crosses that produce transformed adults derived from single fertile injected animals.

In our experiments with the *attB*-P[acman]-Ap$^{R}$ vector alone, we typically injected 150 to 300 ng of DNA together with about 500-700 ng of ϕC31 mRNA, resulting in an integration efficiency of 20-30%. We varied the DNA concentration to explore the lower limit that permits

reasonable transformation efficiencies. With the ~12.9 kb *attB*-P[acman]-Ap$^R$ plasmid, we obtained an integration efficiency of 43% at a DNA concentration of 150 ng/µl, 21% at 75 ng/µl, but only ~4% (one transgenic animal) at 37.5 ng/µl. We reasoned that the molar concentration of DNA molecules was critical. Hence, to obtain transgenic animals for a 129 kb *attB*-P[acman] construct (ten-fold larger than the vector alone), we aimed at injecting 500-750 ng/µl of plasmid DNA (a ten-fold higher DNA concentration) together with 250-500 ng/µl of integrase mRNA. This was not always possible, since large DNA plasmids do not always tend to go in solution very well.

Correct integration events were identified by genomic PCR. Three PCR reactions were performed for the presence of the *attP* (primers attP-F and attP-R, 168 bp), *attL* (primers attB-F and attP-R, 163 bp) and *attR* (primers attP-F and attB-R, 289 bp) sites. Primers are listed in Table S3. PCR parameters were as follows: a denaturation cycle (94°C, 10'), 40 amplification cycles (94°C, 30"; 64°C, 30"; 72°C, 1') and a post-amplification cycle (72°, 10'). Correct integration events were identified by loss of the *attP* PCR product (specific for the original docking site) and the appearance of *attL* and *attR* PCR products (specific for the integration event).

### *Drosophila* mutations and rescue experiments

The following mutant chromosomes for rescue experiments were obtained: *FRT42D, sec6Δ20* and *FRT42D, sec6Δ175* (*15*) (gifts from Ulrich Tepass); *eps15 e*$^{75}$ and *eps15 Δ29* (Tong-Wey Koh and HJB, unpublished); *sec8 Δ1* (*16*) (gift from David Featherstone); *E(spl)*$^{b32.2}$, *e* (gift from Eric Lai), *FRT82B, E(Spl)*$^{b32.2}$, *P{ry*$^+$*,gro*$^+$*]* (gift from Kwang-Wook Choi), *Df(3R)Espl*$^{22}$, *tx*$^1$, *Df(3R)Espl*$^{22}$, *P{ry*$^+$*,gro*$^+$*]* and *Df(3R)P709, st* (gifts from Christos Delidakis)

(*17*); *FRT42D ,DScam*$^{20}$, *FRT42D, DScam*$^{21}$ and *FRT42D, DScam*$^{33}$ (*18*) (gifts from Dietmar Schmucker); *P{PZ}Ten-m*$^{05309}$ (*19*) and *P{SUPor-P}Ten-m*$^{KG00101}$, *ry*$^{506}$ (*9*); *FRT80B, sens*$^{E1}$ and *FRT80B, sens*$^{E2}$ (*20*); *Df(2L) tsh*$^{8}$ (gift from Richard Mann); *dap160 Δ1 (4-141)* and *Df(2L) burK1* (*21*); *P{lacW}l(2)k09022*$^{k09022}$ (*22*); *Dα7 PΔD5* (*23*). Mutant chromosomes for *drp1* and corresponding rescue experiment have been described (*24*). All mutant and rescue chromosomes were crossed in a double balanced background using *yw; L/CyO; D/TM6B, P{w*$^{+mC}$*=iab-2(1.7)lacZ}6B, Tb*$^{1}$ or *yw; L/CyO; D/TM3, Sb*. Double balanced mutant chromosomes were crossed to double balanced rescue chromosomes to generate balanced stocks containing both mutant and rescue chromosome. If homozygous rescue was not obtained, stocks were crossed to other mutations for rescue in a trans-heterozygous condition.

**SUPPLEMENTARY FIGURE LEGENDS**

**Figure S1. Multiple cloning site of P[acman] and primer design for gap-repair into P[acman]. a,** The multiple cloning site (MCS) of P[acman] contains 8-bp restriction sites for *AscI* and *PacI* for cloning of homology arms. *BamHI* is used to linearize constructs between the homology arms to create the gap to be repaired (Fig.1B). *AscI* and *EcoRI* are used to verify correct gap repair using restriction fingerprinting. Sites for sequencing primers (M13-F, T7, T3, M13-R) and primers for colony PCR (MCS-F and -R) are indicated. **b,** A desired genomic fragment (grey), consisting of exons (boxes) and introns (lines), is contained within a genomic clone, P1 bacteriophage or BAC. Four primer sets are designed for the DNA fragment to be retrieved. Primer sets 1 and 2, incorporating appropriate restriction sites for cloning, are used to PCR amplify 500-bp homology arms, a left arm (LA) and a right arm (RA). Primer set 3 (5'-Check-R and 3'-Check-F) is used with vector specific primers (MCS-F and MCS-R) to screen by PCR for correct recombination at the left end (MCS-F and 5'-check-R) and the right end (3'-Check-F and MCS-R). Primer set 4 (LA-Seq-F and RA-Seq-R) is used to sequence across the junctions to confirm correct retrieval of the desired fragment).

# SUPPLEMENTARY FIGURE 1

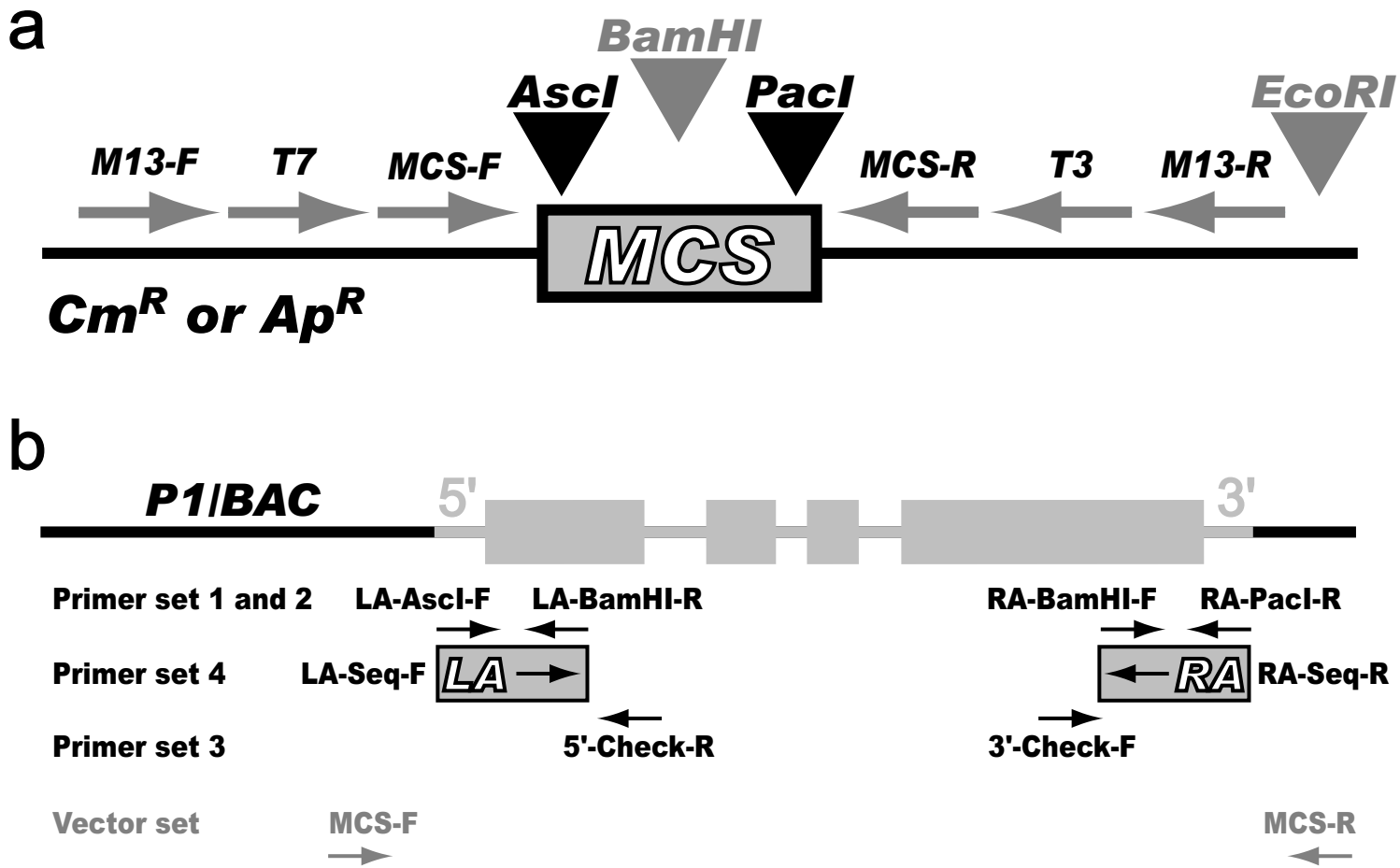


**Figure S2. Gap-repair into P[acman]: Two alternatives. a.** BAC DNA ($Cm^R$) is isolated from the recombination deficient ($rec^-$) DH10B strain and transformed into the tetracycline resistant ($Tc^R$) DY380 strain (*3*), which contains genome-borne recombineering functions ($rec^+$). Recombineering functions in the DY380 strain are repressed at 30°C and induced at 42°C. Induced and uninduced cells are transformed with linearized ampicillin resistant P[acman]-$Ap^R$ ($P^{GAP}$). Uninduced cells serve as a negative control. Potential recombinants containing gap-repaired P[acman] ($P^{REP}$) are selected ($Tc^R$, $Ap^R$). **b.** The recombination deficient ($rec^-$) DH10B strain containing BAC DNA ($Cm^R$) is transformed with tetracycline resistant ($Tc^R$) mini-lambda (*2*), a mobile element carrying all required recombineering functions ($rec^+$). Recombineering functions are repressed at 30°C and induced at 42°C. Induced and uninduced cells are transformed with linearized ampicillin resistant P[acman]-$Ap^R$ ($P^{GAP}$). Uninduced cells serve as a negative control. Potential recombinants containing gap-repaired P[acman] ($P^{REP}$) are selected ($Tc^R$, $Ap^R$). **c.** In either approach, potential recombinants are identified by colony PCR, the recombinant construct is isolated and transformed into the recombination deficient, copy inducible ($copy^+$) EPI300 host and selected ($Ap^R$). Transfer of a potential recombinant ($P^{REP}$) into copy inducible EPI300 cells is detected by colony PCR. The copy number of the plasmid is induced with arabinose, and plasmid DNA is isolated. Restriction fingerprinting and DNA sequencing are used to confirm correct recombination. The plasmid DNA is ready for transgenesis.

# SUPPLEMENTARY FIGURE 2

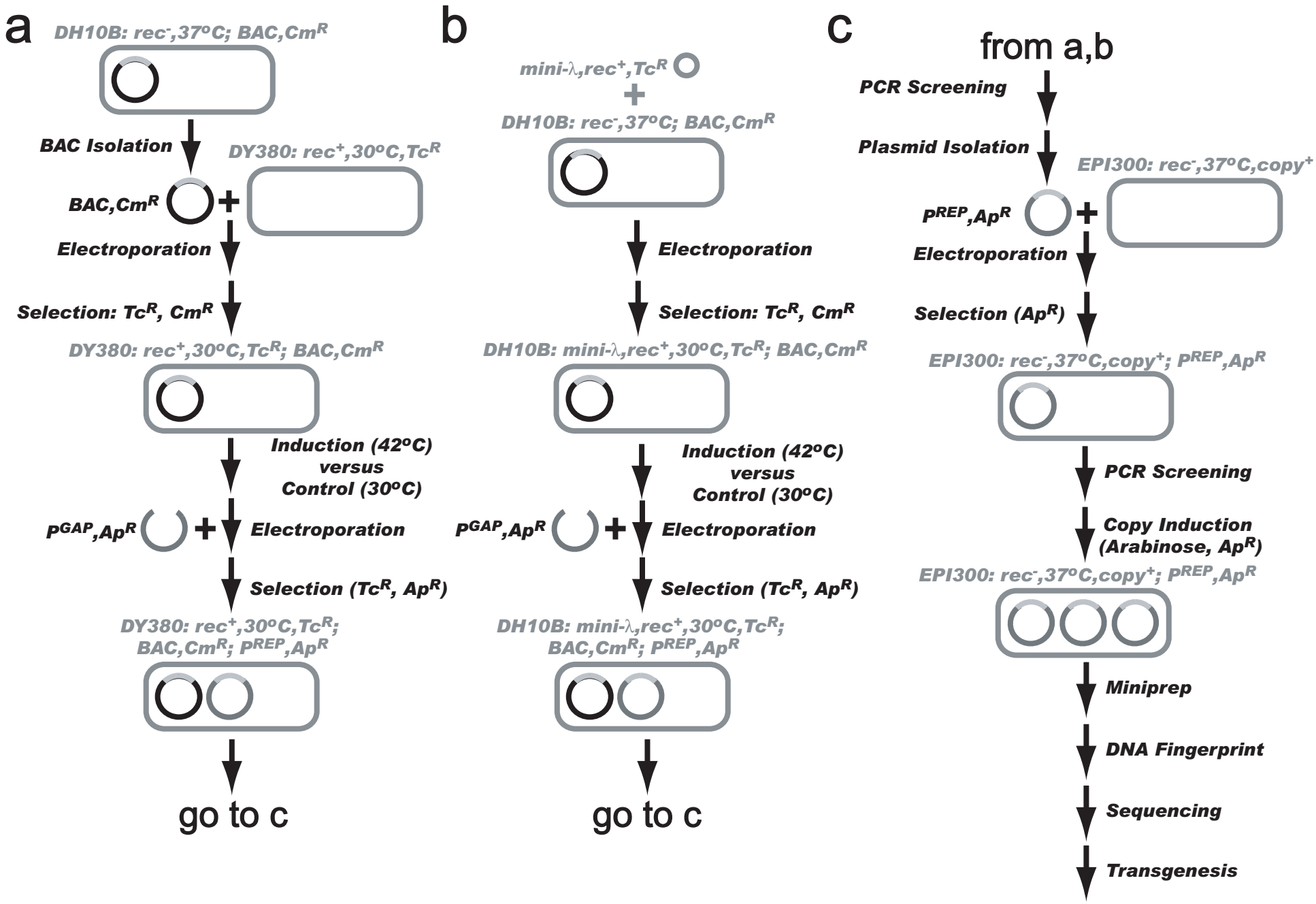


**Figure S3. Integration of *attB*-P[acman]-Ap$^R$ into different docking sites results in different *white*$^+$ phenotypes: Position effects. a,** y w control. **b-f,** *attB*-P[acman]-Ap$^R$ integrated in six different docking sites: **b,** VK6 on X; **c,** VK38 on X; **d,** VK1 on II; **e,** VK2 on II; **f,** VK5 on III; **g,** VK13 on III.

# SUPPLEMENTARY FIGURE 3

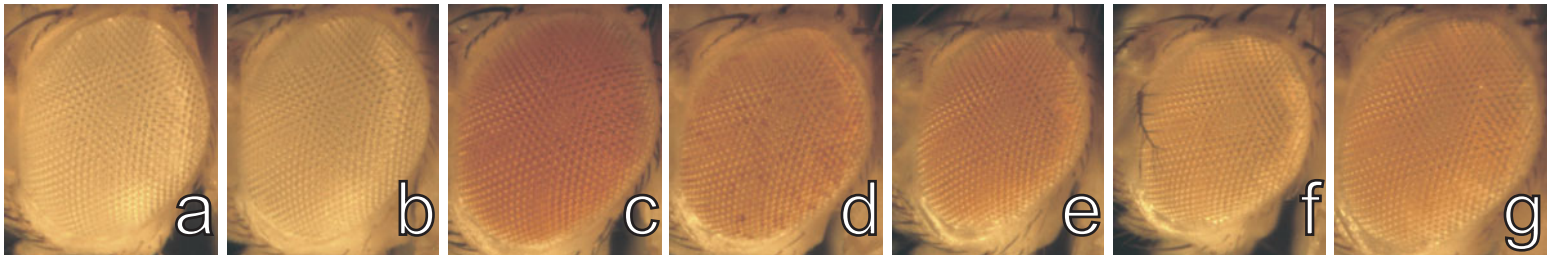
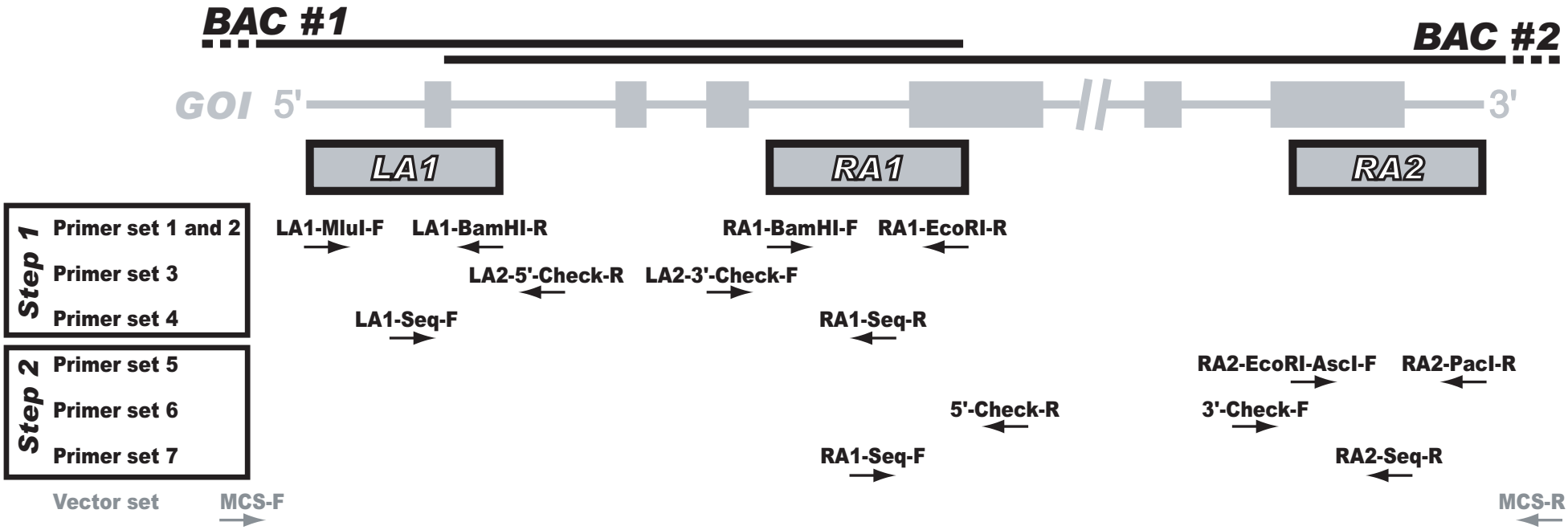


**Figure S4. Primer design for serial gap-repair into P[acman].** To clone a genomic fragment containing a large gene of interest (GOI) contained within two overlapping BACs by serial gap-repair, seven primer sets are designed: four for step 1 and three for step 2. Primer sets 1, 2 and 5 are used to PCR amplify LA1, RA1 and RA2, located at the left end, the region of overlap of the two BACs, and right end of the desired DNA fragment, respectively. The homology arms are cloned into the MCS of P[acman] such that they are separated by restriction sites for linearization of the construct: a *BamHI* site between LA1 and RA1 and an *AscI* site between RA1 and RA2. The *AscI* site used to clone LA1 is destroyed during the procedure. For each step, two additional primer sets are used to screen for correct recombination events at the junctions after gap-repair and for sequencing across the junctions to confirm retrieval of the desired fragment. Primer sets 3 and 4 are used in step 1, and primer sets 6 and 7 are used in step 2.

# SUPPLEMENTARY FIGURE 4



## SUPPLEMENTARY TABLE LEGENDS

**Table S1. Genomic locations of docking sites.** Insertion name, VK line number; Chr, chromosomal location assigned by genetic mapping; Location, genomic insertion site determined by inverse PCR; Polytene band, cytological location determined by BLAST (*13*); Gene, Intra- (YES) or intergenic (NOT) insertion; Info, additional information.

## Supplementary Table 1: Genomic locations of *attP* docking sites

| TRANSFORMATION | | | | | |
|---|---|---|---|---|---|
| **Insertion Name** | **Chr** | **Location** | **Polytene** | **Gene** | **Info** |
| **VK00001** | **2** | 2R: 18279750 (+) | **59D3** | **NO** | |
| **VK00002** | **2** | 2L: 8133094 (-) | **28E7** | **YES** | |
| **VK00003a** | **2** | 2L: 10220182 (-) | **31B1** | **YES** | Double Insert |
| **VK00003b** | | 2L: 21639162 (+) | **40A4** | **NO** | |
| **VK00005** | **3** | 3L: 17900987 (-) | **75B1** | **YES** | |
| **VK00006** | **X** | X: 20689962 (-) | **19E7** | **NO** | |
| **VK00007** | **3** | 3R: 81372 (+) | **82A1** | **NO** | |
| **VK00008** | **2** | 2R: 8443624 (+) | **50A3** | **NO** | |
| **VK00010** | **2** | none | **NA** | **NO** | Quasimodo |
| **VK00011** | **2** | 2L: 21951742 (-) | **40E4** | **NO** | Both flanks hit lvk |
| **VK00012** | **2** | 2L: 4683203 (+) | **25A3** | **YES** | |
| **VK00013** | **3** | 3L: 19153237 (-) | **76A2** | **NO** | |
| **VK00014** | **2** | 2R: 2285107 (-) | **43A2** | **NO** | |
| **VK00015** | **2** | none | **NA** | **NO** | Diver |
| **VK00016** | **2** | 2R: 5964343 (-) | **47C4** | **NO** | |
| **VK00018** | **2** | 2R: 11361979 (+) | **53B1** | **NO** | |
| **VK00019** | **3** | 3L: 11654656 (-) | **68D2** | **YES** | |
| **REMOBILIZATION** | | | | | |
| **Insertion Name** | **Chr** | **Location** | **Polytene** | **Gene** | **Info** |
| **VK00020** | **3** | 3R: 26368401 (+) | **99F8** | **YES** | |
| **VK00021** | **3** | 3L: 19713047 (+) | **76C5** | **YES** | |
| **VK00022** | **2** | 2R: 16689522 (+) | **57F5** | **NO** | |
| **VK00023** | **3** | 3L: 13170566 (+) | **70A2** | **NO** | |
| **VK00024** | **4** | 4: 79131 (-) | **102A3** | **NO** | |
| **VK00026** | **3** | 3R: 21662155 (+) | **96F3** | **NO** | |
| **VK00027** | **3** | 3R: 12878600 (-) | **89E11** | **NO** | |
| **VK00028** | **3** | 3R: 16375389 (-) | **92F1** | **YES** | |
| **VK00029** | **3** | AABU01002142: 94181 (+) | **80 (3L het)** | **NO** | |
| **VK00030** | **2** | 2R: 9203330 (-) | **50E1** | **YES** | |
| **VK00031** | **3** | 3L: 2376116 (+) | **62E1** | **NO** | |
| **VK00032** | **3** | 3R: 4301078 (-) | **85A2** | **NO** | |
| **VK00033** | **3** | 3L; 6402205 (+) | **65B2** | **NO** | |
| **VK00036** | **3** | 3L: 13802829 (+) | **70C4** | **YES** | |
| **VK00037** | **2** | 2L: 1584486 (+) | **22A3** | **YES** | |
| **VK00038** | **X** | X: 5497935 (-) | **5B8** | **YES** | |
| **VK00039** | **3** | 3L: 10463229 (-) | **67E4** | **NO** | |
| **VK00040** | **3** | 3R: 8265934 (-) | **87B10** | **YES** | |

**Table S2. Efficiency of ϕC31 mediated transgenesis with *attB*-P[acman].** Integration efficiencies using ϕC31 were compared between published materials, including the *attB* containing plasmid pUAST-B and the *attP* containing docking element 2R (*14*); and our materials, *attB*-P-Ap$^{R}$ and different *piggyBac-y*$^{+}$*-attP* docking elements. A total of 7 different *piggyBac-y*$^{+}$*-attP* docking elements were tested: 2 on the X (VK6 and VK38), 2 on the 2$^{nd}$ (VK1 and VK2), 2 on the 3$^{rd}$ (VK5 and VK13) and one on the 4$^{th}$ (VK24). Similar efficiencies were obtained for both plasmids and for different 'docking' elements, except for VK24, which seems to be not receptive.

| Supplementary Table 2: phiC31 integration effciencies | | | |
|---|---|---|---|
| **Plasmid** | **Docker** | **Chr** | **Efficiency** |
| **pUAST-B** | 2R | 2 | 27% |
| **attB-P[acman]-Ap$^{R}$** | 2R | 2 | 34% |
| **attB-P[acman]-Ap$^{R}$** | VK6 | X | 33% |
| **attB-P[acman]-Ap$^{R}$** | VK38 | X | 15% |
| **attB-P[acman]-Ap$^{R}$** | VK1 | 2 | 24% |
| **attB-P[acman]-Ap$^{R}$** | VK2 | 2 | 63% |
| **attB-P[acman]-Ap$^{R}$** | VK5 | 3 | 32% |
| **attB-P[acman]-Ap$^{R}$** | VK13 | 3 | 33% |
| **attB-P[acman]-Ap$^{R}$** | VK24 | 4 | 0% |

**Table S3. Primers used for plasmid construction, colony PCR, docking site mapping, and integration verification.** A: Primers used for P[acman] plasmid construction. B: Primers used for colony PCR, genomic integration PCR and genomic inverse PCR.

## Supplementary Table 3: Primers used for plasmid construction and PCR

### A: Primers used for P[acman] plasmid construction

| Primer name | Primer Sequence (5' to 3') |
|---|---|
| MCS-Cloning-F | CGGGATCCGCTAGCAAAAAAACCGGTGTAAAACGACGGCCAGTGAG |
| MCS-Cloning-R | CGGGATCCACTAGTAAAAAAATGCATAAAAAAGAATTCGGAAACAGCTATGACCATG |
| P5-F | CCAATGCATGCTGCTGCTCTAAACGACGCATTTC |
| P5-R | GGACTAGTCTGGACTACGTGGGTCTG |
| P3-F | CCTAGCTAGCTCACGATGAGAATGGCCAGAC |
| P3-R | AGGTCACCGGTTGCGTACTCGCAAATTATTAAAAATAAAAC |
| Amp-Chl-swap-F-pBS | ATCACTACCGGGCGTATTTTTTGAGTTATCGAGATTTTCAGGAGCTAAGGAAGCTAAACATTCAAATATGTATCCGCTC |
| Amp-Chl-swap-R-pBS | CTTATTCAGGCGTAGCAACCAGGCGTTTAAGGGCACCAATAACTGCCTTAAAAAAAAGAGTTGGTAGCTCTTGATC |
| P[acman]-Neo-SalI-R | AGCGGATGAATGGCAGAAATTCGATGATAAGCTGTCAAACATGAGAATTGGGCCTGGTGATGATGGCGGGATC |
| P3-Neo-R | ACGGGACCACCTTATGTTATTTCATCATGGTCTGGCCATTCTCATCGTGATCAGAAGAACTCGTCAAGAAGG |
| pRpsl-Neo-Screen-F | AGTATGCCGTGTTCGTCTGAC |
| P[acman]-attB-SalI-R | AGCGGATGAATGGCAGAAATTCGATGATAAGCTGTCAAACATGAGAATTGGTCGACGATGTAGGTCACGGTC |
| P3-attB-R | ACGGGACCACCTTATGTTATTTCATCATGGTCTGGCCATTCTCATCGTGATCGACATGCCCGCCGTGACCGTC |
| attB-Screen-F | GCGTACTCCACCTCACCCATC |
| attP-shorter-NotI-F | AAGGAAAAAAGCGGCCGCCTTCACGTTTTCCCAGGTC |
| attP-XmaI-R | TCCCCCCGGGTCGCGCTCGCGCGACTGACG |

### B: Primers used for colony PCR, genomic integration PCR and genomic inverse PCR

| Primer name | Primer Sequence (5' to 3') |
|---|---|
| MCS-F | TTTAAACCTCGAGCGGTCCGTTATC |
| MCS-R | CTAAAGGGAACAAAAGCTGGGTAC |
| attP-F | CTTCACGTTTTCCCAGGTCAGAAG |
| attP-R | GTCGCGCTCGCGCGACTGACGGTC |
| attB-F | GTCGACGATGTAGGTCACGGTC |
| attB-R | TCGACATGCCCGCCGTGACCGTC |
| HY.5.F | AAGTAACAAAACTTTTATCGAAT |
| HY.5.R | TAAACCTCGATATACAGACC |
| HY.3.F | CCTAAATGCACAGCGACGGAT |
| HY.3.R | GTGAGGCGTGCTTGTCAATG |

**Table S4. Primers used for single gap-repair or PCR, and double gap-repair.** A: Primers used for single gap-repair. B: Primers used for PCR. C: Primers used for double gap-repair.

# Supplementary Table 4: Primers used for gap-repair or PCR of the different genes

## A: Primers used for single gap-repair

| Gap-repair | Gene | Fragment | Length | Primer Set | Primer name | Homology | Primer Sequence (5' to 3') |
|---|---|---|---|---|---|---|---|
| 1 | ***sens2*** | sens2-22 | **22,334** | 1 | sens2-22-LA-AscI-F | **500 bp** | AGGCGCGCCATTTATAGCTTCGGTGGTG |
| | | | | | sens2-22-LA-BamHI-R | | CGCGGATCCACATGGCTATCAGTTATCAG |
| | | | | 2 | sens2-22-RA-BamHI-F | **500 bp** | CGCGGATCCCTGCAAATGTTGTTACAAC |
| | | | | | sens2-22-RA-AsiSI-R | | AAGGAAAAAAGCGATCGCTGGCAACAAGCGGCGCATGCAG |
| | | | | 3 | sens2-22-5'-Check-R | | TGCCCGGTTGTAATCAAGAAC |
| | | | | | sens2-22-3'-Check-F | | ATTGATTCCGCCTCCGATGTC |
| | | | | 4 | sens2-22-LA-Seq-F | | CTGCGCTATAATTCCCTACTG |
| | | | | | sens2-22-RA-Seq-R | | TTCGTGTGATGCAGACTGCAG |
| 2 | ***sens2*** | sens2-29 | **28,937** | 1 | sens2-29-LA-AscI-F | **476 bp** | AGGCGCGCCATCAGCTATCATTTCCATGAAC |
| | | | | | sens2-29-LA-BamHI-R | | CGCGGATCCTCGCATGACTGAAACGATTTG |
| | | | | 2 | sens2-29-RA-BamHI-F | **470 bp** | CGCGGATCCGCTTTAAATTAGTTGTTAGCTC |
| | | | | | sens2-29-RA-AsiSI-R | | AAGGAAAAAAGCGATCGCGACTTAGCAGGAGGTTTTAAATG |
| | | | | 3 | sens2-29-5'-Check-R | | GGACTGGAGATGGCAATGGAG |
| | | | | | sens2-29-3'-Check-F | | TGAATGCGCCTGCCGAAACTC |
| | | | | 4 | sens2-29-LA-Seq-F | | GATAATTCGGCAGTTAGCTAG |
| | | | | | sens2-29-RA-Seq-R | | TGGTCGGCCCATGATGAGTTG |
| 3 | ***sens2*** | sens2-39 | **38,937** | 1 | sens2-39-LA-AscI-F | **500** | AGGCGCGCCCCCGGGGGAGCAACTGTGCGGCTTC |
| | | | | | sens2-39-LA-BamHI-R | | CGCGGATCCACCAAGTGGTTTGAAGTGCCAAC |
| | | | | 2 | sens2-39-RA-BamHI-F | **490** | CGCGGATCCGTTTTCCTTTCGCCACTCCCAG |
| | | | | | sens2-39-RA-AsiSI-R | | AAGGAAAAAAGCGATCGCCCGCGGTGGGAGGCCATCCTG |
| | | | | 3 | sens2-39-5'-Check-R | | ATCTGACGTGTCACGCGAGTG |
| | | | | | sens2-39-3'-Check-F | | AATAATGCTCCTCTGGCTCTC |
| | | | | 4 | sens2-39-LA-Seq-F | | ACTTGCCTTGCCTTATCTATG |
| | | | | | sens2-39-RA-Seq-R | | TGGCCCACTGCGTGGTAATTC |
| 4 | ***CG10805*** | CG10805-S | **9,685** | 1 | CG10805-S-LA-AscI-F | **500** | AGGCGCGCCGCCCTATCCCAGAACACAG |
| | | | | | CG10805-S-LA-BamHI-R | | CGCGGATCCGCGAGGCGGCGCTGTTGGCCAG |
| | | | | 2 | CG10805-S-RA-BamHI-F | **500** | CGCGGATCCAACCGGTGCCTGGGCGATCTG |
| | | | | | CG10805-S-RA-AsiSI-R | | AAGGAAAAAAGCGATCGCTTGATTTCCTCCCTTAAC |
| | | | | 3 | CG10805-S-5'-Check-R | | CACCCTTCCGGCATCCACGTC |
| | | | | | CG10805-S-3'-Check-F | | AGAGCAGCAGCGTCCGCAATG |
| | | | | 4 | CG10805-S-LA-Seq-F | | ACGCTGGACGTGATGACCAAG |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | CG10805-S-RA-Seq-R | | TTGCTGCTGTGTCTCCAGCTG |
| **5** | ***CG10805*** | CG10805-L | **14,695** | **1** | CG10805-L-LA-AscI-F | **500** | AGGCGCGCCAAAGTGCAAACGTTGAGATAC |
| | | | | | CG10805-L-LA-BamHI-R | | CGCGGATCCGGACACACTGCAACTGGAAC |
| | | | | **2** | CG10805-L-RA-BamHI-F | **500** | CGCGGATCCAGTCGAAAATACTTTGTTG |
| | | | | | CG10805-L-RA-AsiSI-R | | AAGGAAAAAAGCGATCGCCGTATTGTTATTTAACA |
| | | | | **3** | CG10805-L-5'-Check-R | | CGACTCACGTAAGTCCTTCGA |
| | | | | | CG10805-L-3'-Check-F | | AATGAGGCGCTGAAGTGATAG |
| | | | | **4** | CG10805-L-LA-Seq-F | | GCACAGCCATTTGAGTTTGTC |
| | | | | | CG10805-L-RA-Seq-R | | TGATGTTGTCCAGGATCGCAG |
| **6** | ***dap160*** | dap160 | **10,902** | **1** | Dap160-LA-AscI-F | **500** | AGGCGCGCCTTGAATTCAATTACTGACTAAATC |
| | | | | | Dap160-LA-BamHI-R | | CGCGGATCCTTCAGGCCGTGCACATAACCAAG |
| | | | | **2** | Dap160-RA-BamHI-F | **500** | CGCGGATCCATCGAACATCTACTGGGGAAG |
| | | | | | Dap160-RA-PacI-R | | ACCTTAATTAAAAATGAAAATGAATGAAAAGTATC |
| | | | | **3** | Dap160-5'-Check-R | | CCAGCGGTTATGGCAACAGTG |
| | | | | | Dap160-3'-Check-F | | GAAGAACAGCAAGAGCAGGTC |
| | | | | **4** | Dap160-LA-Seq-F | | GGGTGTTCTTTCACCAGCAAC |
| | | | | | Dap160-RA-Seq-R | | GCCCGTGAAGTTCTCCGAATC |
| **7** | ***sens*** | sens-S | **12,108** | **1** | sens-S-LA-AscI-F | **500** | AGGCGCGCCAATCTTAAGGCCATTAGCAAT |
| | | | | | sens-S-LA-BamHI-R | | CGCGGATCCCACAAATGATTGATGATG |
| | | | | **2** | sens-S-RA-BamHI-F | **500** | CGCGGATCCAGGAGCAGGGGCAGCAGCATC |
| | | | | | sens-S-RA-PacI-R | | ACCTTAATTAAACCAAAACATATCTGATAGTAAG |
| | | | | **3** | sens-S-5'-Check-R | | CTAAAGACACACAATTTATAG |
| | | | | | sens-S-3'-Check-F | | ACAGATAGAGGGAGACGGAAG |
| | | | | **4** | sens-S-LA-Seq-F | | TATAATTGTGCGCAGACGCAG |
| | | | | | sens-S-RA-Seq-R | | ATGCGACAGGTTCCAGGATTC |
| **8** | ***sens*** | sens-L | **18,108** | **1** | sens-L-LA-AscI-F | **500** | AGGCGCGCCAGATTCTTTAAAGAAACCTC |
| | | | | | sens-L-LA-BamHI-R | | CGCGGATCCATGCACAAAATCGAGTATCCCTG |
| | | | | **2** | sens-L-RA-BamHI-F | **500** | CGCGGATCCTTCGATTGGAGTGGATTGGAG |
| | | | | | sens-L-RA-PacI-R | | ACCTTAATTAATGCCGTTCAGTAAACTTCTC |
| | | | | **3** | sens-L-5'-Check-R | | AGTTTAAGAGTTCAAGCCGAG |
| | | | | | sens-L-3'-Check-F | | TGGTCTTGAGGTTCTGGCGAC |
| | | | | **4** | sens-L-LA-Seq-F | | CGAGCACATGACGCTGCTCAC |
| | | | | | sens-L-RA-Seq-R | | TCCACCTGTAGCATGCTACTC |
| **9** | ***Dalpha7*** | Dalpha7 | **29,380** | **1** | Dalpha7-LA-AscI-F | **517** | AGGCGCGCCCGCCACTCAACTCAAGTGAA |
| | | | | | Dalpha7-LA-BamHI-R | | CGCGGATCCCGACTTTCGTACCATGGTTTT |
| | | | | **2** | Dalpha7-RA-BamHI-F | **518** | CGCGGATCCAAACGGCGGAGATACAGACTT |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | Dalpha7-RA-PacI-R | | ACCTTAATTAAGTGGCCCCCTGTTTTCTC |
| | | | | 3 | Dalpha7-5'-Check-R | | GGCAAAATGGGTCAACAAAC |
| | | | | | Dalpha7-3'-Check-F | | TCAAATGTTGGGAAAAATGC |
| | | | | 4 | Dalpha7-LA-Seq-F | | GACGCACATCTGTTCGAGAG |
| | | | | | Dalpha7-RA-Seq-R | | GCGATGAGACATCACACCAG |
| **10** | ***Drp*** | Drp | **9,350** | **1** | Drp-LA-AscI-F | **597** | AGGCGCGCCTTTCCAAATTAGATTTTAACTGGTATC |
| | | | | | Drp-LA-BamHI-R | | CGCGGATCCTTGTTAAATGTTCCCCCTAC |
| | | | | **2** | Drp-RA-BamHI-F | **596** | CGCGGATCCCGCCAAGTGTTGCTCATCAAATGG |
| | | | | | Drp-RA-PacI-R | | ACCTTAATTAAAGCGGAGAACACCAATAAATC |
| | | | | **3** | Drp-5'-Check-R | | TCTGTATAAGTCGGATAAGG |
| | | | | | Drp-3'-Check-F | | TTCTAAATACATGGCCAACGG |
| | | | | **4** | Drp-LA-Seq-F | | ATGACGTTGCCTTCGTAAAC |
| | | | | | Drp-RA-Seq-R | | TAAATTTAGCGCACAGGATGC |
| **11** | ***Sec8*** | sec8-S | **4,852** | **1** | Sec8-S-LA-AscI-F | **531** | TGGCGCGCCAATCCCAAATGTTACTCCTC |
| | | | | | Sec8-S-LA-BamHI-R | | CGCGGATCCGTGGGCTTAATGCACATGAT |
| | | | | **2** | Sec8-S-RA-BamHI-F | **518** | GCGGGATCCTACTCAATGCGTTGCTGGAG |
| | | | | | Sec8-S-RA-PacI-R | | ACCTTAATTAAGCGAGCACTCAAACCCTGTT |
| | | | | **3** | Sec8-S-5'-Check-R | | CTATCGTTTGGATTAGTGATG |
| | | | | | Sec8-S-3'-Check-F | | GCATGGAGCCAGACGTAAGAG |
| | | | | **4** | Sec8-S-LA-Seq-F | | TAGATGCAACTACACCGTTAC |
| | | | | | Sec8-S-RA-Seq-R | | GAGGGCTTGGCGCCTAGTATG |
| **12** | ***Sec8*** | sec8-L | **12,050** | **1** | Sec8-L-LA-AscI-F | **451** | AGGCGCGCCATATGTATATGTATAAATGTAAG |
| | | | | | Sec8-L-LA-BamHI-R | | CGCGGATCCATCGAGATCTCTATGTCCATAC |
| | | | | **2** | Sec8-L-RA-BamHI-F | **500** | CGCGGATCCTCCTCATACGTCTGCTTGTG |
| | | | | | Sec8-L-RA-PacI-R | | ACCTTAATTAAGCTCCAGTCGTAGCGAAGCAATTG |
| | | | | **3** | Sec8-L-5'-Check-R | | TCTTGTTTAAGCCCAATGCTC |
| | | | | | Sec8-L-3'-Check-F | | TAAGAAATGTGCTGGAGTTAG |
| | | | | **4** | Sec8-L-LA-Seq-F | | GTACAAGCTGAATAATAGCTC |
| | | | | | Sec8-L-RA-Seq-R | | ACGAATCGACATGTCAACATG |
| **13** | ***Eps15*** | Eps15-S | **10,755** | **1** | Eps15-S-LA-AscI-F | **500** | TTGGCGCGCCGGGGAATATTCCAATATTTATACCTAG |
| | | | | | Eps15-S-LA-BamHI-R | | CGCTACCTAATATTCTGGATCCGCGCTTGTGTGTGTGC |
| | | | | **2** | Eps15-S-RA-BamHI-F | **500** | ACACAAGCGCGGATCCAGAATATTAGGTAGCGGTAGCG |
| | | | | | Eps15-S-RA-PacI-R | | GCCTTAATTAAAAGTGATGCCGTTTGGATCTG |
| | | | | **3** | Eps15-S-5'-Check-R | | CTCCTCGTGCTCAAGG |
| | | | | | Eps15-S-3'-Check-F | | CCTCCCACACTTTCTCG |
| | | | | **4** | Eps15-S-LA-Seq-F | | GGTTTGAGTGTGTGCG |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | Eps15-S-RA-Seq-R | | CTGAGCCGTTGAAGC |
| **14** | ***Eps15*** | Eps15-L | **11,905** | **1** | Eps15-L-LA-AscI-F | **500** | TTGGCGCGCCCCGGACCCTAACAGATC |
| | | | | | Eps15-L-LA-BamHI-R | | CGCTACCTAATATTCTGGATCCTACCCTACTACCAGGAGG |
| | | | | **2** | Eps15-L-RA-BamHI-F | **500** | GGTAGTAGGGTAGGATCCAGAATATTAGGTAGCGGTAGCG |
| | | | | | Eps15-L-RA-PacI-R | | GCCTTAATTAAAAGTGATGCCGTTTGGATCTG |
| | | | | **3** | Eps15-L-5'-Check-R | | CTCCTCGTGCTCAAGG |
| | | | | | Eps15-L-3'-Check-F | | CCTCCCACACTTTCTCG |
| | | | | **4** | Eps15-L-LA-Seq-F | | ACGGTTTTCTGCCAGC |
| | | | | | Eps15-L-RA-Seq-R | | GCGCCTTTTATCCAGG |
| **15** | ***Sec6*** | Sec6-L | **11,478** | **1** | Sec6-L-LA-AscI-F | **481** | AGGCGCGCCGAAAACAATATTTTGGTATTTGTC |
| | | | | | Sec6-L-LA-BamHI-R | | CGCGGATCCGACATGTCCGATAGCGACGAG |
| | | | | **2** | Sec6-L-RA-BamHI-F | **500** | CGCGGATCCATGGGCAACACCATATCTAATC |
| | | | | | Sec6-L-RA-PacI-R | | ACCTTAATTAATTGAGATTAGGTGTGCCGTTCTG |
| | | | | **3** | Sec6-L-5'-Check-R | | GGAGGATTTGACGGGAATCTC |
| | | | | | Sec6-L-3'-Check-F | | GTACCAGGTACTCGGATTCAC |
| | | | | **4** | Sec6-L-LA-Seq-F | | ATTGAACTATAAATGCTGCAG |
| | | | | | Sec6-L-RA-Seq-R | | TGGTCACCGTGAACTTCACTG |
| **16** | ***grp*** | grp | **29,839** | **1** | grp-LA-AscI-F | **478** | AGGCGCGCCACCTGGACAAAGGGTCAAAGGCCAC |
| | | | | | grp-LA-BamHI-R | | CGCGGATCCCGGATGATTTCAGCTATTGGCTTG |
| | | | | **2** | grp-RA-BamHI-F | **500** | CGCGGATCCTTACCGTTAGTACCGTTATGTTC |
| | | | | | grp-RA-PacI-R | | ACCTTAATTAACAAAATGGAAATACACACGAGTAG |
| | | | | **3** | grp-5'-Check-R | | TACACCAGATATTGCTAGAG |
| | | | | | grp-3'-Check-F | | TGCTGCAGTGATATTGCTATG |
| | | | | **4** | grp-LA-Seq-F | | GTGTACGGGGCACACCTTCAC |
| | | | | | grp-RA-Seq-R | | CATCAAACTCGAATCGGAGTC |
| **17** | ***miR-9a*** | miR-9a | **20,078** | **1** | miR-9a-LA-AscI-F | **500** | AGGCGCGCCGAAGAATGTAGATTTTAGTACTG |
| | | | | | miR-9a-LA-BamHI-R | | CAGCCGGGACACTGAAACATGGATCCACTGAAAATTTTCATCCTGTAC |
| | | | | **2** | miR-9a-RA-BamHI-F | **500** | GTACAGGATGAAAATTTTCAGTGGATCCATGTTTCAGTGTCCCGGCTG |
| | | | | | miR-9a-RA-PacI-R | | ACCTTAATTAACAGCAAAATATTCGCTTCTC |
| | | | | **3** | miR-9a-5'-Check-R | | GAAGACATTTCATCTAGTCAC |
| | | | | | miR-9a-3'-Check-F | | CCCGGGCGAAGGGCAGCCAAG |
| | | | | **4** | miR-9a-LA-Seq-F | | ATGTATACCTTAGGCCATATG |
| | | | | | miR-9a-RA-Seq-R | | TTAATCACCGGACTCTCCATC |
| **18** | ***miR-4*** | miR4 | **4,459** | **1** | miR-4-LA-AscI-F | **500** | AGGCGCGCCTACACCTTCTACAAACATGTG |
| | | | | | miR-4-LA-BamHI-R | | CACCCAAAAATGATTGAACAAGGATCCTCTTCATTTCATTTATATG |
| | | | | **2** | miR-4-RA-BamHI-F | **500** | CATATAAATGAAATGAAGAGGATCCTTGTTCAATCATTTTTGGGTG |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | miR-4-RA-PacI-R | | ACCTTAATTAAGGCGGCCATGTCAATGCAGAATG |
| | | | | 3 | miR-4-5'-Check-R | | ATATTATCCATCACACATCTC |
| | | | | | miR-4-3'-Check-F | | TAGCTTGTGCATCAATTCGTC |
| | | | | 4 | miR-4-LA-Seq-F | | TGACAGCAGCGGACTGACAAG |
| | | | | | miR-4-RA-Seq-R | | TGTAGAAATCAACTCCACTAG |
| **19** | ***bancal*** | bancal | **39,523** | **1** | bancal-LA-AscI-F | **500** | AGGCGCGCCTAGAAACGGAAAGACAAAC |
| | | | | | bancal-LA-BamHI-R | | CTTCTCAGGGGCTTTTTGTCGGATCCTTCCCGAGGTAATGAAGTTG |
| | | | | **2** | bancal-RA-BamHI-F | **500** | CAACTTCATTACCTCGGGAAGGATCCGACAAAAAGCCCCTGAGAAG |
| | | | | | bancal-RA-PacI-R | | ACCTTAATTAAACCTTTGGAAAAACAAAGCCGTTG |
| | | | | **3** | bancal-5'-Check-R | | ACTATAGTTTTGCACTCTCTG |
| | | | | | bancal-3'-Check-F | | TGTTCACCTGAACAACTTCAG |
| | | | | **4** | bancal-LA-Seq-F | | CGTCCAAGCTGTTCGCTAGTC |
| | | | | | bancal-RA-Seq-R | | GGACCTTACTAGCTTATACTC |
| **20** | ***E(Spl)-C*** | E(Spl)-C | **77,692** | **1** | E(Spl)-C-LA-AscI-F | **500** | AGGCGCGCCCATAAACTGCACATGGTTGGCAG |
| | | | | | E(Spl)-C-LA-BamHI-R | | CAGAGTCCCGGACAGCCCAAATGGATCCGCCCAGAAGTTCGAGGAGTG |
| | | | | **2** | E(Spl)-C-RA-BamHI-F | **500** | CACTCCTCGAACTTCTGGGCGGATCCATTTGGGCTGTCCGGGACTCTG |
| | | | | | E(Spl)-C-RA-PacI-R | | ACCTTAATTAACTTCAATGGCAATCTGGAAG |
| | | | | **3** | E(Spl)-C-5'-Check-R | | AGTGCGCTGTACTTCATGTTC |
| | | | | | E(Spl)-C-3'-Check-F | | CTCTTCACTGGCCGTGCTTAG |
| | | | | **4** | E(Spl)-C-LA-Seq-F | | TCTGTTAGCAATGCCTTATTC |
| | | | | | E(Spl)-C-RA-Seq-R | | ATCAAAGCATCACGGCTGCTG |
| **21** | ***Brd-C*** | Brd-C | **37,205** | **1** | Brd-C-LA-AscI-F | **500** | AGGCGCGCCCGGAGCCAAAAGCCCGGGCAAC |
| | | | | | Brd-C-LA-BamHI-R | | GACCCATTAAAATCGACCCCCGGATCCAAGTGAACGAGGTTAC |
| | | | | **2** | Brd-C-RA-BamHI-F | **500** | GTAACCTCGTTCACTTGGATCCGGGGGTCGATTTTAATGGGTC |
| | | | | | Brd-C-RA-PacI-R | | ACCTTAATTAATTAGGACTACGCGAATCTGAAAG |
| | | | | **3** | Brd-C-5'-Check-R | | GTAATAATATAGTAGCTCCAC |
| | | | | | Brd-C-3'-Check-F | | ATGCACCAGACCCTAGAAGAG |
| | | | | **4** | Brd-C-LA-Seq-F | | GGTAAACACAGTAATGAGCAC |
| | | | | | Brd-C-RA-Seq-R | | CCATTTACCATGACTTGACTG |
| **22** | ***Dscam*** | Dscam-1 | **73,316** | **1** | Dscam-1-LA-AscI-F | **500** | AGGCGCGCCTGAGGGGGCGGACTCCAAGGAC |
| | | | | | Dscam-1-LA-BamHI-R | | GTTTCAGTTTCAGTTTCAATTTGGATCCGTTGGACGACTCCTTCGGTG |
| | | | | **2** | Dscam-1-RA-BamHI-F | **500** | CACCGAAGGAGTCGTCCAACGGATCCAAATTGAAACTGAAACTGAAAC |
| | | | | | Dscam-1-RA-PacI-R | | ACCTTAATTAACGGCTTGAATGGTCATACCAAAC |
| | | | | **3** | Dscam-1-5'-Check-R | | TTAGTAAACTCACCTGGGTGC |
| | | | | | Dscam-1-3'-Check-F | | AAGCCGTGTTAGGTGGGATAC |
| | | | | **4** | Dscam-1-LA-Seq-F | | GGAGCAAGTGTACCACCAGAC |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | Dscam-1-RA-Seq-R | | AAGCGGCAACAATACCGCTGC |
| 23 | ***Dscam*** | Dscam-2 | **102,316** | **1** | Dscam-2-LA-AscI-F | **500** | AGGCGCGCCATTTTCGATTCTAGCGATAAC |
| | | | | | Dscam-2-LA-BamHI-R | | CAATGGCAGCATAGCGGTAGGGGATCCAAGAAAGGATTGAGTGCGGGCAT |
| | | | | **2** | Dscam-2-RA-BamHI-F | **500** | ATGCCCGCACTCAATCCTTTCTTGGATCCCCTACCGCTATGCTGCCATTG |
| | | | | | Dscam-2-RA-PacI-R | | ACCTTAATTAATTGGACTAAAACTCTTATTTC |
| | | | | **3** | Dscam-2-5'-Check-R | | AAACGTCGTAGCAACGCCAAG |
| | | | | | Dscam-2-3'-Check-F | | TATACTCGAATTGGATAACTG |
| | | | | **4** | Dscam-2-LA-Seq-F | | AACTTGTGATTAGGCATGTAC |
| | | | | | Dscam-2-RA-Seq-R | | GTGGTTAACCAATATTCGTAG |
| **24** | ***Tsh*** | Tsh-1 | **28,419** | **1** | Tsh-1-LA-AscI-F | **667** | AGGCGCGCCCCAAACACACGCAACGGGAAAG |
| | | | | | Tsh-1-LA-BamHI-R | | CGCGGATCCCTGTTATAGATATGTTGCATTTC |
| | | | | **2** | Tsh-1-RA-BamHI-F | **640** | CGCGGATCCCTTTCGAGAAGTAGCAGCAAC |
| | | | | | Tsh-1-RA-PacI-R | | ACCTTAATTAAAAATGCAAATCGAAGCCAGTG |
| | | | | **3** | Tsh-1-5'-Check-R | | GTTGTGATTGCTAATCGCATC |
| | | | | | Tsh-1-3'-Check-F | | ACAATTTGAGTGAAGGCGTAG |
| | | | | **4** | Tsh-1-LA-Seq-F | | TTCACACATAGTGAGTCCTAG |
| | | | | | Tsh-1-RA-Seq-R | | AGCTATCGTTGTCTCCGCTTC |
| **25** | ***Tsh*** | Tsh-2 | **86,419** | **1** | Tsh-2-LA-AscI-F | **599** | AGGCGCGCCTCTTCGTTTGGCTGCCAGGCAG |
| | | | | | Tsh-2-LA-BamHI-R | | CGCGGATCCCTCGCTTGGTTTCGGAATCTG |
| | | | | **2** | Tsh-2-RA-BamHI-F | **631** | CGCGGATCCAAATAAAGGCACCTACTCCAG |
| | | | | | Tsh-2-RA-PacI-R | | ACCTTAATTAAATTAAAGCCACCATCAGAATTG |
| | | | | **3** | Tsh-2-5'-Check-R | | TACTGTCCTGATGTAGGGTAG |
| | | | | | Tsh-2-3'-Check-F | | GCAGTGTTGTCGTGTCTCATG |
| | | | | **4** | Tsh-2-LA-Seq-F | | GCAGCAGACATTATAGCCAAC |
| | | | | | Tsh-2-RA-Seq-R | | TCTTCCTCTCATACGGCATTG |
| **B: Primers used for PCR** | | | | | | | |
| **PCR** | **Gene** | **Fragment** | **Length** | **Primer Set** | **Primer name** | **Homology** | **Primer Sequence (5' to 3')** |
| **26** | ***Sec6*** | Sec6-S | **2,813** | **PCR** | Sec6-S-AscI-F | | AGGCGCGCCCCATACTTAGGTCGAATCCT |
| | | | | | Sec6-S-PacI-R | | GGTTTAATTAACGACCGTTACAACACTACAT |
| | | | | **Seq** | Sec6-S-Seq-1F | | CATCTTCAACGTAGATGCCAG |
| | | | | | Sec6-S-Seq-1R | | CTTGGGCTGCTTGTGCAGTTC |
| | | | | | Sec6-S-Seq-2F | | GCTTGGAGGGCAATGAATATG |
| | | | | | Sec6-S-Seq-2R | | TTCCAGCGCTTTGAGATGTTC |
| | | | | | Sec6-S-Seq-3F | | CAGTAACATTGCCGTGGACAC |
| | | | | | Sec6-S-Seq-3R | | ACAGTCGCTTGGATAGCAGTG |
| **C: Primers used for double gap-repair** | | | | | | | |

| Gap-repair | Gene | Fragment | Length | Primer Set | Primer name | Homology | Primer Sequence (5' to 3') |
|---|---|---|---|---|---|---|---|
| **27** | ***Ten-m*** | ten-m | **20,000** | **1** | Ten-m-LA1-MluI-F | **500** | CGACGCGTCATCAGCCGATCGCCAAATTTTG |
| | | | | | Ten-m-LA1-BamHI-R | | GAGCCAAGCCCTGGGCACCTCTGGATCCAAGTTCGAAAGGGACGTTGTG |
| | | | | **2** | Ten-m-RA1-BamHI-F | **354** | CACAACGTCCCTTTCGAACTTGGATCCAGAGGTGCCCAGGGCTTGGCTC |
| | | | | | Ten-m-RA1-EcoR-R | | GGAATTCGGAGTCTGCCAATGAAAAAAAAGAC |
| | | | | **3** | Ten-m-LA2-5'-Check-R | | CCTCCAGTGCAACTACTACGC |
| | | | | | Ten-m-LA2-3'-Check-F | | CAAAACAGGAGCCAGCGGTAG |
| | | | | **4** | Ten-m-LA1-Seq-F | | ATGGAGTCCATTTGGTGGATC |
| | | | | | Ten-m-RA1-Seq-R | | TCGAATTATCATCATGGCATC |
| | | | **132,959** | **5** | Ten-m-RA2-EcoRI-AscI-F | **500** | GGAATTCGGCGCGCCGTATATGCCACTTGTGTGGATATTTG |
| | | | | | Ten-m-RA2-PacI-R | | ACCTTAATTAAGCTGGGAACTTTAACTTCATC |
| | | | | **6** | Ten-m-5'-Check-R | | GAACGTTTGACACCCGCTCGC |
| | | | | | Ten-m-3'-Check-F | | TGTTTGGTGAGCAATCGATTG |
| | | | | **7** | Ten-m-RA1-Seq-F | | GATGCCATGATGATAATTCGA |
| | | | | | Ten-m-RA2-Seq-R | | GAAATGAAACTCACACCATTC |

**SUPPLEMENTARY REFERENCES**

1. P. Liu, N. A. Jenkins, N. G. Copeland, *Genome Res.* **13**, 476 (2003).

2. D. L. Court *et al.*, *Gene* **315**, 63 (2003).

3. E. C. Lee *et al.*, *Genomics* **73**, 56 (2001).

4. J. Wild, Z. Hradecna, W. Szybalski, *Genome Res.* **12**, 1434 (2002).

5. U. J. Kim *et al.*, *Genomics* **34**, 213 (1996).

6. E. Frengen *et al.*, *Genomics* **58**, 250 (1999).

7. Y. Zhang, J. P. Muyrers, J. Rientjes, A. F. Stewart, *BMC. Mol. Biol.* **4**, 1 (2003).

8. A. C. Groth, E. C. Olivares, B. Thyagarajan, M. P. Calos, *Proc. Natl. Acad. Sci. U. S. A* **97**, 5995 (2000).

9. H. J. Bellen *et al.*, *Genetics* **167**, 761 (2004).

10. X. Li, N. Lobo, C. A. Bauser, M. J. Fraser, Jr., *Mol. Genet. Genomics* **266**, 190 (2001).

11. S. T. Thibault *et al.*, *Nat. Genet.* **36**, 283 (2004).

12. A. M. Handler, R. A. Harrell, *Insect Mol. Biol.* **8**, 449 (1999).

13. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, *J. Mol. Biol.* **215**, 403 (1990).

14. A. C. Groth, M. Fish, R. Nusse, M. P. Calos, *Genetics* **166**, 1775 (2004).

15. S. Beronja *et al.*, *J. Cell Biol.* **169**, 635 (2005).

16. F. L. Liebl, K. Chen, J. Karr, Q. Sheng, D. E. Featherstone, *BMC. Biol.* **3**, 27 (2005).

17. C. Delidakis, A. Preiss, D. A. Hartley, S. rtavanis-Tsakonas, *Genetics* **129**, 803 (1991).

18. T. Hummel *et al.*, *Neuron* **37**, 221 (2003).

19. S. Baumgartner *et al.*, *Cell* **87**, 1059 (1996).

20. R. Nolo, L. A. Abbott, H. J. Bellen, *Cell* **102**, 349 (2000).

21. T. W. Koh, P. Verstreken, H. J. Bellen, *Neuron* **43**, 193 (2004).

22. A. C. Spradling *et al.*, *Genetics* **153**, 135 (1999).

23. A. Fayyazuddin, M. A. Zaheer, P. R. Hiesinger, H. J. Bellen, *PLoS. Biol.* **4**, e63 (2006).

24. P. Verstreken *et al.*, *Neuron* **47**, 365 (2005).