Split Selectable Markers

Nathaniel Jillette[1]*, Menghan Du[1,2]*, Albert Cheng[1,2,3,4]

1. The Jackson Laboratory for Genomic Medicine, Farmington, CT 06032, USA
2. Department of Genetics and Genome Sciences, University of Connecticut Health Center, Farmington, CT 06030, USA
3. Institute for Systems Genomics, University of Connecticut Health Center, Farmington, CT 06030, USA
4. Correspondence: albert.cheng@jax.org

*These authors contributed equally.

Abstract

Selectable markers are widely adopted in transgenesis and genome editing for selecting engineered cells with desired genotype but are limited in choices. We present here split selectable markers each allowing for selection of multiple "unlinked" transgenes in the context of lentivirus-mediated transgenesis as well as CRISPR/Cas-mediated biallelic knock-ins. Future development of split selectable markers may support enrichment or selection of "hyper-engineered" cells containing tens of transgenes or genetic modifications.

Main

Selectable markers, such as antibiotic resistance or fluorescent protein genes, are often used in genetic engineering to isolate cells with desired genotypes [1]. However, there are a limited number of well-characterized antibiotic resistance genes for use in eukaryotic cells and a limited number of fluorescent proteins whose spectra can be unambiguously differentiated by commonly used equipment. Researchers often run into the problem of not having enough choices of selectable markers if they are to incorporate multiple transgenes into a cell. On the other hand, selection with multiple antibiotics at the same time is often harsh to cells. "Selectable marker recycling" may provide a work-around, however, requiring multiple rounds of transgenesis, selection and removal of selection markers [2]. To allow multiple transgenes to be selected by one selection scheme, we have created split antibiotics resistance and fluorescent protein

genes wherein a gene encoding an antibiotic resistance or fluorescent protein is split into two or more segments fused to inteins ("markertrons") that can be rejoined by protein trans-splicing [3] (Fig 1a). Each markertron is inserted onto a transgenic vector carrying a specific transgene. Delivery of transgenic vectors containing a set of markertrons yields cells harboring a subset or a complete set of the marketrons. Only cells containing a complete set of markertrons produce a fully reconstituted marker protein via protein splicing and thus passes through selection while cells with partial sets of markertrons are eliminated, achieving co-selection of cells containing all intended transgenes.

We started out with engineering 2-markertron intein-split resistance (Intres) genes for double transgenesis. Since flanking residues and local protein folding can affect efficiency of intein-mediated trans-splicing, we set out to identify split points in each of the four commonly used antibiotic resistance genes compatible with two well-characterized split inteins derived from *Npu*DnaE [4,5] and *Ssp*DnaB [6]. To facilitate assessment of the effectiveness of double transgenic selection, we cloned markertrons onto lentiviral vectors expressing TagBFP2 or mCherry fluorescent proteins as test transgenes (Fig 1b). Viral preparations were transduced into U2OS cells, which were then split into replicate plates with non-selective or selective media. Following appropriate passages for antibiotics selection, the two cell cultures were analyzed by flow cytometry. For hygromycin (Hygro) resistance gene, one "native" *Ssp*DnaB split point (G200:S201) with flanking residues "GS" and one "native" *Npu*DnaE split point (Y89:C90) with "YC" residues were tested. Both enabled successful selection when both N- and C-markertrons were transduced yielding >99% BFP+ mCherry+ double transgenic cells in selected cultures compared to <10% double-positive cells in non-selected culture (Fig 1c; Plasmid pairs 3,4 and 5,6). Cells transduced with either of the two markertrons did not survive hygromycin selection. In contrast, double transgenesis with conventional full-length non-split hygromycin vectors only allow for ~20% enrichment of BFP+ mCherry+ cells (Plasmid pairs 97,98). We screened three addition potential split points (52S:53C),(240A:241C), and (292R:293C) for *Npu*DnaE with the obligatory cysteine residue on the C-extein junction and a residue on the N-extein junction that supported substantial trans-splicing activities in a previous report [7]. We also incorporated six additional *Npu*DnaE split points by inserting an "artificial" cysteine on the C-extein junction to support splicing at ectopic sites yielding additional split points. In total,

eight out of eleven split points tested supported hygromycin selection (Fig 1c). Similarly, for puromycin (Puro) (Fig 1d), neomycin (Neo) (Fig 1e) and blasticidin (Blast) (Fig 1f) resistance genes, we identified four, two, and one functional Intres pair(s), respectively. In all of these cases, cells transduced with either markertrons did not survive selection, while cells transduced with both yielded >95% double transgenic cells in selective cultures compared to <50% in non-selective cultures with the exception of Blasticidin(102) Intres, achieving lower but still significant enrichment of 91% double transgenic cells (Fig 1c~f). Details of the split points of Intres genes and plasmids are presented in Supplementary fig 1~4 and Supplementary table 1. To facilitate adoption of Intres markers, we created Gateway-compatible lentiviral vectors for convenient restriction-ligation-independent LR clonase recombination of transgenes [8] (Supplementary fig 5). We tested the functionality of these vectors by recombining TagBPF2 and mCherry, respectively to the N- and C-Intres vectors and found robust selection of double transgenic cells (Supplementary fig 5b). One potential utility of Intres vectors is to install different fluorescent markers in cells to label different cellular compartments. To explore such utility, we cloned in NLS-GFP and LifeAct-mScarlet [9], which label nucleus and F-actin, respectively, by Gateway recombination to conventional full length (FL) non-split hygromycin selectable vectors or 2-markertron hygromycin Intres vectors and transduced cells with either sets of plasmids, followed by antibiotic selection (Supplementary fig 5c). The sample transduced with non-split selectable plasmids contained both singly and doubly labelled cells, while cell transduced with Intres plasmids were all doubly labelled (Supplementary fig 5c).

To test whether split fluorescent markers can be used for transgene selection, we screened for *Npu*DnaE split points for mScarlet fluorescent protein (Supplementary fig 6 and 7a) and identified four split points allowing for >96% enrichment of double transgenic cells and three other split points enabling >60% enrichment of double transgenic cells in mScarlet-gated population, compared to <20% double transgenic cells in non-gated population (Supplementary fig 7b).

With the split points identified for 2-markertron Intres genes, we set out to engineer higher degree split markers. We tested combinations of splits points to partition a marker gene into three or more markertrons to allow for co-selection of more than two “unlinked” transgenes with one antibiotic (Fig 2a

and b). To identify pairs of split points that would allow such “Intres chain”, we cloned 3-split markertrons into three lentiviral vectors each carrying one of three fluorescent transgenes TagBFP2, EGFP, or mCherry, that will allow us to assess effectiveness of selection by flow cytometry (Fig 2c). Since hygromycin resistance gene is the longest and provides the most split points for testing, we focused on engineering 3-markertron hygromycin Intres. We tested two 3-markertron hygromycin Intres using two intervening *Npu*DnaE inteins, two using *Npu*DnaE for the first intein and *Ssp*DnaB for the second intein, as well as two using *Ssp*DnaB for the first intein and *Npu*DnaE for the second intein (Fig 2d). Five of these six 3-markertron hygromycin Intres enabled >97% and with the remaining one enabling 80% triple transgenic selection in hygromycin-selected cultures compared to <15% triple transgenic cells in non-selected cultures. Samples with leave-one-out transduction did not yield any viable cells after hygromycin selection while cells transduced with non-split hygromycin vectors yielded only 7% triple transgenic cells after selection. To facilitate the use of 3-markertron Intres, we created Gateway compatible lentiviral vectors with these markers (Supplementary fig 8a). Three sets of these vectors were each tested by recombining TagBFP (as transgene 1), EGFP (as transgene 2) and mCherry (as transgene 3) into the N-, M-, and C-Intres Gateway destination vectors and used to transduce U2OS cells, which were then split and cultured in hygromycin selection or non-selective media (Supplementary fig 8b). Two weeks after selection, cells were analyzed by flow cytometry. All three sets of 3-markertron hygromycin Intres plasmids support triple transgenic cell selection of >99% compared to <25% in the non-selected cultures (Supplementary fig 8c).

A further potential application of split selectable markers is to facilitate genome engineering and editing via the CRISPR/Cas system [10]. Although gene knockout based on NHEJ-mediated insertions/deletions (indels) occur at high frequency, precise editing and knock-in based on homology directed repair (HDR) using exogenous repair templates are inefficient [11]. We tested whether split selectable markers can be used to select for cells with CRISPR-mediated biallelic knock-in at the *AAVS1* locus [12]. We constructed targeting constructs with homology arms flanking the target site, and splice acceptor-2A peptide to trap the markertrons within intron one of the host gene *PPP1R12C*. However, we did not obtain any live cells after CRISPR/Cas knock-in experiments in HEK293T cells using these targeting constructs and two

weeks of antibiotic selection (Data not shown). We suspected that the endogenous promoter of the host gene *PPP1R12C* might not drive sufficient expression of markertrons to reconstitute enough antibiotic resistance protein to counteract actions of the antibiotics. We thus tested an alternative strategy to express Intres markertrons by TetO promoter whose activity can be titrated by doxycycline (dox) concentration. To allow comparison of Intres-mediated biallelic selection versus full-length (FL) non-split selectable markers, we implemented several different targeting construct designs. First, we drive expression of a full-length (FL) resistance gene (e.g., Hygro) together with rtTA under a constitutive EF1a promoter and a separate test Intres (e.g., Blast Intres) under a dox-inducible TetO promoter (Supplementary fig 9b, Plasmids 109 and 110). This allows comparison of full length and split selectable markers within the same constructs. To allow fair comparison of full length versus split markers driven by the same TetO promoter, we constructed two similar plasmids 107 and 108 (cf. Plasmids 109 and 110), wherein full-length antibiotic resistance gene (Blast) is placed downstream of the TetO promoter. To enable single-cell quantification of biallelic targeting and to demonstrate the feasibility of incorporating two transgenes into two *AAVS1* alleles, we appended EGFP and mScarlet fluorescent genes downstream of the test split or non-split markers via self-cleaving 2A peptide. Similarly, to test Hygro Intres, we swapped the EF1a and TetO-driven markers so that FL Hygro or Hygro Intres were placed downstream of TetO and FL Blast downstream of EF1a (Supplementary fig 9c and d; Plasmids 111~114). We co-transfected pX330-AAVS1 (Plasmid 106) containing Cas9 and sgRNA targeting *AAVS1*, and the different pairs of targeting constructs (TC) to HEK293T cells, split into triplicate doxycycline-containing media without antibiotics, with blasticidin, or with hygromycin at the subsequent passages. Two weeks after selection, we analyzed the cultures for biallelic targeting by flow cytometric measurement of GFP and RFP fluorescence (Supplementary fig 9e). As expected, non-selected cultures harbored small fraction (<1%) of biallelic knock-in GFP+/RFP+ cells (Supplementary Fig 9e; Selection = None). Selection of antibiotics where corresponding FL antibiotic resistance genes were present on targeting constructs yielded < 30% biallelic knock-in cells (Supplementary fig 9e; Blast: TC a,c,d; Hygro: TC a,b,c). In contrast, selection of antibiotics where corresponding Intres are present on the targeting constructs yielded 75% (Supplementary fig 9e; Blast Intres: TC b) and 88% (Supplementary fig 9e; Hygro Intres: TC d) biallelic knock-in cells. Selection for an additional two weeks allowed split Blast and Hygro TCs to achieve 96.5%

and 97.0% biallelic knock-in, respectively (Supplementary fig 9f and g). We next tested bilallelic engineering in KOLF2-C1 human induced pluripotent stem cells (hiPSCs), which are karyotypically normal with a stable diploid genome [13] (Fig 3). The full-length non-split Blast targeting constructs (Fig 3a) and 2-split Blast Intres targeting constructs (Fig 3b) were tested for selection of biallelically modified clones. Purified Cas9 proteins were complexed with synthetic sgRNA to form Cas9 ribonucleoprotein (RNP) and co-nucleofected with the targeting constructs into KOLF2-C1, followed by dox-induction and antibiotic selection. Surviving colonies were picked into separate wells for establishing single cell clones. Genotyping PCR revealed that targeting using non-split Blast resistance gene generated only 8% biallelic clones, while targeting using Blasticidin Intres yielded exclusively (100%) biallelically modified clones (Fig 3c and d), showing both fluorescent signals (Fig 3e) indicative of the targeting by each targeting construct at the two alleles of AAVS in these hiPSCs.

In this study, we have engineered split antibiotic resistance and fluorescent protein genes that can allow selection for two or more “unlinked” transgenes. By inserting unnatural residues at selectable markers, we showed that novel high-efficiency split points could be utilized, expanding the positions available for engineering. We demonstrated that split selectable markers could be incorporated into lentiviral vectors or gene targeting constructs in CRISPR/Cas9 genome editing experiments for positive selection of cells with double transgenesis or biallelic knock-ins. By combining two or more splits points, we showed that 3- and 4-split markers could be generated to allow higher degree transgenic selection. Future development of even higher-degree split selectable markers may enable “hyper-engineering” of cells containing tens of transgenes or targeted knock-ins.

Figure Legends

Figure 1. Split selectable marker for antibiotic co-selection of two separate transgenic vectors. (a) The coding sequence of a selectable marker is split into an N-terminal fragment (Mar_N) and a C-terminal fragment (ker_C) and separately cloned upstream of an N-terminal fragment of a split intein (Int_N) and downstream of a C-terminal fragment of the split intein (Int_C), respectively, on two different vectors each carrying a different transgene. These vectors are delivered to cells yielding sub-populations of cells

containing either one of the vectors or both of the vectors. Only cells with both vectors expressing the two intein-split selectable marker fragments ("markertrons") undergo protein trans-splicing to reconstitute a full-length selectable marker, allowing specific selection and enrichment of the double transgenic cells. (b) To screen for split points compatible for inteins for an antibiotic resistance gene, we identified potential split points according to the junctional requirement for the type of intein tested, then cloned the corresponding N-terminal and C-terminal fragments to the split intein scaffolds on lentiviral vectors equipped with TagBFP2 or mCherry fluorescent proteins, which serve as our test transgenes to evaluate selection efficiency. These are delivered into cells via lentiviral transduction. The cells were then split into replicate plates, one subjected to antibiotic selection while the other maintained in non-selective media. Following antibiotic selection, the replicate cultures were analyzed by flow cytometry. (c) 2-markertron hygromycin (Hygro) intein-split resistance (Intres) genes. Top schematics shows the split points tested for hygromycin resistance gene. The last residue of the N-terminal fragment is indicated on top of the lollipops. Circle lollipops represent split points using *Npu*DnaE intein while square lollipops represent those using *Ssp*DnaB intein. Crossed-out and shaded lollipops indicate split pairs that failed to endow cells with hygromycin resistance. The column plot below shows the percentage of double-transgenic cells (BFP+ mCherry+) in the non-selected (white columns) and the selected cultures (blue columns) analyzed by flow cytometry. (d) 2-markertron puromycin (Puro) Intres genes. Top schematics show split points tested for puromycin resistance genes while bottom column plots show percentage of double transgenic cells in the non-selected (white columns) and the selected cultures (brown columns). (e) 2-markertron neomycin (Neo) resistance genes. Top schematics show split points tested for neomycin resistance genes while bottom column plots show percentage of double transgenic cells in the non-selected (white columns) and the selected cultures (orange columns). (f) 2-markertron blasticidin (Blast) Intres genes. Top schematics show split points tested for blasticidin resistance gene while bottom column plot shows percentages of double transgenic cells in the non-selected (white columns) and the selected (cyan columns) cultures.

Figure 2. Multi-split selectable markers for co-selection of three or more transgenic vectors. (a) A selectable marker is partitioned into three fragments (M_1, M_2 and M_3). The first marker fragment (M_1) is fused upstream of the N-terminal fragment of the first split intein (I_{N1}). The second marker fragment (M_2) is

fused downstream of the C-terminal fragment of the first split intein (I_{C1}) and upstream of the N-terminal fragment of the second split intein (I_{N2}). The third marker fragment (M_3) is fused downstream of the C-terminal fragment of the second split intein (I_{C2}). The first split intein catalyzes the joining of M_1 to M_2 while the second split intein catalyzes the joining of M_2 to M_3, effectively reconstituting the full selectable marker. (b) A design of a *k*-split selectable marker via an "intein chain" mechanism. Similar to the 3-split scenario, the selectable marker is partitioned into *k* fragments, and are reconstituted through protein trans-splicing mediated by intervening split inteins. (c) Split points identified from 2-split selectable markers were used in combination to produce 3-split selectable markers. The corresponding fragments were cloned into lentiviral vectors to result in the 3-split selectable marker structure and a reporter fluorescent transgene per vector. Cells were then transduced with viruses prepared from these vectors, split into selective or non-selective media. After appropriate selection period, the cultures were analyzed by flow cytometry. (d) 3-markertron hygromycin (Hygro) Intres. Top schematics shows the split points tested for hygromycin resistance gene, with residue numbers of the last amino acid of the N-terminal fragments indicated above circle or square lollipops, representing *Npu*DnaE and *Ssp*DnaB inteins, respectively. Six 3-markertron hygromycin Intreses were tested, each indicated with a numbered line with circle or square indicating the two split points used for each case. Column plot below shows the percentage of triple transgenic (BFP+ GFP+ mCherry+) cells from the non-selective (white columns) and selective (blue columns) cultures for the 3-markertron hygromycin Intres indicated by the numbers below.

Figure 3. Intres for selection of biallelic CRISPR knock-in in human induced pluripotent stem cells. (a) Plasmids 107 and 108 contains AAVS homology arms, and Dox-inducible full length (FL) Blasticidin (Blast) and EGFP (plasmid 107) or mScarlet (plasmid 108) separated by a self-cleaving 2A peptide sequence. (b) Plasmids 109 and 110 contains AAVS homology arms, and Dox-inducible N-split Blast markertron and EGFP (plasmid 109) or C-split Blast markertron and mScarlet (plasmid 110) separated by a self-cleaving 2A peptide sequence. Cas9 ribonucleoprotein was formed *in vitro* by complexing purified Cas9 proteins with synthetic sgRNA targeting AAVS locus, introduced to hiPSCs by nucleofection, subsequently subjected to dox-induction and antibiotic selection, and finally cloned by colony picking and analyzed for allelic modification. (c) Genotyping PCR for EGFP- and mScarlet-inserted AAVS alleles of single iPSC clones generated by CRISPR editing using non-split targeting constructs (TCs) 107 and 108.

(d) Genotyping PCR for EGFP- and mScarlet-inserted AAVS alleles of single iPSC clones generated by CRISPR editing using Intres TCs 109 and 110. (e) Representative fluorescent microscopic images of hiPSC colonies from the indicated clones derived from the Intres CRISPR experiments.

Acknowledgment

Research reported in this publication was partially supported by the Jackson Laboratory internal grants, National Human Genome Research Institute (1R01HG009900) and National Cancer Institute (P30CA034196). KOLF2-C1 cells were a gift from Bill Skarnes and were derived from the HipSci consortium.

References

1. Shearer, R.F. & Saunders, D.N. Experimental design for stable genetic manipulation in mammalian cell lines: lentivirus and alternatives. *Genes to cells : devoted to molecular & cellular mechanisms* **20**, 1-10 (2015). doi:10.1111/gtc.12183
2. Abuin, A. & Bradley, A. Recycling selectable markers in mouse embryonic stem cells. *Molecular and cellular biology* **16**, 1851-1856 (1996).
3. Shah, N.H. & Muir, T.W. Inteins: Nature's Gift to Protein Chemists. *Chemical science* **5**, 446-461 (2014). doi:10.1039/c3sc52951g
4. Zettler, J., Schütz, V. & Mootz, H.D. The naturally split Npu DnaE intein exhibits an extraordinarily high rate in the protein trans-splicing reaction. *FEBS letters* **583**, 909-914 (2009).
5. Iwai, H., Züger, S., Jin, J. & Tam, P.-H. Highly efficient protein trans-splicing by a naturally split DnaE intein from Nostoc punctiforme. *FEBS letters* **580**, 1853-1858 (2006).
6. Sun, W., Yang, J. & Liu, X.Q. Synthetic two-piece and three-piece split inteins for protein trans-splicing. *The Journal of biological chemistry* **279**, 35281-35286 (2004). doi:10.1074/jbc.M405491200
7. Cheriyan, M., Pedamallu, C.S., Tori, K. & Perler, F. Faster protein splicing with the Nostoc punctiforme DnaE intein using non-native extein residues. *The Journal of biological chemistry* **288**, 6202-6211 (2013). doi:10.1074/jbc.M112.433094
8. Chee, J. & Chin, C. Gateway cloning technology: Advantages and drawbacks. *Cloning Transgenes* **4**, 138 (2015).
9. Bindels, D.S. et al. mScarlet: a bright monomeric red fluorescent protein for cellular imaging. *nature methods* **14**, 53 (2017).
10. Wang, H., La Russa, M. & Qi, L.S. CRISPR/Cas9 in genome editing and beyond. *Annual review of biochemistry* **85**, 227-264 (2016).
11. Peng, R., Lin, G. & Li, J. Potential pitfalls of CRISPR/Cas9-mediated genome editing. *The FEBS journal* **283**, 1218-1231 (2016).
12. Oceguera-Yanez, F. et al. Engineering the AAVS1 locus for consistent and scalable transgene expression in human iPSCs and their differentiated derivatives. *Methods* **101**, 43-55 (2016).
13. Streeter, I. et al. The human-induced pluripotent stem cell initiative—data resources for cellular genetics. *Nucleic acids research* **45**, D691-D697 (2016).

Figure 1

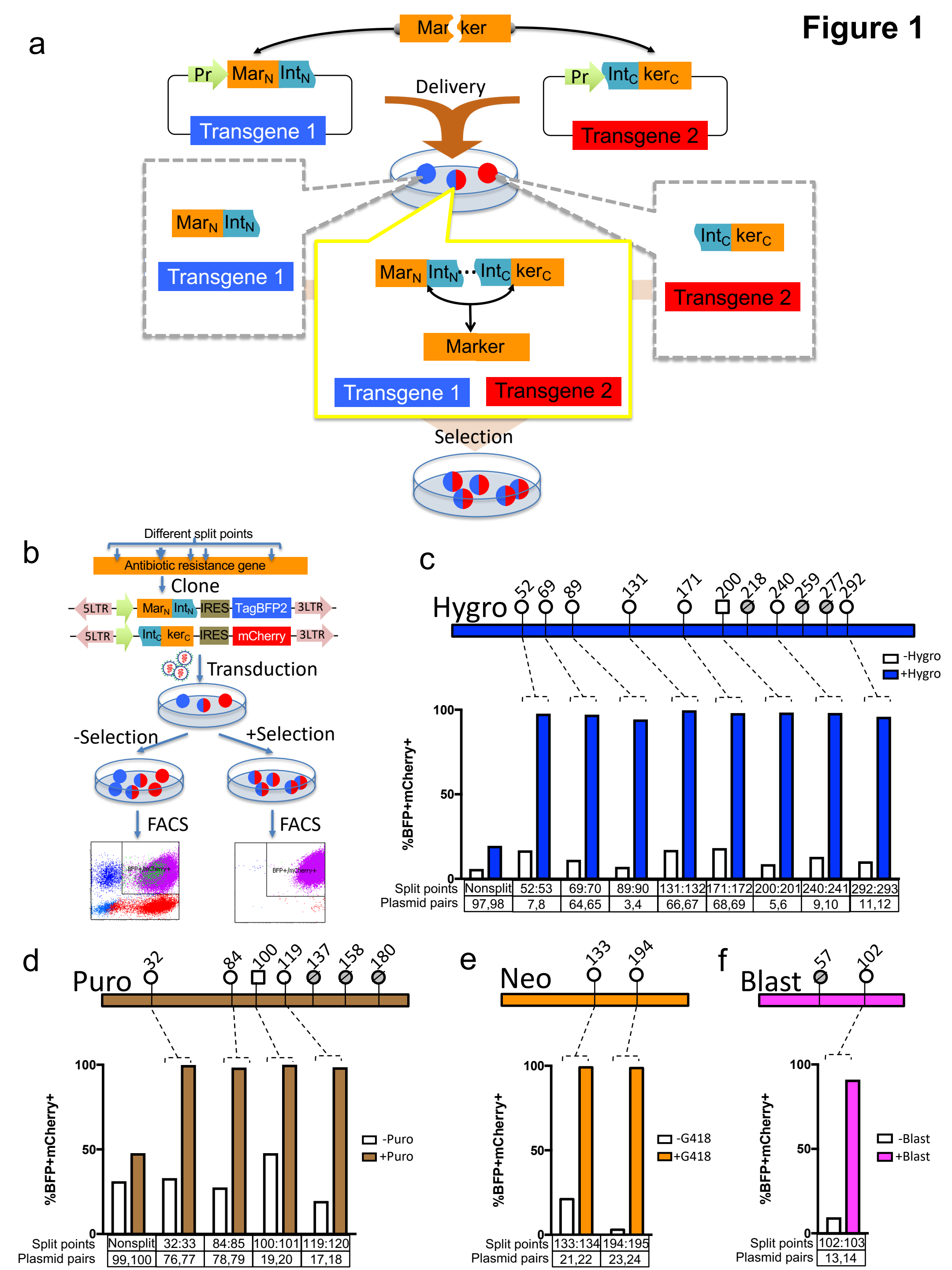

a
Mar ker
Pr
MarN IntN
Transgene 1
Delivery
Pr
IntC kerC
Transgene 2
MarN IntN
Transgene 1
MarN IntN ··· IntC kerC
Marker
Transgene 1
Transgene 2
IntC kerC
Transgene 2
Selection
b
Different split points
Antibiotic resistance gene
Clone
5LTR
MarN IntN
IRES
TagBFP2
3LTR
5LTR
IntC kerC
IRES
mCherry
3LTR
Transduction
-Selection
+Selection
FACS
FACS
BFP+/mCherry+
BFP+/mCherry+
c
Hygro
52 69 89 131 171 200 218 240 259 277 292
-Hygro
+Hygro
%BFP+mCherry+
100
50
0
Split points | Nonsplit | 52:53 | 69:70 | 89:90 | 131:132 | 171:172 | 200:201 | 240:241 | 292:293
Plasmid pairs | 97,98 | 7,8 | 64,65 | 3,4 | 66,67 | 68,69 | 5,6 | 9,10 | 11,12
d
Puro
32 84 100 119 137 158 180
%BFP+mCherry+
100
50
0
-Puro
+Puro
Split points | Nonsplit | 32:33 | 84:85 | 100:101 | 119:120
Plasmid pairs | 99,100 | 76,77 | 78,79 | 19,20 | 17,18
e
Neo
133 194
%BFP+mCherry+
100
50
0
-G418
+G418
Split points | 133:134 | 194:195
Plasmid pairs | 21,22 | 23,24
f
Blast
57 102
%BFP+mCherry+
100
50
0
-Blast
+Blast
Split points | 102:103
Plasmid pairs | 13,14

Figure 2

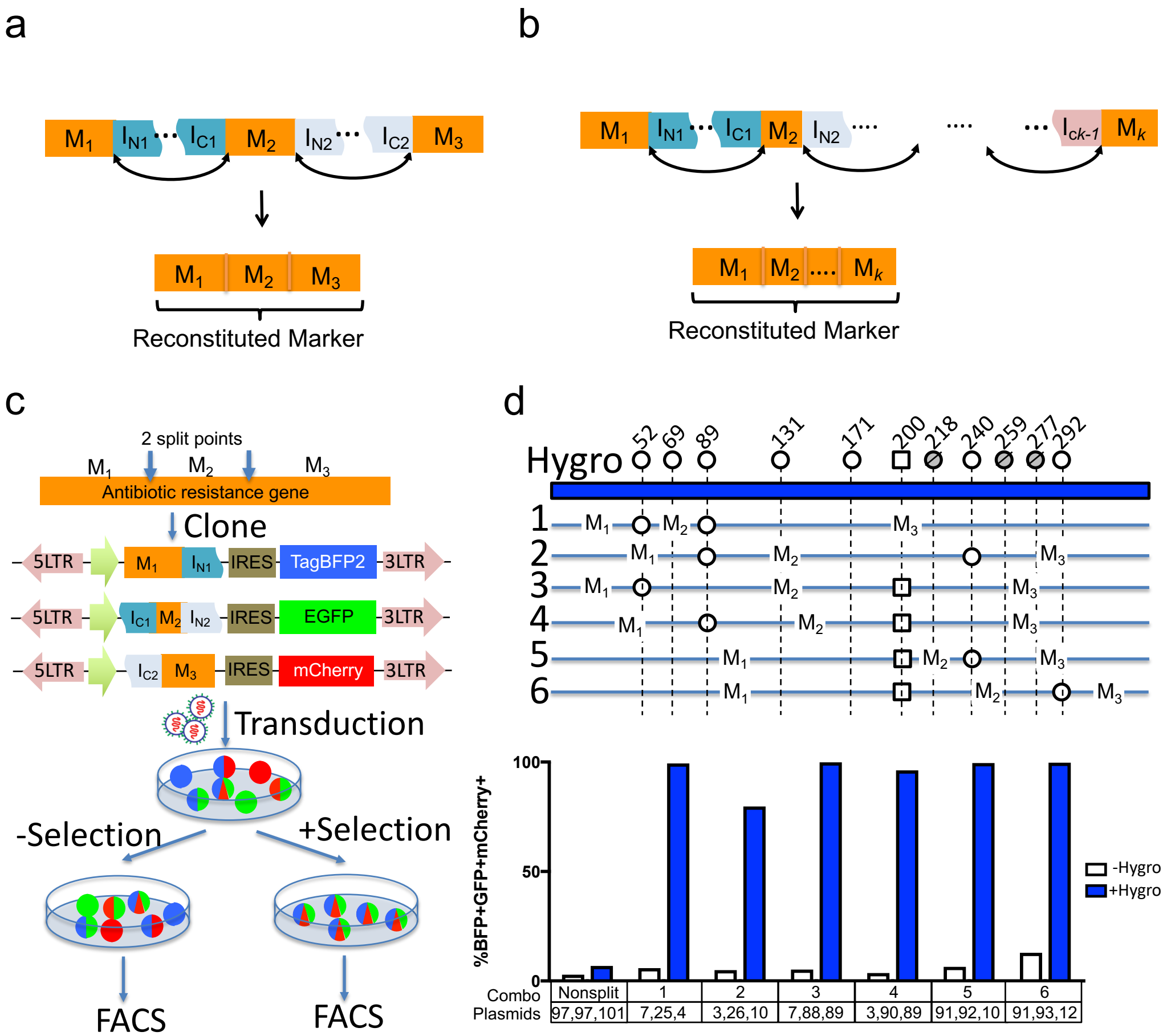
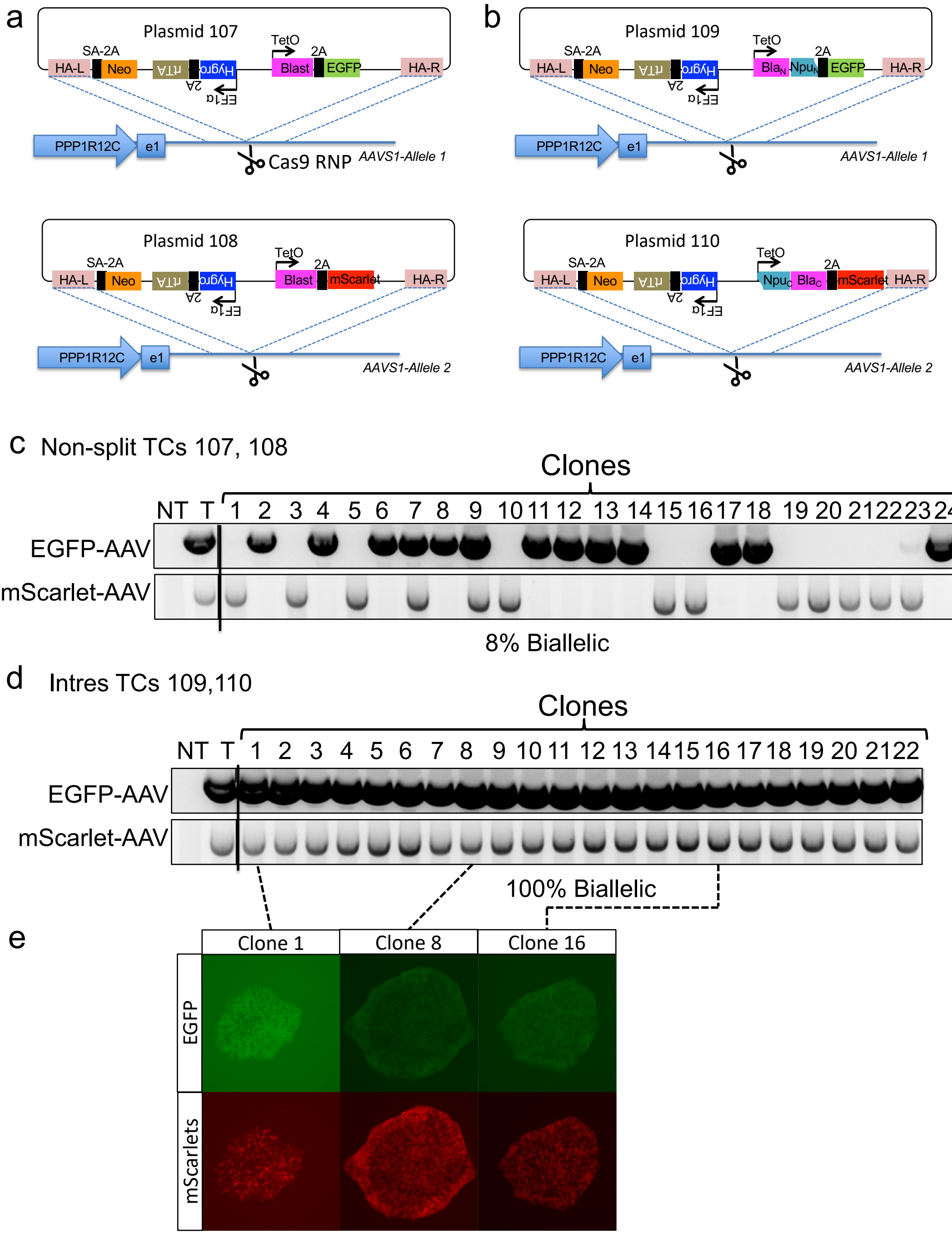
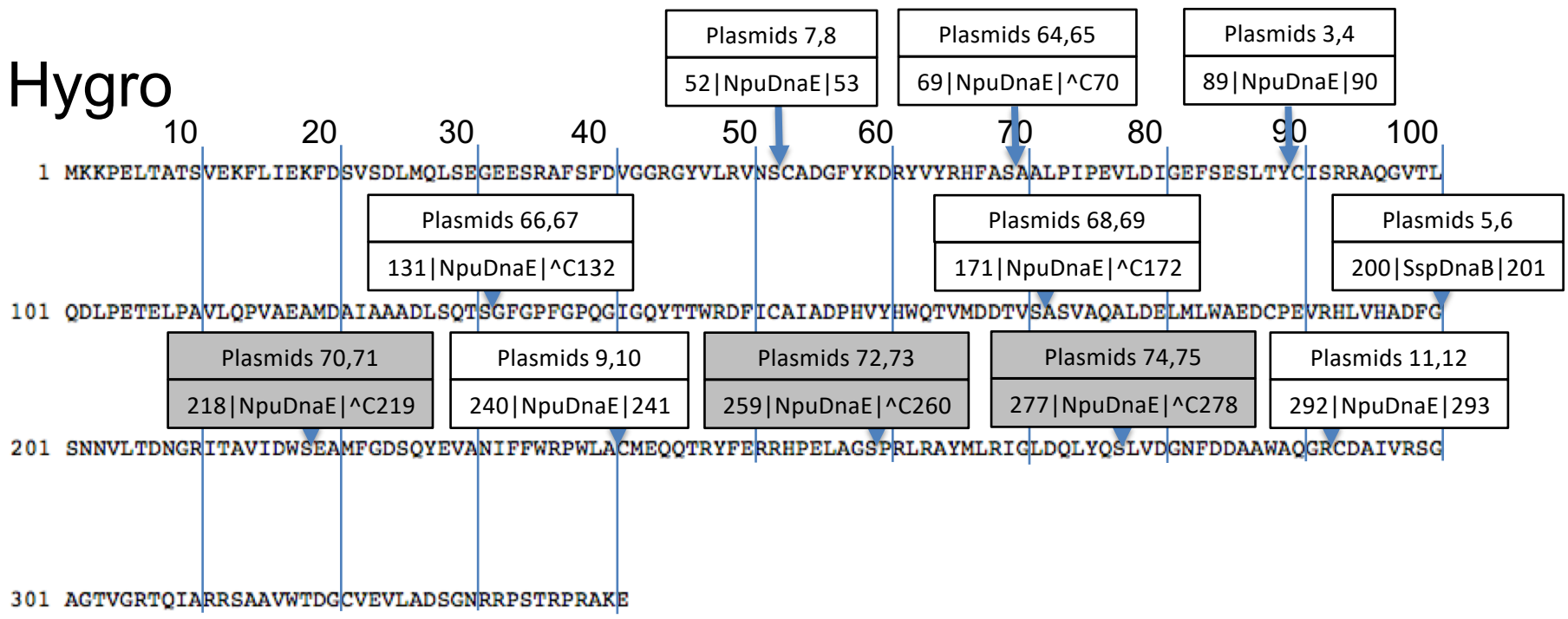

a
M1 IN1 IC1 M2 IN2 IC2 M3
M1 M2 M3
Reconstituted Marker
b
M1 IN1 IC1 M2 IN2 .... Ick-1 Mk
M1 M2 .... Mk
Reconstituted Marker
c
2 split points
M1 M2 M3
Antibiotic resistance gene
Clone
5LTR M1 IN1 IRES TagBFP2 3LTR
5LTR IC1 M2 IN2 IRES EGFP 3LTR
5LTR IC2 M3 IRES mCherry 3LTR
Transduction
-Selection
+Selection
FACS
FACS
d
Hygro
52 69 89 131 171 200 218 240 259 277 292
1 M1 M2 M3
2 M1 M2 M3
3 M1 M2 M3
4 M1 M2 M3
5 M1 M2 M3
6 M1 M2 M3
%BFP+GFP+mCherry+
100
50
0
-Hygro
+Hygro
Combo Nonsplit 1 2 3 4 5 6
Plasmids 97,97,101 7,25,4 3,26,10 7,88,89 3,90,89 91,92,10 91,93,12

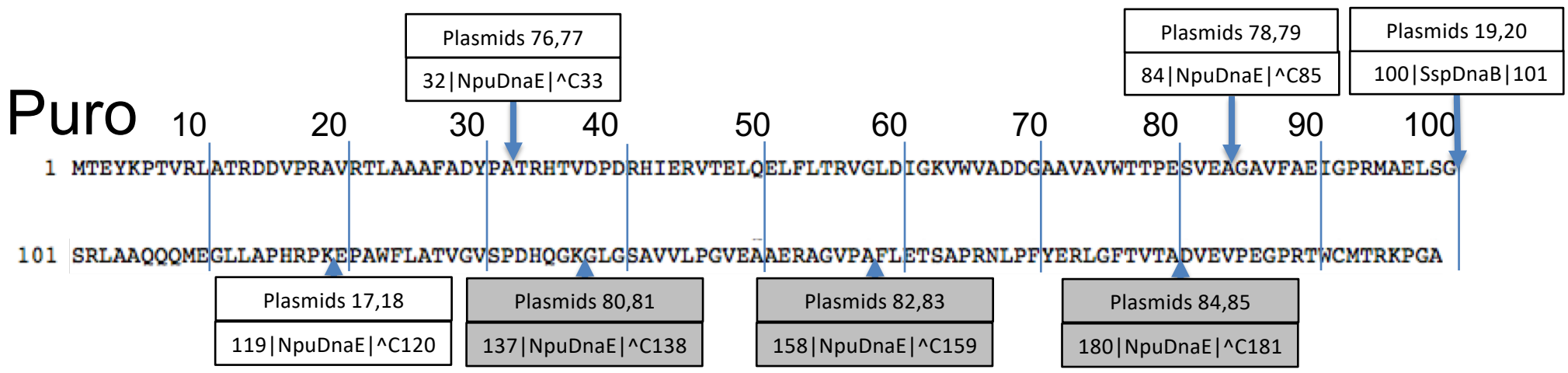
Figure 3


a
Plasmid 107
SA-2A
TetO
2A
HA-L
Neo
ftTA
Hygro
Blast
EGFP
HA-R
2A
EF1α
PPP1R12C
e1
Cas9 RNP
AAVS1-Allele 1
b
Plasmid 109
SA-2A
TetO
2A
HA-L
Neo
ftTA
Hygro
Bla_N
Npu_N
EGFP
HA-R
2A
EF1α
PPP1R12C
e1
AAVS1-Allele 1
Plasmid 108
SA-2A
TetO
2A
HA-L
Neo
ftTA
Hygro
Blast
mScarlet
HA-R
2A
EF1α
PPP1R12C
e1
AAVS1-Allele 2
Plasmid 110
SA-2A
TetO
2A
HA-L
Neo
ftTA
Hygro
Npu_C
Bla_C
mScarlet
HA-R
2A
EF1α
PPP1R12C
e1
AAVS1-Allele 2
c Non-split TCs 107, 108
Clones
NT T 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
EGFP-AAV
mScarlet-AAV
8% Biallelic
d Intres TCs 109,110
Clones
NT T 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22
EGFP-AAV
mScarlet-AAV
100% Biallelic
e
Clone 1
Clone 8
Clone 16
EGFP
mScarlets

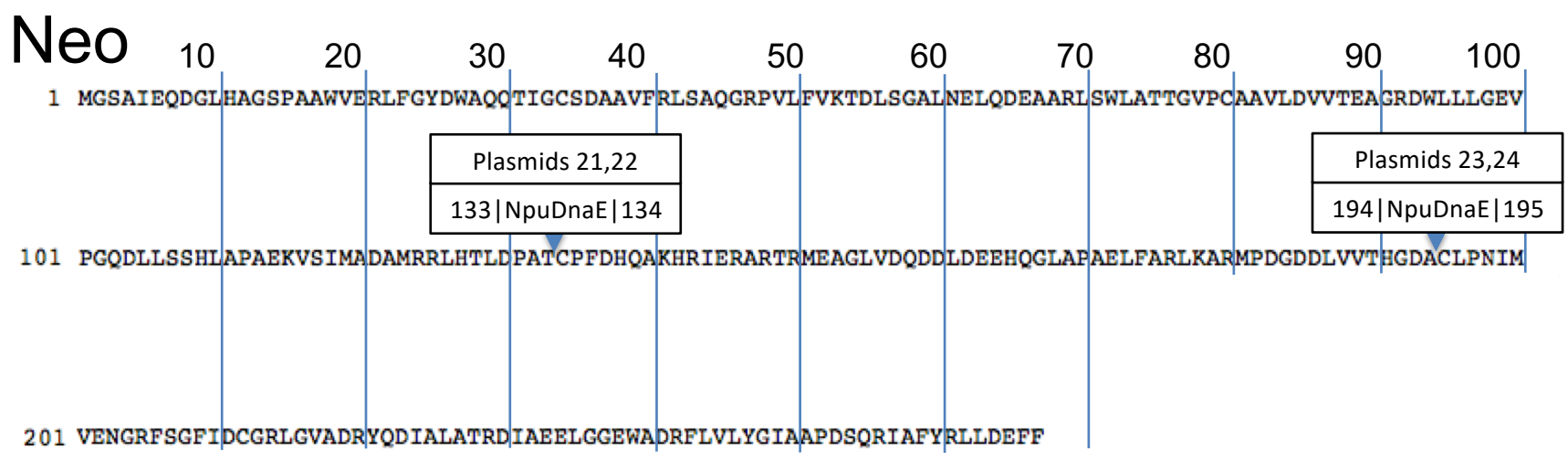
Supplementary Figure 1



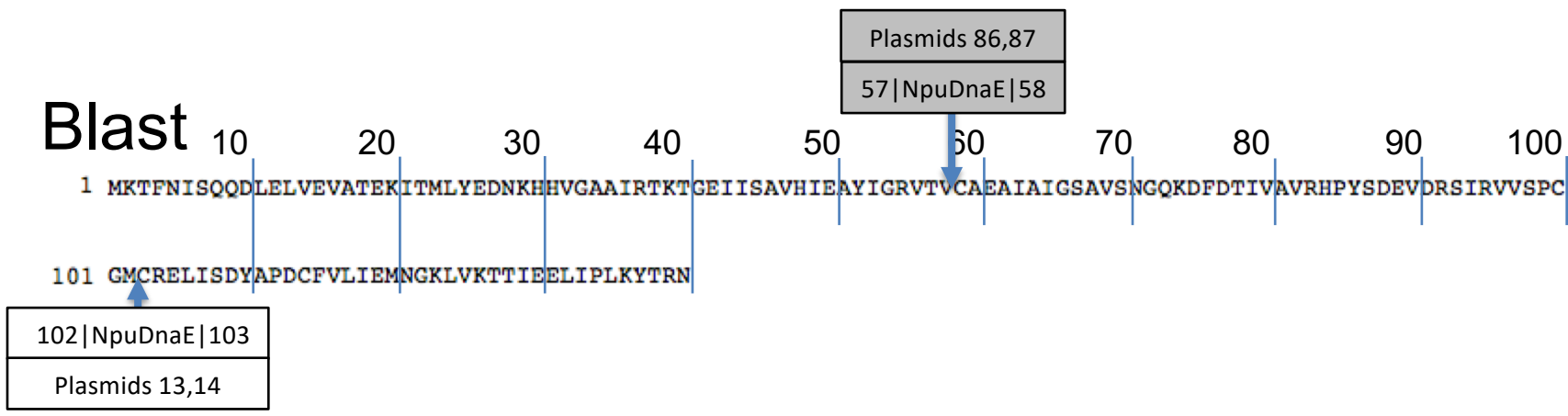
Supplementary Figure 2



Supplementary Figure 3

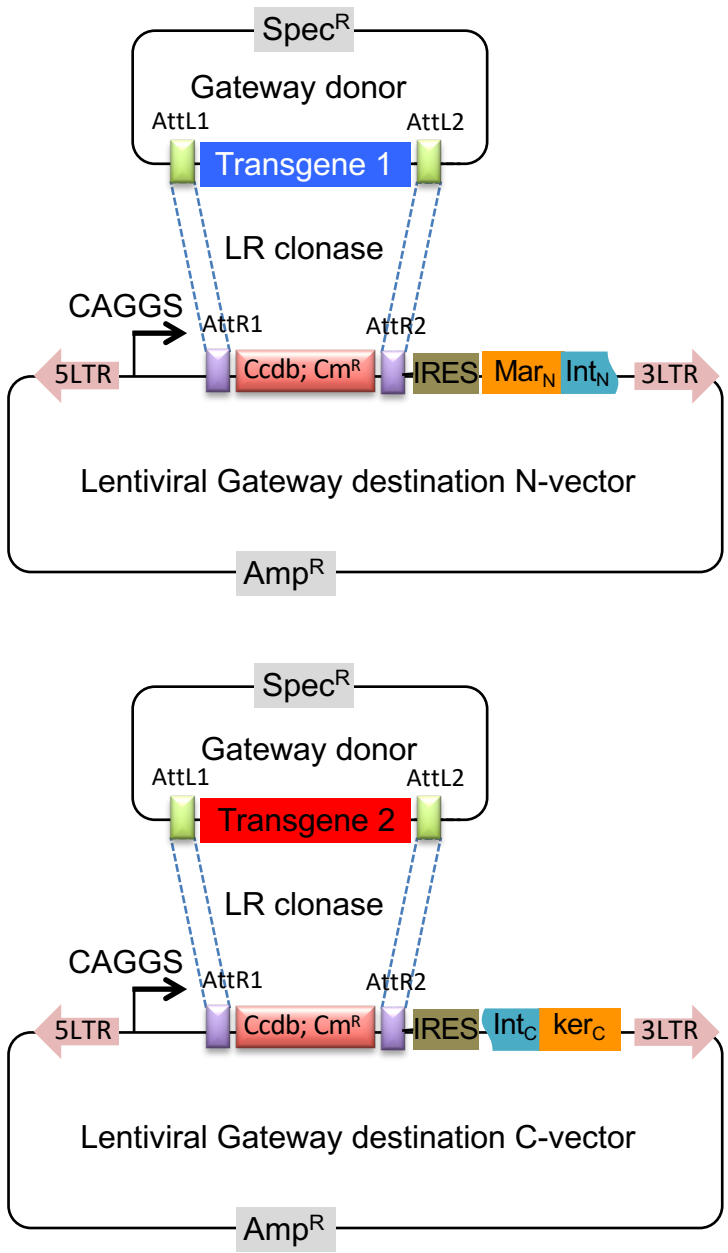


Supplementary Figure 4

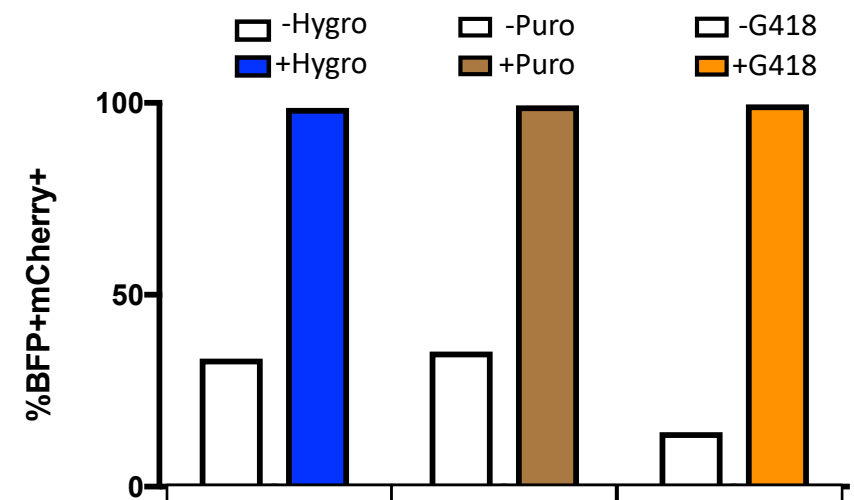


Supplementary Figure 5

a

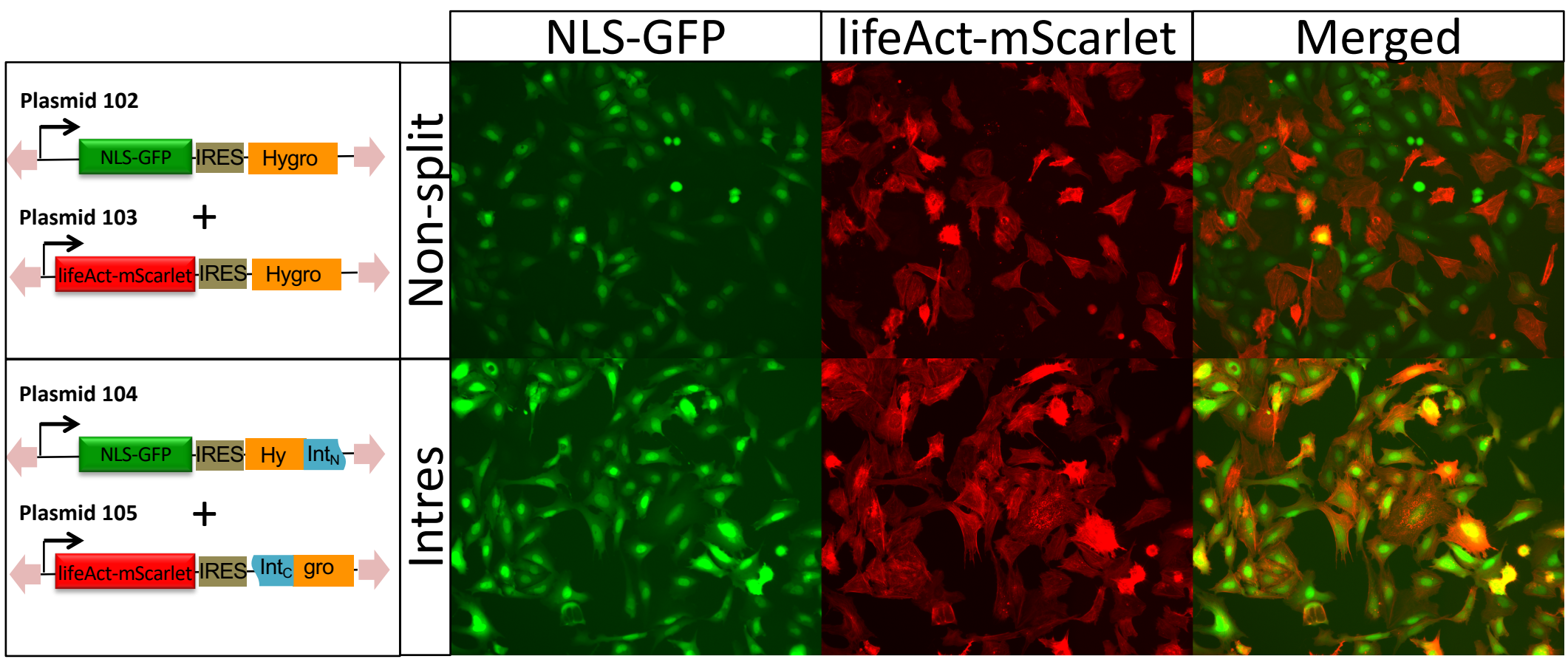


b

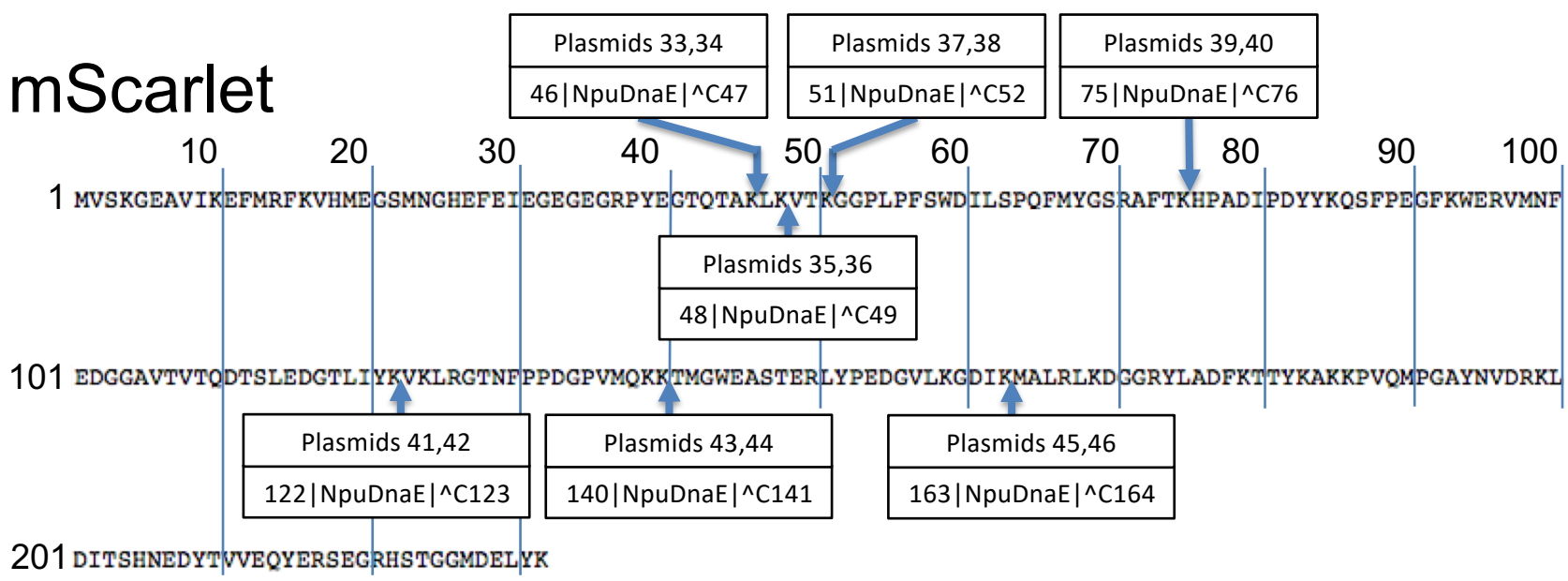


Resistance	Hygro	Puro	Neo
Dest. Plasmids	49,50	53,54	57,58
Final Plasmids	51,52	55,56	59,60

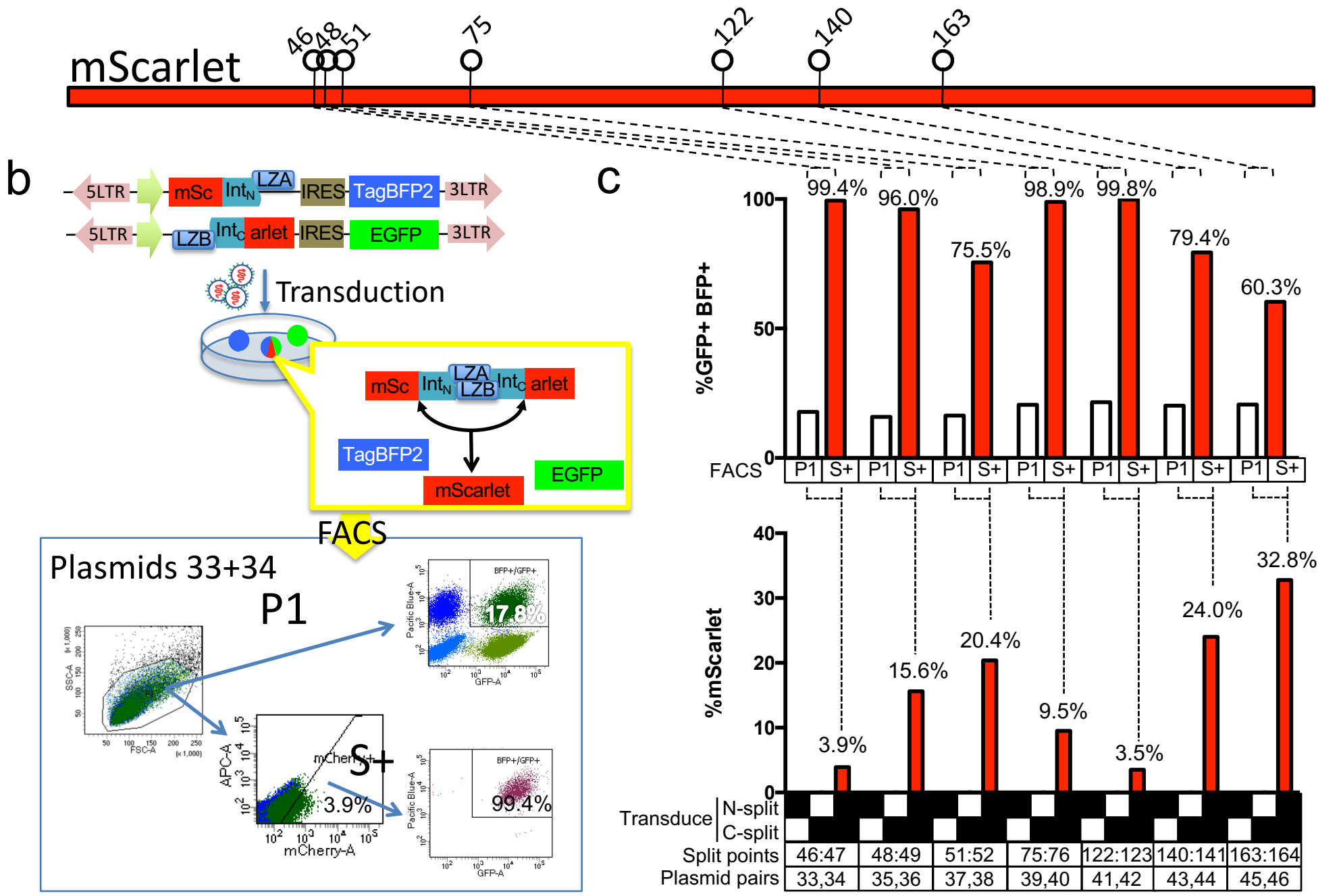
c



Supplementary Figure 6

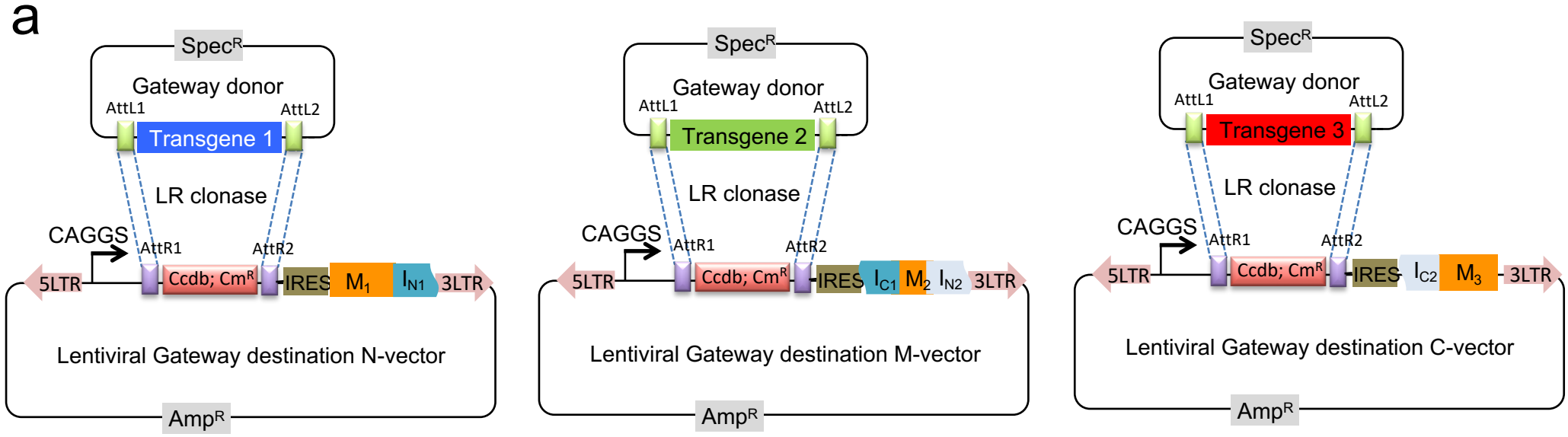


Supplementary Figure 7

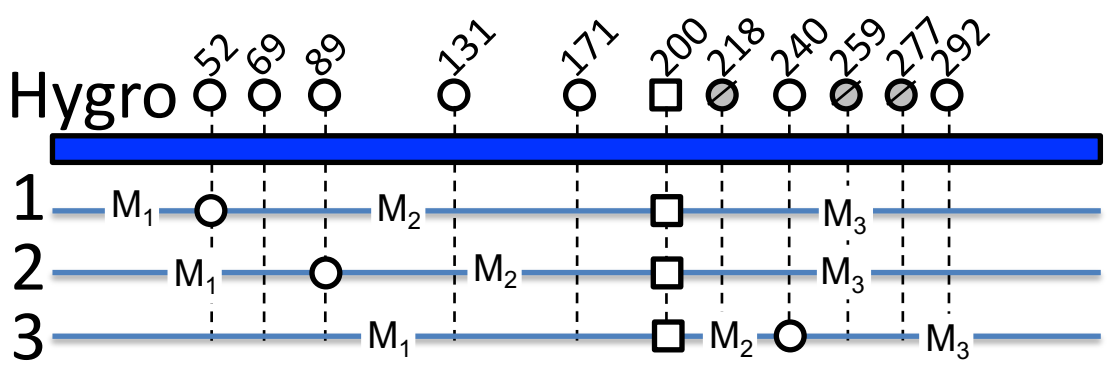


Supplementary Figure 8

a



b

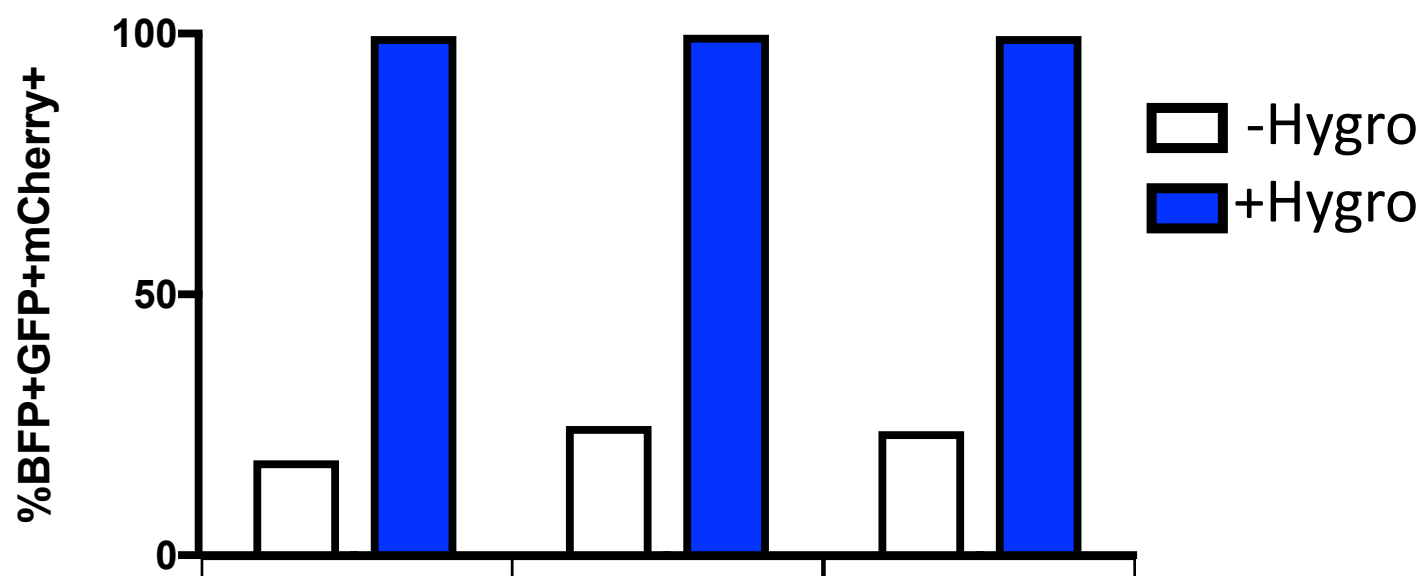


Transgene 1 = TagBFP
Transgene 2 = EGFP
Transgene 3 = mCherry

LR clonase

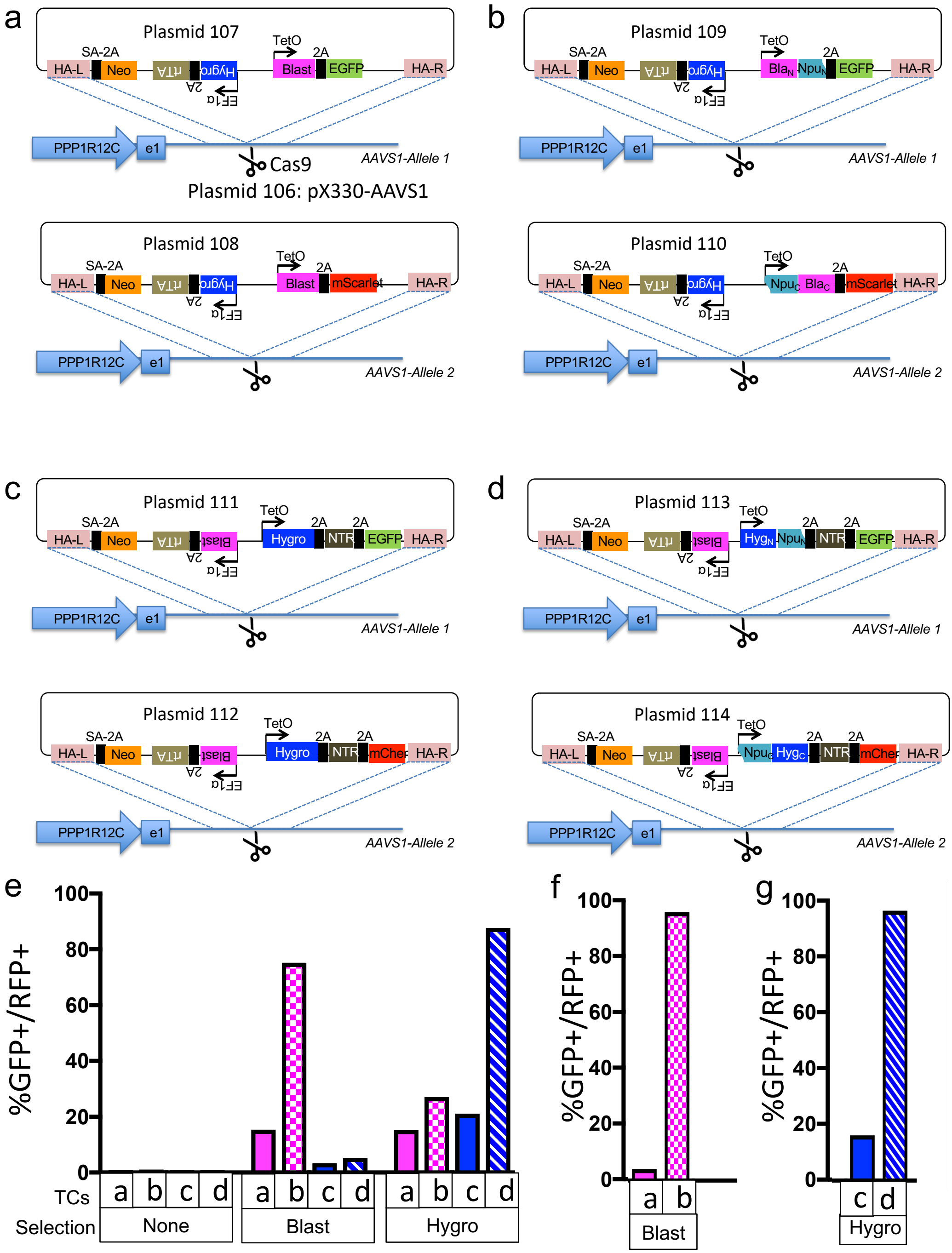
Dest. Plasmids	Final Plasmids
115,116,117	118,119,120
121,122,123	124,125,126
127,128,129	130,131,132

c



Combo	1	2	3
Dest. Plasmids	115,116,117	121,122,123	127,128,129
Final Plasmids	118,119,120	124,125,126	130,131,132

Supplementary Figure 9



Split Selectable Markers

Supplementary Information

Supplementary Figure Legends

Supplementary figure 1. Split points for hygromycin resistance protein. Amino acid sequence of hygromycin resistance protein is presented with clouds labeling the split points characterized in this study. Within the label, the top row indicates the plasmid numbers corresponding to Supplementary Table 1. The bottom row indicates the residue number of the last amino acid in the N-terminal fragment, the species of the intein used, and the residue number of the first amino acid in the C-terminal fragment. “^C” indicates an insertion of a Cysteine.

Supplementary figure 2. Split points for puromycin resistance protein. Amino acid sequence of puromycin resistance protein is presented with clouds labeling the split points characterized in this study. Within the label, the top row indicates the plasmid numbers corresponding to Supplementary Table 1. The bottom row indicates the residue number of the last amino acid in the N-terminal fragment, the species of the intein used, and the residue number of the first amino acid in the C-terminal fragment. “^C” indicates an insertion of a Cysteine.

Supplementary figure 3. Split points for neomycin resistance protein. Amino acid sequence of neomycin resistance gene is presented with clouds labeling the split points characterized in this study. Within the label, the top row indicates the plasmid numbers corresponding to Supplementary Table 1. The bottom row indicates the residue number of the last amino acid in the N-terminal fragment, the species of the intein used, and the residue number of the first amino acid in the C-terminal fragment.

Supplementary figure 4. Split points for blasticidin resistance protein. Amino acid sequence of blasticidin resistance gene is presented with clouds labeling the split points characterized in this study. Within the label, the top row indicates the plasmid numbers corresponding to Supplementary Table 1. The bottom row indicates the residue number of the last amino acid in the N-terminal fragment, the species of the intein used, and the residue number of the first amino acid in the C-terminal fragment.

Supplementary figure 5. Gateway-compatible lentiviral destination vectors with 2-markertron Intres markers. (a) Gateway-compatiable lentiviral destination vector kits for each split Intres marker consists of an N-vector and C-vector. N-vector contains viral LTRs, CAGGS promoter, gateway destination cassette

AttL, ccdB gene, chloramphenicol resistance gene that allows LR clonase-mediated recombination of Gateway donor vector carrying transgenes, followed by internal ribosomal entry site (IRES) that allows polycistronic expression of the N-markertron. Similarly, C-vector contains the C-markertron and allows recombination of another transgene. (b) TagBFP2 (as transgene 1) and mCherry (as transgene 2) were cloned into the 2-markertron Intres plasmids by Gateway recombination and delivered to cells by lentiviral transduction, followed by antibiotic selection and flow cytometry analysis. Column plot shows the percentage of BFP+mCherry+ double-positive cell in the selected culture from the 2-markertron hygromycin (Hygro, blue column), puromycin (Puro, brown column), and neomycin (Neo, orange column) experiments versus their corresponding non-selective cultures (white columns). (c) NLS-GFP (as transgene 1) that labels nucleus with GFP fluorescence and lifeAct-mScarlet (as transgene 2) that labels F-actin with mScarlet fluorescence were recombined into lentiviral vectors expressing full non-split hygromycin resistance gene or lentiviral vectors with 2-markertron hygromycin Intres genes and used to transduce U2OS cells to make dual-label cells. Representative fluorescence microscopic images show GFP, mScarlet and merged channels of cells after hygromycin selection for two weeks.

Supplementary figure 6. Split points for mScarlet fluorescent protein. Amino acid sequence of mScarlet gene is presented with clouds labeling the split points characterized in this study. Within the label, the top row indicates the plasmid numbers corresponding to Supplementary Table 1. The bottom row indicates the residue number of the last amino acid in the N-terminal fragment, the species of the intein used, and the residue number of the first amino acid in the C-terminal fragment. “^C” indicates an insertion of a Cysteine.

Supplementary figure 7. Split mScarlet for fluorescence-mediated co-selection of two separate transgenic vectors. (a) 2-markertron mScarlet proteins. Top schematics shows the split points tested for mScarlet. The last residue of the N-terminal fragment is indicated on top of the lollipops. (b) To screen for NpuDnaE intein-compatible split points for mScarlet, we identified potential split points according to the junctional requirement for *Npu*DnaE intein, then cloned the corresponding N-terminal and C-terminal fragments to the split inteins scaffolds on lentiviral vectors equipped with TagBFP2 or EGFP fluorescent proteins, which serve as our test transgenes to evaluate the selection efficiency. These are delivered into cells via lentiviral transduction. Cells with both lentiviruses contain the necessary protein splicing

machinery and mScarlet fragments to reconstitute the full-length mScarlet fluorescent protein, as well as express both TagBFP2 and EGFP transgenes. Cells were subjected to FACS analysis. Boxed schematic shows an example of FACS analysis of the plasmid pair 33+34. P1 population was gated for forward scatter and side scatter for live single cells. From those, 17.8% of cells are double positive for TagBFP2 and EGFP transgenes. When the P1 cells were further gated for mScarlet-positive (mCherry channel), 99.4% of cells are double positive for TagBFP2 and EGFP. (c) The column plot below shows the percentage of mScarlet-positive cells of each of the indicated split points. The column plot above shows the percentage of TagBFP2+ EGFP+ cells among the P1 cells (white columns) and the mScarlet-positive subset of P1 cells (red columns).

Supplementary figure 8. Gateway-compatible lentiviral destination vectors with 3-markertron hygromycin Intres genes. (a) Gateway-compatiable lentiviral destination vector with viral LTRs, CAGGS promoter, gateway destination cassette AttL, ccdB gene, chloramphenicol resistance gene that allows LR clonase-mediated recombination of Gateway donor vector carrying transgenes, followed by internal ribosomal entry site (IRES) that allows polycistronic expression of the each of the three 3-split hygromycin markertrons. (b) TagBFP2 (as transgene 1) and EGFP (as transgene 2) and mCherry (as transgene 3) were cloned into the 3-split Intres plasmids by Gateway recombination and delivered to cells by lentiviral transduction, followed by antibiotic selection and flow cytometry analysis. (c) Column plot shows the percentage of BFP+GFP+mCherry+ triple-positive cells in the hygromycin selected (blue columns) versus their corresponding non-selective cultures (white columns).

Supplementary figure 9. Intres markers allow enrichment of biallelic targeted cells from CRISPR/Cas-mediated knock-in experiments. Targeting construct pairs containing homology arms for AAVS1 safe harbor locus were designed to contain full length (FL) non-split or split Intres markers and tested for ability to enrich for biallelic targeted cells via antibiotic selection. (a) Plasmids 107 and 108 contains FL Neomycin (Neo) resistance gene driven by endogenous PPP1R12C promoter at the AAVS1 locus, FL Hygromycin (Hygro) gene and rtTA Dox-respsonsive transactivator driven by an EF1a promoter, as well as FL Blasticidin (Blast) expressed as well as EGFP (plasmid 107) and mScarlet (plasmid 108) from a dox-inducible TetO promoter. Plasmid 106 contains Cas9 and an sgRNA targeting the AAVS locus. 2A: self-cleaving 2A peptides. Plasmids 106, 107 and 108 were co-transfected into HEK293T cells, split, and

passaged in dox-containing hygromycin, blasticidin or non-selective media for two weeks, and analyzed by flow cytometry to assay efficiency of biallelic targeting. (b) Plasmids 109 and 110 contain similar structure as Plasmids 107 and 108, but having split Blast Intres in place of the FL Blast. (c) Plasmids 111 and 112 contain an EF1a-driven FL Blast and TetO-driven FL Hygro, nitroreductase (NTR), fluorescent protein (EGFP or mCherry) separated by 2A peptides. (d) Plasmids 113 and 114 are similar to Plasmids 111 and 112 but with Hygro Intres in place of FL Hygro. (e) Flow cytometry analysis of cells transfected with Plasmid 106 (Cas9+AAVS-sgRNA) and the indicated targeting construct pairs, two weeks after culturing in dox-containing non-selective media (Selection: None), blasticidin selection media (Blast) and hygromycin selection media (Hygro). (f) Flow cytometry analysis of cells transfected with Plasmid 106 (Cas9+AAVS-sgRNA) and the indicated targeting construct pairs, four weeks after culturing in dox-containing non-selective media (Selection: None) or blasticidin selection media (Blast) (g) Flow cytometry analysis of cells transfected with Plasmid 106 (Cas9+AAVS-sgRNA) and the indicated targeting construct pairs, four weeks after culturing in dox-containing non-selective media (Selection: None) or hygromycin selection media (Hygro).

Split Selectable Markers

Online Methods

Cloning

To generate a test plasmid for each markertron, we first generated a Gateway donor plasmid containing its ORF and then recombined into lentiviral destination vector with TagBFP2 (Plasmid 94: pLX-DEST-IRES-TagBFP2), EGFP (Plasmid 95: pLX-DEST-IRES-EGFP), or mCherry (Plasmid 96: pLX-DEST-IRES-mCherry) reporters, which were derived from pLX302 (Gift from David Root; Addgene: http://www.addgene.org/25896/) by removing Puromycin resistance gene and inserting IRES-fluorescent genes downstream of the Gateway cassette. The markertron-ORF Gateway donor plasmids were generated either by a nested fusion PCR procedure to combine intein with the coding sequence of fragments of the selectable marker followed by insertion into the pCR8-GW-TOPO plasmid by sequence- and ligation-independent cloning (SLIC) [1], or PCR-amplifying the relevant fragment of the selectable marker followed by insertion into "scaffold" plasmids (Plasmids 27~32) containing the intein sequences by SLIC. DNA sequences encoding inteins were codon optimized for *Homo sapiens,* and synthesized as GBlock (IDT), with AC1947GB encoding *Npu*DnaE intein, AC1949GB encoding *Ssp*DnaB intein. Selectable marker fragments were amplified from plasmids containing these markers. Plasmids created in this study are listed in Supplementary Table 1 with links to Addgene for plasmid sharing and GenBank sequence files.

HEK293T and U2OS Cell Culture

All cells were cultivated in Dulbecco's modified Eagle's medium (DMEM) (Sigma) with 10% fetal bovine serum (FBS)(Lonza), 4% Glutamax (Gibco), 1% Sodium Pyruvate (Gibco) and penicillin-streptomycin (Gibco). Incubator conditions were 37°C and 5% CO2.

Virus Production

A viral packaging mix of pLP1, pLP2, and VSV-G were co-transfected with each lentiviral vector into Lenti-X 293T cells (ClonTech), seeded the day before in 6-well plates at a concentration of 1.2×10^6 cells per well, using Lipofectamine 3000. Media was changed 6h after transfection then incubated overnight.

28 hour post transfection, the media supernatant containing virus was filtered using 45 μM PES filters then stored at -80C until use.

Transduction

The day prior to transduction, U2OS cells were seeded into 12-well plates at a density of $1.5x10^5$ cells per well. Prior to transduction, media was changed to media containing 10 µg/mL polybrene, 1 mL per well. 250 µL of each respective virus (500 µL total for experimental samples with two viruses added) was added to each well and incubated overnight. Media was changed 24-hour post-transduction. Four days post-transduction, cells were split into duplicate plates. Five days post-transduction, media with antibiotics (130µg/mL Hygromycin, 2µg/mL Puromycin, 700µg/mL G418, or 6µg/mL Blasticidin) was added to each respective well of one replicate plate (the other remained under no selection). Antibiotics selection continued for 2 weeks before analysis with flow cytometry.

Flow Cytometry

Cells were trypsinized, suspended in media then analyzed on a LSRFortessa X-20 (BD Bioscience) flow cytometer using FACSDiVa software, version 8, on an HP Z230 workstation. Fifty thousand events were collected each run.

Human iPSC culture and nucleofection

Feeder-free KOLF2-C1 were maintained on plate coated with Synthemax II-SC Substrate (Corning) in StemFlex media (Thermo Fisher Scientific). Subculture was carried out every 4-6 days via Accutase (STEMCELL) detachment method. After plating, 1X RevitaCell supplement (Gibco Life Technologies) was added for 1d to increase cell viability.

For Cas9/gRNA ribonucleoprotein (Cas9/gRNA RNPs) and donor plasmids nucleofection, 4D-Nucleofector™ System (Lonza) was used with the P3 Primary Cell 4D-Nucleofector kit (Lonza). Cells were at 60%-70% confluency at the time of nucleofection. To assemble Cas9/gRNA RNPs, synthetic single guide RNA (Synthego) was resuspended in TE buffer (Synthego) at 2µg/µl, and 8µl of stock solution was mixed with 20µg Cas9 protein before nucleofection. For each reaction, 2×10^6 cells were

collected, resuspended in 100µl complete P3 solution and mixed with pre-assembled Cas9/gRNA RNPs as well as donor plasmids DNA. Doxycycline (Sigma) was added 6d after nucleofection at 5µg/mL. 2d after doxycycline was added, 4µg/mL Blasticidin (Sigma) was applied to select cells with resistance. Surviving single colonies were picked and expanded into Matrigel-coated 24-well plate. If surviving colonies were too large to be manually picked as single colony, cells were replated onto new plate at the density of 2,500 cells per $10cm^2$ plate. Blasticidin treatment continued during single colony expansion in 24-well plates.

For genotyping, genomic DNA was extracted using DNeasy Blood & Tissue Kit (QIAGEN), and PCR was performed using the following primers to identify correctly targeted AAVS1 insertions: (i) EGFP-AAVS (fwd: GCCCGACAACCACTACCTGA, rev: GTGAGTTTGCCAAGCAGTCA), (ii) mScarlet-AAVS (fwd: CTGAGGTCAAGACCACCTACAAG, rev: GTGAGTTTGCCAAGCAGTCA).

References

1. Li, M.Z. & Elledge, S.J. SLIC: a method for sequence-and ligation-independent cloning. *Gene Synthesis: Methods and Protocols*, 51-59 (2012).

Split Selectable Markers

Supplementary Table 1

Plasmids will be available on Addgene. For more information: http://intr.es

Plasmid #*	Plasmid Name	Markertron	Plasmid Filename
3	pLX-Hygro(1-89)-NpuDnaE(N)-IRES-TagBFP2	Hygro(1-89)-NpuDnaE(N)	pLX-[HYGro-NpuDnaE(N)-25-65-85]-IRES-TagBFP2-25-79-47
4	pLX-NpuDnaE(C)-Hygro(90-341)-IRES-mCherry	NpuDnaE(C)-Hygro(90-341)	pLX-[NpuDnaE(C)_hygRO-25-65-86]-IRES-mCherry-25-79-57
5	pLX-Hygro(1-200)-SspDnaB(N)-IRES-TagBFP2	Hygro(1-200)-SspDnaB(N)	pLX-[HYGro(A)-SspDnaB(N)-25-65-72]-IRES-TagBFP2-25-79-42
6	pLX-SspDnaB(C)-Hygro(201-341)-IRES-mCherry	SspDnaB(C)-Hygro(201-341)	pLX-[SspDnaB(C)_hygRO(A)-25-81-23]-IRES-mCherry-25-91-56
7	pLX-Hygro(1-52)-NpuDnaE(N)-IRES-TagBFP2	Hygro(1-52)-NpuDnaE(N)	pLX-[HYGro(B)-NpuDnaE(N)-25-105-15]-IRES-TagBFP2-25-117-16
8	pLX-NpuDnaE(C)-Hygro(53-341)-IRES-mCherry	NpuDnaE(C)-Hygro(53-341)	pLX-[NpuDnaE(C)_hygRO(B)-25-105-16]-IRES-mCherry-25-117-22
9	pLX-Hygro(1-240)-NpuDnaE(N)-IRES-TagBFP2	Hygro(1-240)-NpuDnaE(N)	pLX-[HYGro(C)-NpuDnaE(N)-25-125-46]-IRES-TagBFP2-MD1-73-3
10	pLX-NpuDnaE(C)-Hygro(241-341)-IRES-mCherry	NpuDnaE(C)-Hygro(241-341)	pLX-[NpuDnaE(C)_hygRO(C)-25-105-18]-IRES-mCherry-25-117-23
11	pLX-Hygro(1-292)-NpuDnaE(N)-IRES-TagBFP2	Hygro(1-292)-NpuDnaE(N)	pLX-[HYGro(D)-NpuDnaE(N)[KozakDel3]-25-105-19]-IRES-TagBFP2-25-117-18
12	pLX-NpuDnaE(C)-Hygro(293-341)-IRES-mCherry	NpuDnaE(C)-Hygro(293-341)	pLX-[NpuDnaE(C)_hygRO(D)-25-105-20]-IRES-mCherry-25-117-24
13	pLX-Blast(1-102)-NpuDnaE(N)-IRES-TagBFP2	Blast(1-102)-NpuDnaE(N)	pLX-[Blast(B)-NpuDnaE(N)-25-119-7]-IRES-TagBFP2-MD1-63-6
14	pLX-NpuDnaE(C)-Blast(103-140)-IRES-mCherry	NpuDnaE(C)-Blast(103-140)	pLX-[NpuDnaE(C)_Blast(B)-25-119-8]-IRES-mCherry-MD1-63-9
17	pLX-Puro(1-119)-NpuDnaE(N)-IRES-TagBFP2	Puro(1-119)-NpuDnaE(N)	pLX-[PuroKC2(N)-NpuDnaE(N)-25-131-28"]-IRES-TagBFP2-25-133-5
18	pLX-NpuDnaE(C)-Puro(insCys;120-199)-IRES-mCherry	NpuDnaE(C)-Puro(insCys;120-199)	pLX-[NpuDnaE(C)_PuroKC2(C)-25-131-35"]-IRES-mCherry-25-133-13
19	pLX-Puro(1-100)-SspDnaB(N-S0)-IRES-TagBFP2	Puro(1-100)-SspDnaB(N-S0)	pLX-[Puro(A)_SspDnaB(N-S0)-26-3-37]-IRES-TagBFP2-26-9-14
20	pLX-SspDnaB(C-S0)-Puro(101-199)-IRES-mCherry	SspDnaB(C-S0)-Puro(101-199)	pLX-[SspDnaB(C-S0)_Puro(A)-26-1-19]-IRES-mCherry-26-7-26
21	pLX-Neo(1-133)-NpuDnaE(N)-IRES-TagBFP2	Neo(1-133)-NpuDnaE(N)	pLX-[Neo(A)-NpuDnaE(N)-25-125-1]-IRES-TagBFP2-MD1-73-1

22	pLX-NpuDnaE(C)-Neo(134-267)-IRES-mCherry	NpuDnaE(C)-Neo(134-267)	pLX-[NpuDnaE(C)_Neo(A)-25-119-2]-IRES-mCherry-F79-4
23	pLX-Neo(1-194)-NpuDnaE(N)-IRES-TagBFP2	Neo(1-194)-NpuDnaE(N)	pLX-[Neo(B)-NpuDnaE(N)-25-125-4]-IRES-TagBFP2-MD1-73-2
24	pLX-NpuDnaE(C)-Neo(195-267)-IRES-mCherry	NpuDnaE(C)-Neo(195-267)	pLX-[NpuDnaE(C)_Neo(B)-25-125-6]-IRES-mCherry-MD1-73-4
25	pLX-NpuDnaE(C)_Hygro(53-89)-NpuDnaE(N)-IRES-GFP	NpuDnaE(C)_Hygro(53-89)-NpuDnaE(N)	pLX-[NpuDnaE(C)_HygroBA_NpuDnaE(N)-25-155-79]-IRES-GFP-26-3-21
26	pLX-NpuDnaE(C)_Hygro(53-239)-NpuDnaE(N)-IRES-GFP	NpuDnaE(C)_Hygro(53-239)-NpuDnaE(N)	pLX-[NpuDnaE(C)_HygroAC_NpuDnaE(N)-25-155-83]-IRES-GFP-26-3-22
27			pCR8-BsaI->ccdbCam<-BsaI-NpuDnaE(N)-MD1-68-15
28			pCR8-NpuDnaE(C)_BsaI->ccdbCam<-BsaI-MD1-68-18
31			pCR8-BsaI->ccdbCam<-BsaI-SspDnaB(N-S0)-25-135-18
32			pCR8-SspDnaB(C-S0)_BsaI->ccdbCam<-BsaI-25-155-41
33	pLX-mScarlet(1-46)-NpuDnaE(N)_LZA-IRES-TagBFP2	mScarlet(1-46)-NpuDnaE(N)_LZA	pLX-[mScarlet(1-46)-NpuDnaE(N)_LZA-26-103-38]-IRES-TagBFP2-26-107-16
34	pLX-LZB_NpuDnaE(C)-mScarlet(insCys;47-232)-IRES-TagBFP2	LZB_NpuDnaE(C)-mScarlet(insCys;47-232)	pLX-[LZB_NpuDnaE(C)-mScarlet(insCys;47-232)-26-103-59]-IRES-GFP-26-107-33
35	pLX-mScarlet(1-48)-NpuDnaE(N)_LZA-IRES-TagBFP2	mScarlet(1-48)-NpuDnaE(N)_LZA	pLX-[mScarlet(1-48)-NpuDnaE(N)_LZA-26-103-39]-IRES-TagBFP2-26-107-17
36	pLX-LZB_NpuDnaE(C)-mScarlet(insCys;49-232)-IRES-GFP	LZB_NpuDnaE(C)-mScarlet(insCys;49-232)	pLX-[60-pCR8-LZB_NpuDnaE(C)-mScarlet(insCys;49-232)]-IRES-GFP-26-107-34
37	pLX-mScarlet(1-51)-NpuDnaE(N)_LZA -IRES-TagBFP2	mScarlet(1-51)-NpuDnaE(N)_LZA	pLX-[mScarlet(1-51)-NpuDnaE(N)_LZA-26-103-40]-IRES-TagBFP2-26-107-18
38	pLX-LZB_NpuDnaE(C)-mScarlet(insCys;52-232)-IRES-GFP	LZB_NpuDnaE(C)-mScarlet(insCys;52-232)	pLX-[LZB_NpuDnaE(C)-mScarlet(insCys;52-232)-26-103-61]-IRES-GFP-26-107-35
39	pLX-mScarlet(1-75)-NpuDnaE(N)_LZA-IRES-TagBFP2	mScarlet(1-75)-NpuDnaE(N)_LZA	pLX-[mScarlet(1-75)-NpuDnaE(N)_LZA-26-103-41]-IRES-TagBFP2-26-107-19
40	pLX-LZB_NpuDnaE(C)-mScarlet(insCys;76-232)-IRES-GFP	LZB_NpuDnaE(C)-mScarlet(insCys;76-232)	pLX-[LZB_NpuDnaE(C)-mScarlet(insCys;76-232)-26-103-62]-IRES-GFP-26-107-36
41	pLX-mScarlet(1-122)-NpuDnaE(N)_LZA-IRES-TagBFP2	mScarlet(1-122)-NpuDnaE(N)_LZA	pLX-[mScarlet(1-122)-NpuDnaE(N)_LZA-26-103-44]-IRES-TagBFP2-26-107-22
42	pLX-LZB_NpuDnaE(C)-mScarlet(insCys;123-232)-IRES-GFP	LZB_NpuDnaE(C)-mScarlet(insCys;123-232)	pLX-[LZB_NpuDnaE(C)-mScarlet(insCys;123-232)-26-103-65]-IRES-GFP-26-107-39
43	pLX-mScarlet(1-140)-NpuDnaE(N)_LZA-IRES-TagBFP2	mScarlet(1-140)-NpuDnaE(N)_LZA	pLX-[24-pCR8-mScarlet(1-140)-NpuDnaE(N)_LZA]-IRES-TagBFP2-26-107-24
44	pLX-LZB_NpuDnaE(C)-mScarlet(insCys;141-232)-IRES-GFP	LZB_NpuDnaE(C)-mScarlet(insCys;141-232)	pLX-[LZB_NpuDnaE(C)-mScarlet(insCys;141-232)-26-103-66]-IRES-GFP-26-107-40
45	pLX-mScarlet(1-163)-NpuDnaE(N)_LZA-IRES-TagBFP2	mScarlet(1-163)-NpuDnaE(N)_LZA	pLX-[48-pCR8-mScarlet(1-163)-NpuDnaE(N)_LZA]-IRES-TagBFP2-26-107-25
46	pLX-LZB_NpuDnaE(C)-mScarlet(insCys;164-232)-IRES-GFP	LZB_NpuDnaE(C)-mScarlet(insCys;164-232)	pLX-[LZB_NpuDnaE(C)-mScarlet(insCys;164-232)-26-103-68]-IRES-GFP-26-107-41
47	pCR8-TagBFP2	TagBFP2	pCR8-TagBFP2-25-119-29

48	pCR8-mCherry	mCherry	pCR8-mCherry-25-121-14
49	pLX-DEST-IRES-Hygro(1-89)-NpuDnaE(N)	Hygro(1-89)-NpuDnaE(N)	pLX-DEST-IRES-HYGro-NpuDnaE(N)-25-105-7
50	pLX-DEST-IRES-NpuDnaE(C)-Hygro(90-341)	NpuDnaE(C)-Hygro(90-341)	pLX-DEST-IRES-NpuDnaE(C)_hygRO-25-105-9
51	pLX-[TagBFP2]-IRES-Hygro(1-89)-NpuDnaE(N)	Hygro(1-89)-NpuDnaE(N)	pLX-[TagBFP2-25-119-29]-IRES-HYGro-NpuDnaE(N)-MD1-68-1
52	pLX-[mCherry]-IRES-NpuDnaE(C)-Hygro(90-341)	NpuDnaE(C)-Hygro(90-341)	pLX-[mCherry-25-121-14]-IRES-NpuDnaE(C)_hygRO-MD1-68-3
53	pLX-DEST-IRES-Puro(1-119)-NpuDnaE(N)	Puro(1-119)-NpuDnaE(N)	pLX-DEST-IRES-PuroKC2(N)-NpuDnaE(N)-26-113-44
54	pLX-DEST-IRES-NpuDnaE(C)-Puro(120-199)	NpuDnaE(C)-Puro(insCys;120-199)	pLX-DEST-IRES-NpuDnaE(C)_PuroKC2(C)-26-23-75
55	pLX-[TagBFP2]-IRES-Puro(1-119)-NpuDnaE(N)	Puro(1-119)-NpuDnaE(N)	pLX-[TagBFP2-25-119-29]-IRES-PuroKC2(N)-NpuDnaE(N)-26-121-18
56	pLX-[mCherry]-IRES-NpuDnaE(C)-Puro(120-199)	NpuDnaE(C)-Puro(insCys;120-199)	pLX-[mCherry-25-121-14]-IRES-NpuDnaE(C)_PuroKC2(C)-26-107-60
57	pLX-DEST-IRES-Neo(1-194)-NpuDnaE(N)	Neo(1-194)-NpuDnaE(N)	pLX-DEST-IRES-Neo(B)-NpuDnaE(N)-26-103-6
58	pLX-DEST-IRES-NpuDnaE(C)-Neo(195-267)	NpuDnaE(C)-Neo(195-267)	pLX-DEST-IRES-NpuDnaE(C)_Neo(B)-26-45-8
59	pLX-[TagBFP2]-IRES-Neo(1-194)-NpuDnaE(N)	Neo(1-194)-NpuDnaE(N)	pLX-[TagBFP2-25-119-29]-IRES-Neo(B)-NpuDnaE(N)-26-107-55
60	pLX-[mCherry]-IRES-NpuDnaE(C)-Neo(195-267)	NpuDnaE(C)-Neo(195-267)	pLX-[mCherry-25-121-14]-IRES-NpuDnaE(C)_Neo(B)-26-107-56
64	pLX-Hygro(1-69)-NpuDnaE(N)-IRES-TagBFP2	Hygro(1-69)-NpuDnaE(N)	pLX-[Hygro69-NpuDnaE(N)-26-61-17]-IRES-TagBFP2-26-69-5
65	pLX-NpuDnaE(C)-Hygro(^C;70-341)-IRES-mCherry	NpuDnaE(C)-Hygro(^C;70-341)	pLX-[NpuDnaE(C)_HygroSC69-26-61-30]-IRES-mCherry-26-69-21
66	pLX-Hygro(1-131)-NpuDnaE(N)-IRES-TagBFP2	Hygro(1-131)-NpuDnaE(N)	pLX-[Hygro131-NpuDnaE(N)-26-61-18]-IRES-TagBFP2-26-69-6
67	pLX-NpuDnaE(C)-Hygro(^C;132-341)-IRES-mCherry	NpuDnaE(C)-Hygro(^C;132-341)	pLX-[NpuDnaE(C)_HygroSC131-26-61-31]-IRES-mCherry-26-69-22
68	pLX-Hygro(1-171)-NpuDnaE(N)-IRES-TagBFP2	Hygro(1-171)-NpuDnaE(N)	pLX-[Hygro171-NpuDnaE(N)-26-61-19]-IRES-TagBFP2-26-69-7
69	pLX-NpuDnaE(C)-Hygro(^C;172-341)-IRES-mCherry	NpuDnaE(C)-Hygro(^C;172-341)	pLX-[NpuDnaE(C)_HygroSC171-26-61-32]-IRES-mCherry-26-69-23
70	pLX-Hygro(1-218)-NpuDnaE(N)-IRES-TagBFP2	Hygro(1-218)-NpuDnaE(N)	pLX-[Hygro218-NpuDnaE(N)-26-61-20]-IRES-TagBFP2-26-69-8
71	pLX-NpuDnaE(C)-Hygro(^C;219-341)-IRES-mCherry	NpuDnaE(C)-Hygro(^C;219-341)	pLX-[NpuDnaE(C)_HygroSC218-26-61-33]-IRES-mCherry-26-69-24
72	pLX-Hygro(1-259)-NpuDnaE(N)-IRES-TagBFP2	Hygro(1-259)-NpuDnaE(N)	pLX-[Hygro259-NpuDnaE(N)-26-65-45]-IRES-TagBFP2-26-69-9
73	pLX-NpuDnaE(C)-Hygro(^C;260-341)-IRES-mCherry	NpuDnaE(C)-Hygro(^C;260-341)	pLX-[NpuDnaE(C)_HygroSC259-26-61-34]-IRES-mCherry-26-69-25
74	pLX-Hygro(1-277)-NpuDnaE(N)-IRES-TagBFP2	Hygro(1-277)-NpuDnaE(N)	pLX-[Hygro277-NpuDnaE(N)-26-65-46]-IRES-TagBFP2-26-69-10
75	pLX-NpuDnaE(C)-Hygro(^C; 278-341)-IRES-mCherry	NpuDnaE(C)-Hygro(^C;278-341)	pLX-[NpuDnaE(C)_HygroSC277-26-61-35]-IRES-mCherry-26-69-26
76	pLX-Puro(1-32)-NpuDnaE(N)-IRES-TagBFP2	Puro(1-32)-NpuDnaE(N)	pLX-[PuroAC32-NpuDnaE(N)-26-61-24]-IRES-TagBFP2-26-69-12

77	pLX-NpuDnaE(C)-Puro(^C;33-199)-IRES-mCherry	NpuDnaE(C)-Puro(^C;33-199)	pLX-[NpuDnaE(C)_PuroAC32-26-61-37]-IRES-mCherry-26-69-28
78	pLX-Puro(1-84)-NpuDnaE(N)-IRES-TagBFP2	Puro(1-84)-NpuDnaE(N)	pLX-[PuroAC84-NpuDnaE(N)-26-61-25]-IRES-TagBFP2-26-69-13
79	pLX-NpuDnaE(C)-Puro(^C;85-199)-IRES-mCherry	NpuDnaE(C)-Puro(^C;85-199)	pLX-[NpuDnaE(C)_PuroAC84-26-61-38]-IRES-mCherry-26-69-29
80	pLX-Puro(1-137)-NpuDnaE(N)-IRES-TagBFP2	Puro(1-137)-NpuDnaE(N)	pLX-[PuroKC3(N)-NpuDnaE(N)-25-131-29"]-IRES-TagBFP2-25-133-6
81	pLX-NpuDnaE(C)-Puro(^C;138-199)-IRES-mCherry	NpuDnaE(C)-Puro(^C;138-199)	pLX-[NpuDnaE(C)_PuroKC3(C)-25-131-36"]-IRES-mCherry-25-133-14
82	pLX-Puro(1-158)-NpuDnaE(N)-IRES-TagBFP2	Puro(1-158)-NpuDnaE(N)	pLX-[PuroAC158-NpuDnaE(N)-26-61-26]-IRES-TagBFP2-26-69-14
83	pLX-NpuDnaE(C)-Puro(^C;159-199)-IRES-mCherry	NpuDnaE(C)-Puro(^C;159-199)	pLX-[NpuDnaE(C)_PuroAC158-26-61-39]-IRES-mCherry-26-69-30
84	pLX-Puro(1-180)-NpuDnaE(N)-IRES-TagBFP2	Puro(1-180)-NpuDnaE(N)	pLX-[PuroAC180-NpuDnaE(N)-26-61-27]-IRES-TagBFP2-26-69-15
85	pLX-NpuDnaE(C)-Puro(^C;181-199)-IRES-mCherry	NpuDnaE(C)-Puro(^C;181-199)	pLX-[NpuDnaE(C)_PuroAC180-26-61-40]-IRES-mCherry-26-69-31
86	pLX-Blast(1-58)-NpuDnaE(N)-IRES-TagBFP2	Blast(1-58)-NpuDnaE(N)	pLX-[Blast(A)-NpuDnaE(N)-25-119-5]-IRES-TagBFP2-MD1-63-5
87	pLX-NpuDnaE(C)-Blast(59-140)-IRES-mCherry	NpuDnaE(C)-Blast(59-140)	pLX-[NpuDnaE(C)_Blast(A)-25-119-6]-IRES-mCherry-MD1-63-8
88	pLX-NpuDnaE(C)-HygroBA-SspDnaB(N-S0)-IRES-EGFP	NpuDnaE(C)-Hygro(53-200)-SspDnaB(N-S0)	pLX-[NpuDnaE(C)_HygroBA_SspDnaB(N-S0)-26-37-59]-IRES-GFP-26-43-59
89	pLX-SspDnaB(C-S0)-Hygro(201-341)-IRES-mCherry	SspDnaB(C-S0)-Hygro(201-341)	pLX-[SspDnaB(C-S0)_Hygro(A)-26-1-17]-IRES-mCherry-26-7-24
90	pLX-NpuDnaE(C)-Hygro(90-200)-SspDnaB(N-S0)-IRES-EGFP	NpuDnaE(C)-Hygro(90-200)-SspDnaB(N-S0)	pLX-[NpuDnaE(C)_HygroAA_SspDnaB(N-S0)-26-37-60]-IRES-GFP-26-43-60
91	pLX-Hygro(1-200)-SspDnaB(N-S0)-IRES-TagBFP2	Hygro(1-200)-SspDnaB(N-S0)	pLX-[Hygro(A)_SspDnaB(N-S0)-26-1-10]-IRES-TagBFP2-26-7-15
92	pLX-SspDnaB(C-S0)-Hygro(201-240)-NpuDnaE(N)-IRES-EGFP	SspDnaB(C-S0)-Hygro(201-240)-NpuDnaE(N)	pLX-[SspDnaB(C-S0)_HygroAC_NpuDnaE(N)-26-37-61]-IRES-GFP-26-43-61
93	pLX-SspDnaB(C-S0)-Hygro(201-292)-NpuDnaE(N)-IRES-EGFP	SspDnaB(C-S0)-Hygro(201-292)-NpuDnaE(N)	pLX-[SspDnaB(C-S0)_HygroAD_NpuDnaE(N)-26-37-62]-IRES-GFP-26-43-62
94	pLX-DEST-IRES-TagBFP2		pLX-DEST-IRES-TagBFP2-25-65-63,25-131-25
95	pLX-DEST-IRES-EGFP		pLX-DEST-IRES-GFP-25-81-21
96	pLX-DEST-IRES-mCherry		pLX-DEST-IRES-mCherry-25-65-64,25-131-26
97	pLX-Hygro-IRES-TagBFP2	Non-split Hygro	pLX-[Hygro-24-127-20]-IRES-TagBFP2-25-91-23
98	pLX-Hygro-IRES-mCherry	Non-split Hygro	pLX-[Hygro-24-127-20]-IRES-mCherry-25-117-21
99	pLX-Puro-IRES-TagBFP2	Non-split Puro	pLX-[Puro-24-127-21]-IRES-TagBFP2-25-117-13
100	pLX-Puro-IRES-mCherry	Non-split Puro	pLX-[Puro-24-127-21]-IRES-mCherry-25-91-24
101	pLX-Hygro-IRES-EGFP	Non-split Hygro	pLX-[Hygro-24-127-20]-IRES-GFP-27-61-1
102	pLX-NLS_GFP-IRES-Hygro	Non-split Hygro	pLX-[NLSGFP-6-29-B1]-IRES-Hygro-28-73-13

103	pLX-LifeAct_mCherry-IRES-Hygro	Non-split Hygro	pLX-[LifeAct_mScarlet_N1-26-69-64]-IRES-Hygro-27-131-30
104	pLX-NLS_GFP-IRES-Hygro(1-89)-NpuDnaE(N)	Hygro(1-89)-NpuDnaE(N)	pLX-[NLSGFP-6-29-B1]-IRES-HYGro-NpuDnaE(N)-28-73-11
105	pLX-LifeAct_mScarlet-IRES- NpuDnaE(C)-Hygro(90-341)	NpuDnaE(C)-Hygro(90-341)	pLX-[LifeAct_mScarlet_N1-26-69-64]-IRES-NpuDnaE(C)_hygRO-28-73-12
106	pX330-AAVS1		PX330-[sg377-AAVTargeting]-26-63-13
107	pAAVS1-Nst-EF1aHygro2ArtTA3(-)_TetO-Blast-P2A-EGFP	Non-split Blast	pAAVS1-Nst-EF1aHygro2ArtTA3(-)_pCW-[Blast-P2A-EGFP-28-33-46]-28-123-12
108	pAAVS1-Nst-EF1aHygro2ArtTA3(-)_TetO-Blast-P2A-mScarlet	Non-split Blast	pAAVS1-Nst-EF1aHygro2ArtTA3(-)_pCW-[Blast-P2A-mScarlet-28-33-48]-28-123-13
109	pAAVS1-Nst-EF1aHygro2ArtTA3(-)_TetO-Blast(1-102)_NpuDnaE(N)-P2A-EGFP	Blast(1-102)-NpuDnaE(N)	pAAVS1-Nst-EF1aHygro2ArtTA3(-)_pCW-[Blast(B)_NpuDnaE(N)-P2A-EGFP-28-33-47]-28-123-14"
110	pAAVS1-Nst-EF1aHygro2ArtTA3(-)_TetO-NpuDnaE(C)_Blast(103-140)-P2A-mScarlet	NpuDnaE(C)-Blast(103-140)	pAAVS1-Nst-EF1aHygro2ArtTA3(-)_pCW-[NpuDnaE(C)_Blast(B)-P2A-mScarlet-28-33-49]-28-123-15
111	pAAVS1-Nst-EF1aBlast2ArtTA3(-)_TetO-Hygro-P2A-NTR-E2A-EGFP	Non-split Hygro	pAAVS1-Nst-EF1aBlast2ArtTA3(-)_pCW-[Hygro-P2A-NTR-E2A-EGFP-26-21-33]-28-123-16
112	pAAVS1-Nst-EF1aBlast2ArtTA3(-)_TetO-Hygro-P2A-NTR-E2A-mCherry	Non-split Hygro	pAAVS1-Nst-EF1aBlast2ArtTA3(-)_pCW-[Hygro-P2A-NTR-E2A-mCherry-26-37-81]-28-123-17
113	pAAVS1-Nst-EF1aBlast2ArtTA3(-)_TetO-Hygro(1-89)-NpuDnaE(N)-P2A-NTR-E2A-EGFP	Hygro(1-89)-NpuDnaE(N)	pAAVS1-Nst-EF1aBlast2ArtTA3(-)_pCW-[Hygro(A)_NpuDnaE(N)-P2A-NTR-E2A-EGFP-26-21-31]-28-123-18
114	pAAVS1-Nst-EF1aBlast2ArtTA3(-)_TetO-NpuDnaE(C)-Hygro(90-341)-P2A-NTR-E2A-mCherry	NpuDnaE(C)-Hygro(90-341)	pAAVS1-Nst-EF1aBlast2ArtTA3(-)_pCW-[NpuDnaE(C)_Hygro(A)-P2A-NTR-E2A-mCherry-26-21-55]-28-123-19
115	pLX-Hygro(1-89)_NpuDnaE(N)_LZA-IRES-TagBFP2	Hygro(1-89)-NpuDnaE(N)-LZA	pLX-[Hygro(A)-NpuDnaE(N)_LZA-26-87-79]-IRES-TagBFP2-26-91-10
116	pLX-LZB_NpuDnaGEP(C)_Hygro(90-200)_SspDnaB(N-S0)-IRES-GFP	LZB-NpuDnaGEP(C)-Hygro(90-200)-SspDnaB(N-S0)	pLX-[LZB-NpuDnaGEP(C)-Hygro(90-200)-SspDnaB(N-S0)-27-93-56]-IRES-GFP-27-97-1
117	pLX-SspDnaB(C-S0)_Hygro(201-240)_NpuDnaE(N)_LZA-IRES-GFP	SspDnaB(C-S0)-Hygro(201-240)-NpuDnaE(N)-LZA	pLX-[SspDnaB(C-S0)_HygroAC_NpuDnaE(N)-LZA-27-93-57]-IRES-GFP-27-97-2
118	pLX-LZB_NpuDnaGEP(C)_Hygro(241-341)-IRES-mCherry	LZB-NpuDnaGEP(C)-Hygro(241-341)	pLX-[LZB_NpuDnaGEP(C)_Hygro(C)-26-141-42]-IRES-mCherry-26-145-22