

1 **An efficient CRISPR-based strategy to insert small and large fragments of DNA**
2 **using short homology arms**

3

4 Oguz Kanca^{1,2}, Jonathan Zirin³, Jorge Garcia-Marques⁴, Shannon Knight³, Donghui
5 Yang-Zhou³, Gabriel Amador³, Hyunglok Chung^{1,2}, Zhongyuan Zuo^{1,2}, Liwen Ma¹,
6 Yuchun He^{1,5}, Wen-Wen Lin¹, Ying Fang¹, Ming Ge¹, Shinya Yamamoto^{1,2,6,7}, Karen
7 L. Schulze^{1,5}, Yanhui Hu³, Allan C. Spradling⁸, Stephanie E. Mohr³, Norbert
8 Perrimon³, Hugo J. Bellen^{1,2,5,6,7*}

9 ¹Department of Molecular and Human Genetics, Baylor College of Medicine,
10 Houston, TX USA

11 ²Jan and Dan Duncan Neurological Research Institute, Texas Children's Hospital,
12 Houston, TX USA

13 ³Howard Hughes Medical Institute and Department of Genetics, Harvard Medical
14 School, Boston, MA USA

15 ⁴Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, VA USA

16 ⁵Howard Hughes Medical Institute, Baylor College of Medicine, Houston, TX USA

17 ⁶Program in Developmental Biology, Baylor College of Medicine, Houston, TX USA

18 ⁷Department of Neuroscience, Baylor College of Medicine, Houston, TX USA

19 ⁸Department of Embryology, Howard Hughes Medical Institute, Carnegie Institution
20 for Science, Baltimore, MD USA

21

22

23 Keywords: Cas9, CRISPR, ssDNA donors, short homology arms, S2 cells

24

25 *Hugo J. Bellen: Author for correspondence: hbellen@bcm.edu

26

27

28 **Abstract**

29 We previously reported a CRISPR-mediated knock-in strategy into introns of
30 *Drosophila* genes, generating an *attP-FRT-SA-T2A-GAL4-polyA-3XP3-EGFP-FRT-*
31 *attP* transgenic library for multiple uses (Lee et al., 2018b). The method relied on
32 double stranded DNA (dsDNA) homology donors with ~1 kb homology arms. Here,
33 we describe three new simpler ways to edit genes in flies. We create single stranded
34 DNA (ssDNA) donors using PCR and add 100 nt of homology on each side of an
35 integration cassette, followed by enzymatic removal of one strand. Using this
36 method, we generated GFP-tagged proteins that mark organelles in S2 cells. We
37 then describe two dsDNA methods using cheap synthesized donors flanked by 100
38 nt homology arms and gRNA target sites cloned into a plasmid. Upon injection,
39 donor DNA (1 to 5 kb) is released from the plasmid by Cas9. The cassette integrates
40 efficiently and precisely *in vivo*. The approach is fast, cheap, and scalable.

41

42

43 **Introduction**

44

45 A main goal of the *Drosophila* Gene Disruption Project (GDP) is to create
46 genetic tools that facilitate an integrated approach to analyze the function of each
47 gene in detail. This involves assessment of the loss of function phenotype,
48 identification of the cells that express the gene, determination of the subcellular
49 protein localization, selective removal of the transcript or protein in any tissue, the
50 ability to perform immunoprecipitation of the protein and its interacting proteins or
51 DNA, rescue of the induced fly mutant phenotypes with fly or human cDNAs and
52 assessment of the consequences of amino acid variants *in vivo*.

53 These elegant and precise manipulations are made possible by the
54 integration of a Swappable Integration Cassette (SIC) in the gene of interest (GOI)
55 using transposon mediated integration [*Minos*-mediated Integration Cassette
56 (MiMIC) (Venken et al., 2011; Nagarkar-Jaiswal et al., 2015a)] or homologous
57 recombination mediated by CRISPR, a technique we named CRIMIC (CRISPR-
58 mediated Integration Cassette) (Zhang et al., 2014; Diao et al., 2015; Lee et al.,
59 2018b). A SIC is typically flanked with *attP* sites and can be replaced using
60 Recombination Mediated Cassette Exchange (Bateman et al., 2006; Venken et al.,
61 2011). The CRIMIC variety of SIC currently used by the GDP is an artificial exon
62 consisting of *attP-FRT-SA-T2A-GAL4-polyA-3XP3-EGFP-FRT-attP* inserted in a
63 coding intron (intron flanked by two coding exons) of the GOI (Lee et al., 2018b).
64 This insert typically creates a severe loss of function allele and generates a GAL4
65 protein that is expressed in the target gene's spatial and temporal expression pattern
66 (Diao et al., 2015; Gnerer et al., 2015; Lee et al., 2018b). The resulting GAL4 can
67 then be used to drive a *UAS-nuclear localization signal (NLS)::mCherry* to determine
68 which cells express the gene or a *UAS-membrane (CD8)::mCherry* to outline the cell
69 projections (Brand and Perrimon, 1993; Shaner et al., 2004). Alternatively, a *UAS-*
70 *GOI cDNA* can be used to test for rescue of the loss of function phenotype induced
71 by the insertion cassette. This provides a means for rigorous quality assessment of
72 the genetic reagent and, when combined with mutant and/or truncated forms of the
73 *UAS-GOI cDNA*, facilitates structure-function analysis. In addition, a *UAS-human-*
74 *homologue cDNA* of the GOI permits humanization of the flies and assessment of

75 human variants (Bellen and Yamamoto, 2015; Kanca et al., 2017; Şentürk and
76 Bellen, 2018; Chao et al., 2017; Yoon et al., 2017).

77 The SIC can also be replaced by an artificial exon that consists of *SA-Linker-*
78 *EGFP-FIAsH-StrepII-TEV-3xFlag-Linker-SD*, abbreviated *SA-GFP-SD*, which adds
79 an integral Green Fluorescent Protein (GFP) and other tags to the gene product
80 (Venken et al., 2011; Nagarkar-Jaiswal et al., 2015a). This tag does not disrupt
81 protein function in 75% of cases examined and permits the determination of the
82 subcellular protein localization (Venken et al., 2011; Nagarkar-Jaiswal et al., 2015a;
83 Yoon et al., 2017; Lee et al., 2018b) as well as removal of the protein in any tissue
84 using specific GAL4 drivers (Jenett et al., 2012) to drive a DeGradFP protein that
85 leads to polyubiquitination and degradation of the protein of interest (Caussinus et
86 al., 2011; Nagarkar-Jaiswal et al., 2015a; Lee et al., 2018a). The GFP tag can also
87 be used as an epitope for immunoprecipitation to determine interaction partners of
88 the tagged protein (Neumüller et al., 2012; Zhang et al., 2013; David-Morrison et al.,
89 2016; Yoon et al., 2017).

90 Additionally, SICs can be replaced by other RMCE vectors that enable
91 integration of additional binary or tertiary systems (e.g. *SA-T2A-LexA*, *SA-T2A-split*
92 *GAL4*; Gnerer et al., 2015; Diao et al., 2015) to obtain finer tools to express
93 transgenes in specific cell populations. SICs can also be used to generate
94 conditional alleles of targeted genes (Flip-flop and FLPstop (Fisher et al., 2017;
95 Nagarkar-Jaiswal et al., 2017). Finally, strategies have been developed to convert
96 SICs through genetic crosses rather than by injection (*Trojan Exons* and *Double*
97 *Header*; Nagarkar-Jaiswal et al., 2015b; Diao et al., 2015; Li-Kroeger et al., 2018).

98 Although the above reagents form a powerful toolset, the generation of
99 libraries of many thousands of genes based on these methods is labor-intensive and
100 costly. The cost for reagents and labor for the generation of a single CRIMIC fly line
101 is \$1,000-2,000. Indeed, to create each CRIMIC construct we need to amplify two 1
102 kb homology arms, clone these arms on either site of a SIC in a plasmid, sequence
103 verify the constructs, amplify and inject the DNA with a target-specific gRNA into 600
104 embryos, screen to obtain several independent transgenic flies, establish several fly
105 stocks for each construct, PCR-verify the insertions, and cross each line with *UAS-*
106 *mCherry* to determine expression patterns. Since the production of each transgenic
107 line involves multiple steps, low failure rates at each step accumulate and decrease
108 the overall success rate to ~50%. Given that we are in the process of tagging ~5,000

109 genes that contain suitable introns, it is highly desirable to develop a more efficient,
110 less labor-intensive, and cheaper alternative. One of the main bottlenecks is the
111 production of large (5 kb) SIC homology donor plasmids containing a visible
112 dominant marker and flanked by two ~1kb homology arms to promote homologous
113 recombination (Beumer et al., 2008; 2013; Bier et al., 2018, Lee et al., 2018b). We
114 therefore explored a series of alternative strategies to reduce the construct size and
115 facilitate cloning.

116 Here, we report the development of methods, using either a PCR-generated,
117 single stranded DNA donor (ssDNA drop-in) or a synthesized double stranded
118 homology donor (dsDNA drop-in) that greatly simplify the generation of homology
119 donor constructs and improve the transgenesis rate. We tested both methods *in vivo*
120 in *Drosophila* and targeted the same 10 genes with 5 different constructs to assess
121 transformation efficiency and accuracy of integration. We show that the ssDNA drop-
122 in method works efficiently in *Drosophila* S2R+ cells for constructs that are less than
123 2 kb and we used this method to mark several cellular organelles with GFP tagged
124 proteins. The dsDNA drop-in strategy is based on short homology arms flanking
125 SICs of up to 5 kb. The success rate for tagging the tested genes was 70-80%. The
126 dsDNA drop-in donor vector is easy and cheap to produce, transformation efficiency
127 is high, and the insertions are precise. Hence, these changes significantly decrease
128 the costs of generating a transgenic CRIMIC library and make the CRIMIC technique
129 more accessible to others. We anticipate that the drop-in approaches will also be
130 useful in other species.

131

132 **Results and Discussion**

133

134 **ssDNA homology donors**

135

136 dsDNA homology donors for insertion of large cassettes in *Drosophila*
137 typically require stretches of 500 nt to 1 kb of homology to the target site on either
138 side of the SIC (Rong and Golic, 2000; Beumer et al., 2013; Zhang et al., 2014; Diao
139 et al., 2015; Bier et al., 2018; Lee et al., 2018b). The large size of the homology
140 regions affects cloning efficiency of the donor constructs (Lee et al., 2018b). Single
141 stranded homology donors typically rely on much shorter homology arms (50-100 nt)
142 to successfully integrate short DNA segments in *Drosophila* (~200 nt) based on

143 homology-directed repair (Beumer et al., 2013; Gratz et al., 2013; Wissel et al.,
144 2016; Bier et al., 2018). To facilitate and speed up the preparation of homology
145 donor constructs, we decided to test ssDNA donors for CRISPR-mediated
146 homologous recombination of donors that are 1 to 2 kb.

147 To produce ssDNA homology donors, we established a cloning-free method
148 based on PCR (Higuchi and Ochman, 1989). We generated SIC containing PCR
149 templates that can be amplified with an M13 universal primer-derived primer pair (26
150 nt forward primer and 24 nt reverse primer). Gene-specific 100 nt homology arms
151 are incorporated into these primers as 5' overhangs (left homology arm to the
152 forward primer and reverse complement of right homology arm to the reverse primer)
153 (Figure 1A). One of the primers is phosphorylated at its 5' end. The resulting PCR
154 product contains a 5' phosphorylated strand and a non-phosphorylated strand. When
155 this PCR product is treated with Lambda Exonuclease, a 5'-to-3' nuclease with a
156 preference for 5' phosphorylated DNA (Little, 1967; Mitsis and Kwagh, 1999), the 5'-
157 phosphorylated strand is degraded, leaving the non-phosphorylated strand as a
158 ssDNA homology donor. This PCR-based approach has several advantages. First,
159 the homology arms are included in the ~125 nt primers as 5' overhangs such that a
160 single PCR generates the complete donor without any cloning steps. Second, the
161 same primers can be used to amplify many different SIC templates for the same
162 GOI. Third, the priming sequences and overall construct length do not change
163 between genes. Hence, the same protocol can be applied to create homology
164 donors for different genes. Fourth, the PCR-based method does not require bacterial
165 transformation, eliminating possible rearrangements associated with propagation of
166 DNA in bacteria (Figure 1B). Finally, the risk of inserting PCR-induced mutations in
167 the SIC is negligible when a proofreading polymerase is used.

168 We first tested the efficacy of ssDNA drop-in constructs as a substrate for
169 homology-directed repair in *Drosophila* S2R+ cells stably transfected with Cas9
170 (S2R+-MT::Cas9; Viswanatha et al., 2018). Specifically, we were interested in
171 generating a collection of *Drosophila* S2R+ cells in which different organelles are
172 marked by a protein tagged with a superfolder GFP (sfGFP) in each cell line
173 (Pédélecq et al., 2006). We therefore amplified a SIC consisting of *attP-SA-Linker-*
174 *sfGFP-Linker-SD-attP* (1392 nt including homology arms) using 20 different gene-
175 specific primer pairs (Figure 2A, Supplementary file 3). Each gene-specific ssDNA
176 homology donor was electroporated into Cas9-positive cells along with a

177 corresponding gene-specific gRNA and then subjected to fluorescence-activated cell
178 sorting FACS (Figure 2B).

179 The frequency of GFP+ cells was determined by FACS (Figure 2-
180 supplemental figure 1) to range between ~1-4%. GFP-positive cells were clonally
181 isolated, cultured, and analyzed. We observed that for any given clone, all cells had
182 the same subcellular GFP localization, indicating that they were derived from a
183 single cell and that the insertion was stably integrated. For 19 out of 20 genes
184 targeted we observed GFP signal by FACS (Table 1). For 12 of 19 genes, we could
185 establish GFP+ clones, verify correct integration by PCR, and determine subcellular
186 localization by immunostaining (Table 1, Figure 2C, Figure 2- supplemental figure 2).
187 With the exception of Ref2P, all correct insertions of GFP resulted in fusion proteins
188 with the expected subcellular distribution. The genes for which we observed GFP
189 signal by FACS but could not successfully isolate GFP+ clones tended to be
190 expressed at low levels in S2R+ cells based on modEncode expression profiling
191 (Cherbas et al., 2011) (Table 1). For these genes it is possible that the signal-to-
192 noise ratio for GFP was insufficient to robustly select GFP+ clones, leading to the
193 loss of positive cells in the population.

194 Sequencing of the SIC insertion sites (Table 1) for the 12 cell lines showed
195 that the insertions are precise. Western blotting of cell lysates confirmed that the
196 inserted tags lead to GFP fusion proteins of the expected molecular weights (Figure
197 2- supplemental figure 3). Given the dynamic localization of Polo protein during
198 mitosis (Llamazares et al., 1991) we recorded the Polo subcellular localization
199 pattern in live cells through mitosis. Time-lapse confocal imaging of Polo-GFP
200 showed that the protein is localized to centrosomes, spindle, and midbody during cell
201 division, in agreement with the data obtained by immunofluorescence (Llamazares et
202 al., 1991) or using *polo-GFP* transgenes (Moutinho-Santos et al., 1999) (Video 1).

203 Previously, PCR generated double-stranded constructs containing GFP and
204 an antibiotic resistance gene have been used for homologous recombination in S2
205 cells (Bottcher et al., 2014; Kunzelmann et al., 2016). However, selection with a drug
206 resistance gene was used to enrich the population and GFP integration frequency
207 was about 2% as judged by FACS. We were able to generate GFP protein traps in
208 S2 cells using long ssDNA in up to 4% transfected cells and obtain clones without
209 using drug-based selection (Figure 2, Figure 2-supplemental figure 2).

210 Previous studies by Richardson et al. (Richardson et al., 2016) analyzed the
211 binding dynamics of Cas9 to target sites and showed an increase in homologous
212 recombination efficiency by using homology donors with asymmetric homology arms.
213 Whether the use of asymmetric homology arms will increase the knock-in efficiency
214 remains to be tested for drop-in.

215 In summary, ssDNA drop-in constructs are simple to produce and provided an
216 efficient homology substrate in S2R+ cells for about 60% of the tested genes. The
217 generated cell lines with marked organelles—endoplasmic reticulum (2), recycling
218 endosomes, cis and trans Golgi, kinetochores, lysosomes (2), mitochondria, nucleoli
219 and nuclear envelopes (Table 1)—are listed at the *Drosophila* RNAi Screening
220 Center (DRSC) website (<https://fgr.hms.harvard.edu/crispr-modified-cell-lines>) and
221 available through the Drosophila Genomics Resource Center (DGRC;
222 <https://dgrc.bio.indiana.edu>). These cell lines will be a useful resource for small- and
223 large-scale studies of organelle biogenesis, organelle function, and/or subcellular
224 distribution of organelles or specific fusion proteins, and genetic and
225 pharmacological screens to identify regulators of organelle homeostasis.

226

227 **Integration of ssDNA drop-in donors in the germline**

228

229 We next generated a drop-in SIC for *in vivo Drosophila* transformation that
230 can be amplified using PCR primers. Our initial experiments showed that the
231 empirical size limit for ssDNA production was ~2 kb, significantly smaller than the ~5
232 kb SIC present in the CRIMIC cassettes currently used by the GDP (Lee et al.,
233 2018b). A minimal SIC should contain *attP* sites to enable downstream RMCE
234 applications and a dominant marker that allows detection of positive insertions in the
235 targeted locus. The smallest self-sufficient visible dominant marker in *Drosophila* is
236 *3XP3-EGFP* (Horn et al., 2000). Hence, in order to remain under the size limit, we
237 generated an *attP-3XP3-EGFP-attP* PCR template that is 1242 nt long and can be
238 used for ssDNA donor (1442 nt including two 100 nt homology arms) generation
239 (Figure 3A). The *attP-3XP3-EGFP-attP* cassette is not mutagenic unless inserted in
240 exons and only generates a landing site for RMCE in the targeted region.

241 To test the efficiency of this new ssDNA drop-in construct *in vivo*, we
242 compared the efficiency of the ssDNA drop-in donor (with 100 nt homology arms)
243 and the current CRIMIC dsDNA donor (*attP-FRT-SA-T2A-GAL4-polyA-3XP3-EGFP-*

244 *FRT-attP* with ~1kb homology arms) for the same 10 genes using the same gRNAs
245 by injecting ~400-600 embryos for each gene with each construct. Surviving adults
246 were crossed with *y w* flies as single fly crosses and the numbers of independent
247 targeting events were quantified. We targeted a set of 10 genes with each construct
248 to have a large enough set to reduce the possibility of locus-specific properties
249 skewing the results. We were successful in targeting four genes with both
250 approaches and obtained insertions in *CG5009*, *CG9527*, *Khc*, and *NLaz* with the
251 CRIMIC approach and *CG5009*, *endoB*, *Cp1*, and *Lst* with the ssDNA drop-in
252 construct. We found that the use of CRIMIC constructs was more efficient, as judged
253 by the number of independent fly lines produced per successful gene (Figure 3B).

254 Short ssDNA homology donors were previously used in *Drosophila* to
255 integrate small epitope tags or site-specific recombination sites (Gratz et al., 2013;
256 Wissel et al., 2016). Due to size constraints, these constructs lacked a dominant
257 marker and detection of successful events relied on labor-intensive PCR strategies.
258 More recently, novel methods were developed to generate and use longer ssDNAs
259 (~1000 nt) as homology donors *in vivo* in mice (Miura et al., 2015; Quadros et al.,
260 2017; Lanza et al., 2018). Our results show that longer ssDNA constructs with visible
261 dominant markers integrate in the fly genome *in vivo* albeit with lower efficacy than
262 double stranded homology donors with large homology arms (i.e. CRIMIC donors).

263 In summary, the ssDNA constructs are efficient donors for S2R+ cells for
264 protein tagging by knock-in but for fly transgenesis they are not more efficient than
265 the standard CRIMIC method. Nevertheless, the ease and low cost of producing
266 these constructs and the transformation efficiency in S2 cells may justify their use to
267 integrate GFP tags or landing sites in cultured cells or the germline.

268

269 **dsDNA drop-in donors of <2 kb are efficient homology donors for transgenesis**

270

271 Given that the ssDNA drop-in constructs did not increase the success rate of
272 fly transformation, we next attempted to optimize dsDNA homology donors for
273 production ease and transformation efficiency. One means of increasing the
274 targeting rate is to shorten the SIC, since homologous recombination is dependent
275 on the size of the inserted cassette (Beumer et al., 2013). We used three strategies
276 to shorten SICs (Figure 4A). First, we used a shortened GAL4 construct, referred to
277 as miniGAL4 (1200 nt), which is about half the size of full-length GAL4 (2646 nt) and

278 has about 50% of the transcriptional activity of full-length GAL4 in yeast (Ding and
279 Johnston, 1997). Second, we used a very short dominant marker. Because *3XP3-*
280 *EGFP* is ~1 kb in length, we opted to use an alternative marker that is only ~200 nt
281 long and contains a ubiquitous U6 promotor expressing a gRNA sequence (gRNA1)
282 that does not have a target in the fly genome. This strategy is based on the single-
283 strand annealing pathway of DNA repair to reconstitute expression of a transgenic
284 ubiquitous mCherry that is interrupted by the gRNA1 target sequence (Figure 4-
285 supplemental figure 1) (Garcia Marques et al., 2019). Upon Cas9-mediated double
286 strand break, the modified non-functional *mCherry* gene is repaired and becomes
287 functional, providing a visible marker. This SSA dependent repair reaction occurs in
288 the F1 generation upon stable integration of U6:gRNA1 in the genome (Figure 4-
289 supplemental figure 1). Third, we shortened the polyA sequence from 135 nt to 35 nt
290 (*minipolyA*) (McFarland et al., 2006). This resulted in *attP-SA-T2A-miniGAL4-*
291 *minipolyA-U6gRNA1-attP*, which functions as a gene trap and is 1968 nt in length,
292 including homology arms. We also generated a smaller minimal mutagenic construct,
293 *attP-SA-3XSTOP-minipolyA-U6gRNA1-attP*, that is 715 nt in length including
294 homology arms (Figure 4A). Both dsDNA drop-in constructs are small enough to be
295 commercially synthesized at a low cost (less than \$250).

296 To further improve the insertion efficiency, we decided to induce *in vivo*
297 linearization of the plasmid constructs. Linearization has previously been shown to
298 boost homologous recombination rates in cell culture, zebrafish and mouse
299 transgenesis (Cristea et al., 2013; Hisano et al., 2015; Suzuki et al., 2016; Yao et al.,
300 2017). Hence, use of short 100 nt left and right homology arms flanked by the
301 gRNA1 target sites to linearize the construct *in vivo* upon injection may significantly
302 increase the frequency of homologous recombination (Figure 4B).

303 It is worth noting that higher transgenesis rates have been reported in mice
304 and worms using donor constructs linearized *in vitro* (Paix et al., 2014; 2015; 2017;
305 Dokshin et al., 2018; Yao et al., 2018). However, *in vitro* linearized DNA or direct
306 PCR products are very poor substrates for homologous recombination for germ line
307 transformation in *Drosophila* (Beumer et al., 2008; Bottcher et al., 2014). We
308 therefore opted to use dsDNA donors that are linearized *in vivo* by the Cas9
309 expressed in germ cells.

310 Upon injection of gene-specific gRNA coding plasmids and donor plasmids we
311 isolated multiple transgenic lines for each of the 10 targeted genes. We obtained 65

312 independent transgenic lines for the *attP-SA-3XSTOP-minipolyA-U6gRNA1-attP*
313 construct, and 119 independent transgenic lines for the *attP-SA-T2A-miniGAL4-*
314 *minipolyA-U6gRNA1-attP* construct. We verified the insertion sites by genomic PCR
315 in 7/10 genes for both methods. Analyses of the genomic DNA sequences of the
316 three genes for which we could not verify the integration site revealed that these
317 genes contain an unanticipated variation in the gRNA target site in the isogenized
318 Cas9 injection stocks when compared to the FlyBase reference genome sequence
319 (Santos et al., 2015; Goodman et al., 2018). This suggests that the precise
320 homologous recombination rate is dependent on efficient cutting of the target site
321 and is likely higher than 70%. Our current gRNA design platform considers these
322 strain variations which can be substantial in introns.

323 Because we shortened the *polyA* tail from 135 nt (Okada et al., 1999) to 35 nt
324 (McFarland et al., 2006), we assessed the mutagenic potential of new constructs that
325 contain *minipolyA*. We used Western blotting with an antibody that recognizes the
326 gene product of *CG5009* (*Drosophila* ortholog of *ACOX1*) to compare protein levels
327 in animals homozygous for the CRIMIC allele (with 135 nt polyA tail) with protein
328 levels of animals homozygous for either of the two constructs that carry the
329 *minipolyA*. As shown in Figure 5A, the artificial exon *attP-SA-3XSTOP-minipolyA-*
330 *U6gRNA1-attP* leads to ~40% decrease in protein levels and creates a much less
331 severe allele than the CRIMIC allele. However, the *attP-SA-T2A-miniGAL4-*
332 *minipolyA-U6gRNA1-attP* cassette leads to ~80-90% decrease in protein levels,
333 similar to a CRIMIC insertion in the same locus (Figure 5A). The lower mutagenic
334 efficacy of *attP-SA-3XSTOP-minipolyA-U6gRNA1-attP* may be the result of read
335 through of stop codons, inefficient transcriptional stop at the *minipolyA* sequence or
336 smaller size of the inserted artificial exon. The observation that *attP-SA-T2A-*
337 *miniGAL4-minipolyA-U6gRNA1-attP* is potently mutagenic indicates that the
338 combination of *SA-T2A* with *minipolyA* is stronger than *SA-stopcodon minipolyA*.
339 Alternatively, the increased size of artificial exon in *attP-SA-T2A-miniGAL4-*
340 *minipolyA-U6gRNA1-attP* overcomes the limitation of *attP-SA-3XSTOP-minipolyA-*
341 *U6gRNA1-attP* construct in mutagenesis efficacy.

342 The miniGAL4 construct had never been tested in flies. We therefore
343 compared full length versus miniGAL4 induction of a *UAS-LacZ* reporter by Western
344 blot. As shown in Figure 5B, we observed a ~60% reduction of reporter levels with
345 T2A-miniGAL4 as compared to CRIMIC T2A-GAL4 driven *UAS-LacZ* expression for

346 the *CG5009* locus (Figure 5B). Hence, the *attP-SA-T2A-miniGAL4-minipolyA-*
347 *U6gRNA1-attP* drives lower levels of reporter expression when compared to
348 CRIMIC, which may limit the use of miniGAL4 for genes expressed at low levels.

349 In summary, the decrease of SIC size and linearization increased the
350 transgenesis rate compared to CRIMIC templates. We used novel selection markers,
351 smaller *polyA* tails, and new GAL4 variants to decrease the size of the integrated
352 construct. However, this decrease in size came with trade-offs. The U6-gRNA1
353 dominant marker is very easy to detect and is smaller than previously established
354 dominant markers. However, this marker requires the presence of other transgenes
355 for detection (Figure 4- supplemental figure 1) and the reconstituted dominant
356 marker transgene segregates independently from the targeted gene. These
357 limitations make establishing and maintaining stocks more challenging. Moreover,
358 T2A-miniGAL4 may not be strong enough to drive robust reporter expression in
359 genes with low expression levels.

360

361 **Large double stranded drop-in dsDNA donors (~5 kb) are efficient homology** 362 **donors to integrate CRIMIC-like SICs**

363

364 To avoid the issues raised in the previous section, we designed a strategy to
365 integrate full length CRIMIC cassettes with short homology arms. To achieve this, we
366 tested whether large DNA fragments with short homology arms (100 nt) could be
367 integrated in target genes upon linearization using the gRNA1 *in vivo*. This would
368 allow the use of dominant markers like *3XP3-EGFP* as well as integration of the full
369 length *GAL4* gene with an extended *polyA* tail, i.e. the CRIMIC cassette SIC.
370 Unfortunately, synthesis of a full length 5 kb CRIMIC cassette would be cost
371 prohibitive (>\$1,000) especially for thousands of genes. Hence, we developed a
372 modified cloning strategy in which we first synthesize a *gRNA1 target-100nt*
373 *homology arm-Restriction cassette-100nt homology arm-gRNA1 target* inserted into
374 a pUC57 vector for each target gene (cost is \$80). The SIC containing *attP-FRT-*
375 *T2A-GAL4-polyA-3XP3-EGFP-FRT-attP* is then subcloned directionally into this
376 plasmid in a single straightforward cloning step, replacing the restriction cassette
377 with the SIC of interest (Figure 6A). We refer to these constructs *drop-in int100-*
378 *CRIMIC* constructs. The SIC can be replaced by any other effector (e.g. *SA-GFP-*
379 *SD*) to generate other homology donor constructs.

380 We injected vectors containing the full length CRIMIC cassette for seven of
381 the genes in which we previously successfully inserted dsDNA drop-in cassettes.
382 Using the *drop-in int100 CRIMIC* cassette we obtained multiple knock-in alleles in
383 five genes, as verified by PCR (Figure 6B). Hence, 100 nt homology arms are
384 sufficient to integrate large SICs into target sites. In addition, increasing homology
385 arm length to 200 nt should not increase synthesis costs, as the total length of the
386 construct remains less than 500 nt. Whether use of 200 nt may improve efficiency
387 remains to be tested.

388 To ensure functionality of knock-in alleles, we compared expression patterns
389 of drop-in int100 CRIMIC with the corresponding CRIMIC for *CG5009* (Figure 6C) as
390 well as drop-in int100 CRIMIC with the *T2A-miniGAL4* for *Khc* (Figure 6-
391 supplemental figure 1) by crossing the flies to *UAS-NLS::mCherry* reporter lines. In
392 both cases the constructs lead to mCherry expression in very similar patterns.

393 *In vivo* linearization was previously shown to lead to knock-ins in zebrafish
394 and cell culture even when there is no homology arm, provided that the homology
395 donor and target site are cut concomitantly (Cristea et al., 2013; Auer et al., 2014;
396 Schmid-Burgk et al., 2016; Suzuki et al., 2016). This process is more prone to
397 generating small deletions in the target site and the insertions are non-directional.
398 Recently, a homology independent knock-in method was established for *Drosophila*
399 cell culture and germline transformation (Bosch et al., 2019). By simultaneously
400 cutting the donor construct and target region in the absence of homology arms,
401 Bosch et al. (2019) integrated a CRIMIC cassette in 4 of 11 genes targeted.
402 However, these insertion events were error-prone and non-directional. Note that for
403 drop-in constructs, the presence of short homology arms allows the donor vectors to
404 be synthesized cheaply and introducing larger SICs is accomplished by a single
405 straightforward cloning step, providing a good balance between ease of construct
406 generation and efficient *in vivo* use.

407 In summary, we have developed efficient pipelines for CRISPR knock-in using
408 ssDNA in *Drosophila* cells or a dsDNA approach in the germline. With respect to the
409 dsDNA donors, we significantly improved the overall efficiency compared to the
410 method described in Lee et al. (2018b). The previous method that required cloning
411 large homology arm flanked cassettes has several pitfalls: homology arm PCRs
412 often must be troubleshooted repeatedly; assembly products are often incorrect; and
413 sequencing of the final product is often challenging and needs to be repeated to

414 confirm the construct. Moreover, when the efficacy of fly transformation and the
415 events that can be PCR verified on either side of knock-in region are included, the
416 efficacy of CRIMIC constructs hover around 50% in optimal conditions. With the
417 drop-in cassette strategy, not only is the cloning success rate nearly 100%, but the
418 protocol requires little troubleshooting, dramatically reducing the bench time. Finally,
419 the transformation and verification rate result in a 70-80% success rate. Key features
420 that improved efficiency are: 1) shortening the homology arms to 100 nt, which
421 allows synthesis followed by a simple cloning step, thus eliminating cloning failure
422 and reducing the cost from ~\$300 to ~\$100 per construct; and 2) integration of target
423 gRNA1 sites on either side of the SIC to linearize the donor with Cas9 in the germ
424 cells. In conclusion, the methods we describe are efficient, simple, and precise.

425

426

427 **Materials and Methods**

428

429 **Generation of templates for ssDNA production**

430 The sequences of scaffold vectors can be found in Supplementary file 1. Briefly, to
431 generate the *attP-3XP3-EGFP-attP*, the scaffold vector 3XP3-EGFP cassette was
432 amplified by PCR using the long primers 3XP3-EGFP-RI-for and 3XP3-EGFPPrev-
433 NotI and cloned between the *EcoRI* and *NotI* sites in pCasper3 (Thummel and
434 Pirrotta, 1992). To generate *attP-SA-sfGFP-SD-attP*, a scaffold vector that we
435 named pScaffold was produced by integrating annealed oligonucleotides with
436 sequences M13For-attPfor-SbfI-AvrII-attPrev-M13rev in pCasper3 backbone in
437 *EcoRI-NotI* sites. SA-sfGFP-SD was cloned as a three-fragment ligation in pScaffold
438 with linker-SA (amplified from pDoubleHeader (Li-Kroeger et al., 2018) with primers
439 SA-for-Sbf and Linker-SA-rev_BamHI), sfGFP (amplified from pUAST-NLS-sfGFP-
440 3XMyC-PEST with primers sfGFP-for_BamHI and sfGFP-rev_KpnI), and Linker-SD
441 (amplified from pDoubleHeader with primers Linker-SD-for_KpnI and Linker-SD-
442 rev_NotI).

443

444 **Production of ssDNA drop-in constructs**

445 Gene-specific homology donors were produced by incorporating the homologous
446 sites (regions spanning 100 nt upstream of 3 nt prior to the PAM (protospacer
447 adjacent motif) for left homology arm or 97 nt downstream of that region for the right

448 homology arm (Supplementary file 1) as 5' overhangs to primers with
449 complementarity to the template (26 bases for the forward primer and 24 bases for
450 the reverse primer). The long primers were ordered from IDT (Coralville, Iowa) as
451 Ultramers®. The reverse primer was 5' phosphorylated. Four 50 µl PCR reactions
452 were set up with Q5 hot start high fidelity 2X master mix (NEB #M0494L). PCR
453 conditions were optimized using gradient PCR (Bio-Rad C1000 Touch). The optimal
454 annealing temperature was 70°C. The elongation time was 1 minute for *attP-3XP3-*
455 *EGFP-attP* and *attP-SA-sfGFP-SD-attP*. Resulting PCR amplicons were pooled in
456 two samples and isolated using Qiaquick spin columns (Qiagen #28106) following
457 the manufacturer's protocol. Samples were eluted in 2 X 50µl elution buffer from the
458 kit. Two lambda exonuclease (NEB #M0262L) digestion reactions of 100 µl were set
459 up using 4-6 µg DNA and 8µl enzyme each. Digestion conditions were 37°C for 1
460 hour, followed by 10 minutes at 75°C for heat inactivation. Reaction products were
461 pooled in two samples and ssDNA isolated using NEB Monarch DNA purification kit
462 following the manufacturer's instructions (NEB #T1030L). Samples were isolated in 2
463 X 10µl prewarmed (55°C) elution buffer from the kit and the DNA concentration was
464 measured using NanoDrop One (ThermoFisher Scientific).

465

466 **Generation of GFP Knock-In Cell Lines**

467

468 **Cell Culture and Regular Media**

469 *Drosophila* cells stably expressing Cas9 (S2R+-MT::Cas9; Drosophila Genomics
470 Resource Center cell stock #268; Viswanatha et al. 2018) were cultured in
471 Schneider's *Drosophila* Medium 1X (ThermoFisher Scientific #21720024) with 10%
472 FBS and 1% penicillin/streptomycin (referred to as regular media). We note that this
473 cell line is a derivative of S2R+ NPT005 (DGRC #229) and thus contains an
474 mCherry tag in the *Clic* locus (Neumüller et al., 2012).

475

476 **Conditioned Media**

477 Conditioned media were prepared as previously described by Housden et al. (2017).

478

479 **Cellular Transfection**

480 Electroporation of S2R+-MT::Cas9 cells was performed using a Lonza 4D
481 Nucleofector (Lonza #AAF-1002B) following the manufacturer's protocol. For each
482 transfection, 1 µg of sgRNA (100 ng) and 2 µL of the sfGFP donor (100 ng) were
483 used with 4 x 10⁵ sub-confluent S2R+-MT::Cas9 cells. See Supplementary file 3 for
484 all sgRNA and insert sequences. After electroporation, cells were immediately
485 placed in regular media and health was monitored. Cell cultures were then
486 maintained and expanded for FACS.

487

488 **Isolation of Single-Cell Clones**

489 Prior to FACS, 10⁶ cells were filtered through a 40 µm Falcon Cell Strainer (Corning
490 #21008-949) into 15 mL conical tubes (ThermoFisher Scientific #14-959-70C).

491 Single cells were then FACS-isolated on a BD Aria IIu, based on the presence of
492 GFP with an expression level greater than 2 X 10² (see Figure 2- supplemental
493 figure 1). Single cells were sorted into wells of a 96 well plate (VWR #29444-010)
494 filled with 100 µL of conditioned medium. Cells were observed ~14 days later and
495 any viable clones were expanded.

496

497

498 **Analysis of GFP Knock-In cell lines**

499

500 ***Image Analysis***

501 Multiple clonal cell lines were generated per gene and image analysis was used to
502 measure the expression of the GFP knock-in marker. Several images of each live
503 clone were taken using an InCell Analyzer 6000 automated confocal fluorescence
504 microscope (GE Healthcare Lifesciences) using the dsRed channel to detect
505 mCherry fluorescence present in all cells and the FITC channel to detect GFP
506 fluorescence. Images were analyzed using CellProfiler (version 2.1.1). As mCherry is
507 present throughout each cell, this image was used by CellProfiler to determine the
508 outline of individual cells using the command "Identify Primary Objects." The outlines
509 generated from this step were then applied to the GFP image through the process
510 "Identify Secondary Objects." GFP fluorescence intensity was then measured using
511 the command "Measure Object Intensity," which averages the GFP fluorescence
512 within all of the cellular outlines (see Figure 2- supplemental figure 2). Mean GFP
513 intensities for each clone were noted, and the three clones per gene with the highest
514 means were chosen for molecular analysis. All other clones were pooled and stored

515 for use in a subsequent FACS if needed. Image data were managed using OMERO,
516 as supported by the Harvard Medical School Image Data Management Core.

517

518 ***Genomic DNA Extraction***

519 Genomic DNA was extracted from individual clones using a Quick-DNA MiniPrep Kit
520 (Zymo Research #D3024) according to the manufacturer's protocol. Concentrations
521 were measured using a NanoDrop 8000 Spectrophotometer (ThermoFisher
522 Scientific #ND-8000-GL).

523

524 ***Molecular Analysis***

525 Due to the large size of the insertion, the 5' and 3' integration sites of each insertion
526 were analyzed by PCR amplifying and sequencing a fragment spanning the junction
527 of each end of the insertion with flanking DNA (see Supplementary file 3 for a list the
528 two primer sets per gene). These flanking sites were PCR amplified with High
529 Fidelity Phusion Polymerase (NEB #M0530) using the following program: 1) 98°C for
530 30 sec, 2) 98°C for 10 sec, 3) 56°C for 30 sec, 4) 72°C for 30 sec (35 cycles), 5)
531 72°C for 10 min, 6) 4°C hold. PCR products were excised from a 2% TAE agarose
532 gel and purified using a QIAquick Gel Extraction Kit (Qiagen #28704) and Sanger
533 sequenced at the Dana-Farber/Harvard Cancer Center DNA Resource Core.

534 ***Immunostaining***

535 Cells were fixed in 4% paraformaldehyde in phosphate-buffered saline with 0.1%
536 Triton X-100 for 30 min. A standard staining protocol was used. Primary antibodies
537 were used as indicated: Chicken anti-GFP (1:1000, Abcam ab13970), Mouse anti-
538 Lamin (1:500, DSHB #ADL84.12), Mouse anti-Calnexin99A (1:5, DSHB #Cnx99A 6-
539 2-1), Mouse anti-ATP5A (1:100, Abcam ab14748), Rabbit anti-Arl8 (1:500, DSHB
540 Arl8), Goat anti-GMAP (1:2000, DSHB, #GMAP), Goat anti-Golgin245 (1:2000,
541 DSHB, #Golgin245), Rabbit anti-Ref2P (1:500, Abcam #ab178440), Rabbit anti-Myc-
542 tag (1:1000, Cell Signaling Technology #2278S). Secondary antibodies with Alexa
543 Fluor conjugates and DAPI (Molecular probes #D-1306) were used at 1:1000.
544 Images were obtained using a GE IN Cell 6000 automated confocal microscope with
545 a 60x objective. Time-lapse videos were generated by imaging every 30 sec over a 2
546 hr period.

547 **Western Blotting**

548 Clones that are generated using phase 1 template contained a 5 bp deletion in the
549 splice donor portion of the insert. To determine if full-length fusion proteins were
550 generated, we analyzed the following: S2R+-MT::Cas9 (unmodified control),
551 Actin::GFP (a GFP-positive control), *Rab11* GFP knock-in (no deletion in the insert
552 control), *spin* GFP knock-in (5 bp deletion), two *Lam* GFP knock-in clones (5 bp
553 deletion in the first, 206 bp deletion in the second), *Ref2p* GFP knock-in (5 bp
554 deletion), and *Cnx99a* GFP knock-in (5 bp deletion). Cells were spun down and
555 lysed by resuspension in Pierce RIPA Buffer (ThermoFisher Scientific #89901) and
556 Halt Protease Inhibitor Cocktail, EDTA-Free (ThermoFisher Scientific #87785). Cells
557 were agitated for 30 minutes at 4°C and then centrifuged at 15,000 rpm for 20
558 minutes at 4°C. The lysates (supernatant) were removed and held at -20°C. Protein
559 concentrations were determined using a Pierce BCA Protein Assay Kit
560 (ThermoFisher Scientific #23227) according to the manufacturer's protocol.
561 Appropriate volumes of each lysate were added to 4X Laemmli Sample Buffer
562 (BioRad #1610747), vortexed, held at 100°C for 10 minutes, and then spun down at
563 13,000 rpm for 3 minutes before loading into a Mini-PROTEAN TGX Precast Gel
564 (BioRad #4561095) and running at 100 V for 1 hour. The gel was then transferred
565 onto a PVDF membrane (BioRad #1620177) using the Trans-Blot Turbo Transfer
566 System (BioRad #1704150). After blocking with 5% blocking solution and washing in
567 TBST, the membrane was incubated in the primary antibody rabbit anti-GFP
568 (1:5000; Molecular Probes #A6455) shaking at 4°C overnight. The membrane was
569 then washed four times with TBST and incubated with Donkey anti-rabbit HRP
570 (1:3000; GE Healthcare #NA934) for 1 hour at room temperature, washed with
571 TBST, and prepared for imaging using the SuperSignal West Pico PLUS
572 Chemiluminescent Substrate (ThermoFisher Scientific #34580) according to
573 manufacturer's protocol. The blot was imaged using a ChemiDoc MP Imaging
574 System (see Supplementary Figure 3 top panel; BioRad #17001402). The blot was
575 then stripped with Restore PLUS Western Blot Stripping Buffer (ThermoFisher
576 Scientific #46430), reprobed with mouse anti-tubulin (1:2000; SigmaT #5168) and
577 Sheep anti-mouse HRP (1:3000; GE Healthcare #NXA931), and re-imaged (see
578 Supplementary Figure # middle panel). The blot was then stripped again and
579 reprobed with mouse anti-lamin (1:500; DSHB #ADL84.12) and Sheep anti-mouse

580 Horse Radish Peroxidase (HRP) (1:3000; GE Healthcare #NXA931), and re-imaged
581 (see Supplementary Figure # bottom panel).
582 For western analysis in adult flies, flies were dissected and lysed in 0.1% CHAPS
583 buffer [50mM NaCl, 200mM HEPES, 1mM EDTA and protease inhibitor cocktail
584 (Roche)] Tissue or cell debris were removed by centrifugation. Isolated lysates were
585 subjected to electrophoresis using a 4%–12% gradient SDS-PAGE gel and
586 transferred to Immobilon-FL polyvinylidene difluoride membranes. Loading input was
587 adjusted for protein concentration. Primary antibodies used were as follows:
588 Rabbit anti-ACOX1 (1:1000; HPA021195, Sigma), Rabbit anti- β -galactosidase
589 (1:1000; MP Biomedicals #55976), and Mouse anti-Actin-c4 (1:5000; Millipore Sigma
590 #MAB1501,). Secondary antibodies include Jackson ImmunoResearch HRP
591 conjugated (1:5000). Blots were imaged on a Bio-Rad ChemiDocMP. The intensity of
592 each band was measured and normalized to a loading control using Imagemlab
593 software (Bio-RAD). 3 biological repeats were performed and ordinary one way
594 ANOVA was used to compare expression levels of ACOX1 in different conditions.
595 Two technical repeats were performed for β -galactosidase measurements.

596

597 **Fly injections**

598 ssDNA constructs were injected at 50-100 ng/ μ l concentration with 25ng/ μ l gene
599 specific gRNA encoding pCFD3 vector (Port et al., 2014). *attP-SA-3XStop-*
600 *minipolyA-U6gRNA1-attP* and *attP-SA-T2A-miniGAL4-minipolyA-U6gRNA1-att*
601 dsDNA drop-in constructs were injected at ~150ng/ μ l concentration together with
602 25ng/ μ l gene specific gRNA. dsDNA drop-in int100-CRIMIC constructs were injected
603 at 300-400 ng/ μ l along with 25ng/ μ l gene specific gRNA and 25 ng/ μ l pCFD3-
604 gRNA1. Injections were performed as described in (Lee et al., 2018b). 400-600 *yw;*
605 *iso; attP2(y+){nos-Cas9(v+)}* embryos per genotype were injected. Resulting G0
606 males and females were crossed to *yw* flies as single fly crosses for *3XP3-EGFP*
607 detection and with *actin5C-Cas9; actin5C-GF-gRNA2-FP; actin5C-mcherr-#1-ry* flies
608 for gRNA1 detection (Figure 4- supplemental figure 1B) (Garcia Marques et al.,
609 2019). Up to 5 independent lines were generated per construct per gene. *actin5C-*
610 *GF-#2-FP* is an internal control that would detect non-specific activation of dominant
611 markers. We have not detected GFP in any of the screened flies, showing the
612 specificity of dominant marker detection through gRNA1.

613

614 **PCR validation**

615 PCR primers that flank the integration site were designed for each targeted gene
616 (Supplementary file 3 for primer sequences). These primers were used in
617 combination with insert-specific primers that bind 5' of the inserted cassette in
618 reverse orientation and 3' of the insert in forward orientation (pointing outwards from
619 the insert cassette). 200-800 nt amplicons were amplified from genomic DNA from
620 individual insertion lines through single fly PCR (Gloor et al., 1993) using OneTaq
621 PCR master mix (NEB #M0271L). PCR conditions were denaturation at 95°C for 30
622 seconds, 95°C 30 seconds, 58°C 30 seconds, 68°C 1 minute for 34 cycles and 68°C
623 5 minutes.

624

625 **dsDNA drop-in constructs production**

626 Templates for ordering the dsDNA drop-in constructs can be found in Supplementary
627 file 2. dsDNA drop-in constructs were ordered for production from Genewiz
628 ("ValueGene" option) in pUC57 Kan vector backbone at 4 µg production scale. When
629 lyophilized samples arrived from production, samples were resuspended in 25 µl of
630 ddH₂O at 55°C for 30 minutes. 19 µl was mixed with 1 µl gene-specific gRNA
631 plasmid (25 ng/ul final concentration of gRNA plasmid). The rest was stored at -20°C
632 for back-up purposes.

633

634 **Confocal imaging of transgenic larval brains**

635 Dissection and imaging were performed following the protocols in (Lee et al., 2018b).
636 In brief, fluorescence-positive 3rd instar larvae were collected in 1x PBS solution and
637 then cut in half and inverted to expose the brain. Brains were transferred into 1.5mL
638 centrifuge tubes and fixed in 4% PFA in 1xPBS buffer for 20 minutes. Brains were
639 then washed for 10 minutes three times in 0.2% PBST. Finally, samples were
640 mounted on glass slides with 8µL of VectaShield (VectorLabs #H-1000) and imaged
641 at 20x zoom with a Nikon W1 dual laser spinning-disc confocal microscope.

642

643 **Figure legends**

644

645 **Figure 1. ssDNA pipeline is faster than cloning CRIMIC constructs. (A)**

646 Schematic of PCR-based generation of drop-in ssDNA constructs. Gray boxes,

647 UTRs; orange boxes, coding exons, yellow line, coding introns, black line, outside
648 coding introns and exons (B) Comparison of donor generation pipelines for CRIMIC
649 and PCR-based drop-in ssDNA homology donors. Making ssDNA donors is about
650 5X times faster than making CRIMIC donors.

651

652 **Figure 2. ssDNA homology donors are effective in S2 cells to tag organelles.**

653 (A) Schematic of drop-in cassette encoding for sfGFP artificial exon. Size of the
654 construct including the homology arms is indicated in the right. sfGFP:
655 superfolderGFP; SA: Splice Acceptor of *mhc*; SD: Splice Donor of *mhc*; L: flexible
656 linker that consists of four copies of Gly-Gly-Ser. (B) Diagram of steps to isolate cell
657 clones resulting from successful homologous recombination events. (C) Examples of
658 S2R+ cells with organelles marked with GFP. Left panel, antibody staining; middle
659 panel, GFP signal; right panel, the merge.

660

661 **Figure 3. ssDNA constructs are not as efficient as double stranded CRIMIC**

662 **constructs for fly transformation.** (A) Schematic of drop-in cassette used for fly
663 transformation. Size of the construct including homology arms is indicated on the
664 right. (B) Injection results for the 10 genes selected for comparison of transformation
665 efficacy. Numbers indicate positive events / fertile G0 single fly crosses. Red, no
666 positive events; light green, positive non-confirmed events; dark green, genes with
667 PCR-confirmed events.

668

669 **Figure 4. Double stranded DNA synthetic constructs are efficient for fly**

670 **transformation.** (A) Schematic of synthesized plasmid drop-in donors. mPA
671 indicates *minipolyA*. (B) Schematic of the donor plasmid followed by linearization by
672 Cas9 in germ cells and integration of donors *in vivo*. gRNA^{target} is gene specific
673 gRNA. (C) Injection results for the 10 genes selected for comparison of
674 transformation efficacy. Numbers indicate positive events / fertile G0 single fly
675 crosses. Red, no positive events; light green, positive non-confirmed events; and
676 dark green, genes with PCR-confirmed events.

677

678 **Figure 5. T2A-miniGAL4 gene trap cassette is mutagenic and expresses an**

679 **active GAL4.** (A) Western blot and quantification of the level of ACOX1 protein in
680 flies homozygous for *SA-T2A-miniGAL4-minipolyA*, *SA-3XSTOP-minipolyA* or

681 CRIMIC construct for *CG5009*. **** $P < 0.0001$, * $P < 0.01$. mPA1 and mPA2 indicate
682 two independent lines of *CG5009-SA-3XSTOP-minipolyA*. (B) Western blot of β -
683 galactosidase and quantification of heterozygous flies carrying a copy of the
684 miniGAL4 construct compared to heterozygous flies carrying a CRIMIC.

685

686 **Figure 6. Single step cloning method allows efficient insertion of the CRIMIC**
687 **cassette in coding introns.** (A) Schematic of a single step cloning vector pUC57.
688 LHA Left Homology Arm, RE1 Restriction Enzyme 1, RE2 restriction Enzyme 2, RHA
689 Right Homology Arm. (B) Injection results for the 7 genes selected to estimate
690 transformation efficacy. Numbers indicate positive events / G0 single fly crosses. (C)
691 Third instar larval brain expression domain of *CG5009* as determined by crossing
692 conventional CRIMIC or drop-in int100-CRIMIC flies to *UAS-NLS-mCherry* reporter
693 lines. Scale bar is 100 μ m.

694

695 **Table 1. Summary of ssDNA drop-in mediated homologous recombination in**
696 **S2R+ cells**

697

698 **Figure2- Supplementary Figure 1. FACS data of control cells (left) and ssDNA**
699 **knock-in cells (right).** All cell lines express mCherry::Clic, which is present in the
700 parental cell line and thus in these derivatives. Single cell clones with GFP
701 expression levels greater than 2×10^2 were retained.

702

703 **Figure2- Supplementary Figure 2. Detection of subcellular localization of GFP**
704 **tagged proteins in S2 cells.** Left panels, confocal fluorescence detection of GFP
705 fusion proteins in single-cell isolated clones. The specific proteins tagged by GFP
706 knock-in are indicated. Center panels, confocal fluorescence detection of
707 mCherry::Clic, which is present in the parental cell line and thus in these derivatives.
708 Right panels, merged image.

709

710 **Figure2- Supplementary Figure 3. Western blot analysis of tagged proteins**
711 **observed in S2R+ cells.** Western blot analysis of selected clones using anti-GFP
712 antibody shows expected protein sizes.

713

714 **Figure4- Supplementary Figure 1. Schematic and crossing scheme for CRISPR**
715 **gRNA-based dominant marker strategy.** (A) Schematics of the action of gRNA1-
716 based detection of dominant marker. (B) Crossing scheme for screening the
717 transgenics using *U6gRNA1* as dominant marker. The crosses for a second
718 chromosome insertion are shown; for insertions on other chromosomes, an
719 appropriate balancer stock would be used in the second and third crosses.

720

721 **Figure 6- Supplementary Figure 1. Comparison of expression domain obtained**
722 **by T2A-miniGAL4 and drop-in int100-CRIMIC.** Third instar larval brain expression
723 domain of *Khc* as determined by crossing conventional T2A-miniGAL4 or drop-in
724 int100-CRIMIC flies to *UAS-NLS-mCherry* reporter lines. Scale bars are 100µm.

725

726 **Supplementary file 3. List of oligos and ultramers used in this study.**

727

728 **Video 1.** Polo-sfGFP knock-in in S2R+ cells through ssDNA drop-in faithfully reports
729 dynamic localization of Polo throughout the cell cycle.

730

731 **Acknowledgements**

732

733 The Drosophila Gene Disruption Project is supported by NIH NIGMS R01GM067858
734 to H.J.B. We thank the Harvard Medical School Image Data Management Core for
735 support with the OMERO platform and the Harvard Medical School Division of
736 Immunology's Flow Cytometry Facility for the use of their FACS equipment. We also
737 thank Raghuvir Viswanatha, Justin Bosch, Denise Lanza and Jason Heaney for
738 helpful discussions, Robert Levis for critical reading and editing of the manuscript
739 and Tzumin Lee for fly stocks. Development of the tagged cell line resource at the
740 Drosophila RNAi Screening Center (DRSC) was supported by NIH ORIP R24
741 OD019847 (PI: N.P., Co-PI: A. Simcox, Co-I: S.E.M.). Additional relevant support for
742 the DRSC includes NIH NIGMS R01 GM084947 and P41 GM132087 (PI: N.P., Co-I:
743 S.E.M.). N.P. A.C.S. and H.B. are investigators of Howard Hughes Medical Institute.

744

745 **References**

746 Auer TO, Duroure K, De Cian A, Concordet JP, Del Bene F. Highly efficient
747 CRISPR/Cas9-mediated knock-in in zebrafish by homology-independent DNA
748 repair. *Genome Res.* 2014 Jan 1;24(1):142–53.

- 749 Bateman JR, Lee AM, Wu C-T. Site-specific transformation of *Drosophila* via phiC31
750 integrase-mediated cassette exchange. *Genetics*. *Genetics*; 2006
751 Jun;173(2):769–77. PMID: PMC1526508
- 752 Bellen HJ, Yamamoto S. Morgan's legacy: fruit flies and the functional annotation of
753 conserved genes. *Cell*. 2015 Sep 24;163(1):12–4. PMID: PMC4783153
- 754 Beumer KJ, Trautman JK, Bozas A, Liu J-L, Rutter J, Gall JG, et al. Efficient gene
755 targeting in *Drosophila* by direct embryo injection with zinc-finger nucleases. *Proc.*
756 *Natl. Acad. Sci. U.S.A. National Academy of Sciences*; 2008 Dec
757 16;105(50):19821–6. PMID: PMC2604940
- 758 Beumer KJ, Trautman JK, Mukherjee K, Carroll D. Donor DNA Utilization During
759 Gene Targeting with Zinc-Finger Nucleases. *G3 (Bethesda). G3: Genes,*
760 *Genomes, Genetics*; 2013 Apr 9;3(4):657–64. PMID: PMC3618352
- 761 Bier E, Harrison MM, O'Connor-Giles KM, Wildonger J. Advances in Engineering the
762 Fly Genome with the CRISPR-Cas System. *Genetics*. *Genetics*; 2018
763 Jan;208(1):1–18. PMID: PMC5753851
- 764 Bottcher R, Hollmann M, Merk K, Nitschko V, Obermaier C, Philippou-Massier J, et
765 al. Efficient chromosomal gene modification with CRISPR/cas9 and PCR-based
766 homologous recombination donors in cultured *Drosophila* cells. *Nucleic Acids*
767 *Res*. 2014 Jun 19;42(11):e89–9.
- 768 Bosch, J. A., Colbeth, R., Zirin, J., & Perrimon, N. (2019). Gene knock-ins in
769 *Drosophila* using homology-independent insertion of universal donor plasmids,
770 16, 29923–47. <http://doi.org/10.1101/639484>
771
- 772 Brand AH, Perrimon N. Targeted gene expression as a means of altering cell fates
773 and generating dominant phenotypes. *Development*. 1993 Jun;118(2):401–15.
- 774 Caussinus E, Kanca O, Affolter M. Fluorescent fusion protein knockout mediated by
775 anti-GFP nanobody. *Nat. Struct. Mol. Biol. Nature Publishing Group*; 2011 Dec
776 11;19(1):117–21.
- 777 Chao, H.-T., Davids, M., Burke, E., Pappas, J. G., Rosenfeld, J. A., McCarty, A. J., et
778 al. (2017). A Syndromic Neurodevelopmental Disorder Caused by De Novo
779 Variants in EBF3. *American Journal of Human Genetics*, 100(1), 128–137.
780 <http://doi.org/10.1016/j.ajhg.2016.11.018>
781
- 782 Cherbas L, Willingham A, Zhang D, Yang L, Zou Y, Eads BD, et al. The
783 transcriptional diversity of 25 *Drosophila* cell lines. *Genome Res*. 2011 Feb
784 1;21(2):301–14.
- 785 Cristea S, Freyvert Y, Santiago Y, Holmes MC, Urnov FD, Gregory PD, et al. In vivo
786 cleavage of transgene donors promotes nuclease-mediated targeted integration.
787 *Biotechnol. Bioeng. John Wiley & Sons, Ltd*; 2013 Mar;110(3):871–80.

788 David-Morrison G, Xu Z, Rui Y-N, Charng W-L, Jaiswal M, Yamamoto S, et al. WAC
789 Regulates mTOR Activity by Acting as an Adaptor for the TTT and Pontin/Reptin
790 Complexes. *Dev. Cell.* 2016 Jan 25;36(2):139–51. PMID: PMC4730548

791 Diao F, Ironfield H, Luan H, Diao F, Shropshire WC, Ewer J, et al. Plug-and-play
792 genetic access to drosophila cell types using exchangeable exon cassettes. *Cell*
793 *Rep.* 2015 Mar 3;10(8):1410–21. PMID: PMC4373654

794 Ding WV, Johnston SA. The DNA binding and activation domains of Gal4p are
795 sufficient for conveying its regulatory signals. *Mol. Cell. Biol. American Society for*
796 *Microbiology (ASM);* 1997 May;17(5):2538–49. PMID: PMC232103

797 Dokshin GA, Ghanta KS, Piscopo KM, Mello CC. Robust Genome Editing with Short
798 Single-Stranded and Long, Partially Single-Stranded DNA Donors in
799 *Caenorhabditis elegans.* *Genetics.* 2018 Nov 6;210(3):781–7.

800 Fisher YE, Yang HH, Isaacman-Beck J, Xie M, Gohl DM, Clandinin TR. FlpStop, a
801 tool for conditional gene control in *Drosophila.* *Elife. eLife Sciences Publications*
802 *Limited;* 2017 Feb 17;6:796. PMID: PMC5342825

803 Garcia Marques J, Yang C-P, Espinosa-Medina I, Mok K, Koyama M, Lee T.
804 Unlimited Genetic Switches for Cell-Type-Specific Manipulation. *Neuron. Elsevier*
805 *Inc;* 2019 Aug 9;:1–20.

806 Gloor, G. B., Preston, C. R., Johnson-Schlitz, D. M., Nassif, N. A., Phillis, R. W.,
807 Benz, W. K., et al. (1993). Type I repressors of P element mobility. *Genetics,*
808 *135(1),* 81–95.
809

810 Gnerer JP, Venken KJT, Dierick HA. Gene-specific cell labeling using MiMIC
811 transposons. *Nucleic Acids Res.* 2015 Apr 30;43(8):e56. PMID: PMC4417149

812 Goodman JL, Strelets VB, Marygold SJ, Matthews BB, Millburn G, Trovisco V, et al.
813 FlyBase 2.0: the next generation. *Nucleic Acids Res. Oxford University Press;*
814 *2018 Oct 26;47(D1):D759–65.*

815 Gratz SJ, Cummings AM, Nguyen JN, Hamm DC, Donohue LK, Harrison MM, et al.
816 Genome engineering of *Drosophila* with the CRISPR RNA-guided Cas9 nuclease.
817 *Genetics. Genetics;* 2013 Aug;194(4):1029–35. PMID: PMC3730909

818 Higuchi RG, Ochman H. Production of single-stranded DNA templates by
819 exonuclease digestion following the polymerase chain reaction. *Nucleic Acids*
820 *Res. Oxford University Press;* 1989 Jul 25;17(14):5865. PMID: PMC318226

821 Hisano Y, Sakuma T, Nakade S, Ohga R, Ota S, Okamoto H, et al. Precise in-frame
822 integration of exogenous DNA mediated by CRISPR/Cas9 system in zebrafish.
823 *Nature Publishing Group. Nature Publishing Group;* 2015 Mar 5;5(1):8841.
824 PMID: PMC4350073

825 Horn C, Jaunich B, Wimmer EA. Highly sensitive, fluorescent transformation marker
826 for *Drosophila* transgenesis. *Dev. Genes Evol.* 2000 Dec;210(12):623–9.

- 827 Housden, B. E., Nicholson, H. E., & Perrimon, N. (2017). Synthetic Lethality Screens
828 Using RNAi in Combination with CRISPR-based Knockout in Drosophila Cells.
829 *Bio-Protocol*, 7(3). <http://doi.org/10.21769/BioProtoc.2119>
830
- 831 Jenett A, Rubin GM, Ngo T-TB, Shepherd D, Murphy C, Dionne H, et al. A GAL4-
832 driver line resource for Drosophila neurobiology. *Cell Rep*. 2012 Oct 25;2(4):991–
833 1001. PMID: PMC3515021
- 834 Kanca O, Bellen HJ, Schnorrer F. Gene Tagging Strategies To Assess Protein
835 Expression, Localization, and Function in Drosophila. *Genetics*. 2017
836 Oct;207(2):389–412. PMID: PMC5629313
- 837 Kunzelmann S, Böttcher R, Schmidts I, Förstemann K. A Comprehensive Toolbox
838 for Genome Editing in Cultured Drosophila melanogaster Cells. *G3 (Bethesda)*.
839 *G3: Genes, Genomes, Genetics*; 2016 Jun 1;6(6):1777–85. PMID:
840 PMC4889673
- 841 Lanza DG, Gaspero A, Lorenzo I, Liao L, Zheng P, Wang Y, et al. Comparative
842 analysis of single-stranded DNA donors to generate conditional null mouse
843 alleles. *BMC Biology*; 2018 Jun 15;:1–18.
- 844 Lee P-T, Lin G, Lin W-W, Diao F, White BH, Bellen HJ. A kinase-dependent
845 feedforward loop affects CREBB stability and long term memory formation. *Elife*.
846 2018a Feb 23;7:2743. PMID: PMC5825208
- 847 Lee P-T, Zirin J, Kanca O, Lin W-W, Schulze KL, Li-Kroeger D, et al. A gene-specific
848 T2A-GAL4 library for Drosophila. *Elife*. eLife Sciences Publications Limited; 2018b
849 Mar 22;7:e35574. PMID: PMC5898912
- 850 Li-Kroeger D, Kanca O, Lee P-T, Cowan S, Lee MT, Jaiswal M, et al. An expanded
851 toolkit for gene tagging based on MiMIC and scarless CRISPR tagging in
852 Drosophila. *Elife*. 2018 Aug 9;7. PMID: PMC6095692
- 853 Little JW. An exonuclease induced by bacteriophage lambda. II. Nature of the
854 enzymatic reaction. *J. Biol. Chem*. 1967 Feb 25;242(4):679–86.
- 855 Llamazares S, Moreira A, Tavares A, Girdham C, Spruce BA, Gonzalez C, et al. polo
856 encodes a protein kinase homolog required for mitosis in Drosophila. *Genes Dev*.
857 Cold Spring Harbor Lab; 1991 Dec;5(12A):2153–65.
- 858 McFarland TJ, Zhang Y, Atchaneeyaskul L-O, Francis P, Stout JT, Appukuttan B.
859 Evaluation of a novel short polyadenylation signal as an alternative to the SV40
860 polyadenylation signal. *Plasmid*. 2006 Jul;56(1):62–7.
- 861 Mitsis PG, Kwagh JG. Characterization of the interaction of lambda exonuclease with
862 the ends of DNA. *Nucleic Acids Res*. Oxford University Press; 1999 Aug
863 1;27(15):3057–63. PMID: PMC148530
- 864 Miura H, Gurumurthy CB, Sato T, Sato M, Ohtsuka M. CRISPR/Cas9-based
865 generation of knockdown mice by intronic insertion of artificial microRNA using

866 longer single-stranded DNA. Nature Publishing Group. Nature Publishing Group;
867 2015 Jul 27;;1–11.

868 Moutinho-Santos T, Sampaio P, Amorim I, Costa M, Sunkel CE. In vivo localisation
869 of the mitotic POLO kinase shows a highly dynamic association with the mitotic
870 apparatus during early embryogenesis in *Drosophila*. *Biol. Cell*. 1999
871 Nov;91(8):585–96.

872 Nagarkar-Jaiswal S, Lee P-T, Campbell ME, Chen K, Anguiano-Zarate S, Gutierrez
873 MC, et al. A library of MiMICs allows tagging of genes and reversible, spatial and
874 temporal knockdown of proteins in *Drosophila*. *Elife*. eLife Sciences Publications
875 Limited; 2015 Mar 31;4:e05338. PMID: PMC4379497

876 Nagarkar-Jaiswal S, DeLuca SZ, Lee P-T, Lin W-W, Pan H, Zuo Z, et al. A genetic
877 toolkit for tagging intronic MiMIC containing genes. *Elife*. eLife Sciences
878 Publications Limited; 2015 Jun 23;4:166. PMID: PMC4499919

879 Nagarkar-Jaiswal S, Manivannan SN, Zuo Z, Bellen HJ. A cell cycle-independent,
880 conditional gene inactivation strategy for differentially tagging wild-type and
881 mutant cells. *Elife*. eLife Sciences Publications Limited; 2017 May 31;6:1385.
882 PMID: PMC5493436

883 Neumüller RA, Wirtz-Peitz F, Lee S, Kwon Y, Buckner M, Hoskins RA, et al.
884 Stringent analysis of gene function and protein-protein interactions using
885 fluorescently tagged genes. *Genetics*. *Genetics*; 2012 Mar;190(3):931–40.
886 PMID: PMC3296255

887 Okada A, Lansford R, Weimann JM, Fraser SE, McConnell SK. Imaging cells in the
888 developing nervous system with retrovirus expressing modified green fluorescent
889 protein. *Exp. Neurol*. 1999 Apr;156(2):394–406.

890 Paix A, Folkmann A, Goldman DH, Kulaga H, Grzelak MJ, Rasoloson D, et al.
891 Precision genome editing using synthesis-dependent repair of Cas9-induced DNA
892 breaks. *Proc. Natl. Acad. Sci. U.S.A.* National Academy of Sciences; 2017 Dec
893 12;114(50):E10745–54. PMID: PMC5740635

894 Paix A, Folkmann A, Rasoloson D, Seydoux G. High Efficiency, Homology-Directed
895 Genome Editing in *Caenorhabditis elegans* Using CRISPR-Cas9
896 Ribonucleoprotein Complexes. *Genetics*. *Genetics*; 2015 Sep;201(1):47–54.
897 PMID: PMC4566275

898 Paix A, Wang Y, Smith HE, Lee C-YS, Calidas D, Lu T, et al. Scalable and versatile
899 genome editing using linear DNAs with microhomology to Cas9 Sites in
900 *Caenorhabditis elegans*. *Genetics*. *Genetics*; 2014 Dec;198(4):1347–56. PMID:
901 PMC4256755

902 Pédelacq J-D, Cabantous S, Tran T, Terwilliger TC, Waldo GS. Engineering and
903 characterization of a superfolder green fluorescent protein. *Nat. Biotechnol*.
904 Nature Publishing Group; 2006 Jan;24(1):79–88.

905 Port F, Chen H-M, Lee T, Bullock SL. Optimized CRISPR/Cas tools for efficient
906 germline and somatic genome engineering in *Drosophila*. *Proc. Natl. Acad. Sci*.

- 907 U.S.A. National Academy of Sciences; 2014 Jul 22;111(29):E2967–76. PMID:
908 PMC4115528
- 909 Quadros RM, Miura H, Harms DW, Akatsuka H, Sato T, Aida T, et al. Easi-CRISPR:
910 a robust method for one-step generation of mice carrying conditional and insertion
911 alleles using long ssDNA donors and CRISPR ribonucleoproteins. *Genome*
912 *Biology*; 2017 May 16;;1–15.
- 913 Richardson CD, Ray GJ, DeWitt MA, Curie GL, Corn JE. *Nat. Biotechnol.* Nature
914 Publishing Group; 2016 Jan 20; 34: 339-344.
- 915 Rong YS, Golic KG. Gene targeting by homologous recombination in *Drosophila*.
916 *Science*. American Association for the Advancement of Science; 2000 Jun
917 16;288(5473):2013–8.
- 918 Santos dos G, Schroeder AJ, Goodman JL, Strelets VB, Crosby MA, Thurmond J, et
919 al. FlyBase: introduction of the *Drosophila melanogaster* Release 6 reference
920 genome assembly and large-scale migration of genome annotations. *Nucleic*
921 *Acids Res.* 2015 Jan 15;43(D1):D690–7.
- 922 Schmid-Burgk JL, Höning K, Ebert TS, Hornung V. CRISPaint allows modular base-
923 specific gene tagging using a ligase-4-dependent mechanism. *Nat Commun.*
924 Nature Publishing Group; 2016 Jul 28;7:12338. PMID: PMC4974478
- 925 Shaner NC, Campbell RE, Steinbach PA, Giepmans BNG, Palmer AE, Tsien RY.
926 Improved monomeric red, orange and yellow fluorescent proteins derived from
927 *Discosoma* sp. red fluorescent protein. *Nat. Biotechnol.* Nature Publishing Group;
928 2004 Nov 21;22(12):1567–72.
- 929 Suzuki K, Tsunekawa Y, Hernandez-Benitez R, Wu J, Zhu J, Kim EJ, et al. In vivo
930 genome editing via CRISPR/Cas9 mediated homology-independent targeted
931 integration. *Nature*. Nature Publishing Group; 2016 Dec 1;540(7631):144–9.
932 PMID: PMC5331785
- 933 Şentürk M, Bellen HJ. ScienceDirect Genetic strategies to tackle neurological
934 diseases in fruit flies. *Current Opinion in Neurobiology*. Elsevier Ltd; 2018 Jun
935 1;50:24–32. PMID: PMC5940587
- 936 Thummel C, Pirrotta V. New pCaSpeR P element vectors. *Drosoph Inf Serv*
937 1992;71:150
938
- 939 Venken KJT, Schulze KL, Haelterman NA, Pan H, He Y, Evans-Holm M, et al.
940 MiMIC: a highly versatile transposon insertion resource for engineering
941 *Drosophila melanogaster* genes. *Nat. Methods*. 2011 Sep;8(9):737–43. PMID:
942 PMC3191940
- 943 Viswanatha R, Li Z, Hu Y, Perrimon N. Pooled genome-wide CRISPR screening for
944 basal and context-specific fitness gene essentiality in *Drosophila* cells. *Elife*. eLife
945 Sciences Publications Limited; 2018 Jul 27;7:705. PMID: PMC6063728

946 Wissel S, Kieser A, Yasugi T, Duchek P, Roitinger E, Gokcezade J, et al. A
947 Combination of CRISPR/Cas9 and Standardized RNAi as a Versatile Platform for
948 the Characterization of Gene Function. *G3 (Bethesda)*. *G3: Genes, Genomes,*
949 *Genetics*; 2016 Aug 9;6(8):2467–78. PMID: PMC4978900

950 Yao X, Wang X, Liu J, Hu X, Shi L, Shen X, et al. CRISPR/Cas9 - Mediated Precise
951 Targeted Integration In Vivo Using a Double Cut Donor with Short Homology
952 Arms. *EBioMedicine*. The Authors; 2017 Jun 1;20(C):19–26.

953 Yao X, Zhang M, Wang X, Ying W, Hu X, Dai P, et al. Tild-CRISPR Allows for
954 Efficient and Precise Gene Knockin in Mouse and Human Cells. *Dev. Cell*.
955 Elsevier Inc; 2018 May 21;45(4):526–536.e5.

956 Yoon WH, Sandoval H, Nagarkar-Jaiswal S, Jaiswal M, Yamamoto S, Haelterman
957 NA, et al. Loss of Nardilysin, a Mitochondrial Co-chaperone for α -Ketoglutarate
958 Dehydrogenase, Promotes mTORC1 Activation and Neurodegeneration. *Neuron*.
959 2017 Jan 4;93(1):115–31. PMID: PMC5242142

960 Zhang K, Li Z, Jaiswal M, Bayat V, Xiong B, Sandoval H, et al. The C8ORF38
961 homologue Sicily is a cytosolic chaperone for a mitochondrial complex I subunit.
962 *J. Cell Biol.* Rockefeller University Press; 2013 Mar 18;200(6):807–20. PMID:
963 PMC3601355

964 Zhang X, Koolhaas WH, Schnorrer F. A versatile two-step CRISPR- and RMCE-
965 based strategy for efficient genome engineering in *Drosophila*. *G3 (Bethesda)*.
966 *G3: Genes, Genomes, Genetics*; 2014 Oct 15;4(12):2409–18. PMID:
967 PMC4267936

968

969

970

971

972

973

974

975

976

977

978

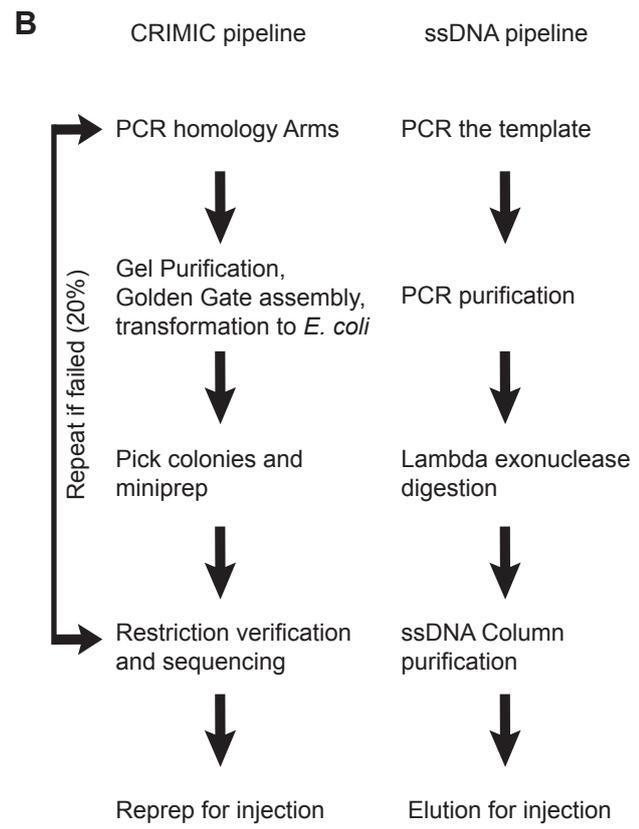
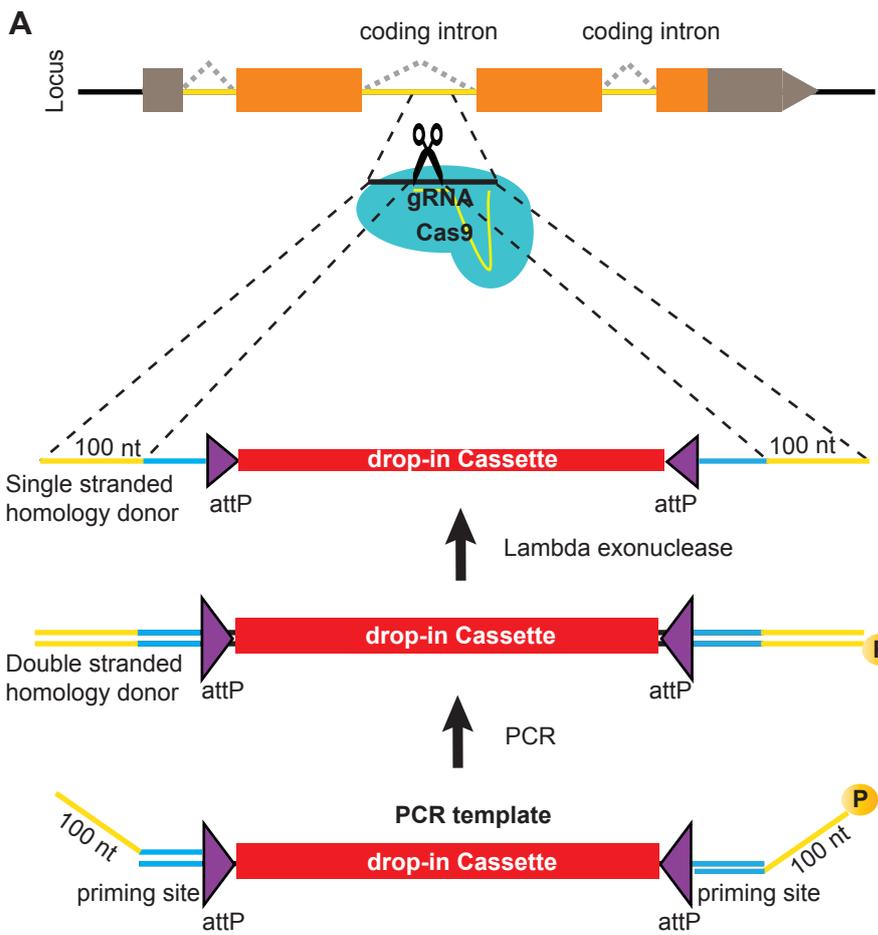
979

980

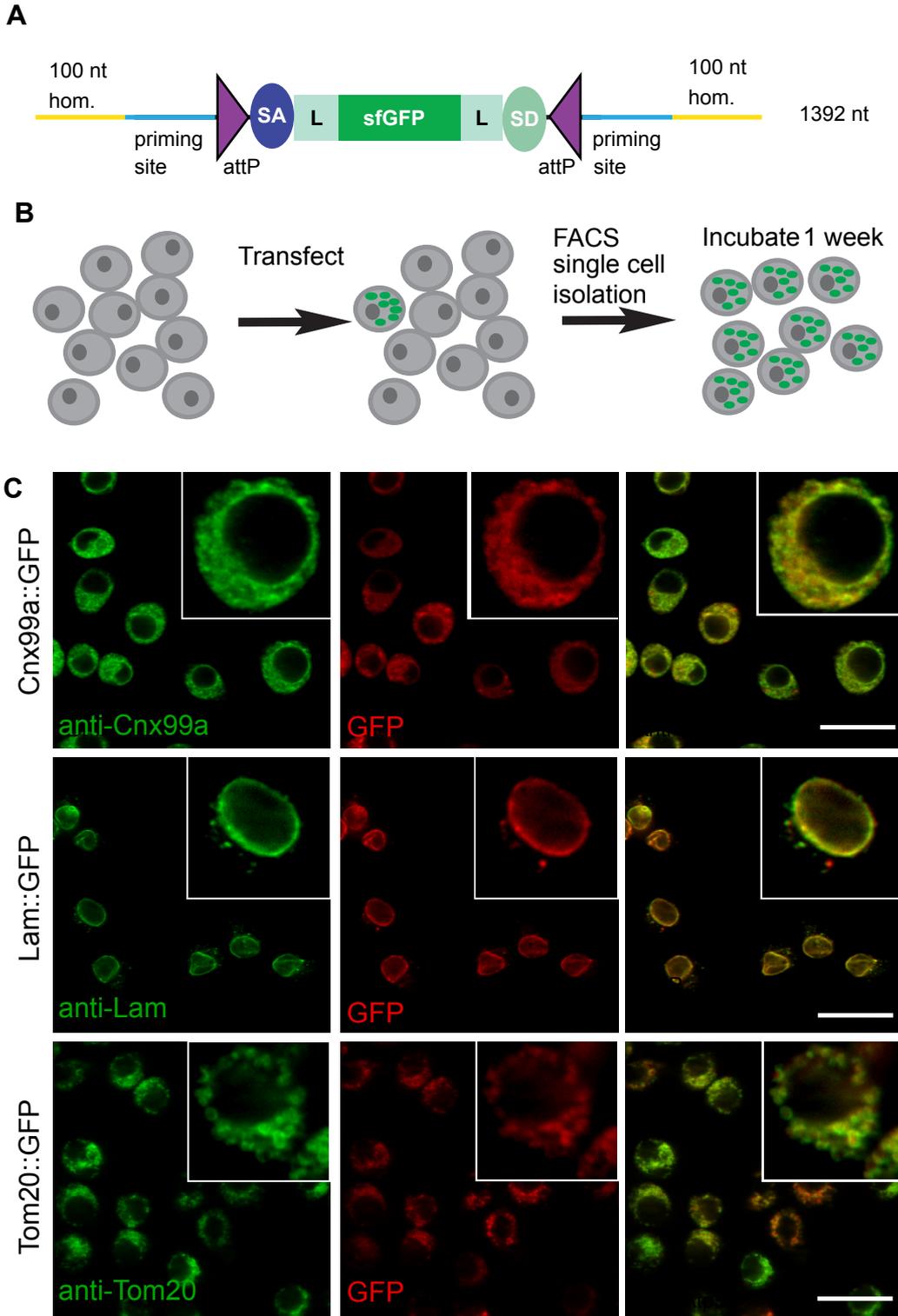
Organelle	Fly protein	clones obtained	population imaged	clones imaged	insertion sequence verified	Immunostained	Correct GFP localization	DGRC stock#	S2R+ expression (modENCODE RPKM)
Autophagosomes	Atg8a	14	✓	✓	N	N	N	N	153
Autophagosomes/aggregates	Ref(2)P	9	✓	✓	Y	a-Ref2P, a-FK2	N	N	138
Endoplasmic reticulum (ER)	Calnexin99A	16	✓	✓	Y	a-Cnx99a	Y	273	235
Endoplasmic reticulum (ER), transitional	Sec23	30	✓	✓	Y	N	*	294	101
Endosomes, early	Rab5	9	✓	✓	N	N	N	N	98
Endosomes, recycling	Rab11	23	✓	✓	Y	N	*	274	302
G-Bodies (cytoplasmic puncta)	Pfk	14	✓	✓	N	N	N	N	23
Golgi (cis-Golgi)	Gmap	10	✓	✓	Y	a-GMAP	Y	276, 277	15
Golgi (trans-Golgi)	Sec71	10	✓	✓	N	N	N	N	8
Golgi (trans-Golgi)	Golgin245	1	✓	✓	Y	a-Golgin245	Y	280	27
Kinetochore	Polo	2	✓	✓	Y	N	Y	275	50
Lipid droplets	Seipin	12	✓	✓	N	N	N	N	9
Lysosomes	spin	2	✓	✓	Y	a-Arl8	Y	293	112
Lysosomes	Arl8	9	✓	✓	Y	a-Arl8	Y	291	78
Mitochondria	Tim17b	3	✓	✓	N	a-ATP5A	N	N	266
Mitochondria	Tom20	17	✓	✓	Y	a-ATP5A	Y	302	117
Nuclear membrane, inner	dLBR	0	✓	N	N	N	N	N	42
Nuclear membrane, inner	Lamin	53	✓	✓	Y	a-Lamin	Y	292	249
Nucleolus	Fibrillarlin	14	✓	✓	Y	a-Fib	Y	278, 279	53
Peroxisomes	Pmp70	6	✓	✓	N	N	N	N	26

* = distribution was as expected, but no antibody available to test by co-stain

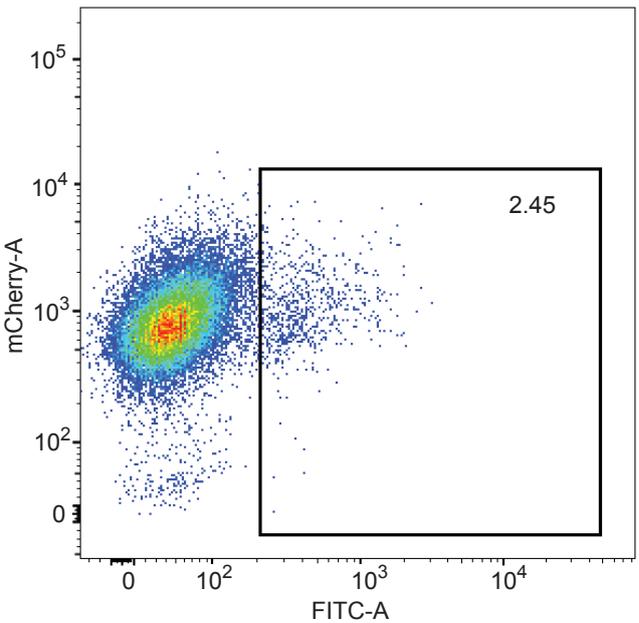
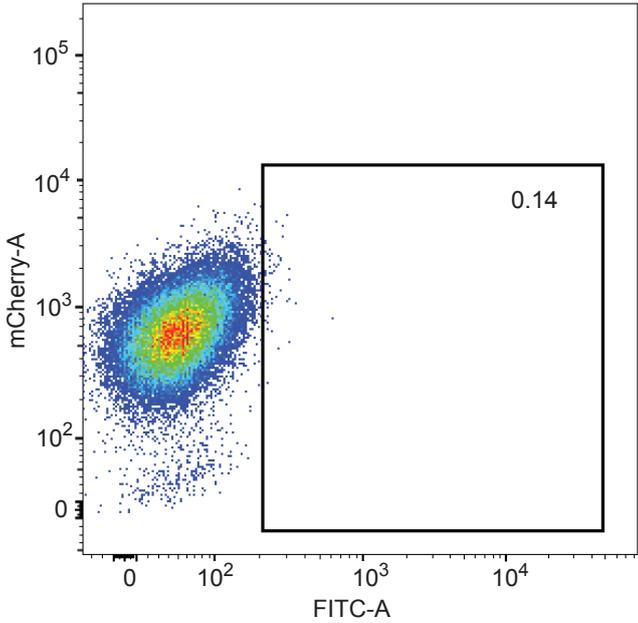
Kanca et al. Figure 1. Comparison of ssDNA pipeline to CRIMIC



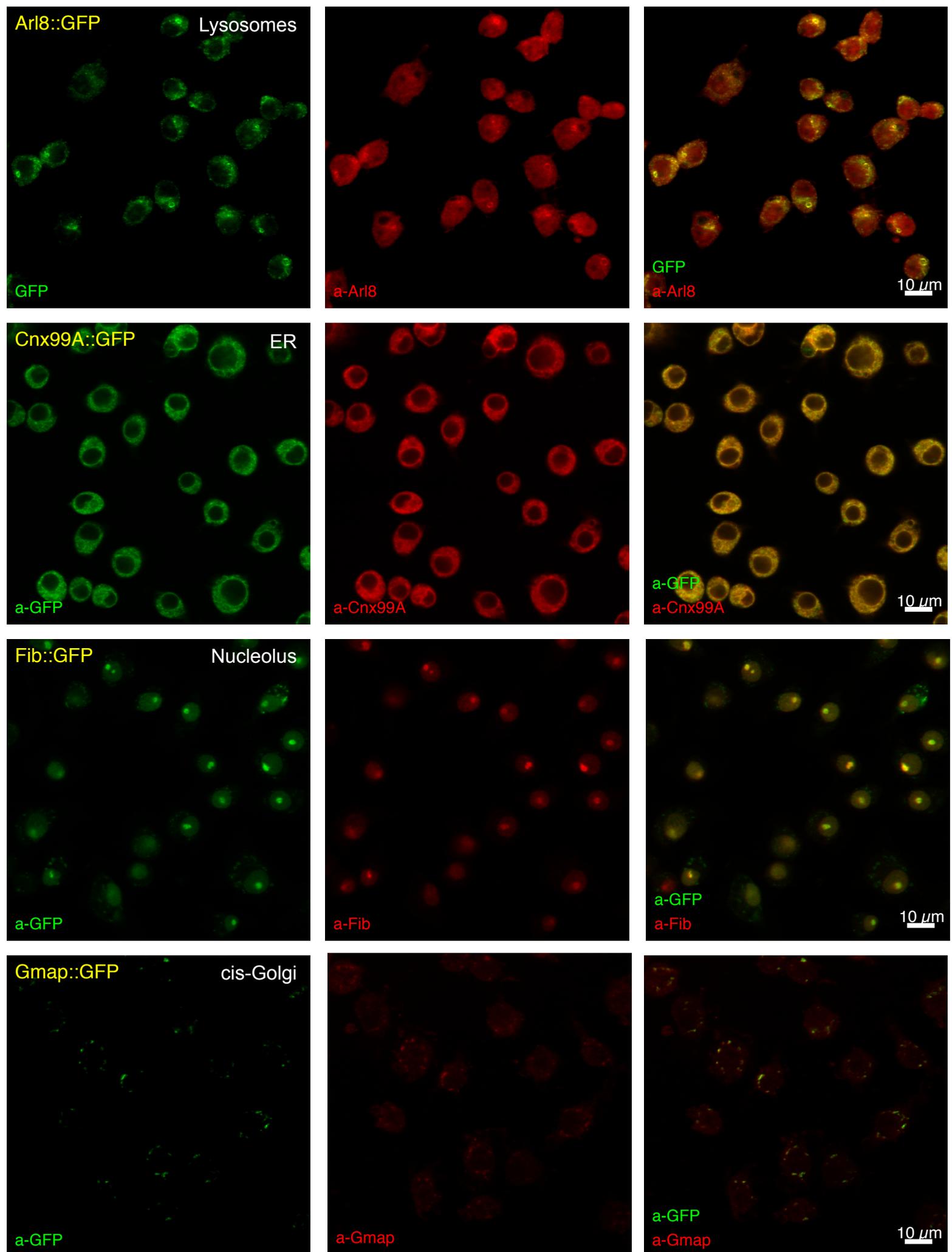
Kanca et al. Figure 2. ssDNA homology donors are effective in S2 cells to tag organelles

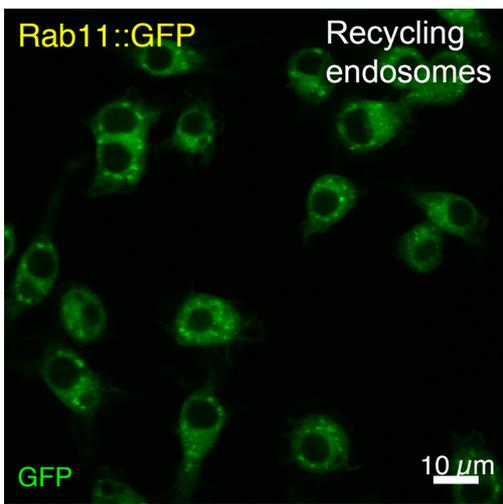
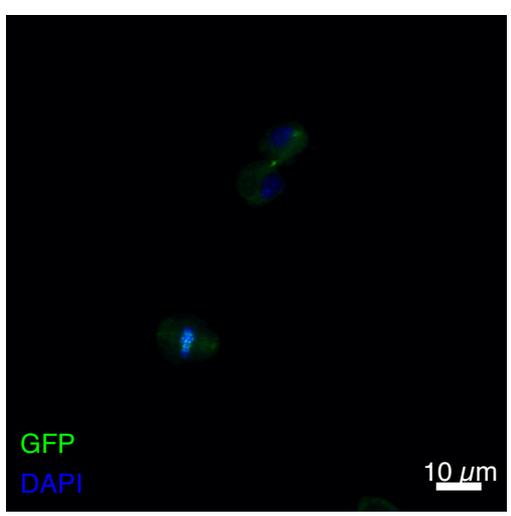
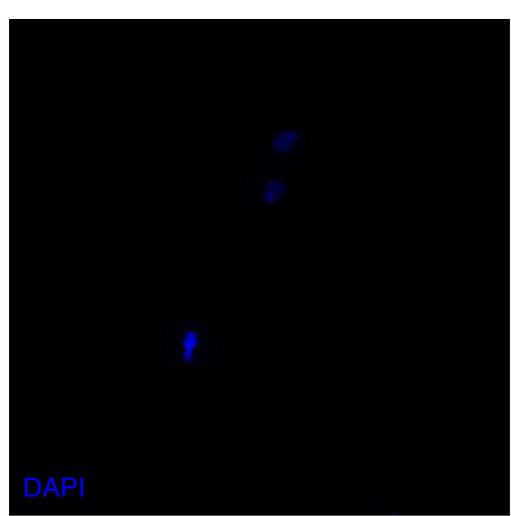
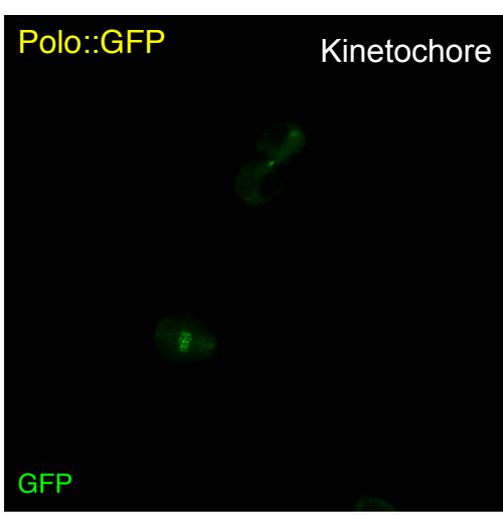
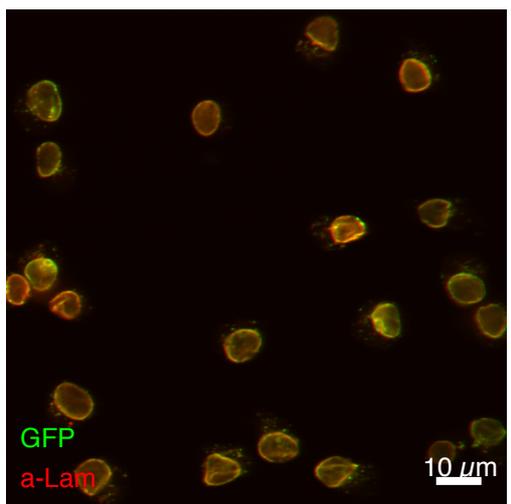
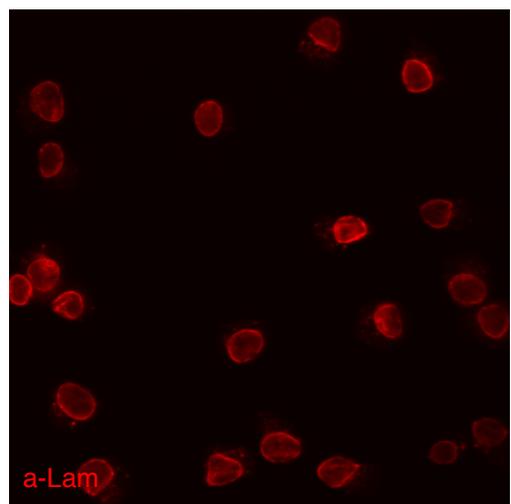
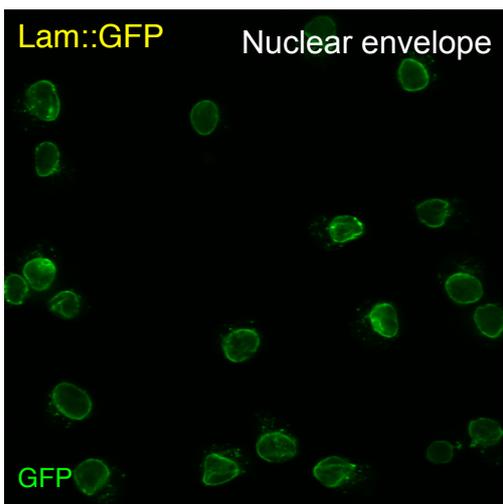
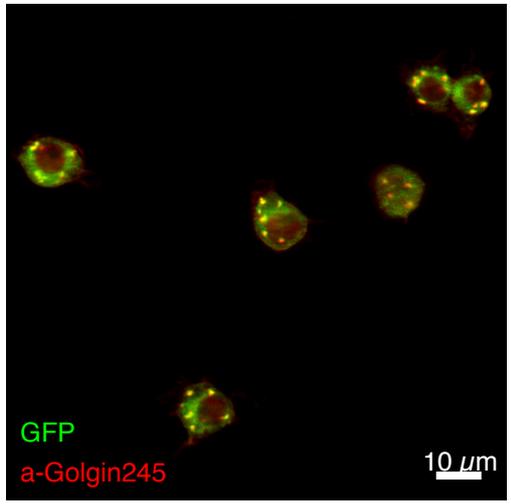
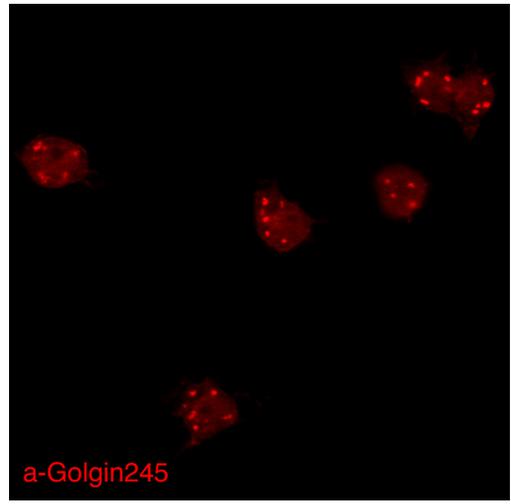
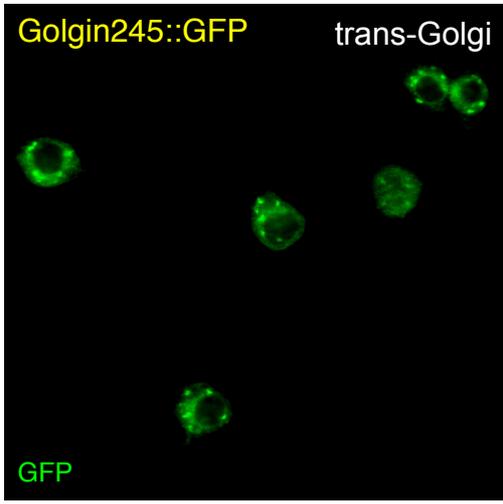


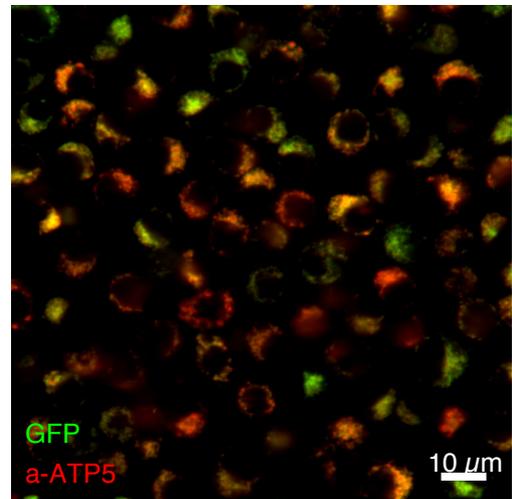
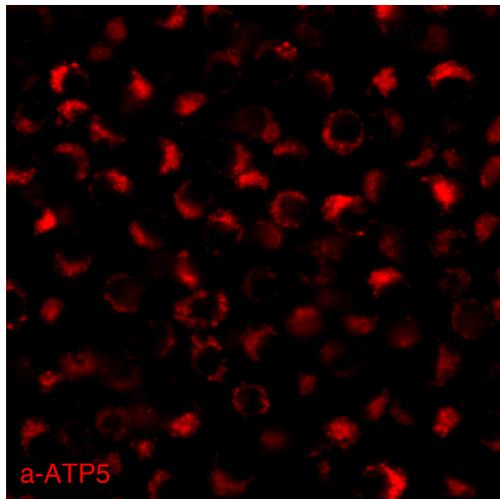
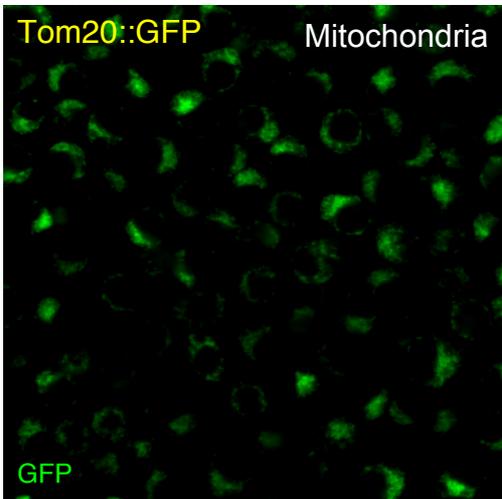
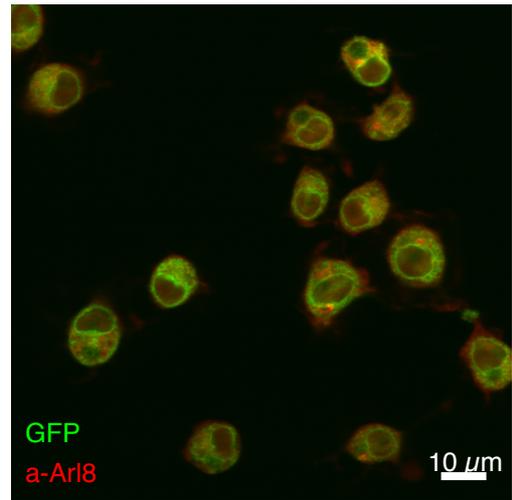
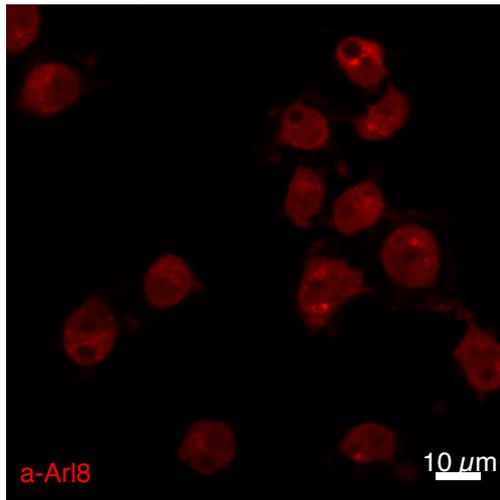
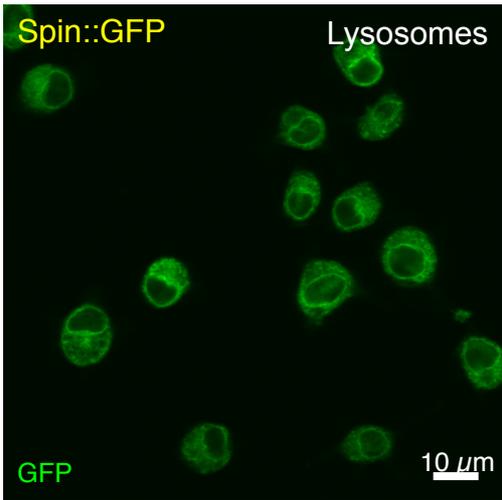
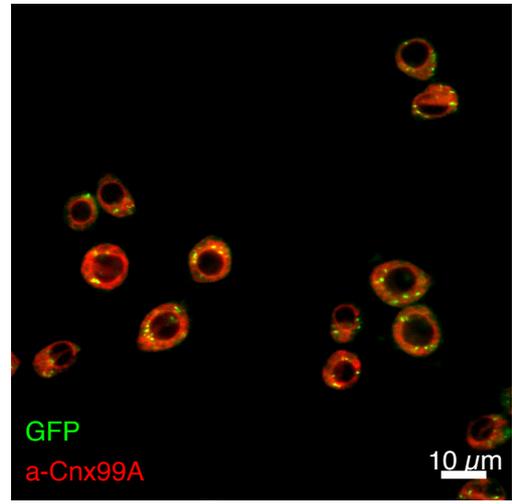
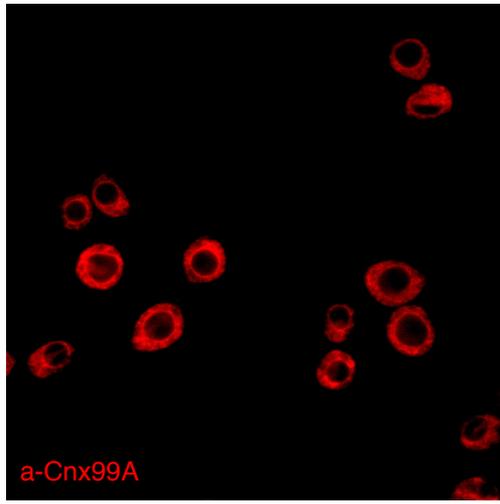
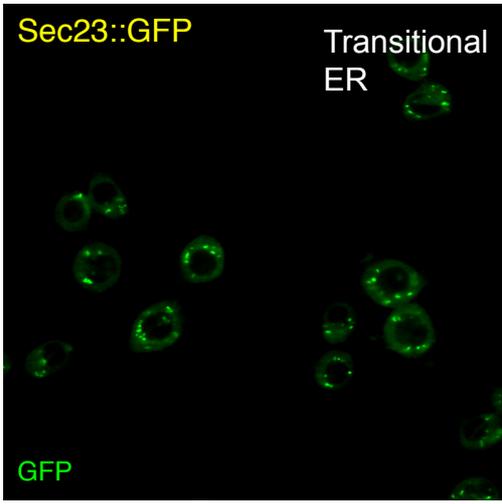
Kanca et al. Figure S1. FACS data of control cells (left) and ssDNA knock-in cells (right).



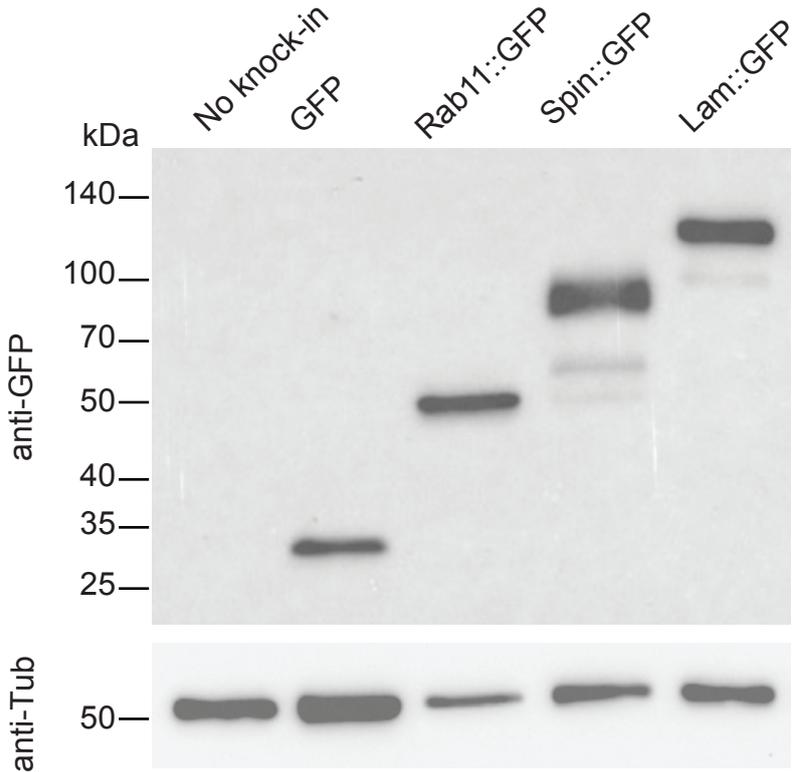
Kanca et al. Figure S2. Detection of subcellular localization of GFP tagged proteins in S2 cells







Kanca et al. Figure S3. Western blot analysis of tagged proteins observed in S2R+ cells.



Kanca et al. Figure 3. Comparison of ssDNA drop-in donor to CRIMIC construct for fly transformation

A

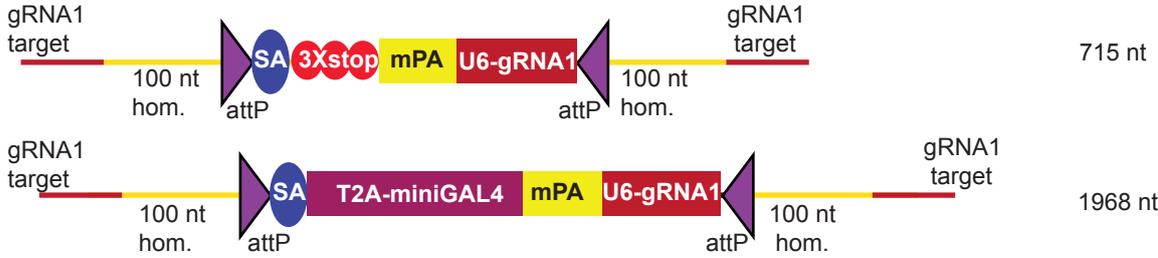


B

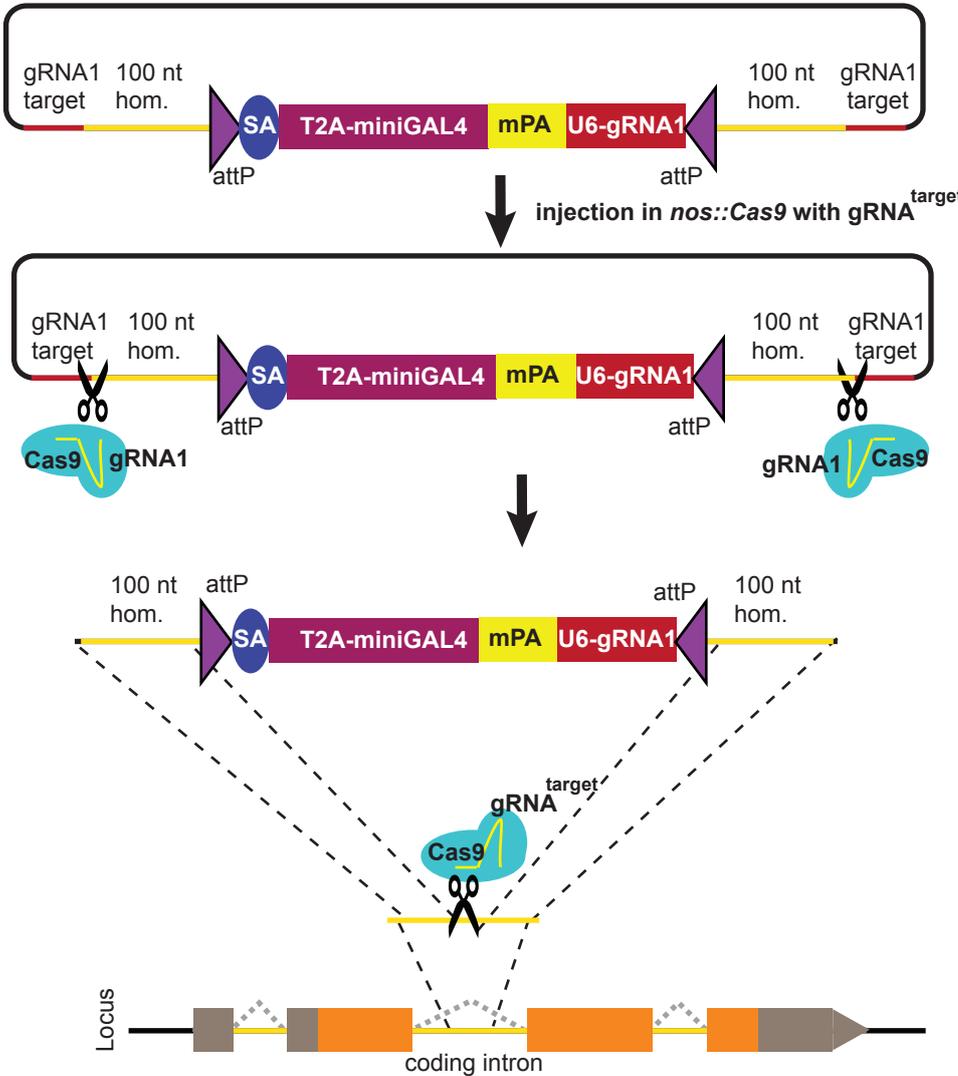
Gene	CRIMIC	attP-3XP3-EGFP-attP_ssDNA
<i>Acox57D</i>	0/ 102 crosses	0/ 120 crosses
<i>CG3376</i>	Failed at cloning	0/ 104 crosses
<i>CG5009</i>	10/ 160 crosses	1/ 82 crosses
<i>CG9527</i>	10/ 225 crosses	0/ 73 crosses
<i>CG17544</i>	0/ 185 crosses	0/ 77 crosses
<i>Cp1</i>	0/ 181 crosses	1/ 86 crosses
<i>endoB</i>	Failed at cloning	3/ 175 crosses
<i>Khc</i>	10/ 229 crosses	0/ 180 crosses
<i>Lst</i>	Failed at cloning	1/ 107 crosses
<i>NLaz</i>	7/ 162 crosses	0/ 84 crosses

Kanca et al. Figure 4. Double stranded synthetic constructs are efficient for fly transformation

A



B



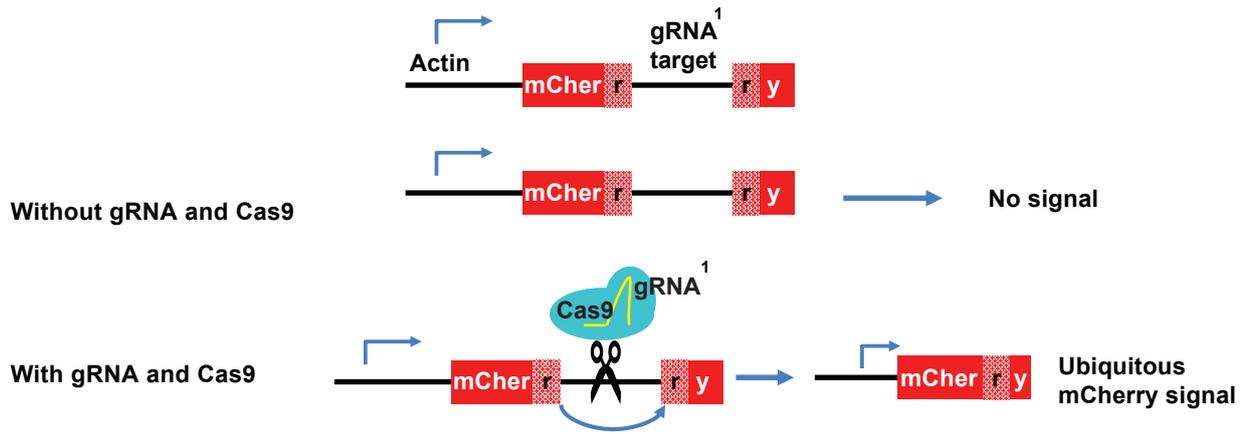
C

Gene	CRIMIC	attP-SA-stop-minipolyA-U6gRNA1-attP	attP-SA-T2A-miniGAL4-mPA-U6gRNA1-attP
<i>Acox57D</i>	0 / 102 crosses	3 / 150 crosses	10 / 208 crosses
<i>CG3376</i>	Failed at cloning	7 / 181 crosses	7 / 206 crosses
<i>CG5009</i>	10 / 160 crosses	12 / 160 crosses	21 / 136 crosses
<i>CG9527</i>	10 / 225 crosses	4 / 130 crosses	6 / 115 crosses
<i>CG17544</i>	0 / 185 crosses	4 / 160 crosses	5 / 198 crosses
<i>Cp1</i>	0 / 181 crosses	12 / 95 crosses	32 / 240 crosses
<i>endoB</i>	Failed at cloning	3 / 61 crosses	6 / 146 crosses
<i>Khc</i>	10 / 229 crosses	3 / 110 crosses	7 / 228 crosses
<i>Lst</i>	Failed at cloning	6 / 102 crosses	8 / 166 crosses
<i>NLaz</i>	7 / 162 crosses	11 / 155 crosses	17 / 161 crosses

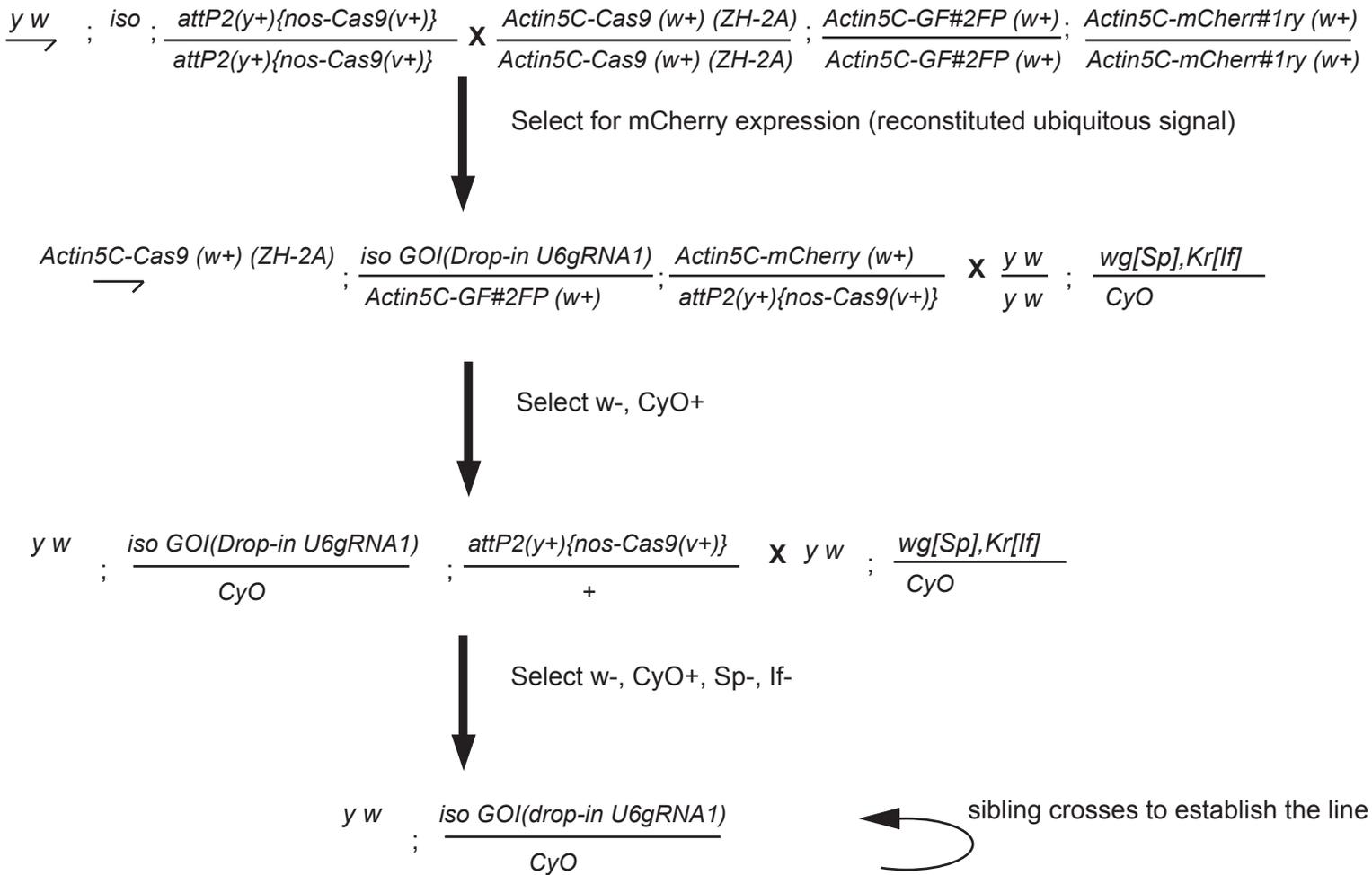
Not verified
 Verified
 Failed

Kanca et al. Figure S4. Schematic and crossing scheme for CRISPR gRNA-based dominant marker strategy.

A

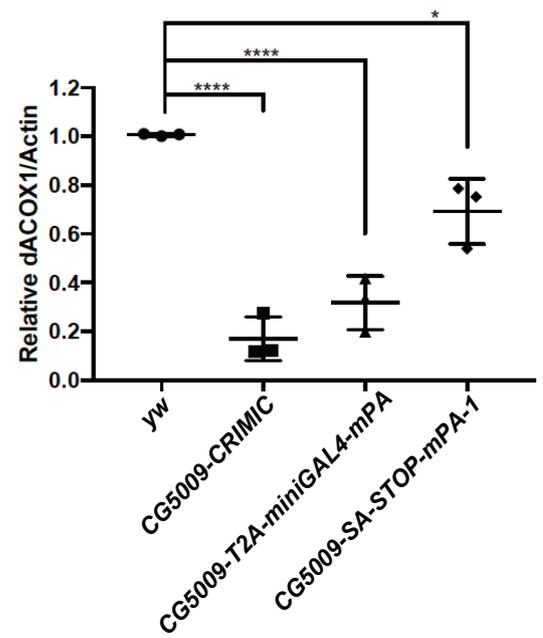
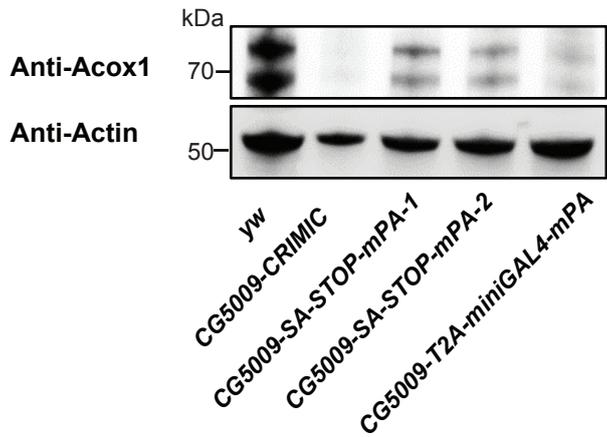


B

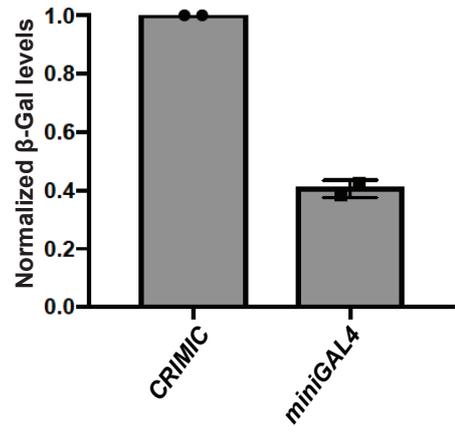
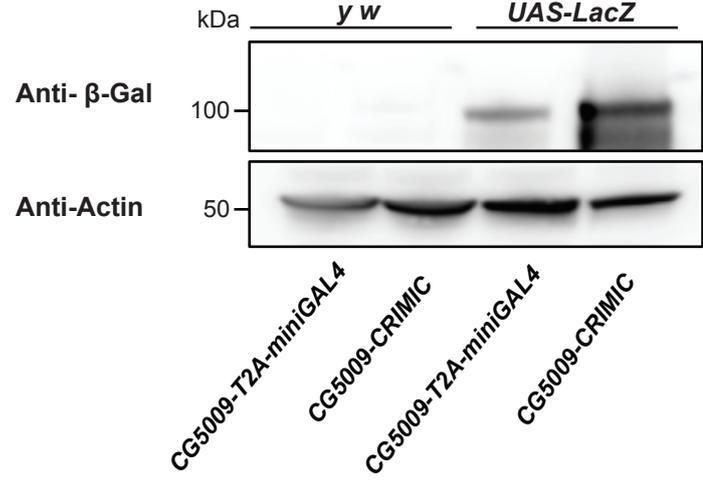


Kanca et al. Figure 5. T2A-miniGAL4 is a mutagen and expresses an active GAL4

A

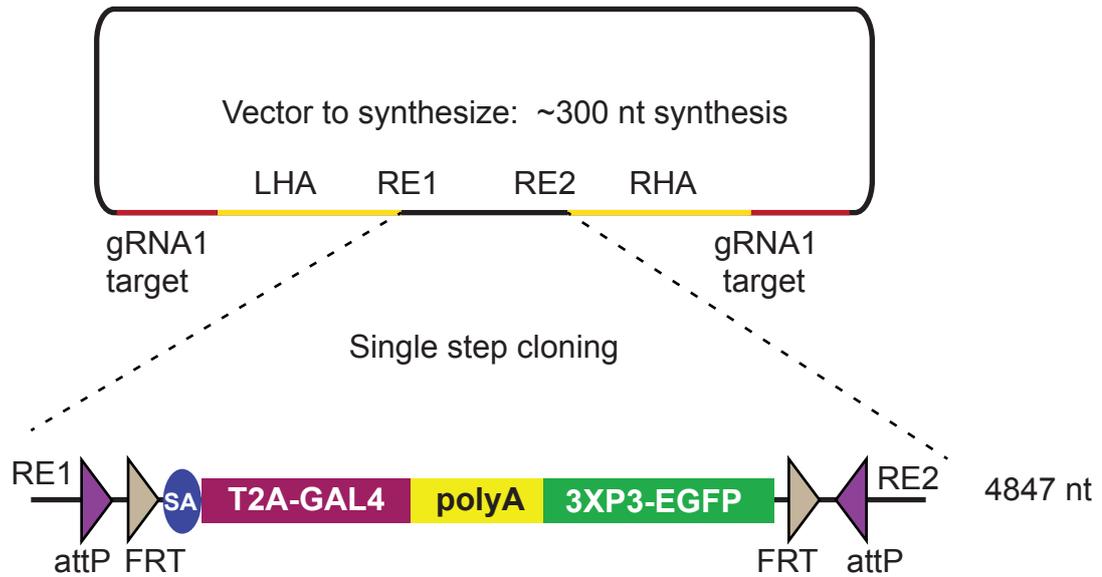


B



Kanca et al. Figure 6. Single step cloning method allows cheaper and efficient insertion of CRIMIC cassette in coding introns.

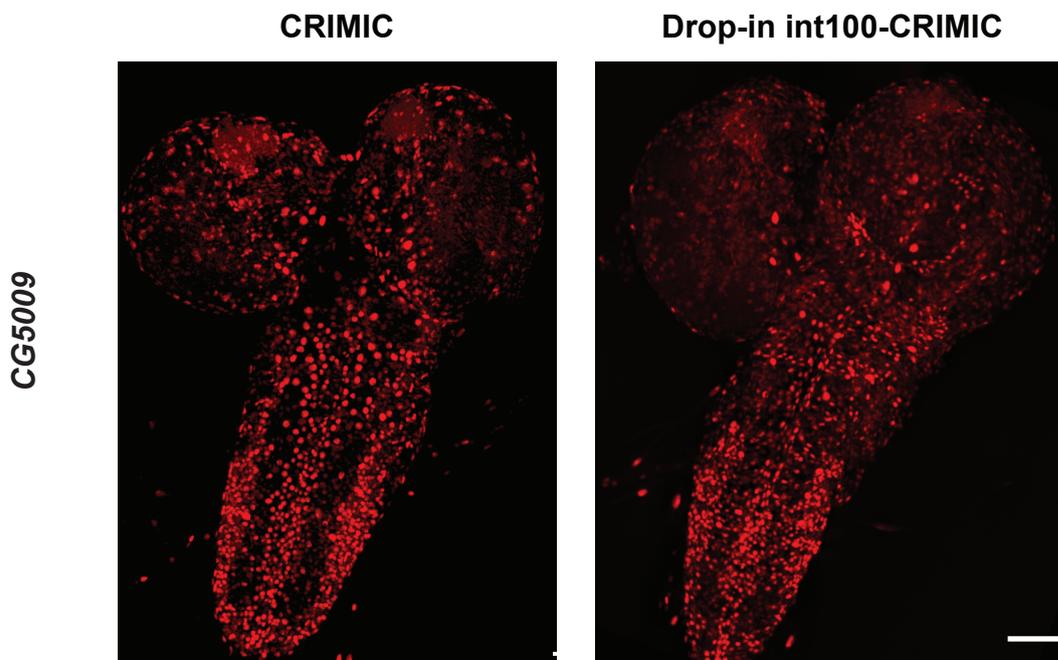
A



B

Gene	CRIMIC	Drop-in int100-CRIMIC
<i>CG5009</i>	10 / 160 crosses	22 / 140 crosses
<i>CG9527</i>	10 / 225 crosses	0 / 55 crosses
<i>Cp1</i>	0 / 181 crosses	3 / 65 crosses
<i>endoB</i>	Failed at cloning	6 / 135 crosses
<i>Khc</i>	10 / 229 crosses	3 / 110 crosses
<i>Lst</i>	Failed at cloning	0 / 90 crosses
<i>NLaz</i>	7 / 162 crosses	3 / 66 crosses

C



Kanca et al. FigureS5. Comparison of expression domain obtained by T2A-miniGAL4 and drop-in int100-CRIMIC.

