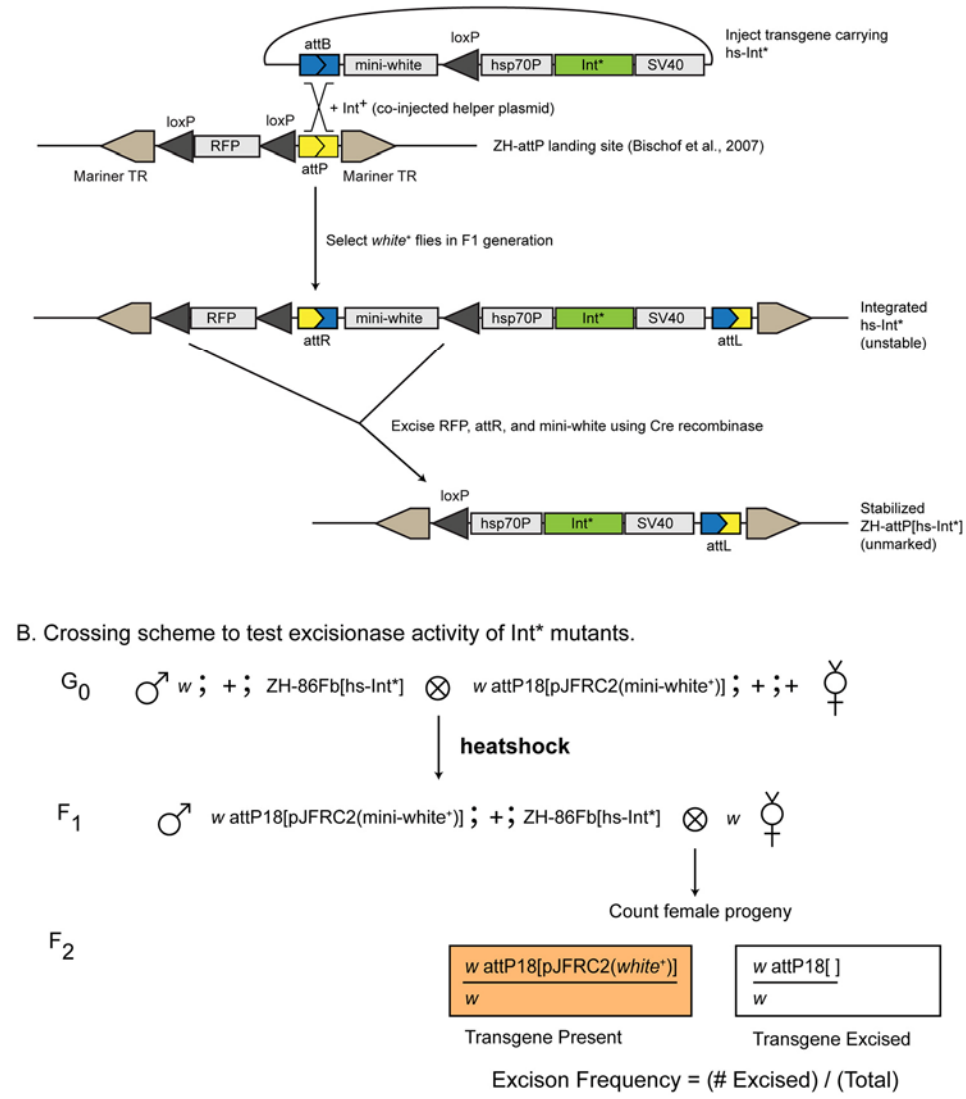


Figure S1. Testing Int* variants for excisionase activity *in vivo*.

Figure S1 Testing Int* variants for excisionase activity *in vivo.* (A) Strategy to stably and site-specifically introduce Int* into the genome: Int* transgene constructs bearing a single loxP site (5′ to hs-Int*) were integrated into the landing sites *ZH-51C* or *ZH-86Fb* (Bischof *et al.* 2007), and integrants were identified by mini-white expression. To prevent Int* from excising itself, sequences between the upstream-most loxP in the landing site and the loxP in the integrated transgene were eliminated by Cre recombinase. The removal of RFP and mini-white leaves the *hs-Int** transgene unmarked, to avoid interference with downstream applications. (B) Crossing scheme to test Int* variants for excisionase activity: Virgins bearing *attP18[JFRC2]* were crossed to males bearing *hs-Int**, and progeny were heat shocked for one hour during the third larval instar. Following eclosion, individual males were crossed to *white* virgins, and the numbers of female progeny that were *mini-white*$^{+}$ (transgene present) and *white* (transgene excised) were counted. The frequency of transgene excision was determined by comparing the number of *white* female progeny to the total number of female progeny. Also see Figure 1D and Table S1.

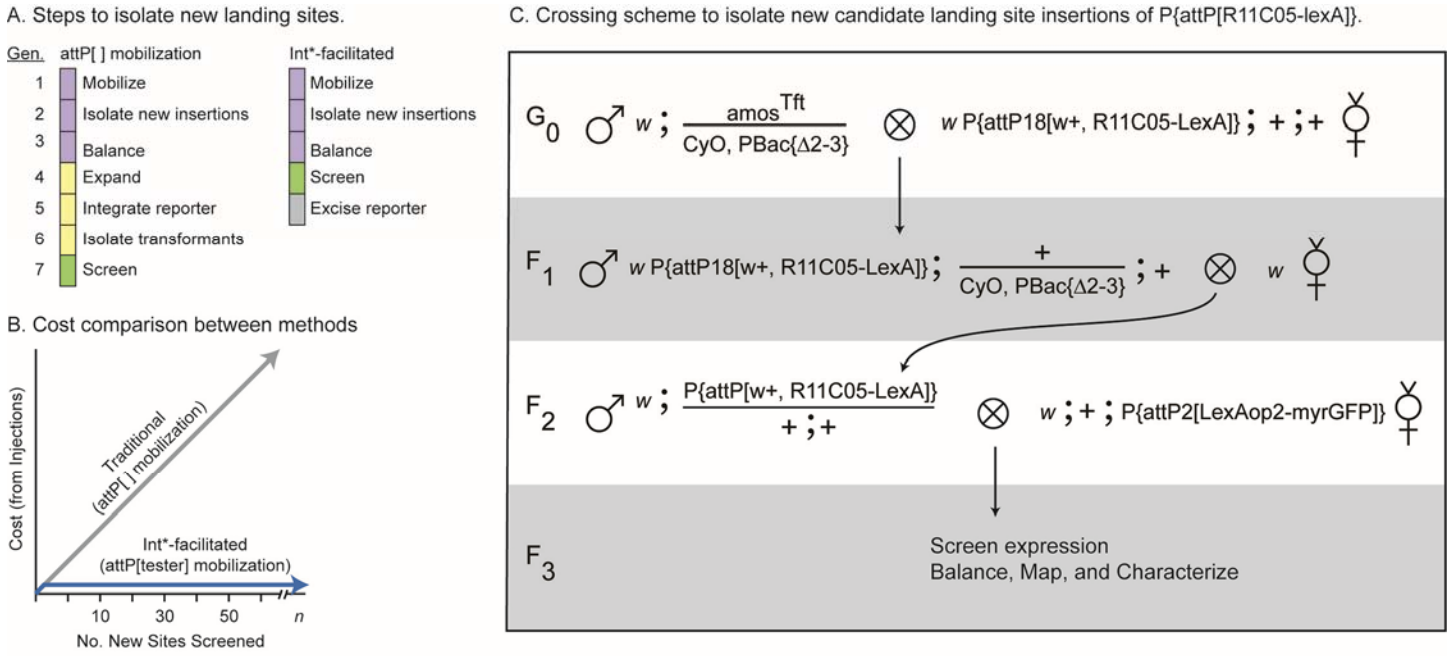


Figure S2. Characterization of Int*-facilitated landing site isolation.

Figure S2 Characterization of Int*-facilitated landing site isolation. (A) Comparison of steps required to isolate new landing sites via traditional or Int*-facilitated methods: For either method, the first three steps are identical. However, the Int*-facilitated method allows new candidate landing sites to be screened as early as the fourth generation. This eliminates the need to maintain and expand lines that will later be rejected, thereby significantly reducing the amount of fly work required to characterize new candidate landing sites. (B) Illustration of the theoretical cost differential for assessing *n* new landing sites by traditional vs. Int*-facilitated methods: Because the traditional method requires a separate injection for each site tested, total injection cost increases linearly with the number of sites assessed. In contrast, Int*-facilitated landing site isolation requires no injections once the tester transgene has been integrated, resulting in a flat cost curve. (C) Crossing scheme to isolate new, autosomal candidate landing site insertions: To mobilize *P{CaryP}attP18[R11C05-lexA]* off the *X* chromosome, males bearing P transposase on *CyO* (*PBac{Δ2-3}*; BL#8201) were crossed to virgins homozygous for *P{CaryP}attP18[R11C05-LexA]* to produce dysgenic males, which were crossed to w^{1118} virgins. The Cy^{+}, *mini-white*$^{+}$ male progeny of this cross represent new insertions of the P-element. (Note: the mobile element can be followed by the mini-white associated with the R11C05-lexA tester or by the mini-yellow allele that is part of the P{CaryP}attP[ ] landing site, though in practice we found it more convenient to track mini-white.)

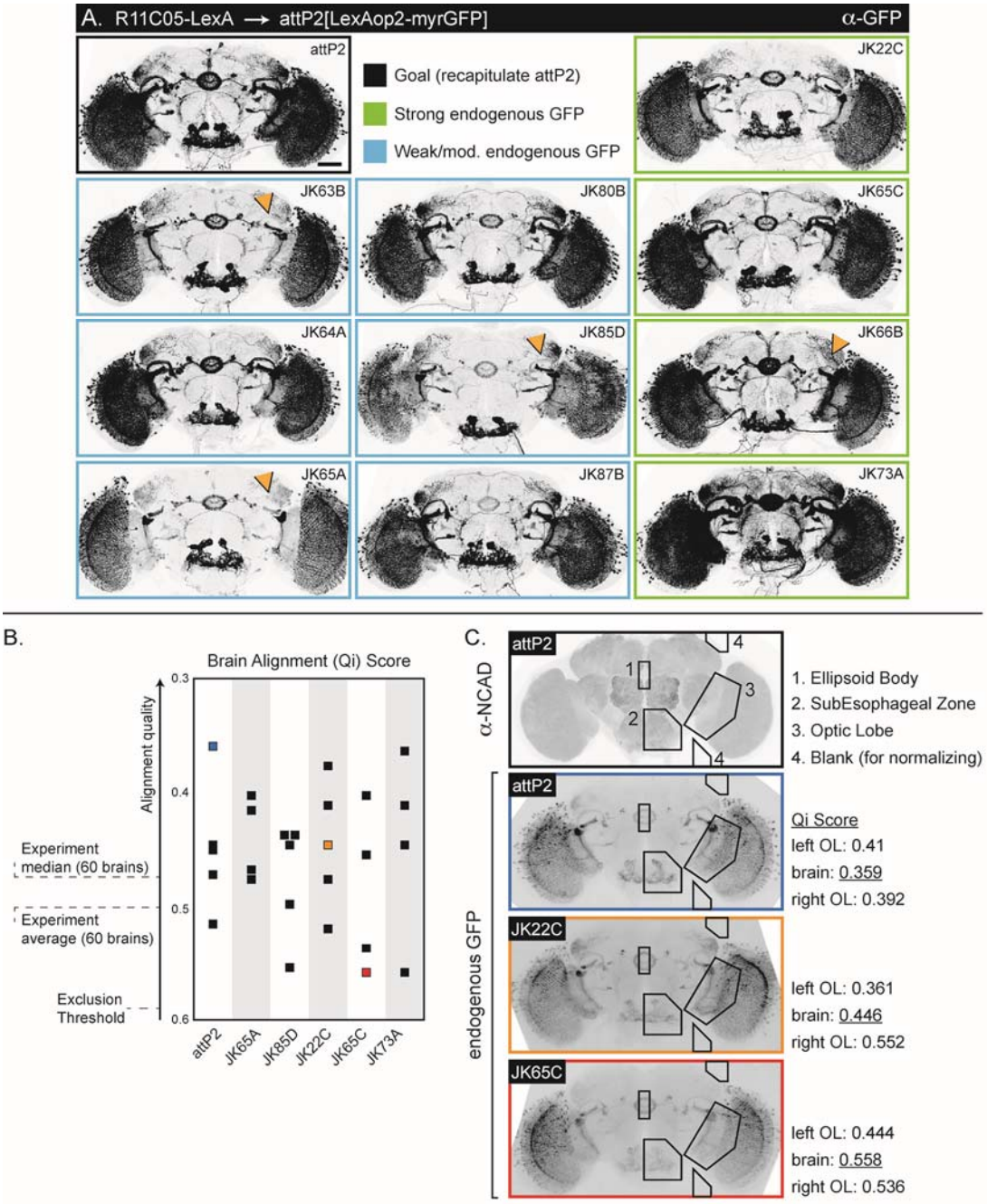


Figure S3. Characterization of R11C05-LexA expression in attP2 and several new candidate landing sites.

Figure S3 Characterization of R11C05-LexA expression in *attP2* and several new candidate landing sites. (A) To further characterize the expression of R11C05-LexA in the best new candidate landing sites, brains were stained with anti-GFP and imaged quantitatively by confocal microscopy. Images are representative samples from each line; the border around each image indicates the strength of native GFP fluorescence, as judged during the visual screen (see Figure 2B and 2C). Orange arrowheads indicate qualitative differences between R11C05-LexA expression in new candidate landing sites vs. *attP2*. (B) Alignment quality of brains presented in Figure 2D: Brains were imaged quantitatively for native GFP fluorescence by confocal microscopy, then stacks were aligned to a reference brain. Samples were assigned alignment (Qi) scores for each optic lobe and the central brain (smaller Qi indicates better alignment), and samples with brain-Qi > 0.59 were excluded from the analysis shown in Figure 2D. Samples represented by the blue, orange, and red squares are shown in panel C, with regions of interest (ROIs) drawn and complete Qi scores. (C) ROIs used for quantitative comparison of R11C05-LexA expression: Top - ROIs superimposed over a maximum projection of the neuropil, stained by anti-N-Cadherin. Bottom - ROIs drawn over three examples of aligned brains; the border color of each image indicates the sample in panel B with the matching color. To quantify signal in each ROI, the mean fluorescence intensity (mFI) of the ellipsoid body (1), subesophageal zone (2), and optic lobe (3) were computed, then the average mFI of the two blank regions (4) was subtracted.

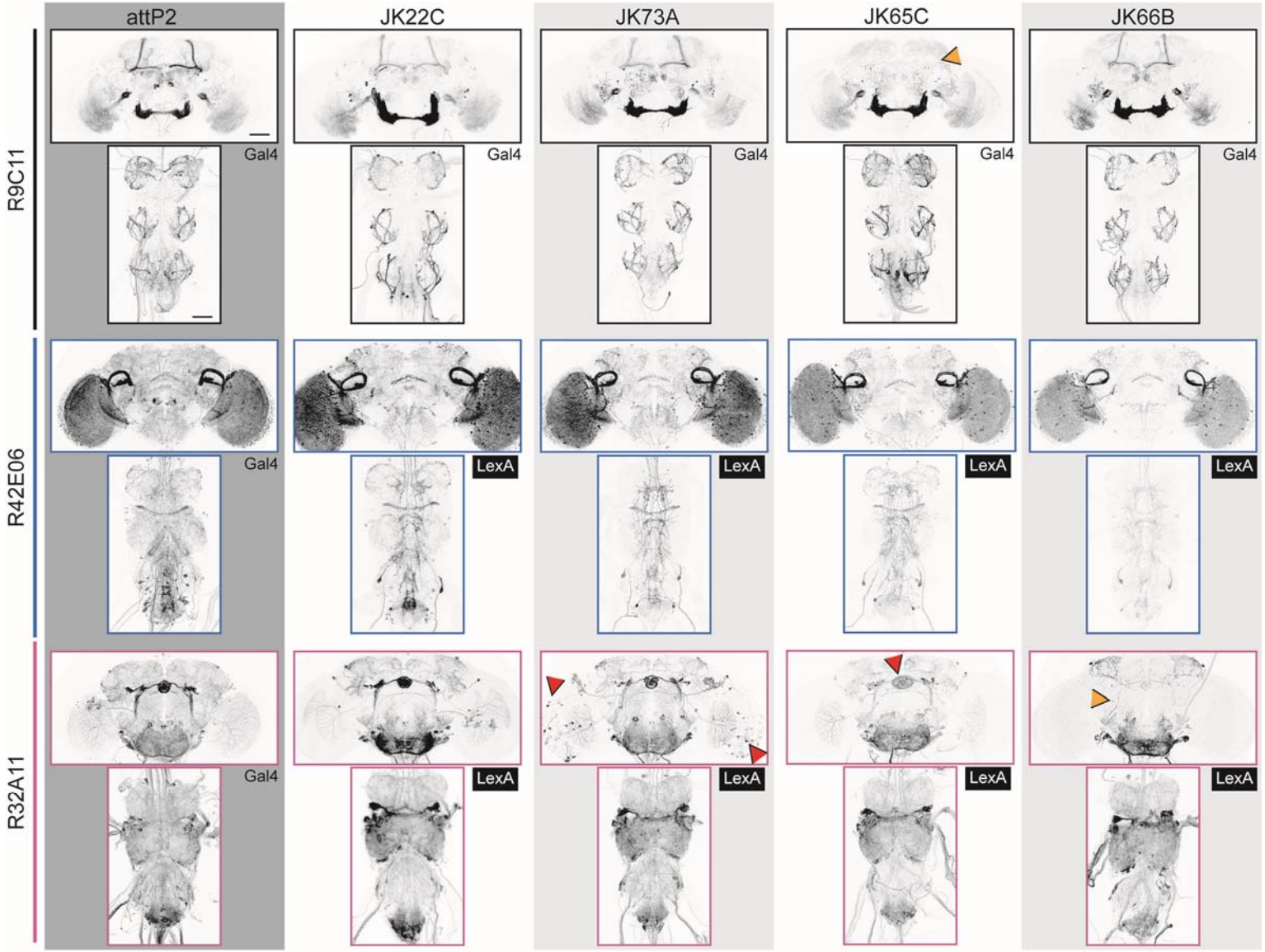
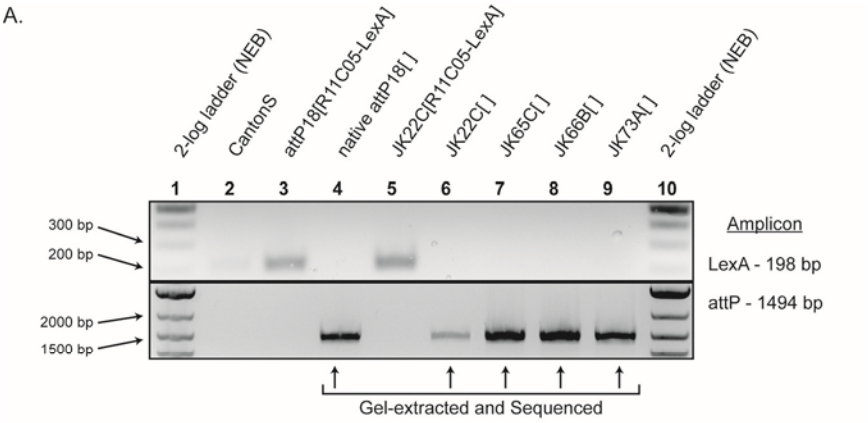


Figure S4. Behavior of additional driver transgenes in a subset of new landing sites.

Figure S4 Behavior of additional driver transgenes in new landing sites: Three additional driver transgenes were integrated into the four best candidate landing sites. These were crossed to *attP2[pJFRC12-UAS-myrGFP]* (for Gal4 transgenes) or *attP2[pJFRC19-LexAop2-myrGFP]* (for LexA transgenes), and immunostained to detect GFP. Generally, the expression pattern of each driver in the new landing sites was similar to that construct's expression in *attP2*, though particular landing site/driver combinations exhibited some deviations. Orange arrowheads indicate reduced or absent expression (relative to *attP2[driver]*); red arrowheads denote expression in novel cell types. Images are maximum projections of representative samples of adult brains and ventral nerve cords. Scale bars, 50 µm.



B.

attP

GTAGTGCCCCAACTGGGGTAACCT **TTG** AGTTCTCTCAGTTGGGGGCGTAG

attP18	...GA GTAGTGCCCCAACTGGGGTAACCT **TTG** AGTTCTCTCAGTTGGGGGCGTAG GG...
JK22C	...GA GTAGTGCCCCAACTGGGGTAACCT **TTG** AGTTCTCTCAGTTGGGGGCGTAG GG...
JK65C	...GA GTAGTGCCCCAACTGGGGTAACCT **TTG** AGTTCTCTCAGTTGGGGGCGTAG GG...
JK66B	...GA GTAGTGCCCCAACTGGGGTAACCT **TTG** AGTTCTCTCAGTTGGGGGCGTAG GG...
JK73A	...GA GTAGTGCCCCAACTGGGGTAACCT **TTG** AGTTCTCTCAGTTGGGGGCGTAG GG...

C.

	Location	Insertion
JK22C	2L: 22C3	5' - *GlyP* | *CG4259* - 3'
JK64A	3L: 64A12	5' UTR of *CG1309*
JK65C	3L: 65C3	5' UTR of *dikar*
JK66B	3L: 66B4	5' - *Ect4* | *CR32360* - 3'
JK73A	3L: 73A9	5' - *Smn* | *nxf2* - 3'
JK80B	3R: 80B1	5' UTR of *CG33170*
JK87B	3R: 87B9	5' UTR of *CG5196*

Figure S5. Sequencing and genomic location of new landing sites.

Figure S5 Molecular characterization of new landing sites. (A) Genomic PCR verification of R11C05-LexA excision. Top – The presence of R11C05-LexA was probed using primers that bind in the LexA coding sequence. Controls that harbor the transgene (lanes 3 and 5) showed amplification, but not native attP18 or the new landing sites following treatment with Int* (lanes 4 and 6-9). Bottom – Amplification of a region containing attP corroborates the absence of the R11C05-LexA transgene. Amplification was successful in samples with an intact attP, but fails in samples with an integrated transgene. Arrows indicate bands that were gel-extracted and sequenced. (B) To confirm the integrity of reconstituted attPs in new landing sites, the PCR products indicated in (A) were sequenced and compared to wild-type attP. The cross-over nucleotides where recombination occurs are indicated in bold. (C) Genomic location of new landing sites: Landing sites were mapped to the genome using splinkerette PCR (Potter and Luo 2010).

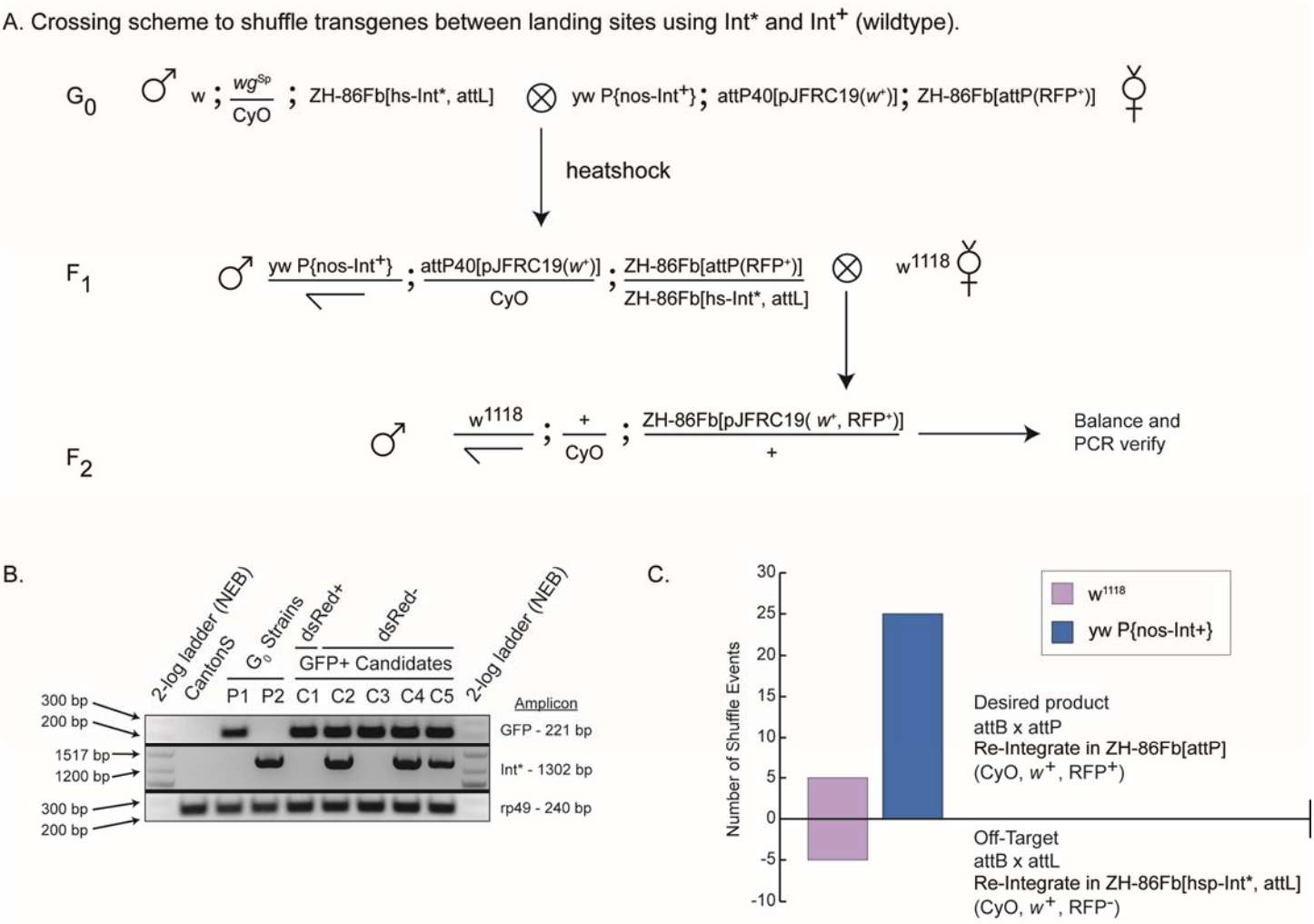


Figure S6. Genetics and integration characteristics of Int*/Int transgene shuffling.

Figure S6 Genetics and integration characteristics of Int*/Int transgene shuffling. (A) Crossing scheme to shuffle pJFRC2 between two autosomal landing sites. In this example, wild-type integrase is supplied by an *X*-linked transgene (Bischof *et al.* 2007). The movement of pJFRC19 from *attP40* (donor on *II*) to *ZH-86Fb* (receiver on *III*) was followed using the transgene's mini-white marker. Flies in the F_2 generation that carried mini-white and *CyO* carried candidate shuffle events; genetic mapping and PCR were used to confirm the presence of pJFRC19 at the receiver site. (B) Molecular characterization of selected shuffle candidates. Top – Genomic PCR confirmed the presence of GFP in five shuffle candidates (C1-C5). P1 corresponds to the maternal G_0 genotype, *w; attP40[pJFRC19(w+)]; ZH86Fb[attP(RFP+)].* P2 corresponds to the paternal G0 genotype, *w; wg^{Sp}/CyO; ZH86Fb[hs-Int*, attL].* Middle – Genomic PCR detected Int* in three shuffle candidates. Int* was not detected in candidate C1, since pJFRC19 shuffled into the receiver landing site (marked by *3xP3-DsRed*). Int* was detected in C2, C4, and C5. This indicates that pJFRC19 re-integrated on the hs-Int* chromosome, which is corroborated by the absence of DsRed expression in these flies. The attL sequence downstream of hs-Int* (see Figure S1A) presented a potential re-integration target due to the relaxed integration site specificity of Int*. The re-integration site of candidate C3 was not determined, though the lack of DsRed strongly suggests this was off-target. Bottom – PCR control with rp49. (C) Wild-type Int enforces canonical attP x attB recombination during transgene shuffling: Flies were scored for the presence of the *ZH-86Fb* landing site marker 3xP3-DsRed to distinguish re-integration at the receiver site from off-target integration (see panel B). When Int* provided both excisionase and integrase activities, half of the recovered candidates lacked DsRed, indicating re-integration at a site other than the receiver site. In contrast, in the presence of wild-type integrase, all candidates re-integrated at receiver landing site.

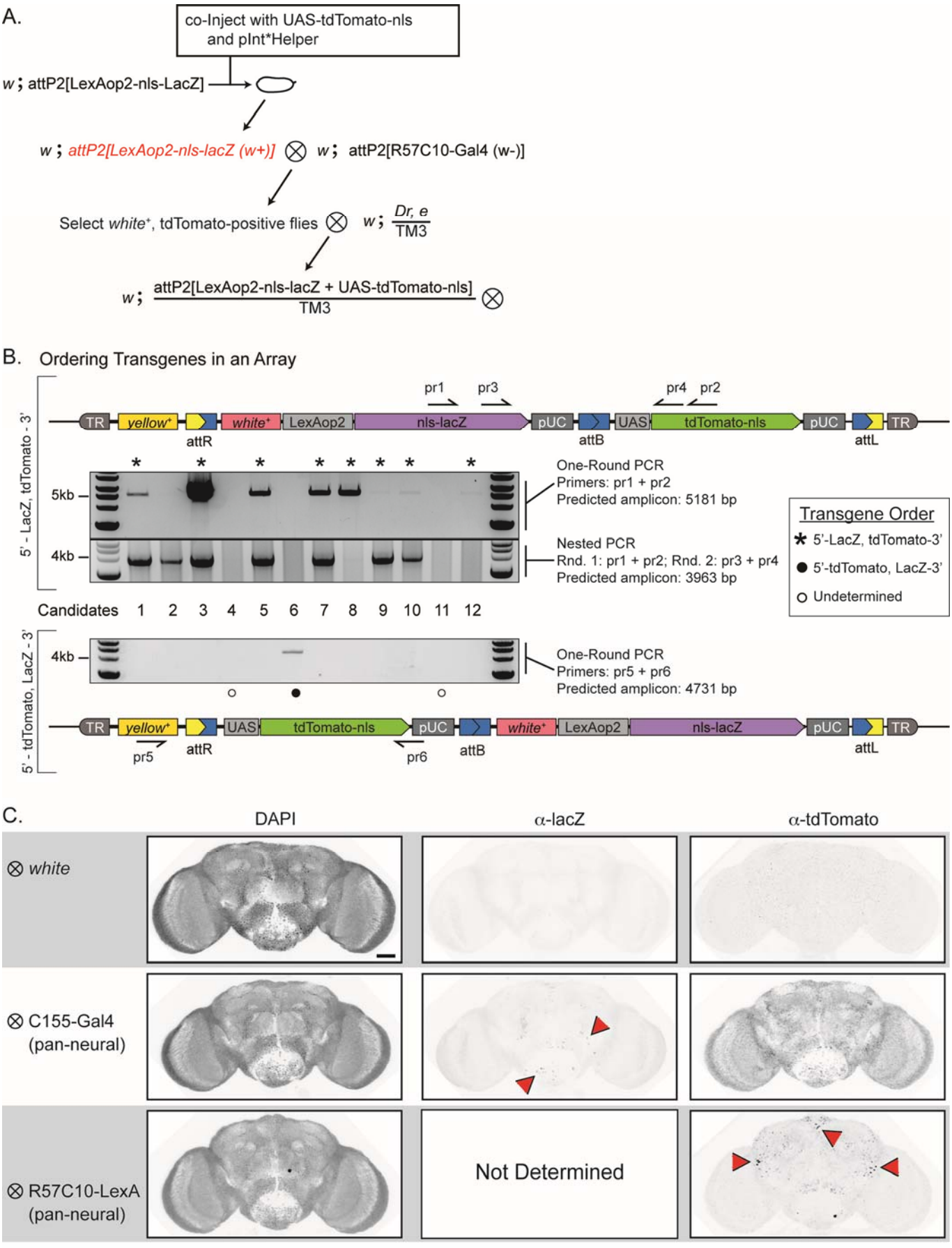


Figure S7. Construction and characterization of a UAS-/LexAop2-reporter transgene array.

Figure S7 Construction and characterization of a UAS-/LexAop2-reporter transgene array. (A) Injection and screening strategy to isolate a transgene array: Embryos carrying the primary transgene *attP2[LexAop2-nls-lacZ]* were co-injected with the secondary transgene UAS-tdTomato-nls and an Int*-expressing plasmid construct. G_0 adults were crossed to flies homozygous for *attP2[R57C10-Gal4]*, which expresses Gal4 pan-neurally. Progeny were screened for the presence of mini-white (the LexAop2-nls-lacZ marker) and fluorescence. Double-positive flies represented candidate arrays. (B) Molecular characterization of candidate UAS-tdTomato-nls/LexAop2-nls-LacZ reporter arrays: The order of transgenes in an array is determined by whether integration of the additional component occurs at attL or attR. Top – Schematic of the locus that results from integrating the secondary transgene at attL, with the secondary transgene (green) downstream of the primary transgene (purple). Genomic PCR (primers indicated above the schematic) showed that 9/12 candidates were in this orientation. Bottom – Schematic of the locus that results from integrating the secondary transgene at attR. Genomic PCR (primers indicated below the schematic) revealed that candidate 6 is in this orientation. Though candidates 4 and 11 were double-positive for tdTomato-nls and nls-LacZ, they are presumed to reflect off-target integrations and were not further characterized. (C) Expression characteristics of *attP2[LexAop2-nls-lacZ + UAS-tdTomato-nls]*: Top – In the absence of Gal4 and LexA drivers, LacZ and tdTomato are undetectable. Middle – LacZ can be detected in a small number of cells (red arrowheads) when the pan-neural Gal4 driver *C155* (*P{GawB}elav*^{C155}) is used to drive tdTomato-nls. Bottom – tdTomato-nls can be detected in a small number of cells in the central brain when the pan-neural driver *attP2[R57C10-LexA]* is used to express nls-LacZ.

Table S1 Excisionase activity of all Int* variants, and accessibility of several genomic locations.

A.

Int*	Excision Frequency	N (vials)	Int* location	Vector	Landing Site	Resident Transgene
E449K	64% ± 8%	10	ZH-86Fb	Gen1	attP18	pJFRC2-UAS-mCD8gfp
E449K	51% ± 11%	10	ZH-86Fb	Gen2	attP18	pJFRC2-UAS-mCD8gfp
E449K, E463K	61% ± 10%	10	ZH-86Fb	Gen1	attP18	pJFRC2-UAS-mCD8gfp
E456K	0.1% ± 0.2%	10	ZH-86Fb	Gen2	attP18	pJFRC2-UAS-mCD8gfp
E449H	40% ± 8%	10	ZH-86Fb	Gen2	attP18	pJFRC2-UAS-mCD8gfp
E449H, E463H	1% ± 2%	10	ZH-86Fb	Gen2	attP18	pJFRC2-UAS-mCD8gfp
E449H, E463G	23% ± 6%	10	ZH-86Fb	Gen2	attP18	pJFRC2-UAS-mCD8gfp
E449G, E463H	3% ± 2%	10	ZH-86Fb	Gen2	attP18	pJFRC2-UAS-mCD8gfp
E449K, E456K, E463K	16% ± 7%	8	ZH-86Fb	Gen1	attP40	pJFRC2-UAS-mCD8gfp

B.

Landing Site	Location	Landing Site Reported	Int*	Int* location	N (vials)	Excision Frequency
attP18	**X - 6C**	**Markstein et al., 2008**	**E449K, E463K**	**ZH-51C**	**10**	**36% ± 8%**
attP18	**X - 6C**	**Markstein et al., 2008**	**E449K, E463K**	**ZH-86Fb**	**10**	**62% ± 16%**
attP40	2L - 25C	Markstein et al., 2008	E449K	ZH-86Fb	6	34% ± 10%
su(Hw)attP5	2R - 51E	Ni et al., 2009	E449K	ZH-86Fb	7	60% ± 16%
attP2	3L - 68A	Groth et al., 2004	E449K	ZH-51C	5	28% ± 10%
VK00005	3L - 75B	Venken et al., 2006	E449K	ZH-51C	7	26% ± 10%
su(Hw)attP1	3R - 87B	Ni et al., 2009	E449K	ZH-51C	7	63% ± 10%
su(Hw)attP2	3R - 92D	Ni et al., 2009	E449K	ZH-51C	7	59% ± 14%

Table S1. The excisionase activity of Int* variants, and accessibility of several genomic locations.

(A) The activity of each Int* variant (column 1) was assayed using the genetic scheme detailed in Figure S1B. The observed activity (column 2) of variants ranges from virtually undetectable to greater than 60%. The fifth column, "Vector," indicates the plasmid in which Int* was cloned. These two plasmids are identical in the core transgene (hsp70P-Int*-SV40), though sequence differences outside this region may influence Int* expression (compare first and second rows). See also Figure 1D. (B) Int* activity assayed at multiple genomic locations: Int* at *ZH-51C* or *ZH-86Fb* were used to excise pJFRC2 from various landing sites on the *X* and third or second chromosomes, respectively. The first two rows (bold) permit a comparison between levels of Int* activity when expressed from *ZH-51C* or *ZH-86Fb*. Int* excised pJFRC2 from landing sites on every major chromosome arm, suggesting that the most (and perhaps all) sites in the genome are accessible to Int*.