A rapid and robust selection procedure for generating drug-selectable marker-free recombinant malaria parasites

Giulia Manzoni[1,2,3§], Sylvie Briquet[1,2,3§], Veronica Risco-Castillo[1,2,3], Charlotte Gaultier[1,2,3], Selma Topçu[1,2,3], Maria Larisa Ivănescu[1,2,3], Jean-François Franetich[1,2,3], Bénédicte Hoareau[4], Dominique Mazier[1,2,3,5], Olivier Silvie[1,2,3*]

[1]Sorbonne Universités, UPMC Univ Paris 06, CR7, Centre d'Immunologie et des Maladies Infectieuses (CIMI-Paris), 75013, Paris, France. [2]INSERM, U1135, CIMI-Paris, 75013, Paris, France. [3]CNRS, ERL 8255, CIMI-Paris, 75013, Paris, France. [4]Sorbonne Universités, UPMC Univ Paris 06, Plateforme de Cytométrie en Flux CyPS, site Pitié-Salpêtrière, Paris, France. [5]Assistance Publique-Hôpitaux de Paris, Groupe Hospitalier Pitié-Salpêtrière, Service de Parasitologie-Mycologie, Paris, France.

[§]GM and SB contributed equally to the work

[*]Correspondence and requests for materials should be addressed to O.S. (olivier.silvie@inserm.fr).

SUPPLEMENTARY INFORMATION

SUPPLEMENTARY FIGURE LEGENDS

Figure S1. GOMO plasmid maps

A. The GOMO-GFP plasmid contains 1) a GFP coding sequence (0.7 kb, green box) under control of the constitutive *PbHSP70* promoter (1.9 kb) 2) a hDHFR-yFCU fusion gene (2.1 kb, dark and light blue boxes), for positive-negative selection 3) a mCherry coding sequence (0.7

kb, red box). Both hDHFR-yFCU and mCherry are placed under control of a single bidirectional *Pbe*EF1α promoter (1.0 kb). The GFP and mCherry reporter genes are followed by the same sequence corresponding to the 3' UTR of *P. berghei DHFR-TS* (1.0 kb, orange box and lollipop), placed in the same orientation, which serves both as a transcription terminator and for excision of the drug resistance cassette after recombination. The hDHFR-yFCU coding sequence is followed by the 3' UTR of *PbHSP70* (0.7 kb, blue lollipop). For homologous recombination gene deletion, two fragments of the target gene can be introduced on each side of the cassettes, in *Sac*II and *Not*I, and *Xho*I and *Kpn*I restriction sites, respectively.

B. The GOMO-GFP-LUC plasmid contains the same elements as the GOMO-GFP vector, except that the GFP coding sequence was replaced by a GFP-Luciferase fusion gene (2.4 kb, light and dark green boxes), placed under control of the *PbeEF1α* promoter (0.6 kb).

Plasmid maps are represented not at scale.

Figure S2. Gating strategy for flow cytometry-assisted sorting of drug-selectable marker-free recombinant parasites

P. berghei ANKA parasites were transfected with a GOMO-GFP construct for *PbP230p* gene replacement. After positive selection with pyrimethamine, GFP^+ $mCherry^+$ parasites (A) were sorted by flow cytometry, and submitted to negative selection with 5-FC for sorting of GFP^+ $mCherry^-$ parasites (B). For gating, singlets were first selected using forward scatter height (FSC-H) and forward scatter area (FSC-A) (left panels). Then small particles and debris and large cells were excluded and the predominant population of red blood cells (RBCs) selected using forward scatter area (FSC-A) and side scatter area (SSC-A) (middle panels). Finally, GFP^+ $mCherry^+$ (A) or GFP^+ $mCherry^-$ (B) infected erythrocytes were gated and sorted (open red and green boxes, respectively).

Figure S3. Alternative integration of the GOMO constructs at the parasite *DHFR/TS* locus

A-B. After transfection with GOMO-GFP (A) or GOMO-GFP-LUC (B) constructs, a double crossover recombination at the 3' end of the *DHFR/TS* locus can occur, due to the presence of the repeated *PbDHFR/TS* 3' element (pink lollipop), resulting in integration of the hDHFR-yFCU and mCherry cassettes into the parasite genome. A second recombination event results in excision of the double cassette and restoration of a WT *DHFR/TS* locus.

C-D. PCR analysis of genomic DNA isolated from *P.berghei* parasites transfected with GOMO-GFP (C) or GOMO-GFP-LUC (D) constructs targeting the *PbP230p* gene, after positive selection with pyrimethamine but before 5-FC exposure. Confirmation of construct integration at the *PbDHFR/TS* locus was achieved with primer combinations specific for 5' or 3' integration.

Figure S4. Alternative integration of the GOMO constructs at the parasite *eEF1α* locus

After transfection with a GOMO-GFP-LUC construct, a double crossover recombination in the endogenous e*EF1α* promoter region may theoretically also occur, due to the presence of the two repeated e*EF1α* promoter elements in the construct, but could not be documented in this study. This would result in integration of the hDHFR-yFCU and GFP-LUC cassettes in the parasite genome. A second recombination event results in excision of the two cassettes and restoration of a WT e*EF1α* locus.

Figure S5. Alternative excision of the GFP-LUC and hDHFR-yFCU cassettes results in drug-selectable marker-free mCherry expressing parasites

A. After integration of a GOMO-GFP-LUC construct at the target gene locus, a second recombination event can occur between the repeated e*EF1α* promoter sequences instead of the *PbDHFR/TS* 3' elements, resulting in excision of the GFP-LUC cassette together with the hDHFR-yFCU selection marker, and persistence of the mCherry cassette.

B. *P. berghei* ANKA parasites were transfected with a GOMO-GFP-LUC construct for *PbP230p* gene replacement. After positive and negative selection, a population of GFP^- $mCherry^+$ parasites was obtained and sorted by FACS, and their genomic DNA analysed by PCR. Confirmation of *PbP230p* gene deletion was achieved with primer combinations specific for 5' or 3' integration. Excision of the GFP-LUC cassette was confirmed using a primer combination specific for the 5' excision event (5' excised).

Figure S6. Replacement strategy to generate drug-selectable marker-free Δ*slarp*-GFP and Δ*slarp*-GFP-LUC *P. berghei* ANKA parasites

The wild-type (WT) genomic locus of *P. berghei SLARP* (PBANKA_090210) was targeted with GOMO-GFP (A) and GOMO-GFP-LUC (B) replacement plasmids containing a 5' and a 3' homologous sequence inserted on each side of the plasmid cassettes. Upon a double crossover event, the *SLARP* gene is replaced by the GFP(-LUC)/hDHFR-yFCU/mCherry triple cassette. Recombination between the *PbDHFR/TS* 3' UTR repeated sequences (pink lollipops) results in excision of the hDHFR-yFCU and mCherry cassettes. Replacement- and wild type- specific test primer combinations and expected PCR fragments (WT, 5' integration, 3' integration and 3' marker excised) are indicated by arrows and lines, respectively. The Southern probe and expected restriction fragments are also shown.

Figure S7. Replacement strategy to generate drug-selectable marker-free Δ*p230p*-GFP and Δ*p230p*-GFP-LUC *P. yoelii* 17XNL parasites

The wild-type (WT) genomic locus of *P. yoelii P230p* (PY03857) was targeted with GOMO-GFP (A) and GOMO-GFP-LUC (B) replacement plasmids containing a 5' and a 3' homologous sequence inserted on each side of the plasmid cassettes. Upon a double crossover event, the *P230p* gene is replaced by the GFP(-LUC)/hDHFR-yFCU/mCherry triple cassette. Recombination between the *PbDHFR/TS* 3' UTR repeated sequences (pink lollipops) results in excision of the hDHFR-yFCU and mCherry cassettes. Replacement- and wild type-specific test primer combinations and expected PCR fragments (WT, 5′ integration, 3′ integration and 3' marker excised) are indicated by arrows and lines, respectively.

SUPPLEMENTARY MOVIE LEGENDS

Supplementary movie 1

Movie showing the end of sporozoite transformation and the beginning of liver stage development of a Δ*p230p*-GFP *P. berghei* parasite inside a HepG2 cell. Time labels are expressed in hours:minutes. Bar, 10 μm.

Supplementary movie 2

Movie showing the end of sporozoite transformation and the beginning of liver stage development of a Δ*p230p*-GFP *P. berghei* parasite inside a HepG2 cell. Time labels are expressed in hours:minutes. Bar, 10 μm.

Supplementary movie 3

Movie showing two intracellular Δ*slarp*-GFP *P. berghei* parasites inside HepG2 cells. The parasites do not develop and disappear at 12:45 and 17:45, respectively. Time labels are expressed in hours:minutes. Bar, 10 μm.

Supplementary movie 4

Movie showing an intracellular Δ*slarp*-GFP *P. berghei* parasite inside a HepG2 cell. The parasite does not develop and disappears at 10:45. Time labels are expressed in hours:minutes. Bar, 10 μm.

Supplementary movie 5

Movie showing two intracellular Δ*slarp*-GFP *P. berghei* parasites inside HepG2 cells. The parasites do not develop and disappear at 9:05 and 15:25, respectively. Time labels are expressed in hours:minutes. Bar, 10 μm.

Table S1. List of oligonucleotides.

Experiment	Oligonucleotide name	Sequence 5´ → 3´
Assembly of GOMO plasmids	eEF1aBidirPromFor	CGGggtaccCGctcgagCGcctaggTATAAAATTTTTATTTATTTATAAGC
	eEF1aBidirPromRev	CGcaattgCGtgatcaCGgtcgacTATAAAATTTTTATTTATTTATAAGC
	mCherryFor	CCGcctaggAAAATGGTGAGCAAGGGCGAGG
	mCherryRev	CCGctcgagTTACTTGTACAGCTCGTCCATG
	PbDHFRutr1For	ATAAGAATgcggccgcagtGAATTCTGTTCGTTTTTCTTATTTATATATTTATACC
	PbDHFRutr1Rev	CGCAATTGTGAAATTAATAAAATAAAATACATATCCCTC
	PbDHFRutr2For	ACGCGTCGACTGTTCGTTTTTCTTATTTATATATTTATACC
	PbDHFRutr2Rev	GGGGTACCACTCTCGAGTGAAATTAATAAAATAAAATACATATCCCTC
	HSPutrFor	CGCAATTGATTATTGTTCTGTACTTCTTTTGTGAATAG
	HSPutrRev	ATAAGAATgcggccgcagtGAATTCCAAATACCAATAATACCGTTTGGAGAATTG
	HSPpromFor	ATAAGAATGCGGCCGCACAGTGTATATTCCCTCAGTTTTCAAATGG
	HSPpromRev	CCGCCTAGGGTAATTGTAATTTATTGGGATAATAATGTTGG
	eEF1aPromFor	ATAAGAATGCGGCCGCTCGACGATGCTTGTAGATGAGTTAAGC
	eEF1aPromRev	CCGCCTAGGTATAAAATTTTTATTTATTTATAAGC
	GFPFor	CCGCCTAGGAAAATGAGTAAAGGAGAAGAACTTTTCACTGG
	GFPRev	CCCAAGCTTACTGCCTCCACTGCCTCCACTGCCTCCTTTGTATAGTTCATCCATGCCATGTG
	hDHFRFor	CCGTGATCAggtggaggtggaagtgctagc
	hDHFRRev	CCGGAATTCttaaacacagtagtatctgtcaccaaag
	yFCUFor	CGGgtcgacAAAATGGTTGGTTCGCTAAACTGC
	yFCURev	CCGggatccACCTCCTCCGCTAGCATCATTCTTC
Insertion of homologous sequences for gene knockout	PbP230p5'For	TCCCCGCGGATTTTTATTTTTTTCTCGGTTTGCGAAAGG
	PbP230p5'rev	ATAAGAATGCGGCCGCTTAACATCAGTTATCCCTCTTGTTATAACG
	PbP230p3'For	CCGCTCGAGATTGTTTTAGCTTGGCGATTTCTGTGTGTG
	PbP230p3'rev	GGGGTACCTGCAACTTATACTAGTGAAGTATCACTGTG
	PyP230p5'For	TCCCCGCGGAAGGAATGACATAATAAAAACCGAAAACC
	PyP230p5'rev	ATAAGAATgcggccgcACATCAGTTACCTCTCTTTTTATAGCGTC
	PyP230p3'For	CCGCTCGAGTTGGGTTCGTGTTTTGCGATTTCTTTGG
	PyP230p3'rev	GGGGTACCTGTATCTTTTTTTAAATGTGCGATTTGTCG
	PbSLARP5'For	TCCCCGCGGACCTATATTTAATACATTTACACCTCCACG
	PbSLARP5'rev	ATAAGAATGCGGCCGCtcaTCTCAAAAACATAGGAACGTGCCTTGG
	PbSLARP3'For	CCGCTCGAGCTACTCAAGTATGTTATATTTATGCATGCC
	PbSLARP3'rev	GGGGTACCGAAAATATCTAAGAAAATCATGTAACGACG
Parasite genotyping	P1 = HSPpromTestRev	TAATAATTGAGTCTTTAGTAACGAATTGCC
	P2 = eEF1aPromTestRev	TAAGCATAAAGAGCTCGAAAAGAATTAAGC
	P3 = mCherryTestFor	ATCGTGGAACAGTACGAACGCGCCGAGG

P4 = mCherryTestRev	TCACCTTCAGCTTGGCGGTCTGG
P5 = GFPTestFor	GATGGAAGCGTTCAACTAGCAGACC
P6 = PbDHFRTestFor	TTCATCACGATAAAATAAATATGGATATGG
P7 = PbDHFRTestRev	CGTATTCAAATACCTGTATTTTTATTCCTC
P8 = HSPutrTestFor	AAAGTTAAAAAATGATCGGTGTAGTTGG
P9 = PbP230pTest5'For	TTTATATGACTTAACGGCAATGCCAAAGG
P10 = PbP230pTest3'Rev	AAGTGTTGCAAATATATTACACATGTCATG
P11 = PbP230pTestWTFor	TCCATTTGATAAATTTGTGATAAAATGTCC
P12 = PbP230pTestWTRev	TTATGTGTTTCTAGTTCTTTATCTGTTCC
P13 = PyP230pTest5'For	TCTTAATGATTCGTAGATGCTAAGAATTGG
P14 = PyP230pTest3'Rev	ATGCACTATATAAAGATAATGAGAAATGGG
P15 = PyP230pTestWTFor	AAACACATAATGATGATAATATCCTGAACC
P16 = PyP230pTestWTRev	TCACAAACATAATTTGGAGTTATTTGATCC
P17 = PbSLARPTest5'For	ACACATCCTAGATATACAACTTTCATCAAG
P18 = PbSLARPTest3'Rev	TTTTAAATTATGAAATAAAATGCCCACACC
P19 = PbSLARPTestWTFor	GAGACATATCAAATAATTACTACATACCACC
P20 = PbSLARPTestWTRev	GGGGTTCATAATTATATTTTCATTAGGGTCC

Restriction sites in the primer sequences are underlined.

Figure S1

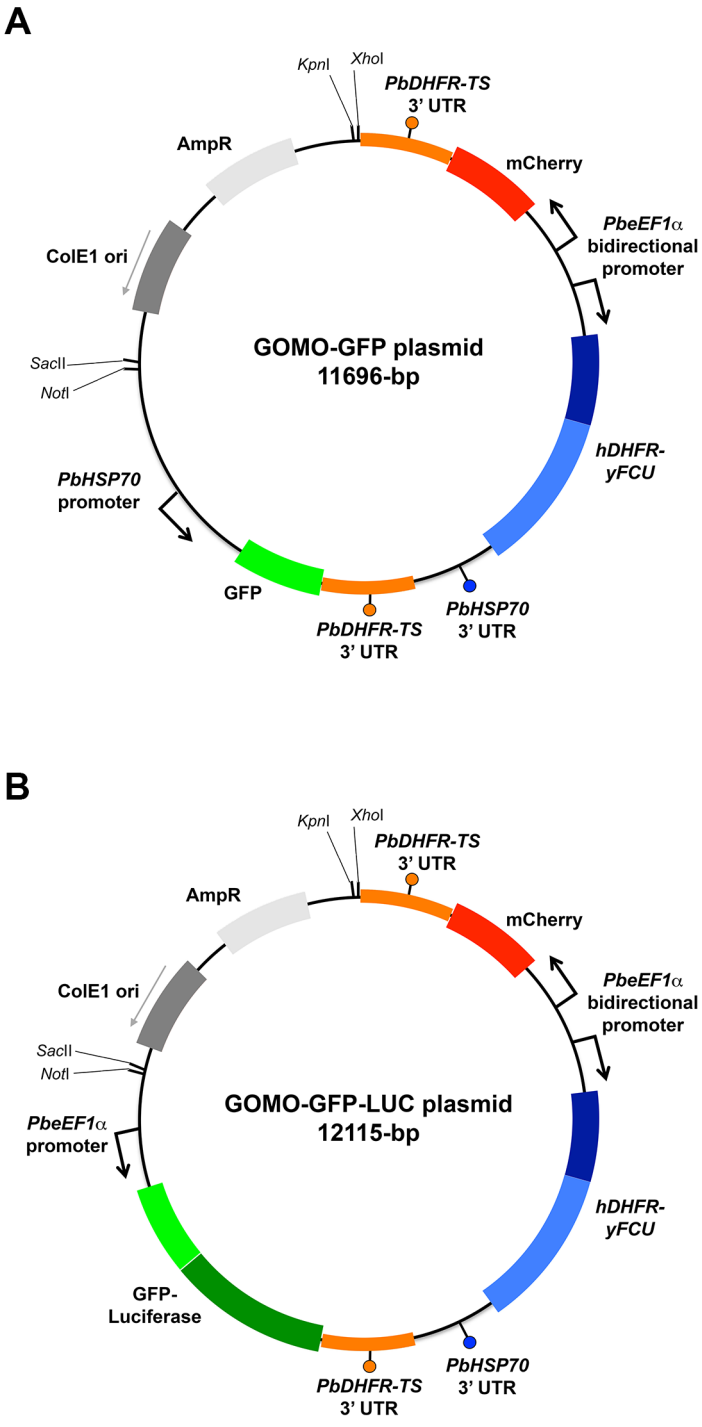

A
KpnI
XhoI
PbDHFR-TS
3' UTR
mCherry
PbeEF1α
bidirectional
promoter
hDHFR-
yFCU
PbHSP70
3' UTR
PbDHFR-TS
3' UTR
GFP
PbHSP70
promoter
NotI
SacII
ColE1 ori
AmpR
GOMO-GFP plasmid
11696-bp
B
KpnI
XhoI
PbDHFR-TS
3' UTR
mCherry
PbeEF1α
bidirectional
promoter
hDHFR-
yFCU
PbHSP70
3' UTR
PbDHFR-TS
3' UTR
GFP-
Luciferase
PbeEF1α
promoter
NotI
SacII
ColE1 ori
AmpR
GOMO-GFP-LUC plasmid
12115-bp

Figure S2

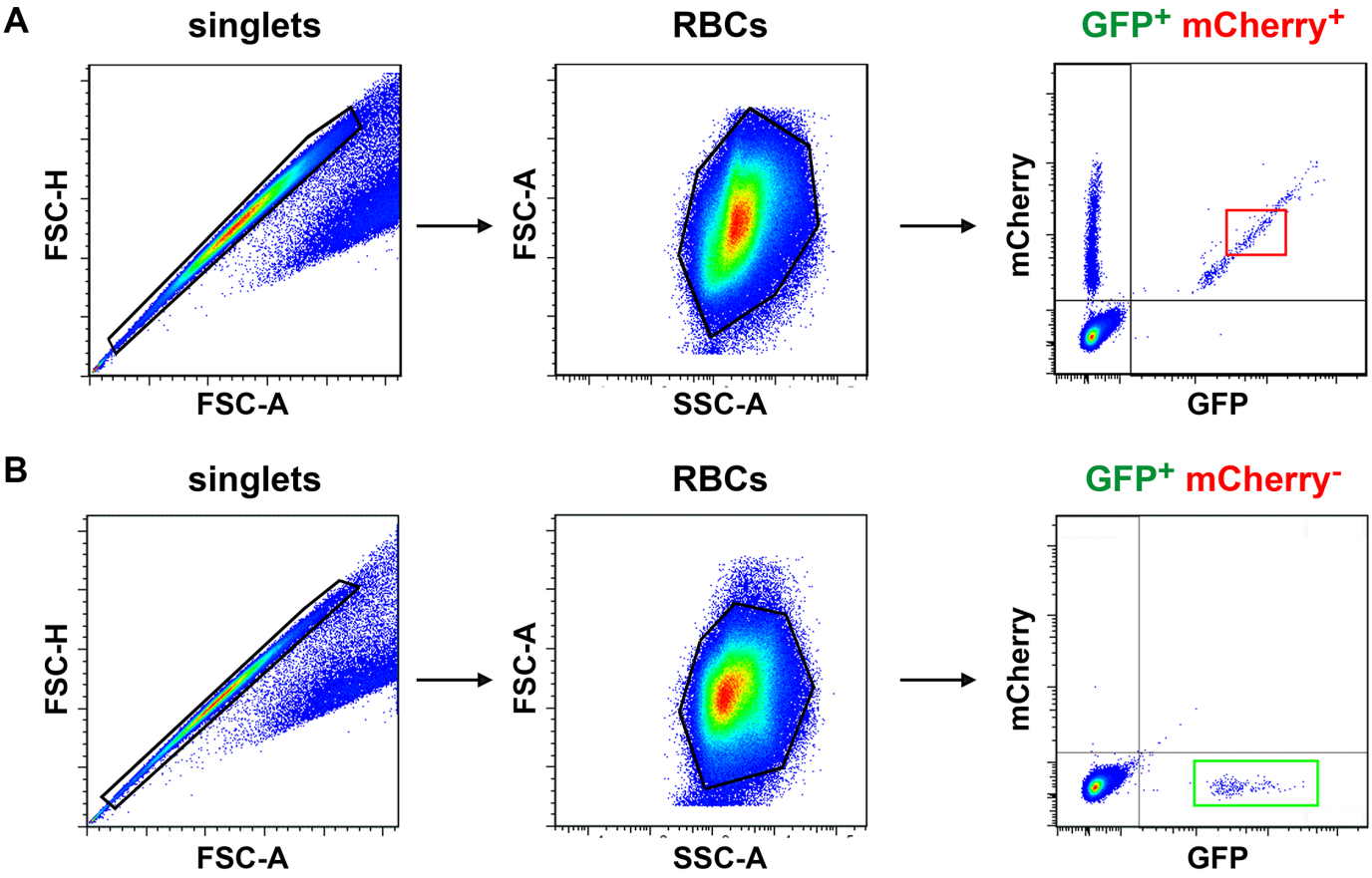

A
singlets
RBCs
GFP+ mCherry+
FSC-H
FSC-A
FSC-A
SSC-A
mCherry
GFP
B
singlets
RBCs
GFP+ mCherry-
FSC-H
FSC-A
FSC-A
SSC-A
mCherry
GFP

Figure S3

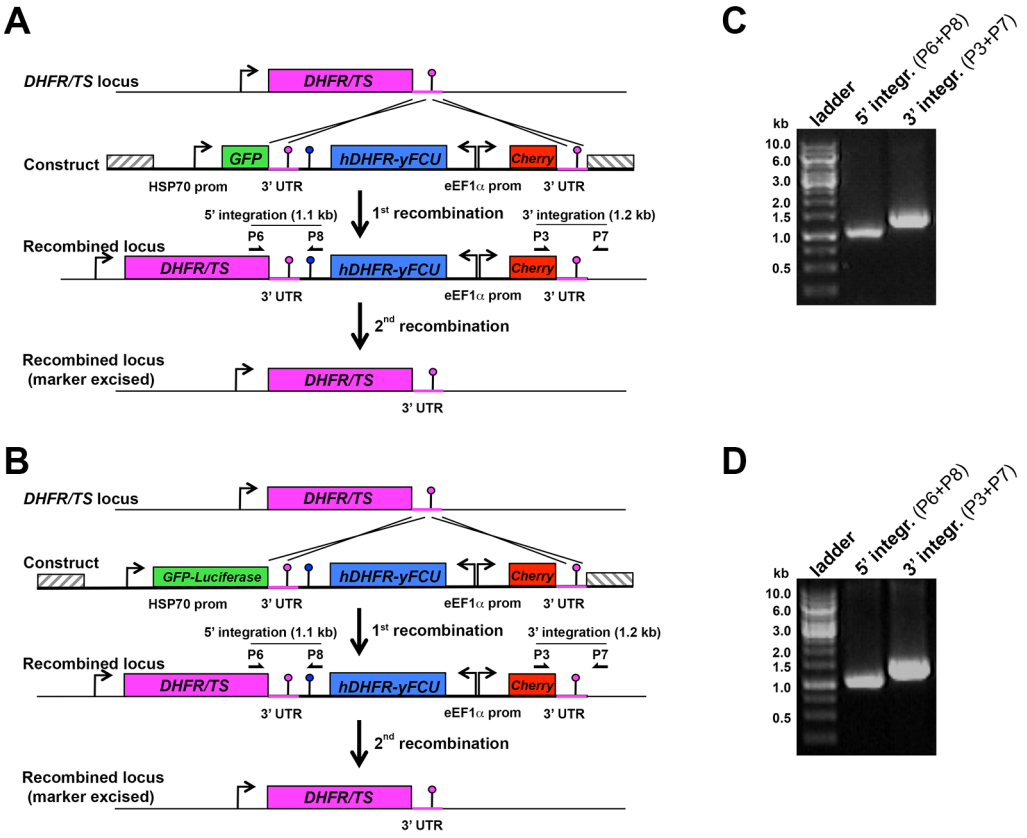

A
DHFR/TS locus
DHFR/TS
Construct
GFP
hDHFR-yFCU
Cherry
HSP70 prom
3' UTR
eEF1α prom
3' UTR
5' integration (1.1 kb)
1st recombination
3' integration (1.2 kb)
Recombined locus
P6
P8
P3
P7
DHFR/TS
hDHFR-yFCU
Cherry
3' UTR
eEF1α prom
3' UTR
2nd recombination
Recombined locus
(marker excised)
DHFR/TS
3' UTR
B
DHFR/TS locus
DHFR/TS
Construct
GFP-Luciferase
hDHFR-yFCU
Cherry
HSP70 prom
3' UTR
eEF1α prom
3' UTR
5' integration (1.1 kb)
1st recombination
3' integration (1.2 kb)
Recombined locus
P6
P8
P3
P7
DHFR/TS
hDHFR-yFCU
Cherry
3' UTR
eEF1α prom
3' UTR
2nd recombination
Recombined locus
(marker excised)
DHFR/TS
3' UTR
C
ladder
5' integr. (P6+P8)
3' integr. (P3+P7)
kb
10.0
6.0
3.0
2.0
1.5
1.0
0.5
D
ladder
5' integr. (P6+P8)
3' integr. (P3+P7)
kb
10.0
6.0
3.0
2.0
1.5
1.0
0.5

Figure S4

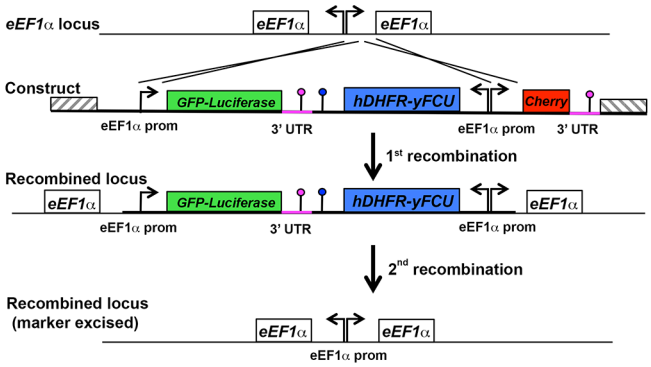

eEF1α locus
eEF1α
eEF1α
Construct
GFP-Luciferase
hDHFR-yFCU
Cherry
eEF1α prom
3' UTR
eEF1α prom
3' UTR
1st recombination
Recombined locus
eEF1α
GFP-Luciferase
hDHFR-yFCU
eEF1α
eEF1α prom
3' UTR
eEF1α prom
2nd recombination
Recombined locus
(marker excised)
eEF1α
eEF1α
eEF1α prom

Figure S5

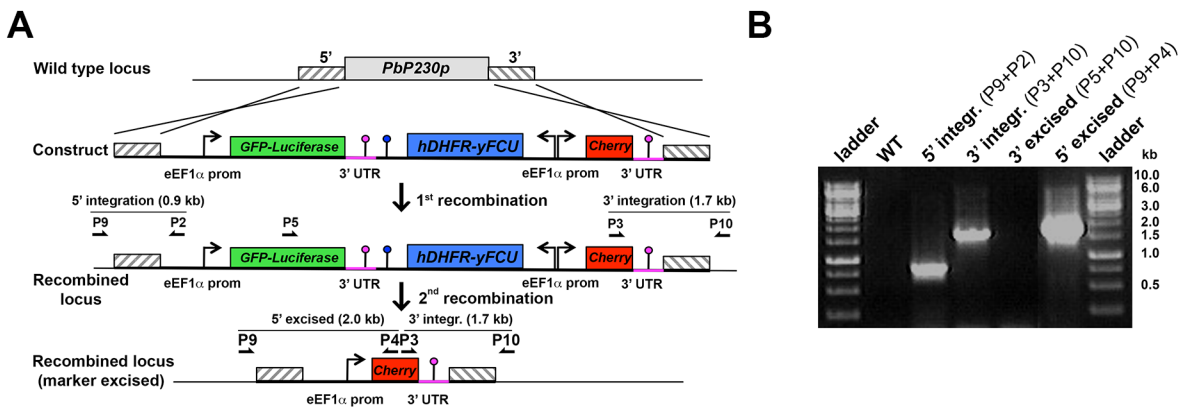

A
Wild type locus
5'
PbP230p
3'
Construct
GFP-Luciferase
hDHFR-yFCU
Cherry
eEF1α prom
3' UTR
eEF1α prom
3' UTR
1st recombination
5' integration (0.9 kb)
3' integration (1.7 kb)
P9
P2
P5
P3
P10
Recombined locus
2nd recombination
5' excised (2.0 kb)
3' integr. (1.7 kb)
P4
Recombined locus (marker excised)
B
ladder
WT
5' integr. (P9+P2)
3' integr. (P3+P10)
3' excised (P5+P10)
5' excised (P9+P4)
ladder
kb
10.0
6.0
3.0
2.0
1.5
1.0
0.5

Figure S6

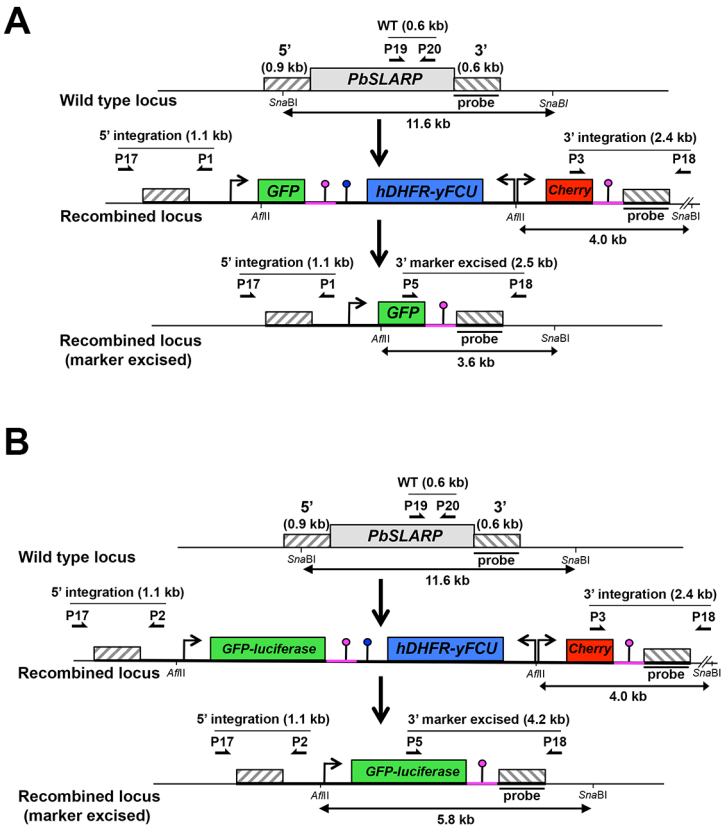
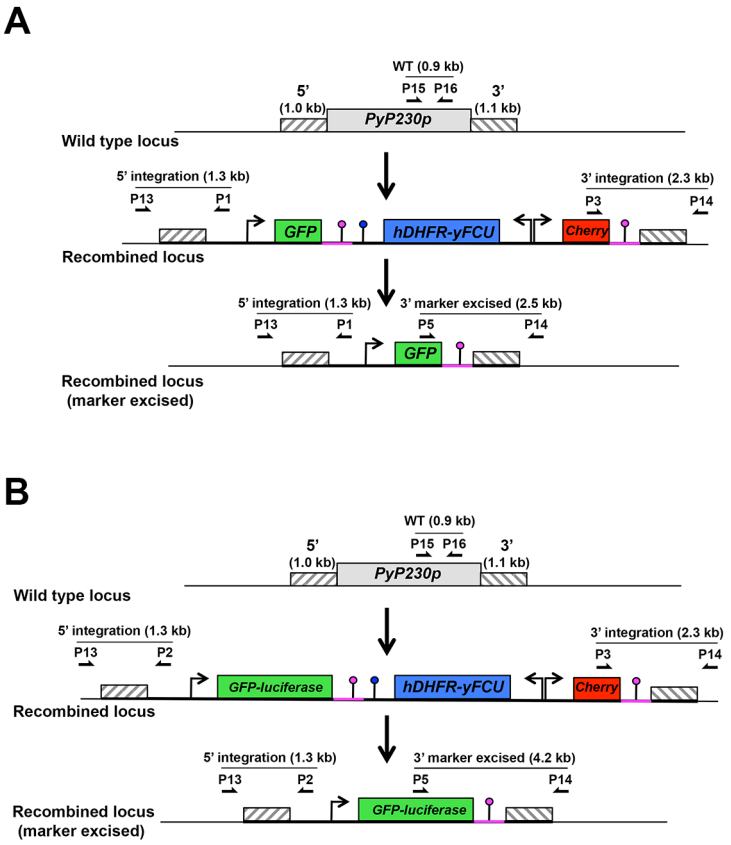

A
WT (0.6 kb)
5'
(0.9 kb)
P19 P20
3'
(0.6 kb)
PbSLARP
Wild type locus
SnaBI
probe
SnaBI
11.6 kb
5' integration (1.1 kb)
P17
P1
3' integration (2.4 kb)
P3
P18
GFP
hDHFR-yFCU
Cherry
Recombined locus
AflII
AflII
probe
SnaBI
4.0 kb
5' integration (1.1 kb)
3' marker excised (2.5 kb)
P17
P1
P5
P18
GFP
Recombined locus
(marker excised)
AflII
probe
SnaBI
3.6 kb
B
WT (0.6 kb)
5'
(0.9 kb)
P19 P20
3'
(0.6 kb)
PbSLARP
Wild type locus
SnaBI
probe
SnaBI
11.6 kb
5' integration (1.1 kb)
P17
P2
3' integration (2.4 kb)
P3
P18
GFP-luciferase
hDHFR-yFCU
Cherry
Recombined locus
AflII
AflII
probe
SnaBI
4.0 kb
5' integration (1.1 kb)
3' marker excised (4.2 kb)
P17
P2
P5
P18
GFP-luciferase
Recombined locus
(marker excised)
AflII
probe
SnaBI
5.8 kb

Figure S7


A
WT (0.9 kb)
5'
(1.0 kb)
P15 P16
3'
(1.1 kb)
PyP230p
Wild type locus
5' integration (1.3 kb)
P13
P1
3' integration (2.3 kb)
P3
P14
GFP
hDHFR-yFCU
Cherry
Recombined locus
5' integration (1.3 kb)
3' marker excised (2.5 kb)
P13
P1
P5
P14
GFP
Recombined locus
(marker excised)
B
WT (0.9 kb)
5'
(1.0 kb)
P15 P16
3'
(1.1 kb)
PyP230p
Wild type locus
5' integration (1.3 kb)
P13
P2
3' integration (2.3 kb)
P3
P14
GFP-luciferase
hDHFR-yFCU
Cherry
Recombined locus
5' integration (1.3 kb)
3' marker excised (4.2 kb)
P13
P2
P5
P14
GFP-luciferase
Recombined locus
(marker excised)