Gene knock-ins in *Drosophila* using homology-independent insertion of universal donor plasmids

Justin A. Bosch[1], Ryan Colbeth[1], Jonathan Zirin[1], Norbert Perrimon[1,2]

[1] Department of Genetics, Harvard Medical School, Boston, MA, 02115, USA
[2] Howard Hughes Medical Institute, Boston, MA, 02115, USA

Short Running Title: Homology-independent knock-ins in flies

Keywords: CRISPR-Cas9, knock-in, homology-independent, gene function, CRISPaint

100-word article summary: We report a fast and simple genomic knock-in method in *Drosophila* to insert large DNA elements into any target gene. Using CRISPR-Cas9 and non-homologous end joining (NHEJ), an entire donor plasmid is inserted into the genome without the need for homology arms. We demonstrate its usefulness in cultured cells to fluorescently tag endogenous proteins and in the fly germ line to generate heritable insertions that disrupt gene function and can act as expression reporters.

Corresponding author:
Norbert Perrimon
Harvard Medical School
77 Avenue Louis Pasteur
Dept. of Genetics, NRB 336
Boston, MA 02115
617-432-7672

Email: perrimon@genetics.med.harvard.edu

Abstract

Targeted genomic knock-ins are a valuable tool to probe gene function. However, knock-in methods involving homology-directed repair (HDR) can be laborious. Here, we adapt the mammalian CRISPaint homology-independent knock-in method for *Drosophila melanogaster*, which uses CRISPR/Cas9 and non-homologous end joining (NHEJ) to insert universal donor plasmids into the genome. This method is a simple and fast alternative to HDR for certain strategies such as C-terminal tagging and gene disruption. Using this method in cultured S2R+ cells, we efficiently tagged four endogenous proteins with the bright fluorescent protein mNeonGreen, thereby demonstrating that an existing collection of CRISPaint universal donor plasmids is compatible with insect cells. In addition, we inserted the transgenesis marker *3xP3-RFP* into seven genes in the fly germ line, producing heritable loss of function alleles that were isolated by simple fluorescence screening. Unlike in cultured cells, indels always occurred at the genomic insertion site, which prevents predictably matching the insert coding frame to the target gene. Despite this effect, we were able to isolate *T2A-Gal4* insertions in four genes that serve as in vivo expression reporters. Finally, we apply this fast knock-in method to uncharacterized small open reading frame (smORF) genes. Therefore, homology-independent insertion is a useful genome editing technique in *Drosophila* that will better enable researchers to dissect gene function.

Introduction

Insertion of DNA into the animal genome is a powerful method to study gene function. This approach is multipurpose, and can be used, for example, to alter gene function, assay gene expression, visualize protein localization, and purify endogenous proteins (AUER AND DEL BENE 2014; SINGH *et al.* 2015). Furthermore, the ability to insert large DNA elements such as promoters, protein coding sequences, or entire genes into the genome offers researchers endless options for genome modification. *Drosophila melanogaster* is an excellent animal model to analyze gene function because of its many genetic tools, fast generation time, and in vivo analysis (VENKEN *et al.* 2016; KORONA *et al.* 2017; BIER *et al.* 2018).

Two commonly used methods in *Drosophila* to insert DNA into endogenous genes are transposable DNA elements and homology directed-repair (HDR). Transposable elements insert randomly in the genome (BELLEN *et al.* 2011) and thus cannot be used to target a user-specified gene. In contrast, HDR is used to insert DNA into a precise genomic location by homologous recombination (BIER *et al.* 2018). Circular plasmids are commonly used as donor DNA for HDR because they can carry a large DNA insert (≤10kb) and homology arms corresponding to the target locus are added by molecular cloning. However, the design and construction of plasmid donors for each gene can be laborious and prone to troubleshooting. As a cloning-free alternative, synthesized single-stranded DNA (ssDNA) with short homology arms (~50-100 bp each) (BIER *et al.* 2018) or long ssDNAs of ≤2kb (QUADROS *et al.* 2017) can be used as donors. However, ssDNA donors are limited to small insertions, and, like plasmid donors, must be designed and generated for each gene that is targeted by HDR. The burden of generating unique donor DNAs for HDR makes this method inefficient for targeting many genes in parallel, such as hits from a genetic or proteomic screen, or the same gene in divergent species. Therefore, there is a need for simpler and more scalable alternatives to knock-in large DNA elements into the *Drosophila* genome.

Large DNA elements can also be inserted in target genes without homology arms, known as homology-independent insertion (CRISTEA *et al.* 2013; MARESCA *et al.* 2013; AUER *et al.* 2014; KATIC *et al.* 2015; LACKNER *et al.* 2015; SCHMID-BURGK *et al.* 2016; SUZUKI *et al.* 2016; KATOH *et al.* 2017; KUMAGAI *et al.* 2017; ZHANG *et al.* 2018; GAO *et al.* 2019). In this method, simultaneous cutting of a circular donor plasmid and a genomic target-site by a nuclease results in integration of linearized insert DNA into the genomic cut site by non-homologous end joining (NHEJ). Such donor plasmids are "universal" because they lack gene-specific homology arms, allowing insert sequences (e.g. GFP) to be targeted to any genomic location. Despite this advantage, homology-independent insertion is less precise than HDR. For example, donor DNA can insert in two directions, insertion and deletion mutations (indels) at the integration site can affect the insert translation frame, the entire plasmid can integrate, and inserts can form concatemers. Nevertheless, for some targeting strategies such as C-terminal protein tagging or gene disruption, homology-independent insertion is a fast, simple, and effective alternative to HDR. While homology-independent insertion has been successfully applied in human cell lines (CRISTEA *et al.* 2013; MARESCA *et al.* 2013; LACKNER *et al.* 2015; SCHMID-BURGK *et al.* 2016; KATOH *et al.* 2017; ZHANG *et al.* 2018), mouse somatic cells (SUZUKI *et al.* 2016; GAO *et al.* 2019), zebrafish (AUER *et al.* 2014), *C. elegans* (KATIC *et al.* 2015), and *Daphnia* (KUMAGAI *et al.* 2017), it has not yet been utilized in *Drosophila*.

Here, we use the CRISPaint strategy to show that homology-independent insertion functions effectively in *Drosophila*. First, we inserted an mNeonGreen

universal donor plasmid into the C-terminus of four target proteins in S2R+ cells and used Puromycin-resistance to select for in-frame insertions, demonstrating that this is an efficient method to produce cell lines with tagged endogenous proteins. Second, we inserted a *T2A-Gal4 3xP3-RFP* universal donor plasmid into the fly germ line and isolated knock-in lines for seven genes by simple fluorescent screening. By targeting insertions to 5' coding sequence, we demonstrate that this is an effective method to disrupt gene function and generate loss of function fly strains. Furthermore, by screening insertions for Gal4 expression, we identified four that are in vivo reporters of their target gene. Finally, we use this knock-in method to quickly generate insertions in uncharacterized small open reading frame (smORF) genes.

Materials and Methods

Plasmid cloning

*pCFD3-frame_selector_*(*0, 1, or 2*) plasmids (Addgene #127553-127555, DGRC #1482-1484) were cloned by ligating annealed oligos encoding sgRNAs that target the CRISPaint target site (SCHMID-BURGK *et al.* 2016) into *pCFD3* (PORT *et al.* 2014), which contains the *Drosophila U6:3* promoter.

Additional sgRNA-encoding plasmids were generated by the TRiP (https://fgr.hms.harvard.edu/) or obtained from Filip Port (PORT *et al.* 2015). sgRNA plasmids targeting CDS close to the stop codon were *GP07595* (*Act5c*) (Addgene #130278, DGRC #1492), *GP07596* (*His2Av*), *GP07609* (*alphaTub84B*), and *GP07612* (*Lam*). sgRNA plasmids targeting CDS close to the start codon were *GP06461* (*wg*), *GP02894* (*FK506-bp2*), *GP05054* (*alphaTub84B*), *GP00225* (*esg*), *GP00364* (*Myo1a*), *GP00400* (*btl*), *GP00583* (*Mhc*), *GP01881* (*hh*), *GP03252* (*Desat1*), *GP05302* (*ap*), *pFP545* (*ebony*), and *pFP573* (*ebony*). These sgRNAs were cloned into *pCFD3*, with the exception of those targeting *esg*, *Myo1a*, *btl*, and *Mhc*, which were cloned into *pl100* (KONDO AND UEDA 2013).

pCRISPaint-T2A-Gal4-3xP3-RFP (Addgene #127556, DGRC #1481) was constructed using Gibson assembly (E2611, NEB) of three DNA fragments: 1) *Gal4-SV40-3xP3-RFP* was PCR amplified from *pHD-Gal4-DsRed* (Xu et al. 2019 in preparation, (GRATZ *et al.* 2014); 2) linear plasmid backbone generated by digesting *pWalium10-roe* (PERKINS *et al.* 2015) with AscI/SacI; and 3) a synthesized double-stranded DNA fragment (gBlock, IDT) encoding the CRISPaint target site, linker sequence, T2A, and ends that overlap the other two fragments.

pCRISPaint-3xP3-RFP (Addgene # 130276, DGRC #1490) was constructed by digesting *pCRISPaint-T2A-Gal4-3xP3-RFP* with Not1/BamHI to remove *T2A-Gal4* sequences and using Gibson to add a gBlock with a multiple cloning site for future manipulation.

pCRISPaint-3xP3-GFP (Addgene # 130277, DGRC #1491) was constructed by Gibson assembly of PCR-amplified backbone from *pCRISPaint-3xP3-RFP* and PCR-amplified EGFP sequence from pM37 (LEE *et al.* 2018).

pCRISPaint-T2A-ORF-3xP3-RFP donor plasmids (Addgene #127557-127565, DGRC #1485-1489,1493-1497) were cloned by PCR amplifying the ORFs and Gibson cloning into *CRISPaint-T2A-Gal4-3xP3-RFP* cut with NheI/KpnI. ORF sequences were amplified from templates as follows: *sfGFP* [amplified from *pUAS-TransTimer* (HE *et al.* 2019)], *LexGAD* [amplified from p*CoinFLP-LexGAD/Gal4* (BOSCH *et al.* 2015)], *QF2* amplified from Addgene #80274, *Cas9-T2A-GFP* (amplified from template kindly provided by Raghuvir Viswanatha, Cas9 from *Streptococcus pyogenes*), *FLPo* (amplified from Addgene #24357), *Gal80* (amplified from Addgene #17748), *Nluc*

154 (amplified from Addgene #62057), *Gal4DBD*, (amplified from Addgene #26233), and
155 *p65* (amplified from Addgene #26234).
156 *pCRISPaint-sfGFP-3xP3-RFP* (Addgene #127566, DGRC #1486) was cloned by
157 PCR amplifying *sfGFP* coding sequence and Gibson cloning into *CRISPaint-T2A-Gal4-*
158 *3xP3-RFP* cut with NotI/KpnI.
159 *pCFD5-frame_selectors_0,1,2* (Addgene #131152) was cloned by PCR
160 amplifying two fragments encoding the three sgRNAs that cut the CRISPaint target site
161 and Gibson cloning into *pCFD5* (PORT AND BULLOCK 2016). This plasmid allows
162 expression of sgRNAs that are separated by tRNA sequences, and thus *pCFD5-*
163 *frame_selectors_0,1,2* expresses all three frame-selector sgRNAs simultaneously.
164 *pLHA-T2A-Gal4-3xP3-RFP-RHA_ebony* was cloned by digesting *pCRISPaint-*
165 *T2A-Gal4-3xP3-RFP* with AscI/SacI and purifying insert and backbone DNA together
166 (QIAQuick, Qiagen). Homology arms were amplified from *Drosophila* genomic DNA (*yw*,
167 single fly) using Phusion polymerase (LHA: 1592bp, RHA: 1544bp). Insert, backbone,
168 and both homology arms were assembled by Gibson cloning.
169 See Supplemental Table 7 for oligo and dsDNA sequences and Addgene and
170 DGRC for plasmid sequences.
171
172 **Cell culture**
173
174 *Drosophila* S2R+ cells stably expressing Cas9 and a mCherry protein trap in *Clic*
175 (known as PT5/Cas9) (VISWANATHA *et al.* 2018) were cultured at 25˚C using Schneider's
176 media (21720-024, ThermoFisher) with 10% FBS (A3912, Sigma) and 50 U/ml penicillin
177 strep (15070-063, ThermoFisher). S2R+ cells were transfected using Effectene
178 (301427, Qiagen) following the manufacturer's instructions. Plasmid mixes were
179 composed of sgRNA-expressing plasmids (see above) and *pCRISPaint-mNeon-PuroR*
180 (SCHMID-BURGK *et al.* 2016). Cells were transfected with plasmid mixes in 6-well dishes
181 at $1.8x10^6$ cells/ml, split at a dilution of 1:6 after 3-4 days, and incubated with 2 µg/ml
182 Puromycin (540411 Calbiochem). Every 3-5 days, the media was replaced with fresh
183 Puromycin until the cultures became confluent (~12-16 days). For single-cell cloning
184 experiments, cultures were split 1:3 two days before sorting. Cells were resuspended in
185 fresh media, triturated to break up cell clumps, and pipetted into a cell straining FACS
186 tube (352235 Corning). Single cells expressing mNeonGreen were sorted into single
187 wells of a 96 well plate containing 50% conditioned media 50% fresh media using an
188 Aria-594 instrument at the Harvard Medical School Division of Immunology's Flow
189 Cytometry Facility. Once colonies were visible by eye (3-4 weeks), they were expanded
190 and screened for mNeonGreen fluorescence.
191
192 **Fly genetics and embryo injections**
193
194 Flies were maintained on standard fly food at 25˚C. Fly stocks were obtained
195 from the Perrimon lab collection or Bloomington Stock center (indicated with BL#).
196 Stocks used in this study are as follows: *yw* (Perrimon Lab), *yw/Y hs-hid* (BL8846), *yw;*
197 *nos-Cas9attP40/CyO* (derived from BL78781), *yw;; nos-Cas9attP2* (derived from
198 BL78782), *yw; Sp hs-hid/CyO* (derived from BL7757), *yw;; Dr hs-hid/TM3,Sb* (derived
199 from BL7758), *UAS-2xGFP* (BL6874), *wg1-17/CyO* (BL2980), *wg1-8/CyO* (BL5351),
200 *Df(2L)BSC291/CyO* (BL23676), *Mhc[k10423]/CyO* (BL10995), Df(2L)H20/CyO
201 (BL3180), *Df(2L)ED8142/SM6a* (BL24135), *hh[AC]/TM3 Sb* (BL1749),
202 *Df(3R)ED5296/TM3, Sb* (BL9338), *esgG66/CyO UAS-GFP* (BL67748),
203 *Df(2R)Exel6069/CyO* (BL7551), ywCre; D/TM3, Sb (BL851), Dp(2;1)Sco[rv23];
204 Df(2L)Sco[rv23], b[1] pr[1]/CyO (BL6230).

For embryo injections, each plasmid was column purified (Qiagen) twice, eluted in injection buffer (100 μM NaPO4, 5 mM KCl), and adjusted to 200ng/μl. Plasmids were mixed equally by volume, and mixes were injected into *Drosophila* embryos using standard procedures. For targeting genes on Chr. 2, plasmid mixes were injected into *yw;; nos-Cas9attP2* embryos. For targeting genes on Chr. 3, plasmid mixes were injected into *yw; nos-Cas9attP40/CyO* embryos. Approximately 500 embryos were injected for each targeted gene.

Injected G0 flies were crossed with *yw*. We used *yw/Y hs-hid* to facilitate collecting large numbers of virgin flies by incubating larvae and pupae at 37˚C for 1hr. G1 flies were screened for RFP expression in the adult eye on a Zeiss Stemi SVII fluorescence microscope. G1 RFP+ flies were crossed with the appropriate balancer stock (*yw; Sp hs-hid/CyO* or *yw;; Dr hs-hid/TM3,Sb)*. G2 RFP+ males that were *yellow-* (to remove the *nos-Cas9* transgene) and balancer+ were crossed to virgins of the appropriate balancer stock (*yw; Sp hs-hid/CyO* or *yw;; Dr hs-hid/TM3,Sb*). G3 larvae and pupae were heat shocked at 37˚C for 1hr to eliminate the *hs-hid* chromosome, which generates a balanced stock (e.g. yw; [RFP+]/CyO).

Imaging

S2R+ cells expressing mNeonGreen were plated into wells of a glass-bottom 384 well plate (6007558, PerkinElmer). For fixed cell images, cells were incubated with 4% paraformaldehyde for 30min, washed with PBS with .1% TritonX-100 (PBT) 3x 5min each, stained with 1:1000 DAPI (D1306, ThermoFisher) and 1:1000 phalloidin-TRITC (P1951, Sigma), and washed with PBS. Plates were imaged on an IN Cell Analyzer 6000 (GE) using a 20x or 60x objective. Time-lapse videos of live mNeonGreen expressing single cell cloned lines were obtained by taking an image every minute using a 60x objective. Images were processed using Fiji software.

Wing imaginal discs from 3rd instar larvae were dissected in PBS, fixed in 4% paraformaldehyde, and permeabilized in PBT. For Wg staining, carcasses were blocked for 1hr in 5% normal goat serum (S-1000, Vector Labs) at room temp, and incubated with 1:50 mouse anti-wg (4D4, DSHB) primary antibody and 1:500 anti-mouse 488 (A-21202, Molecular Probes) secondary antibody. Primary and secondary antibody incubations were performed at 4˚C overnight. All carcasses were stained with DAPI and phalloidin-TRITC, and mounted on glass slides with vectashield (H-1000, Vector Laboratories Inc.) under a coverslip. Images of mounted wing discs were acquired on a Zeiss 780 confocal microscope.

Larvae, pupae, and adult flies were imaged using a Zeiss Axio Zoom V16 fluorescence microscope.

Quantification of mNeonGreen expressing S2R+ cells

For FACs-based cell counting, we collected cultures from each gene knock-in experiment before and after puromycin selection. Pre-selection cultures were obtained by collecting of 500ul of culture 3-4 days after transfection. Post-selection cultures were obtained after at least 2 weeks of puromycin incubation. Non-transfected cells were used as a negative control. 100,000 cells were counted for each sample and FlowJo software was used to analyze and graph the data. FSC-A vs GFP-A was plotted and we defined mNeonGreen+ cells by setting a signal intensity threshold where <0.02% of negative controls are counted due to autofluorescence.

For microscopy-based cell counting, the number of mNeonGreen cells was quantified by analyzing confocal images in Fiji using the manual Cell Counter Plugin

256 (model). For transfected cells, 6 fields containing at least 200 cells were quantified (i.e.
257 n=6). For puro-selected cells, 3 fields containing at least 200 cells were quantified (i.e.
258 n=3).
259
260 **Western blotting**
261
262 Single cell-cloned cell lines were grown until confluent and 1ml of resuspended
263 cells was centrifuged at 250g for 10min. The cell pellet was resuspended in 1ml ice cold
264 PBS, re-centrifuged, and the pellet was lysed in 250ul 2x SDS-Sample buffer and boiled
265 for 5min. 10ul was loaded on a 4-20% Mini-Protean TGX SDS-Page gel (4561096,
266 BioRad), transferred to PVDF membrane (IPFL00010, Millipore), blocked in 5% non-fat
267 dry milk, primary blotting using mouse anti-mNeonGreen (1:1000, Chromtek 32F6) or
268 hFAB™ Rhodamine Anti-Actin (12004164 BioRad), and secondary blotting using
269 1:3000 anti-mouse HRP (NXA931, Amersham), imaging using ECL (34580,
270 ThermoFisher) on a ChemiDoc MP Imaging System (BioRad).
271
272 **PCR, sequencing, and sgRNA cutting assays**
273
274 S2R+ cell genomic DNA was isolated using QuickExtract (QE09050, Lucigen).
275 Fly genomic DNA was isolated by grinding a single fly in 50µl squishing buffer (10 mM
276 Tris-Cl pH 8.2, 1 mM EDTA, 25 mM NaCl) with 200µg/ml Proteinase K (3115879001,
277 Roche), incubating at 37˚C for 30 min, and 95˚C for 2 minutes. PCR was performed
278 using Taq polymerase (TAKR001C, ClonTech) when running DNA fragments on a gel,
279 and Phusion polymerase (M-0530, NEB) was used when DNA fragments were
280 sequenced. DNA fragments corresponding to *mNeonGreen* or *T2A-Gal4* insertion sites
281 were amplified using primer pairs where one primer binds to genomic sequence and the
282 other primer binds to the insert. For amplifying non-knock-in sites, we used primers that
283 flank the sgRNA target site. Primer pairs used for gel analysis and/or Sanger
284 sequencing were designed to produce DNA fragments <1kb. Primer pairs used for next-
285 generation sequencing of the insertion site were designed to produce DNA fragments
286 200-280bp. DNA fragments were run on a 1% agarose gel for imaging or purified on
287 QIAquick columns (28115, Qiagen) for sequencing analysis. See Supplemental Table 7
288 for oligo sequences.
289 Sanger sequencing was performed at the DF/HCC DNA Resource Core facility
290 and chromatograms were analyzed using Lasergene 13 software (DNASTAR). Next-
291 generation sequencing was performed at the MGH CCIB DNA Core. Fastq files were
292 analyzed using CRISPresso2 (CLEMENT *et al.* 2019) by entering the PCR fragment
293 sequence into the exon specification window and setting the window size to 10 bases.
294 Quantification of insertion types (seamless, in-frame indel, and frameshift indel) was
295 taken from the allele plot and frameshift analysis outputs of CRISPresso2. The small
296 proportion of "unmodified" reads that were not called by frameshift analysis were not
297 included in the quantification.
298 T7 endonuclease assays (M0302L, NEB) were performed following the
299 manufacturer instructions.
300 Splinkerette sequencing was performed as previously described (POTTER AND
301 LUO 2010). Briefly, genomic DNA was isolated by single fly squishing (described above)
302 and digested using enzyme BstYI (R0523S, NEB) at 60˚C overnight and heat
303 inactivated for 20min at 80˚C. Digested DNA was ligated with annealed splinkerette
304 oligonucleotides overnight at 16˚C. PCR annealing temperatures were 60˚C (round 1)
305 and 64˚C (round 2). Round 2 PCR products were either run on an agarose gel and gel
306 fragments purified using QIAquick Gel Extraction Kit (28704, Qiagen), or purified using

ExoSAP-IT PCR clean up reagent (78200.200.UL, Applied Biosystems). Oligo sequences used for splinkerette sequencing are listed in Supplemental Table 7. 3' splinkerette sequence traces, using the *M13R* sequencing primer, frequently mapped to *vermillion*, even when genomic DNA was isolated from yw flies. We believe this resulted from contamination, since commonly used plasmids and fly lines (e.g. *pCFD3* and *pValium20*) have *M13R* sequence adjacent to the *vermillion+* marker transgene.

Data availability

Plasmids generated in this paper will be available from Addgene and the Drosophila Genomics Resource Center (DGRC). Gal4-expressing fly strains will be available at the Bloomington Drosophila Stock Center (BDSC). Two mNeonGreen-expressing S2R+ cell lines are available at the DGRC (His2Av-mNeonGreen clone B11, DGRC #295; Lamin-mNeonGreen clone D6, DGRC #296). The remaining flies, cells, and sequence data are available from the Perrimon lab on request. Oligo and dsDNA sequences are listed in Supplemental Table 7.

Results

To test if homology-independent insertion works in *Drosophila*, we implemented a strategy known as CRISPaint (SCHMID-BURGK *et al.* 2016). This system is used to insert sequence encoding a protein tag or reporter gene into the coding sequence of an endogenous gene. Although it was originally developed for mammalian cell culture, CRISPaint has several advantages for use in *Drosophila*. First, this system uses CRISPR-Cas9 to induce double-strand breaks (DSBs), which is known to function efficiently in *Drosophila* cultured cells (BOTTCHER *et al.* 2014; VISWANATHA *et al.* 2018) and the germ line (KONDO AND UEDA 2013; REN *et al.* 2013; YU *et al.* 2013; BASSETT *et al.* 2014). Second, its use of a modular frame-selector system makes it simple to obtain insertions that are translated with the target gene. Third, a collection of existing CRISPaint donor plasmids (SCHMID-BURGK *et al.* 2016) containing common tags (e.g. GFP, RFP, Luciferase) are seemingly compatible for expression in *Drosophila*.

The CRISPaint system works by introducing three components into Cas9-expressing cells: 1) a single guide RNA (sgRNA) targeting a genomic locus; 2) a donor plasmid containing an insert sequence; and 3) a frame selector sgRNA targeting the donor plasmid (Figure 1A). This causes cleavage of both the genomic locus and the donor plasmid, leading to the integration of the entire linearized donor plasmid into the genomic cut site by non-homologous end joining (NHEJ). This insertion site destroys both sgRNA target sites and will no longer be cut. While indels can occur at the insertion site due to error-prone NHEJ, a remarkably high percentage of insertions in human cells are seamless (48-86%) (SCHMID-BURGK *et al.* 2016). Therefore, users have some control over their insertion frame by using one of three frame-selector sgRNAs. Importantly, these frame-selector sgRNAs do not target the *Drosophila* genome.

Homology-independent insertion functions efficiently in *Drosophila* S2R+ cells to tag endogenous proteins

To test the CRISPaint method in *Drosophila*, we set out to replicate the findings of (SCHMID-BURGK *et al.* 2016) in cultured S2R+ cells by tagging endogenous proteins at their C-terminus. To accomplish this, we generated plasmids expressing frame-selector sgRNAs (frame 0,1, or 2) under the control of *Drosophila U6* sequences (PORT *et al.* 2015) (Figure 1A). In addition, we generated plasmids expressing sgRNAs that target

358 the 3' coding sequence of *Drosophila* genes. We chose to target *Actin5c*, *His2Av*,
359 *alphaTub84B*, and *Lamin* because these genes are expressed in S2R+ cells (HU *et al.*
360 2017) and encode proteins with known subcellular localization (actin filaments,
361 chromatin, microtubules, nuclear envelope, respectively). For donor plasmid, we used
362 *pCRISPaint-mNeonGreen-T2A-PuroR* (SCHMID-BURGK *et al.* 2016), which contains a
363 frame-selector sgRNA target site upstream of coding sequence for the fluorescent
364 mNeonGreen protein and Puromycin resistance protein (PuroR) linked by a cleavable
365 T2A peptide sequence. Only integration of the donor plasmid in-frame with the target
366 coding sequence should result in translation of mNeonGreen-T2A-PuroR (Figure 1A).
367 We transfected Cas9-expressing S2R+ cells (VISWANATHA *et al.* 2018) with a mix
368 of three plasmids: *pCRISPaint-mNeonGreen-T2A-PuroR* donor, target-gene sgRNA,
369 and the appropriate frame-selector sgRNA (Figure 1A, Supplemental Table 1). As an
370 initial method to detect knock-in events, we used PCR to amplify the predicted insertion
371 sites from transfected cells. Using primers that are specific to the target gene and
372 *mNeonGreen* sequence, we successfully amplified sense-orientation *gene-*
373 *mNeonGreen* DNA fragments for all four genes (Figure 1B). Furthermore, next-
374 generation sequencing of these amplified fragments revealed that 34-50% of sense-
375 orientation insertions are in frame with the target gene (Figure 1B, Supplemental Figure
376 1, Supplemental Table 1). 7-27% of sense-orientation insertions are seamless, which is
377 slightly lower than previously observed (SCHMID-BURGK *et al.* 2016).
378 Next, we quantified in-frame knock-in frequency by measuring mNeonGreen
379 fluorescence in transfected S2R+ cells. Flow cytometry-based cell counting of
380 transfected cells revealed that the number of mNeonGreen+ cells range from 0.19-2.4%
381 (Figure 1C, Supplemental Table 1), in agreement with human cultured cells (SCHMID-
382 BURGK *et al.* 2016). These results were confirmed by confocal analysis of transfected
383 cells, which showed mNeonGreen fluorescence in a small subset of cells (Figure 1C).
384 3.2% (*Act5c*) and 2.4% (*His2Av*) of transfected cells expressed mNeonGreen (Figure
385 1C, Supplemental Table 1), which roughly agreed with flow cytometry cell counting.
386 Finally, mNeonGreen localized to the expected subcellular compartments, most
387 obviously observed by His2Av-mNeonGreen and Lam-mNeonGreen co-localization with
388 the nucleus, and Act5c-mNeonGreen and alphatub-mNeonGreen exclusion from the
389 nucleus (Figure 1C). These results suggest that a significant number of transfected
390 S2R+ cells received in-frame insertion of mNeonGreen at their C-terminus using the
391 CRISPaint homology-independent insertion method.
392 For knock-in cells to be useful in experiments, it is important to derive cultures
393 where most cells, if not all, carry the insertion. Therefore, we enriched for in-frame
394 insertion events using Puromycin selection (Figure 1D). After a two-week incubation of
395 transfected S2R+ cells with Puromycin, flow-cytometry and confocal analysis revealed
396 that 31.7-89.1% of cells expressed mNeonGreen and exhibited correct subcellular
397 localization (Figure 1E, Supplemental Table 1). For *alphaTub84B*, cell counting by flow-
398 cytometry greatly underestimated the number of mNeonGreen+ cells counted by
399 confocal analysis, likely because mNeonGreen expression level was so low. These
400 results demonstrate that Puromycin selection is a fast and efficient method of selecting
401 for mNeonGreen expressing knock-in cells.
402 A subset of cells in Puro-selected cultures had no mNeonGreen expression or
403 unexpected localization (Figure 1E). Since each culture is composed of different cells
404 with independent insertion events, we used FACs to derive single-cell cloned lines
405 expressing mNeonGreen for further characterization (Figure 2A). At least 14 single-cell
406 cloned lines were isolated for each target gene and imaged by confocal microscopy.
407 Within a given clonal culture, each cell exhibited the same mNeonGreen localization
408 (Figure 2B), confirming our single-cell cloning approach and demonstrating that the

409 insertion is genetically stable over many cell divisions. Importantly, while many clones
410 exhibited the predicted mNeonGreen localization, a subset of the clonal cell lines
411 displayed an unusual localization pattern (Figure 2B). For example, 3/21 Act5c-
412 mNeonGreen clones had localization in prominent rod structures, and 12/14 Lamin-
413 mNeonGreen clones had asymmetric localization in the nuclear envelope (Figure 2B).
414 In addition, some clones had diffuse mNeonGreen localization in the cytoplasm and
415 nucleus (Figure 2B).
416 To better characterize the insertions in single cell-cloned lines, we further
417 analyzed three clones per gene (12 total), selecting different classes when possible
418 (correct localization, unusual localization, and diffuse localization) (Figure 2C,
419 Supplemental Table 2). Using PCR amplification of the predicted insertion site (Figure
420 1A, Figure 2D) and sequencing of amplified fragments (Figure 2E, Supplemental Table
421 2), we determined that all clones with correct or unusual mNeonGreen localization
422 contained an in-frame insertion of mNeonGreen with the target gene. In contrast, we
423 were unable to amplify DNA fragments from the expected insertion site in clones with
424 diffuse mNeonGreen localization (Figure 2D). Western blotting of cell lysates confirmed
425 that only clones with in-frame mNeonGreen insertion express fusion proteins that match
426 the predicted molecular weights (Figure 2F). All together, these results suggest that
427 clones with correct mNeonGreen localization are likely to contain an in-frame insert in
428 the correct target gene.
429 S2R+ cells are polyploid (LEE *et al.* 2014), and clones expressing mNeonGreen
430 could bear one or more insertions. Furthermore, indels induced at the non-insertion
431 locus could disrupt protein function. To explore these possibilities, we amplified the non-
432 insertion locus in our single-cell cloned lines and used Sanger and next-generation
433 sequencing to analyze the DNA fragments (Figure 2D, Supplemental Table 2). For each
434 gene, we could find indels occurring at the non-insertion sgRNA cut site. For example,
435 we could distinguish four distinct alleles in clone B11: a 3bp deletion, a 2bp deletion, a
436 1bp deletion, and a 27bp deletion. In addition, we identified an unusual mutation in
437 clone C6, where a 1482bp DNA fragment inserted at the sgRNA cut site, which
438 corresponds to a region from *alphatub84D*. We assume that this large insertion was
439 caused by homologous recombination, since *alphatub84D* and *alphatub84B* share 92%
440 genomic sequence identity (FlyBase, http://flybase.org/). For *Act5c-mNeonGreen* clones
441 A5 and A19, numerous indel sequences were found, suggesting this region has an
442 abnormal number of gene copies. We were unable to amplify a DNA fragment from
443 *Lam-mNeonGreen* D9, despite follow-up PCRs using primers that bind genomic
444 sequence further away from the insertion site (not shown).
445
446 **Loss of function knock-in fly lines by homology-independent insertion in the**
447 **germ line**
448
449 We next explored whether homology-independent insertion could function in the
450 *Drosophila* germ line for the purpose of generating knock-in fly strains. As opposed to
451 antibiotic selection in cultured cells, we wanted to identify transgenic flies using a visible
452 body marker that was not dependent on target gene expression. In addition, we wanted
453 to target insertions to 5' coding sequence to create loss of function alleles by premature
454 protein truncation. Finally, we wanted to test if insertion of a reporter could capture the
455 expression of the target gene. To accomplish these goals, we constructed a new
456 universal donor plasmid called *pCRISPaint-T2A-Gal4-3xP3-RFP* (Figure 3A). This
457 donor contains a frame-selector sgRNA target site upstream of *T2A-Gal4*, which
458 encodes a self-cleaving form of the transcription factor (DIAO AND WHITE 2012). This

459 donor plasmid also contains the transgenesis marker *3xP3-RFP*, which expresses red
460 fluorescence in larval tissues and the adult eye (BERGHAMMER *et al.* 1999) (Figure 3A).
461 To test if homology-independent insertion functions in the germ line and
462 determine its approximate efficiency, we targeted 11 genes using *pCRISPaint-T2A-*
463 *Gal4-3xP3-RFP* (Supplemental Table 3). These genes were selected based on their
464 known loss of function phenotypes, expression patterns, and availability of sgRNA
465 plasmids from the TRiP (https://fgr.hms.harvard.edu/). Plasmid mixes of donor, frame-
466 selector sgRNA, and target gene sgRNA were injected into *nos-Cas9* embryos, and the
467 resulting G0 progeny were crossed to *yw*. G1 progeny were screened for RFP
468 fluorescence in adult eyes, and each RFP+ founder fly was crossed to balancer flies to
469 establish a stable stock. Figure 3C and Supplemental Table 3 shows the integration
470 efficiency results for each gene and Supplemental Table 4 has information on 20 RFP+
471 lines, each derived from a different G0 founder. For injections that produced RFP+
472 animals, the frequency of G0 crosses yielding RFP+ G1 progeny ranges from 5-21%
473 (Figure 3C, Supplemental Table 3). For example, when targeting *ebony* with sgRNA
474 *pFP545*, 3 out of 16 G0 crosses produced ≥1 RFP+ G1 flies. In one month following
475 embryo injection, we obtained balanced RFP+ lines for seven out of 11 genes targeted
476 (64%) (Figure 3C, Supplemental Table 3).
477 Next, we used PCR and sequencing to confirm the insertion sites in our RFP+
478 lines. For each target site, we used two primer pairs to detect the presence of an on-
479 target insertion as well as its orientation (Supplemental Figure 2). Gel images of PCR
480 amplified DNA showed that all 20 of our RFP+ lines contained an insertion in the correct
481 target site (Supplemental Table 4, Supplemental Figure 2), where 13 insertions were in
482 the sense orientation and seven were antisense (Figure 3D). Sequence analysis of PCR
483 fragments confirmed that the insertions were present at the target site (Supplemental
484 Table 4, Supplemental Figure 3). Unexpectedly, all 20 lines contained genomic
485 sequence indels at the insertion site (Supplemental Figure 3), unlike the frequent
486 seamless insertions observed in cultured cells. Germ line indels were also
487 predominantly longer than indels in cultured cells. In particular, a remarkably long
488 1896bp genomic deletion was found at the *hh* insertion site. Regardless, these results
489 demonstrate that all 20 independently derived RFP+ lines contained *pCRISPaint-T2A-*
490 *Gal4-3xP3-RFP* at the target site.
491 Next, we analyzed insertion lines for loss of function phenotypes. Flies with
492 insertions in *wg*, *Mhc*, *hh*, and *esg* were homozygous lethal (Supplemental Table 4),
493 which is consistent with characterized mutations in these genes (FlyBase). Similarly,
494 complementation tests revealed that insertions in *wg*, *Mhc*, *hh*, and *esg* were lethal in
495 trans with loss of function alleles or genomic deletions spanning the target gene
496 (Supplemental Table 4). In addition, homozygous insertions in *hh* and *wg* cause a "lawn
497 of denticles" phenotype in embryos (BEJSOVEC AND MARTINEZ ARIAS 1991) (Figure 3F)
498 and flies with heterozygous insertions in *Mhc* were flightless (MOGAMI *et al.* 1986) (data
499 not shown), all reminiscent of classical allele phenotypes. Four insertions in *ebony*
500 produced flies with dark cuticle pigment when homozygous (Figure 3E, Supplemental
501 Table 4) or in trans with *ebony*^{1} (not shown). In one case (*ebony pFP545 #2),* the
502 insertion was homozygous lethal, but was viable over *ebony*^{1} and trans-het flies
503 exhibited dark cuticle pigment. An insertion in *myo1a* (also known as *Myo31DF*) was
504 viable and flies exhibited genital rotation phenotypes observed in previously reported
505 loss-of-function alleles (SPEDER *et al.* 2006) (Figure 3G). An insertion in *FK506-bp2* was
506 homozygous viable, though this gene has not been well characterized in *Drosophila*. In
507 all, these results indicate that insertion of *pCRISPaint-T2A-Gal4-3xP3-RFP* into 5'
508 coding sequence can disrupt gene function.

509 An important consideration when using genome editing is off-target effects. For
510 example, in addition to on-target insertions, our RFP+ fly lines could contain off-target
511 insertions or other mutations that disrupt non-target genes. We noticed that RFP
512 fluorescence always co-segregated with the target chromosome when balancing
513 insertions (not shown), suggesting that multiple insertions may be uncommon.
514 Furthermore, we did not detect evidence of off-target insertions by splinkerette PCR
515 (POTTER AND LUO 2010), which largely reconfirmed the on-target sites (Supplemental
516 Table 4). However, this analysis was problematic because some sequence traces
517 aligned to the donor plasmid or to genomic sequence on the opposite expected side of
518 the target site, suggesting that insertions could be concatemers. Others have observed
519 concatemers when performing NHEJ knock-ins (AUER *et al.* 2014; KUMAGAI *et al.* 2017),
520 and we found that seven of our RFP+ insertions were head-to-tail concatemers as
521 determined by PCR (Supplemental Figure 4). We could rescue the lethality of our *esg*
522 insertion using a duplication chromosome (stock # BL6230), suggesting there were no
523 other second site lethal mutations on this chromosome. Finally, one of five *ebony*
524 insertions (*ebony pFP545 #2*) was homozygous lethal, suggesting it contains a second
525 site lethal mutation.
526 We did not obtain insertions when targeting *ap*, *alphaTub84B*, *btl*, or *Desat1*.
527 Therefore, we investigated whether the sgRNAs targeting these genes were functional.
528 4 sgRNAs used for germ line knock-ins have an acceptable efficiency score of >5, with
529 the exception being the sgRNA targeting *btl* (Supplemental Table 5). A T7
530 endonuclease assay from transfected cells revealed that sgRNAs targeting *ap*,
531 *alphaTub84B*, *btl* can cut at the target site, whereas the results with *desat1* were
532 inconclusive (Supplemental Figure 5A). As an alternative functional test, we used PCR
533 to detect knock-in events in S2R+ cells transfected with the *pCRISPaint-T2A-Gal4-*
534 *3xP3-RFP* donor plasmid. This showed that sgRNAs targeting *ap*, *alphaTub84B*, *btl* and
535 *desat1* can lead to successful knock-in of *pCRISPaint-T2A-Gal4-3xP3-RFP* and
536 suggests that the sgRNAs are functional (Supplemental Figure 5B). Finally, we
537 sequenced the sgRNA target sites in the *nos-Cas9* fly strains and found a SNP in the *btl*
538 sgRNA binding site (not shown), whereas all 10 remaining sgRNAs had no SNPs in the
539 target site. In summary, we conclude that the sgRNAs targeting *ap*, *alphaTub84B*, *btl*,
540 and *Desat1* are able to induce cleavage at their target site in S2R+ cells, but that the
541 sgRNA targeting *btl* may not function in the germ line using our *nos-Cas9* strains.
542 Collectively, our results demonstrate that the *pCRISPaint-T2A-Gal4-3xP3-RFP*
543 universal donor plasmid inserts into target loci in the fly germ line and can be used to
544 generate loss of function fly lines. Therefore, we generated two simplified universal
545 donor plasmids that were more tailored to this purpose, *pCRISPaint-3xP3-RFP* and
546 *pCRISPaint-3xP3-GFP* (Figure 3H). These donor plasmids contain green or red
547 fluorescent markers and a useful multiple cloning site for adding additional insert
548 sequence.
549
550 **Knock-ins that express Gal4 under the control of the target gene**
551
552 We next determined if any of our 20 RFP+ lines expressed Gal4 from the target
553 gene. Since the *T2A-Gal4* reporter gene in *pCRISPaint-T2A-Gal4-3xP3-RFP* is
554 promoterless, we reasoned that only insertions in the sense orientation and in-frame
555 with the target gene should express Gal4. Nonetheless, we took an unbiased approach
556 by crossing all 20 RFP+ lines to *UAS-GFP* and screening progeny for GFP fluorescence
557 throughout development. From this effort, we identified 6 Gal4-expressing lines (Figure
558 4A, Supplemental Tables 3,4). *wg-T2A-Gal4* (#1 and 4), *Mhc-T2A-Gal4* (#1 and 2), and
559 *Myo1a-T2A-Gal4* #1 insertions are expressed in the imaginal disc, larval muscle, and

560 larval gut (Figure 4B, Supplemental Table 4), respectively, which matches the known
561 expression patterns for these genes. *wg-T2A-Gal4 #1 and #4* insertions are expressed
562 in a distinctive Wg pattern in the wing disc pouch (Figure 4C). *ebony-T2A-Gal4 pFP545*
563 *#2* is expressed in the larval brain and throughout the pupal body, consistent with a
564 previous study (HOVEMANN *et al.* 1998). Interestingly, it is also expressed in the larval
565 trachea, with particularly strong expression in the anterior and posterior spiracles
566 (Figure 4D). Indeed, classical *ebony* mutations are known to cause dark pigment in
567 larval spiracles (BREHME 1941).
568 Next, we compared the Gal4 expression screening results with our sequence
569 analysis of the insertion sites. As expected, all lines that expressed Gal4 had insertions
570 that were in the sense orientation (Supplemental Table 4, Supplemental Figure 2). For
571 example, *ebony-T2A-Gal4 pFP545 #2* contains a 15bp genomic deletion that is
572 predicted to keep *T2A-Gal4* in-frame with *ebony* coding sequence. Similarly, *wg-T2A-*
573 *Gal4* #1 contains an in-frame 45bp deletion and 21bp insertion. Remarkably, *wg-T2A-*
574 *Gal4* #4 contains a frameshift indel (Supplemental Figure 3), yet still expresses Gal4 in
575 the Wg pattern, albeit at significantly lower levels than *wg-T2A-Gal4* #1 (Figure 4C). In
576 addition, *Mhc-T2A-Gal4* lines #1, #2, and *Myo1a-T2A-Gal4* #1, each have indels that
577 put *T2A-Gal4* out of frame with the target gene coding sequence. These findings
578 confirm that our Gal4-expressing lines have *T2A-Gal4* inserted in the sense orientation,
579 but that in-frame insertion with the target gene coding sequence is not necessarily a
580 requirement for Gal4 expression.
581 Artifacts due to homology-independent insertion, such as indels at the insertion
582 site, could conceivably interfere with Gal4 expression. For example, *ebony* expression
583 has not been well studied, so we did not know if *ebony-T2A-Gal4 pFP545* #2 ("*ebony-*
584 *Gal4* CRISPaint") was an accurate reporter. To address this issue, we generated a
585 precise in-frame HDR insertion in *ebony* ("*ebony-Gal4* HDR") using the same insert
586 sequence (*T2A-Gal4-3xP3-RFP*) and target gene sgRNA (*pFP545*). *ebony-Gal4* HDR
587 knock-in fly lines were validated by PCR (Supplemental Figure 6) and *ebony-Gal4* HDR
588 homozygous flies were viable and had dark cuticle pigment (not shown), suggesting on-
589 target knock-in. By crossing CRISPaint and HDR knock-in alleles with *UAS-GFP*, we
590 found that their expression patterns were similar at larval, pupal, and adult stages
591 (Figure 4D,E). However, *ebony-Gal4* CRISPaint had higher levels of RFP fluorescence
592 and expressed Gal4 in the larval anal pad and gut, which is coincident with expression
593 of *3xP3-RFP*. Since this insertion is a concatemer (Supplemental Figure 4,
594 Supplemental Table 4), we speculated that multiple copies of *3xP3-RFP* lead to
595 increased RFP fluorescence and ectopic expression of Gal4. Unfortunately, we could
596 not test this because our efforts to remove *3xP3-RFP* using Cre/*loxP* excision were
597 unsuccessful; progeny expressing Cre and *ebony-Gal4* CRISPaint were apparently
598 lethal, whereas we easily generated RFP-free derivatives of *ebony-Gal4* HDR using the
599 same crossing scheme. In contrast, *Mhc-T2A-Gal4* lines #1, #2, and *Myo1a-T2A-Gal4*
600 #1 do not express in the anal pad, and gut and anal pad expression in *wg-T2A-Gal4* #1
601 likely corresponds to normal *wg* expression in these tissues (TAKASHIMA AND MURAKAMI
602 2001) (Figure 4B). In conclusion, these results suggest that Gal4-expressing CRISPaint
603 insertions can capture the endogenous expression pattern of the target gene, albeit with
604 caveats that we demonstrated (also see discussion).
605 To facilitate the insertion of other sequences into the germ line to generate
606 expression reporters, we generated 10 additional universal donor plasmids (Figure 4F).
607 These include T2A-containing donors with sequence encoding the alternative binary
608 reporters LexGAD, QF2, and split-Gal4, as well as Cas9 nuclease, FLP recombinase,
609 Gal80 repression protein, NanoLuc luminescence reporter, and super-folder GFP. In
610 addition, we generated *pCRISPaint-sfGFP-3xP3-RFP*, which can be used to insert into

611 3' coding sequence, generating a C-terminal GFP fusion protein. Finally, we created a
612 single plasmid expressing all three frame-selector sgRNAs (*pCFD5-frame-*
613 *selectors_0,1,2*). Since indels at the insertion site in the germ line effectively
614 randomizes the coding frame, we reasoned that simultaneous expression of all three
615 sgRNAs would maximize cutting and linearization of a CRISPaint universal donor
616 plasmid.
617
618 **Knock-ins targeting uncharacterized small open reading frame (smORF) genes**
619
620 To further demonstrate that this method can quickly generate knock-in insertions
621 for any gene, we targeted a class of uncharacterized small open reading frame
622 (smORF) genes. smORF genes encode peptides <100 amino acids and have been
623 historically understudied (COUSO AND PATRAQUIM 2017). For example, these genes may
624 have been missed in mutagenesis and proteomics screens because of their small size.
625 In addition, they frequently lack long introns useful for isolating exon trap insertions
626 (BUSZCZAK *et al.* 2007).
627 Using sequence alignment tools (BLASTP, NCBI; DIOPT, www.flyrnai.org/diopt)
628 and gene databases (FlyBase; Alliance of Genome Resources,
629 www.alliancegenome.org/) we identified uncharacterized smORF genes that have clear
630 human homologues and targeted them for knock-in using sgRNAs that cut 5' coding
631 sequence and the *pCRISPaint-T2A-Gal4-3xP3-RFP* donor plasmid. Using the same
632 injection and crossing scheme described above, we isolated 12 insertions in 5 genes as
633 balanced stocks in 1 month (Supplemental Table 6). Diagnostic PCR analysis confirmed
634 that all the insertions integrated in the correct target site (Supplemental Figure 7). These
635 knock-in alleles will be described elsewhere. These results demonstrate that this is a
636 fast method for producing knock-in lines for any gene.
637
638 **Discussion**
639
640 Inserting large DNA elements into the genome by HDR requires a great deal of
641 expertise and labor for the design and construction of donor plasmids. Some groups
642 have developed strategies to improve the efficiency and scale at which homology arms
643 are cloned into donor plasmids (HOUSDEN *et al.* 2014; GRATZ *et al.* 2015), but the root
644 problem still remains. In this study, we showed that homology-independent insertion can
645 be used in *Drosophila* as an alternative to HDR for certain gene targeting strategies in
646 cultured cells and in vivo. With no donor design and construction, knock-in lines are
647 simpler and faster to obtain (see Table 1), making it easier to target multiple genes in
648 parallel.
649 To perform homology-independent insertion in *Drosophila*, we adapted the
650 mammalian CRISPaint system (SCHMID-BURGK *et al.* 2016) because of its user-friendly
651 design. CRISPaint donor plasmids are "universal" because they lack homology arms,
652 and publicly available collections enable researchers to "mix and match" insert
653 sequences and target genes. Indeed, we showed that *pCRISPaint-mNeonGreen-T2A-*
654 *PuroR,* originally used in human cells, functions in *Drosophila* S2R+ cells (Figure 1). To
655 linearize donor plasmids in *Drosophila* in three coding frames, we generated plasmids
656 expressing the three CRISPaint frame-selector sgRNAs under the *Drosophila U6*
657 promoter. Similarly, we also generated a plasmid to express all three sgRNAs
658 simultaneously for maximum cutting efficiency. Therefore, to perform a knock-in
659 experiment, only an sgRNA that targets a gene of interest is needed, which are simple
660 to produce by molecular cloning or ordering. In fact, all target gene sgRNAs plasmids

661 used in this study were obtained from the TRiP facility and an ever-growing number of
662 additional sgRNAs are being generated.
663 Unlike HDR, which can seamlessly insert any DNA into a target genomic
664 location, homology-independent insertion is restricted to specific targeting scenarios.
665 For example, since the entire *pCRISPaint* donor plasmid integrates into the target site, it
666 cannot be used to replace genes, make amino acid changes, or tag proteins at their N-
667 terminus. In addition, potential concatemerization of the insertion (AUER *et al.* 2014;
668 KUMAGAI *et al.* 2017) prevents its use for strategies where a single copy of the insertion
669 is required, such as artificial exon traps that exploit target gene splicing (BUSZCZAK *et al.*
670 2007). Despite these limitations, homology-independent insertion is still well suited to
671 generate loss of function alleles by premature truncation, as well as C-terminal gene
672 tagging when a 3' UTR is included.
673 Using CRISPaint in the fly germ line, we rapidly generated loss of function alleles
674 by inserting the selectable marker *3xP3-RFP* into target gene coding sequence. This
675 method has several advantages over isolating frameshift alleles using CRISPR-induced
676 indels (BIER *et al.* 2018). First, identifying founder flies with RFP fluorescence is easier
677 and faster than PCR genotyping/sequencing for frameshift indels. Second, insertion into
678 5' coding sequence is more likely to disrupt the target gene, whereas frameshift alleles
679 can sometimes retain protein activity (TULADHAR *et al.* 2019). Though, insertions could
680 also have unintended consequences, such as disrupting neighboring gene expression.
681 Third, genetic crosses are simplified by tracking insertion alleles with a fluorescent
682 marker, which could be especially useful in non-*melanogaster* species that do not have
683 balancer chromosomes. Indeed, the combined use of *3xP3-RFP* and *3xP3-GFP*
684 universal donor plasmids (Figure 3H) could facilitate generating double mutant lines.
685 While we did not find evidence for off-target insertions, knock-in lines should be vetted
686 using traditional techniques, including molecular verification of the insertion site,
687 comparing independent insertions, outcrossing to wild type, and performing
688 complementation tests.
689 Using homology-independent insertion for gene tagging requires screening for
690 properly expressed inserts, due to the unpredictable and error-prone nature of NHEJ.
691 For example, donor plasmids can integrate in two orientations and indels at the insertion
692 site can shift the coding frame. In cultured S2R+ cells, we used antibiotic resistance to
693 efficiently select for mNeonGreen protein fusions (Figure 1D). However, such an assay
694 is not feasible in flies, and so we screened germ line *T2A-Gal4* insertions for expression
695 by crossing with *UAS-GFP* (Figure 3B, Figure 4). This yielded Gal4-expressing lines for
696 *ebony*, *wg*, *Mhc*, and *myo1a*. In contrast, we obtained no insertions for *alpthaTub84B*,
697 *btl*, *Desat*, *ap*, and one non-Gal4 expressing insertion each for *FK506-bp2*, *esg*, and *hh*.
698 This highlights the importance of obtaining multiple independently derived germ line
699 insertions to screen for insert expression. Additional steps could be taken to obtain
700 more insertions per gene, such as increasing the number of injected embryos or
701 reattempting with a different sgRNA. Since our knock-in efficiency is roughly similar to
702 HDR (5-21% (Figure 3C) vs. 5-22% (GRATZ *et al.* 2014), 46-88% (PORT *et al.* 2015), and
703 7-42% (GRATZ *et al.* 2015)), the limited number of independent germ line insertions may
704 simply be a constraint of embryo injection-based transgenesis.
705 Sequence analysis of insertion sites in cultured cells and in vivo revealed
706 expected and unexpected results. Single cell cloned S2R+ lines expressing
707 mNeonGreen fusion proteins had insertion sites that were seamless (7/9) or caused an
708 in-frame 3bp deletion (2/9), which is consistent with human cell data (SCHMID-BURGK *et*
709 *al.* 2016). However, all sequenced germ line insertions (20) contained indels at the
710 insertion site, and indels were larger than in S2R+ cells (Supplemental Figure 3). Germ
711 cells are known to differ in their NHEJ mechanisms compared to somatic cells (PRESTON

712 *et al.* 2006; AHMED *et al.* 2015), but it is not clear why this would prevent seamless
713 insertions. In addition, 4/6 of Gal4-expressing insertions were out of frame relative to the
714 target gene (Supplemental Figure 3). We speculate that this could result from ribosome
715 frameshifting (KETTELER 2012), an internal ribosome entry site (IRES) (KOMAR AND
716 HATZOGLOU 2005), alternative open reading frames (altORFs) (MOUILLERON *et al.* 2016),
717 or Gal4 translation initiation from its own start codon. Ultimately, we consider this a
718 fortuitous effect as long as Gal4 is expressed in the correct pattern. Conversely, *wg* #6
719 is in-frame but does not express Gal4, perhaps due to a mutation of Gal4 sequences.
720 Indeed, splinkerette sequencing results from 3/20 of our RFP+ insertions suggest that
721 portions of the donor plasmid may be deleted during insertion (Supplemental Table 4).
722 These findings again highlight the importance of screening insertions for insert
723 expression.
724 For gene-tagged cell lines and fly lines generated using homology-independent
725 insertion, potential artifacts that could impact faithful protein fusion or reporter
726 expression. First, indels at the insertion site could introduce cryptic splice sites, delete
727 an important sequence region, or cause non-optimal codons that impact translation of
728 the fusion protein. Second, bacterial sequences in the inserted plasmid may cause
729 transgene silencing or impact neighboring gene expression (CHEN *et al.* 2004; SUZUKI *et*
730 *al.* 2016). Though, we note that thousands of transgenic fly lines contain bacterial
731 sequences from phiC31 integration (PERKINS *et al.* 2015) with no reports of ill effects.
732 Third, concatemer insertions could affect gene expression in unpredictable ways.
733 Indeed, insertion *ebony-T2A-Gal4 pFP545* #2 was a head-to-tail concatemer
734 (Supplemental Figure 4) and had ectopic Gal4 expression from the *3xP3* enhancer
735 (Figure 4D,E). Finally, as noted above, the insertion frame may be important for
736 expression, since *wg-T2A-Gal4* #4 line is expressed weakly and in a non-faithful patchy
737 pattern in the wing disc. While these are important considerations, we note that 9/9 of
738 our mNeonGreen tagged S2R+ lines were localized properly, and 5/6 Gal4 expressing
739 lines appear to be faithful reporters of their target genes, with the above exceptions
740 noted.
741 Some of our S2R+ mNeonGreen protein fusion lines exhibited unusual or
742 unexpected protein localization. In clones D6 and D9, Lamin-mNeonGreen localized to
743 the nuclear envelope, but in D9 this localization was enriched asymmetrically in the
744 direction of the previous plane of cell division. Since both clones had seamless
745 mNeonGreen insertions, we believe the localization difference is caused by the lack of a
746 non-knock-in locus in D9, whereas D6 contained an in-frame 3bp deletion at the non-
747 knock-in locus, likely retaining wild-type function. We saw a similar pattern for clones A3
748 and A5, both of which had seamless insertions in *Actin5c*, but clone A5 exhibited
749 distinct rod structures. alphaTub84B-mNeonGreen fluorescence and protein levels were
750 low in all cell lines, despite *alphaTub84B* being highly expressed in S2R+ cells (HU *et al.*
751 2017). We speculate that the alphaTub84B-mNeonGreen fusion protein is unstable and
752 previous studies in other organisms have highlighted problems with C-terminal tagging
753 of alpha-Tubulin (CARMINATI AND STEARNS 1997). Similarly, C-terminal tags can disrupt
754 Lamin and Actin function (DAVIES *et al.* 2009; NAGASAKI *et al.* 2017). Therefore, it is
755 important to consider the existing knowledge of the protein when generating C-terminal
756 protein fusions and examine multiple single cell cloned lines. Importantly, this level of
757 scrutiny is also required when using HDR to produce knock-ins. Further validation steps
758 could be done, such as comparing the tagged protein localization to antibody staining of
759 the untagged wild type protein.
760 Additional applications of homology-independent insertion in the *Drosophila* germ
761 line can be imagined. For example, new insert sequences could be added to
762 *pCRISPaint-3xP3-RFP* or *pCRISPaint-3xP3-GFP* starting plasmids. In addition, a

763 modified donor plasmid could be constructed that does not integrate the bacterial
764 plasmid backbone, such as providing donor plasmids as mini-circles (SCHMID-BURGK *et*
765 *al.* 2016; SUZUKI *et al.* 2016), cutting donor plasmid twice to liberate the insert fragment
766 (LACKNER *et al.* 2015; SUZUKI *et al.* 2016; GAO *et al.* 2019), or using PCR amplified
767 inserts (MANNA *et al.* 2019). Protein tagging could be performed in vivo, but the large
768 indels observed in germ line may limit its use to proteins that have C-terminal tails that
769 can be deleted without altering protein function. Regardless, we generated a donor
770 plasmid for in vivo sfGFP C-terminal fusions (Figure 4F). Finally, we note that most of
771 our germ line donor plasmids contain enzyme restriction sites that can be used to insert
772 genomic homology arms by molecular cloning, making them "dual use" reagents for
773 HDR and homology-independent insertion. Indeed, we used this approach to generate
774 the *ebony-Gal4* HDR allele (see Materials and Methods).
775 In summary, we demonstrated that homology-independent insertion can be used
776 as an alternative to HDR in *Drosophila*, enabling researchers to rapidly obtain knock-ins
777 without donor design and construction. The most practical application of this approach
778 is to perform C-terminal protein tagging in cultured cells and gene knockout by insertion
779 in vivo. While we obtained in vivo *T2A-Gal4* gene-tagged insertions, it is low efficiency
780 and thus less appealing compared to HDR. However, the techniques required for
781 screening insertions for *T2A-Gal4* are less specialized than those for constructing donor
782 plasmids, making this trade off potentially attractive for those with less molecular biology
783 expertise or who previously failed using HDR. Finally, our in vivo donor plasmids are
784 immediately useable in other arthropod species because of the *3xP3*-fluorescent marker
785 (BERGHAMMER *et al.* 1999), a testament to the modularity of this knock-in system and the
786 benefits of a community of researchers creating and sharing universal donor plasmids.
787
788 **Acknowledgements**
789
790 We thank Jonathan Schmid-Burgk for advice and the *CRISPaint-mNeonGreen* donor
791 plasmid, Claire Hu and the TRiP for sgRNA design and construction, Ben Ewen-
792 Campen for valuable comments on the manuscript, Stephanie Mohr, Oguz Kanca, and
793 Hugo Bellen for helpful discussions, Raghuvir Viswanatha for the *Cas9-T2A-EGFP*
794 template sequence, Rich Binari for help with mounting embryos on slides, and Cathryn
795 Murphy for general assistance. We also thank the HMS MicRoN (Microscopy
796 Resources on the North Quad) Core and the HMS Department of Immunology's Flow
797 Cytometry Facility for technical assistance. J.A.B. was supported by the Damon Runyon
798 Foundation. This work was supported by NIH grants R01GM084947, R01GM067761,
799 R24OD019847. N.P. is an investigator of the Howard Hughes Medical Institute.
800
801 **References:**
802
803 Ahmed, E. A., H. Scherthan and D. G. de Rooij, 2015 DNA Double Strand Break
804 Response and Limited Repair Capacity in Mouse Elongated Spermatids. Int J
805 Mol Sci 16**:** 29923-29935.
806 Auer, T. O., and F. Del Bene, 2014 CRISPR/Cas9 and TALEN-mediated knock-in
807 approaches in zebrafish. Methods 69**:** 142-150.
808 Auer, T. O., K. Duroure, A. De Cian, J. P. Concordet and F. Del Bene, 2014 Highly
809 efficient CRISPR/Cas9-mediated knock-in in zebrafish by homology-independent
810 DNA repair. Genome Res 24**:** 142-153.
811 Bassett, A. R., C. Tibbit, C. P. Ponting and J. L. Liu, 2014 Mutagenesis and homologous
812 recombination in Drosophila cell lines using CRISPR/Cas9. Biol Open 3**:** 42-49.

Bejsovec, A., and A. Martinez Arias, 1991 Roles of wingless in patterning the larval epidermis of Drosophila. Development 113: 471-485.

Bellen, H. J., R. W. Levis, Y. He, J. W. Carlson, M. Evans-Holm *et al.*, 2011 The Drosophila gene disruption project: progress using transposons with distinctive site specificities. Genetics 188: 731-743.

Berghammer, A. J., M. Klingler and E. A. Wimmer, 1999 A universal marker for transgenic insects. Nature 402: 370-371.

Bier, E., M. M. Harrison, K. M. O'Connor-Giles and J. Wildonger, 2018 Advances in Engineering the Fly Genome with the CRISPR-Cas System. Genetics 208: 1-18.

Bosch, J. A., N. H. Tran and I. K. Hariharan, 2015 CoinFLP: a system for efficient mosaic screening and for visualizing clonal boundaries in Drosophila. Development 142: 597-606.

Bottcher, R., M. Hollmann, K. Merk, V. Nitschko, C. Obermaier *et al.*, 2014 Efficient chromosomal gene modification with CRISPR/cas9 and PCR-based homologous recombination donors in cultured Drosophila cells. Nucleic Acids Res 42: e89.

Brehme, K. S., 1941 The Effect of Adult Body Color Mutations upon the Larva of Drosophila Melanogaster. Proc Natl Acad Sci U S A 27: 254-261.

Buszczak, M., S. Paterno, D. Lighthouse, J. Bachman, J. Planck *et al.*, 2007 The carnegie protein trap library: a versatile tool for Drosophila developmental studies. Genetics 175: 1505-1531.

Carminati, J. L., and T. Stearns, 1997 Microtubules orient the mitotic spindle in yeast through dynein-dependent interactions with the cell cortex. J Cell Biol 138: 629-641.

Chen, Z. Y., C. Y. He, L. Meuse and M. A. Kay, 2004 Silencing of episomal transgene expression by plasmid bacterial DNA elements in vivo. Gene Ther 11: 856-864.

Clement, K., H. Rees, M. C. Canver, J. M. Gehrke, R. Farouni *et al.*, 2019 CRISPResso2 provides accurate and rapid genome editing sequence analysis. Nat Biotechnol 37: 224-226.

Couso, J. P., and P. Patraquim, 2017 Classification and function of small open reading frames. Nat Rev Mol Cell Biol 18: 575-589.

Cristea, S., Y. Freyvert, Y. Santiago, M. C. Holmes, F. D. Urnov *et al.*, 2013 In vivo cleavage of transgene donors promotes nuclease-mediated targeted integration. Biotechnol Bioeng 110: 871-880.

Davies, B. S., L. G. Fong, S. H. Yang, C. Coffinier and S. G. Young, 2009 The posttranslational processing of prelamin A and disease. Annu Rev Genomics Hum Genet 10: 153-174.

Diao, F., and B. H. White, 2012 A novel approach for directing transgene expression in Drosophila: T2A-Gal4 in-frame fusion. Genetics 190: 1139-1144.

Gao, Y., E. Hisey, T. W. A. Bradshaw, E. Erata, W. E. Brown *et al.*, 2019 Plug-and-Play Protein Modification Using Homology-Independent Universal Genome Engineering. Neuron.

Gratz, S. J., C. D. Rubinstein, M. M. Harrison, J. Wildonger and K. M. O'Connor-Giles, 2015 CRISPR-Cas9 Genome Editing in Drosophila. Curr Protoc Mol Biol 111: 31 32 31-20.

Gratz, S. J., F. P. Ukken, C. D. Rubinstein, G. Thiede, L. K. Donohue *et al.*, 2014 Highly specific and efficient CRISPR/Cas9-catalyzed homology-directed repair in Drosophila. Genetics 196: 961-971.

He, L., R. Binari, J. Huang, J. Falo-Sanjuan and N. Perrimon, 2019 In vivo study of gene expression with an enhanced dual-color fluorescent transcriptional timer. Elife 8.

Housden, B. E., S. Lin and N. Perrimon, 2014 Cas9-based genome editing in Drosophila. Methods Enzymol 546: 415-439.

Hovemann, B. T., R. P. Ryseck, U. Walldorf, K. F. Stortkuhl, I. D. Dietzel *et al.*, 1998 The Drosophila ebony gene is closely related to microbial peptide synthetases and shows specific cuticle and nervous system expression. Gene 221**:** 1-9.

Hu, Y., A. Comjean, N. Perrimon and S. E. Mohr, 2017 The Drosophila Gene Expression Tool (DGET) for expression analyses. BMC Bioinformatics 18**:** 98.

Katic, I., L. Xu and R. Ciosk, 2015 CRISPR/Cas9 Genome Editing in Caenorhabditis elegans: Evaluation of Templates for Homology-Mediated Repair and Knock-Ins by Homology-Independent DNA Repair. G3 (Bethesda) 5**:** 1649-1656.

Katoh, Y., S. Michisaka, S. Nozaki, T. Funabashi, T. Hirano *et al.*, 2017 Practical method for targeted disruption of cilia-related genes by using CRISPR/Cas9-mediated, homology-independent knock-in system. Mol Biol Cell 28**:** 898-906.

Ketteler, R., 2012 On programmed ribosomal frameshifting: the alternative proteomes. Front Genet 3**:** 242.

Komar, A. A., and M. Hatzoglou, 2005 Internal ribosome entry sites in cellular mRNAs: mystery of their existence. J Biol Chem 280**:** 23425-23428.

Kondo, S., and R. Ueda, 2013 Highly improved gene targeting by germline-specific Cas9 expression in Drosophila. Genetics 195**:** 715-721.

Korona, D., S. A. Koestler and S. Russell, 2017 Engineering the Drosophila Genome for Developmental Biology. J Dev Biol 5.

Kumagai, H., T. Nakanishi, T. Matsuura, Y. Kato and H. Watanabe, 2017 CRISPR/Cas-mediated knock-in via non-homologous end-joining in the crustacean Daphnia magna. PLoS One 12**:** e0186112.

Lackner, D. H., A. Carre, P. M. Guzzardo, C. Banning, R. Mangena *et al.*, 2015 A generic strategy for CRISPR-Cas9-mediated gene tagging. Nat Commun 6**:** 10237.

Lee, H., C. J. McManus, D. Y. Cho, M. Eaton, F. Renda *et al.*, 2014 DNA copy number evolution in Drosophila cell lines. Genome Biol 15**:** R70.

Lee, P. T., J. Zirin, O. Kanca, W. W. Lin, K. L. Schulze *et al.*, 2018 A gene-specific T2A-GAL4 library for Drosophila. Elife 7.

Manna, P. T., L. J. Davis and M. S. Robinson, 2019 Fast and cloning-free CRISPR/Cas9-mediated genomic editing in mammalian cells. bioRxiv.

Maresca, M., V. G. Lin, N. Guo and Y. Yang, 2013 Obligate ligation-gated recombination (ObLiGaRe): custom-designed nuclease-mediated targeted integration through nonhomologous end joining. Genome Res 23**:** 539-546.

Mogami, K., P. T. O'Donnell, S. I. Bernstein, T. R. Wright and C. P. Emerson, Jr., 1986 Mutations of the Drosophila myosin heavy-chain gene: effects on transcription, myosin accumulation, and muscle function. Proc Natl Acad Sci U S A 83**:** 1393-1397.

Mouilleron, H., V. Delcourt and X. Roucou, 2016 Death of a dogma: eukaryotic mRNAs can code for more than one protein. Nucleic Acids Res 44**:** 14-23.

Nagasaki, A., T. K. S, T. Yumoto, M. Imaizumi, A. Yamagishi *et al.*, 2017 The Position of the GFP Tag on Actin Affects the Filament Formation in Mammalian Cells. Cell Struct Funct 42**:** 131-140.

Perkins, L. A., L. Holderbaum, R. Tao, Y. Hu, R. Sopko *et al.*, 2015 The Transgenic RNAi Project at Harvard Medical School: Resources and Validation. Genetics 201**:** 843-852.

Port, F., and S. L. Bullock, 2016 Augmenting CRISPR applications in Drosophila with tRNA-flanked sgRNAs. Nat Methods 13**:** 852-854.

Port, F., H. M. Chen, T. Lee and S. L. Bullock, 2014 Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in Drosophila. Proc Natl Acad Sci U S A 111**:** E2967-2976.

Port, F., N. Muschalik and S. L. Bullock, 2015 Systematic Evaluation of Drosophila CRISPR Tools Reveals Safe and Robust Alternatives to Autonomous Gene Drives in Basic Research. G3 (Bethesda) 5**:** 1493-1502.

Potter, C. J., and L. Luo, 2010 Splinkerette PCR for mapping transposable elements in Drosophila. PLoS One 5**:** e10168.

Preston, C. R., C. C. Flores and W. R. Engels, 2006 Differential usage of alternative pathways of double-strand break repair in Drosophila. Genetics 172**:** 1055-1068.

Quadros, R. M., H. Miura, D. W. Harms, H. Akatsuka, T. Sato *et al.*, 2017 Easi-CRISPR: a robust method for one-step generation of mice carrying conditional and insertion alleles using long ssDNA donors and CRISPR ribonucleoproteins. Genome Biol 18**:** 92.

Ren, X., J. Sun, B. E. Housden, Y. Hu, C. Roesel *et al.*, 2013 Optimized gene editing technology for Drosophila melanogaster using germ line-specific Cas9. Proc Natl Acad Sci U S A 110**:** 19012-19017.

Schmid-Burgk, J. L., K. Honing, T. S. Ebert and V. Hornung, 2016 CRISPaint allows modular base-specific gene tagging using a ligase-4-dependent mechanism. Nat Commun 7**:** 12338.

Singh, P., J. C. Schimenti and E. Bolcun-Filas, 2015 A mouse geneticist's practical guide to CRISPR applications. Genetics 199**:** 1-15.

Speder, P., G. Adam and S. Noselli, 2006 Type ID unconventional myosin controls left-right asymmetry in Drosophila. Nature 440**:** 803-807.

Suzuki, K., Y. Tsunekawa, R. Hernandez-Benitez, J. Wu, J. Zhu *et al.*, 2016 In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. Nature 540**:** 144-149.

Takashima, S., and R. Murakami, 2001 Regulation of pattern formation in the Drosophila hindgut by wg, hh, dpp, and en. Mech Dev 101**:** 79-90.

Tuladhar, R., Y. Yeu, J. Tyler Piazza, Z. Tan, J. Rene Clemenceau *et al.*, 2019 CRISPR-Cas9-based mutagenesis frequently provokes on-target mRNA misregulation. Nat Commun 10**:** 4056.

Venken, K. J., A. Sarrion-Perdigones, P. J. Vandeventer, N. S. Abel, A. E. Christiansen *et al.*, 2016 Genome engineering: Drosophila melanogaster and beyond. Wiley Interdiscip Rev Dev Biol 5**:** 233-267.

Viswanatha, R., Z. Li, Y. Hu and N. Perrimon, 2018 Pooled genome-wide CRISPR screening for basal and context-specific fitness gene essentiality in Drosophila cells. Elife 7.

Yu, Z., M. Ren, Z. Wang, B. Zhang, Y. S. Rong *et al.*, 2013 Highly efficient genome modifications mediated by CRISPR/Cas9 in Drosophila. Genetics 195**:** 289-291.

Zhang, C., X. He, Y. K. Kwok, F. Wang, J. Xue *et al.*, 2018 Homology-independent multiallelic disruption via CRISPR/Cas9-based knock-in yields distinct functional outcomes in human cells. BMC Biol 16**:** 151.

Figure 1

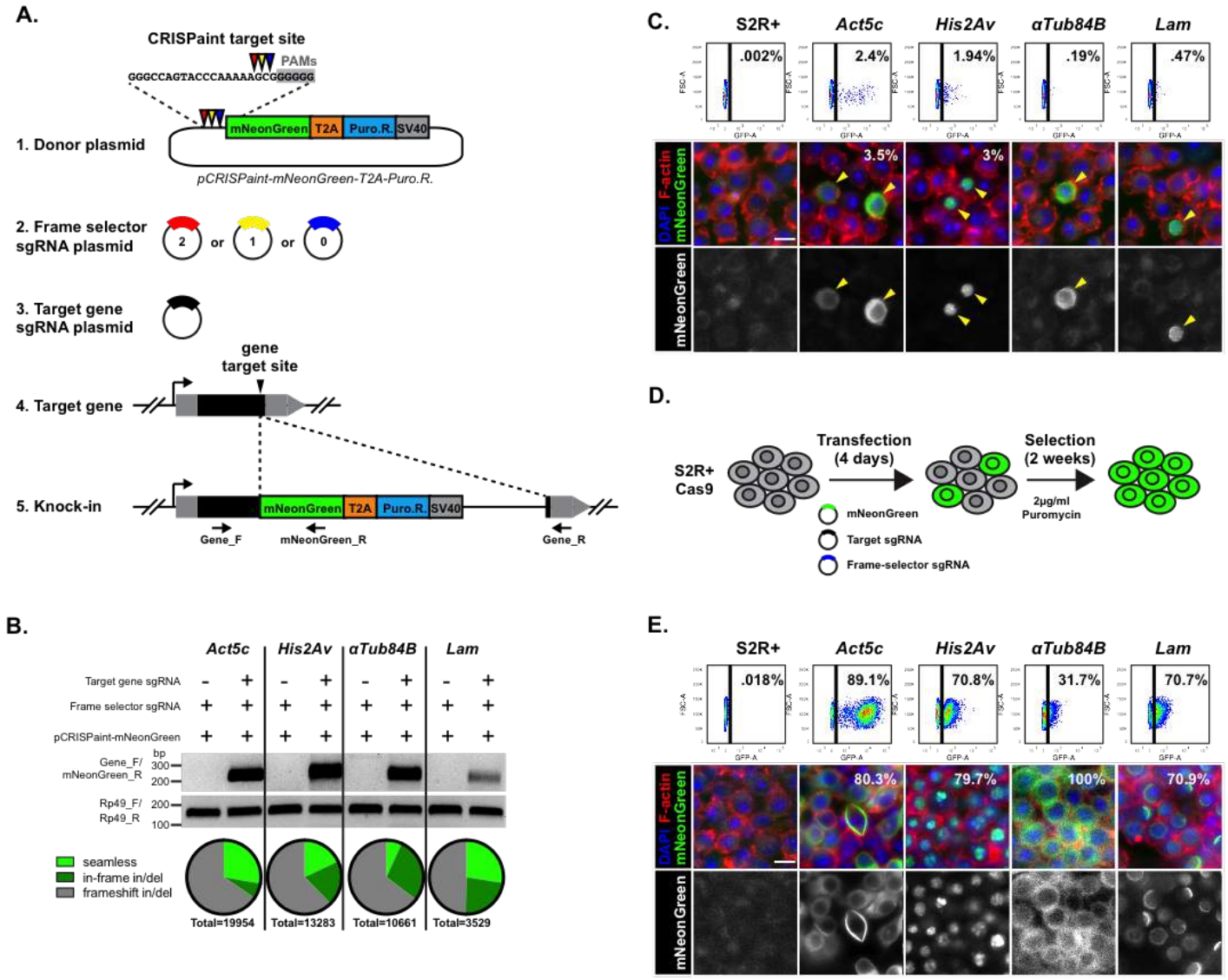
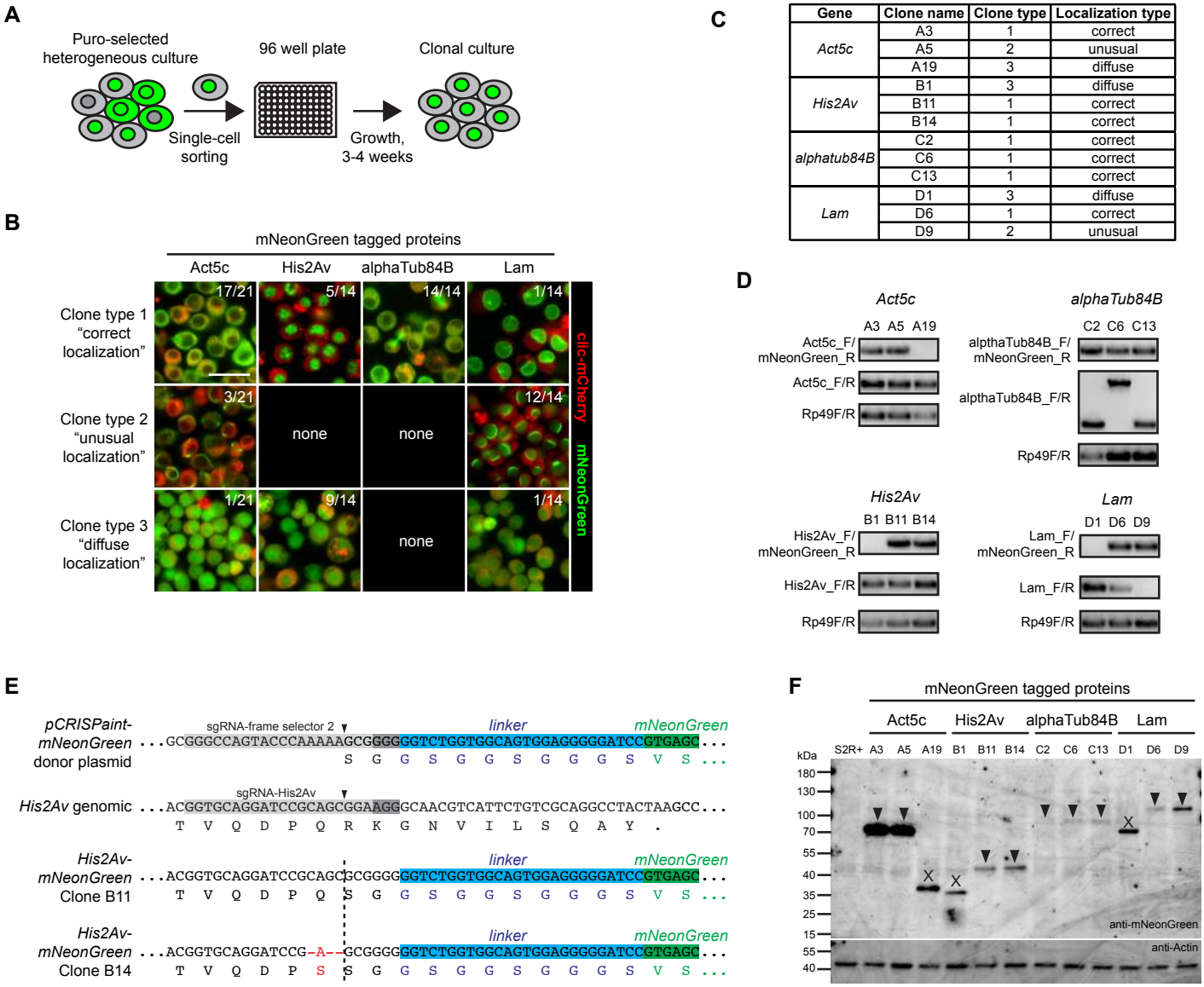


Figure 1 – Knock-in of *mNeonGreen-T2A-PuroR* into Drosophila S2R+ cells using homology-independent insertion. (**A**) Schematic of CRISPaint knock-in approach. *mNeonGreen-T2A-PuroR* is inserted into 3' coding sequence. (**B**) Analysis of knock-in efficiency of transfected cells by diagnostic PCR (DNA gel image) and next-generation sequencing (pie charts). (**C**) Analysis of knock-in efficiency of transfected cells by FACs and confocal microscopy. Numbers indicate percentage of cells with fluorescence. F-actin stained using Phalloidin-TRITC (red), nuclei labeled with DAPI (blue), mNeonGreen signal is in green. Scale bar 10µm. (**D**) Schematic of Puromycin selection of mNeonGreen-expressing cells. (**E**) Analysis of knock-in frequency of puromycin-selected cells using FACs and confocal microscopy. Numbers indicate percentage of cells with green fluorescence. Scale bar 10µm.

Figure 2



Gene	Clone name	Clone type	Localization type
Act5c	A3	1	correct
	A5	2	unusual
	A19	3	diffuse
His2Av	B1	3	diffuse
	B11	1	correct
	B14	1	correct
alphatub84B	C2	1	correct
	C6	1	correct
	C13	1	correct
Lam	D1	3	diffuse
	D6	1	correct
	D9	2	unusual

Figure 2. Analysis of S2R+ mNeonGreen-expressing single-cell cloned lines. (**A**) Schematic of FACS isolation of single-cell clones expressing mNeonGreen. (**B**) Confocal images of live mNeonGreen-expressing cell lines, categorized into three clone types. Numbers indicate the frequency of each clone type for each gene targeted. Images show fluorescence from Clic-mCherry (red) and mNeonGreen (green). Scale bar 25µm. (**C**) Single cell cloned lines retained for further analysis. (**D**) Agarose gel with PCR fragments amplified from knock-in (Gene_F/mNeonGreen_R) and non-knock-in loci (Gene_F/R). Positive control bands were amplified from *Rp49* genomic sequence. (**E**) Example sequence results of *His2Av-mNeonGreen* clones. sgRNA target site and PAM sequence shown as grey bars, Cas9 cut sites shown with arrowheads. (**F**) Western blot detecting mNeonGreen protein fusions. Arrowheads indicate expected molecular weight. X's indicate incorrect molecular weight.

Figure 3

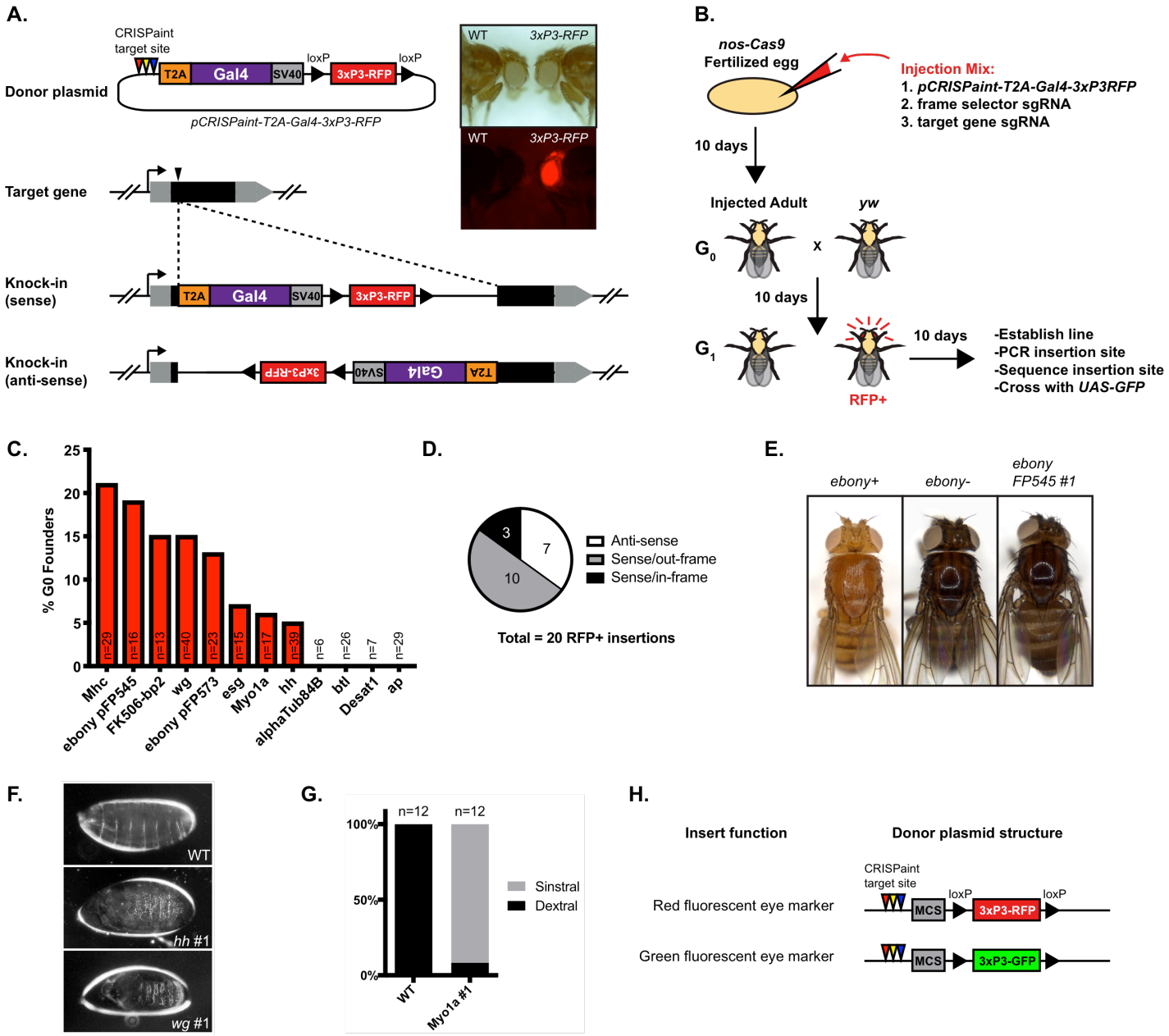


Figure 3 – Germ line knock-ins using homology-independent insertion. (**A**) Schematic of knock-in approach. *pCRISPaint-T2A-Gal4-3xP3-RFP* is inserted into 5' coding sequence. Inset shows images of adult flies with 3xP3-RFP fluorescence in the eye. Top panel is brightfield, bottom panel is fluorescence. (**B**) Schematic of plasmid injections, fly crosses, and analysis of insertions. (**C**) Graph with results of knock-in efficiency for 12 sgRNA target sites and 11 genes. (**D**) Insert orientation and frame of 20 RFP+ fly lines. (**E**) Image of adult flies. Homozygous *ebony-T2A-Gal4 FP545* #1 flies have dark cuticle pigment. (**F**) Darkfield images of *hh* #1 and *wg* #1 homozygous embryo "lawn of denticles" phenotypes. (**G**) Quantification of genital rotation in wild type (*yw*) and a *myo1a* insertion line. Rotation is determined from the direction of wrapping of the adult male spermiduct around the gut from dissected live abdomens. (**H**). Diagram of two simplified universal donor plasmids for germ line knock-ins.

Figure 4

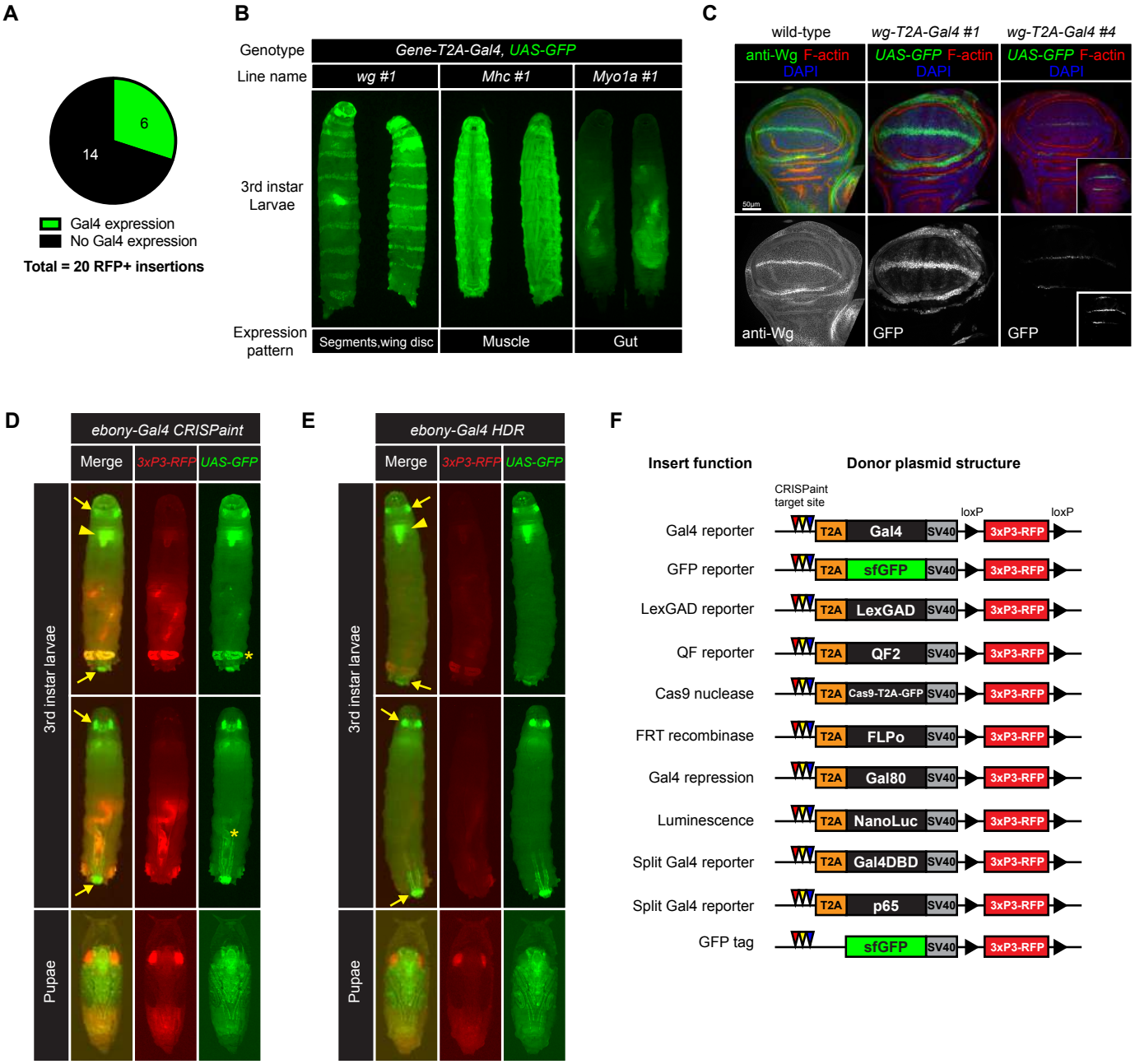


Figure 4 – Germ line insertions that express *T2A-Gal4* under the control of the target gene. (**A**) Proportion of 20 RFP+ knock-in lines that express Gal4. (**B**) Fluorescence images of 3rd instar larvae with indicated genotypes. Two larvae of the same genotype are shown as rotated in different angles. Expression of Gal4 under control of the target gene drives expression of the *UAS-GFP* reporter. (**C**) Confocal images of wing imaginal discs showing protein staining of Wg protein (anti-wg, green) or *UAS-GFP* expression (green). GFP fluorescence was recorded at identical exposure settings for lines *wg-T2A-Gal4* #1 and #4. Inset shows digitally increased GFP signal. Scale bar 50µm. (**D,E**) Fluorescence images of 3rd instar larvae and pupae. Fluorescence is from *3xP3-RFP* (red) and Gal4 expression with *UAS-GFP* (green). (**D**) shows results with *ebony-T2A-Gal4 pFP545* #2 (*ebony-Gal4* CRISPaint) and (**E**) shows results with *ebony-Gal4* HDR. (**F**) Additional universal donor plasmids for insertion into 5' coding sequence.

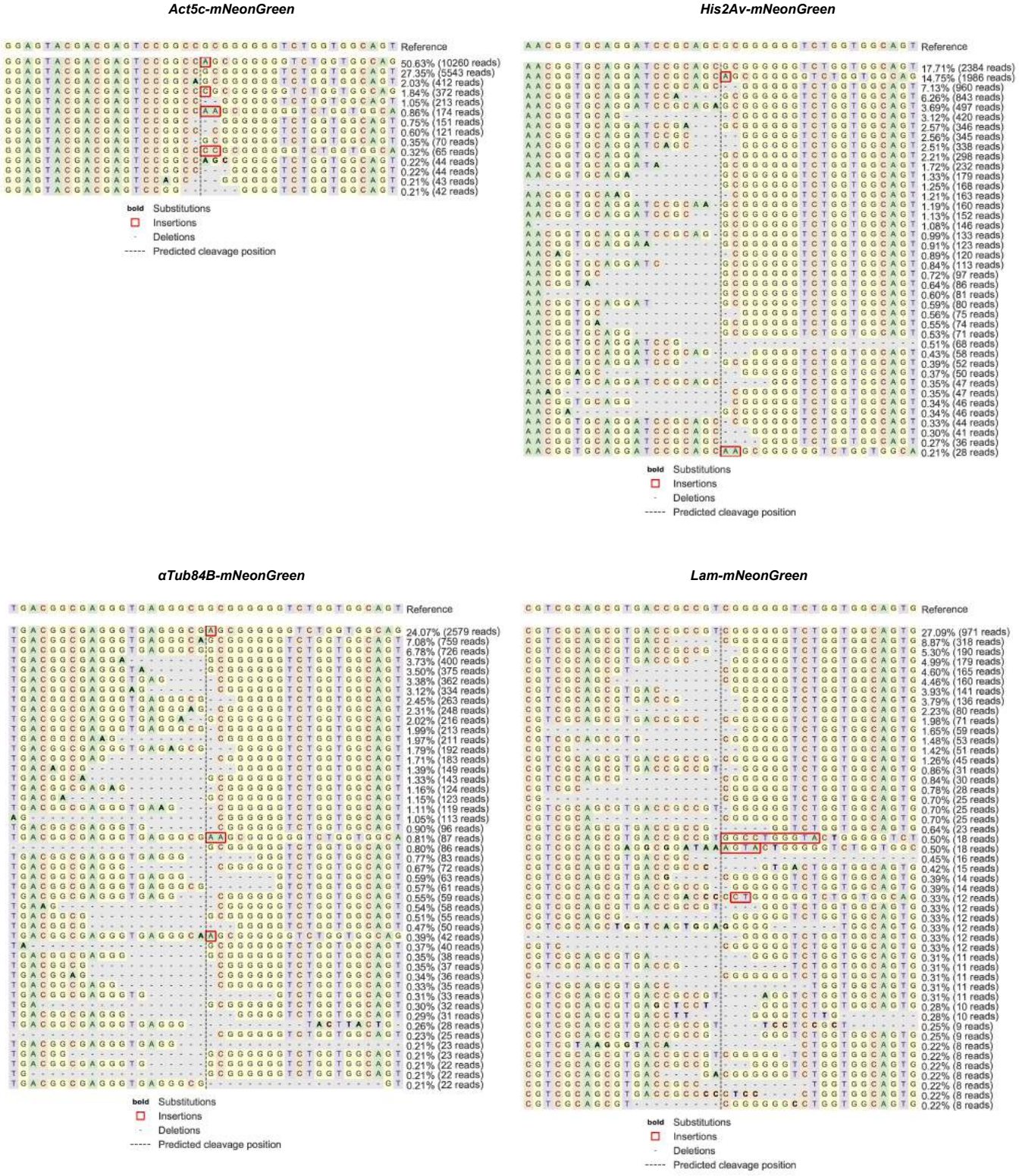
S2R+ gene tagging (e.g. mNeonGreen)	Donor design and construction	sgRNA cloning	Transfection -> single cell clones	Molecular characterization	Screening expression
HDR knock-in	Variable, 1-3 weeks	3 days	2 months	1-2 days (PCR, Gel image)	1 day, confocal microscopy
CRISPaint knock-in	N/A	3 days	2 months	3 days (PCR, Gel image, Sequence)	1 day, confocal microscopy

In vivo KO	Donor design and construction	sgRNA cloning	Injection -> fly stock	Molecular characterization
HDR knock-in	Variable, 1-3 weeks	3 days	1 month	1-2 days (PCR, Gel image)
In/del frameshift	N/A	3 days	1 month	3-7 days (PCR, Gel image, Sequence)
CRISPaint knock-in	N/A	3 days	1 month	3 days (PCR, Gel image, Sequence)

In vivo gene tagging (e.g. T2A-Gal4)	Donor design and construction	sgRNA cloning	Injection -> fly stock	Molecular characterization	Screening expression
HDR knock-in	Variable, 1-3 weeks	3 days	1 month	1-2 days (PCR, Gel image)	Optional though recommended: Cross with UAS-GFP, 5-14 days
CRISPaint knock-in	N/A	3 days	1 month	3 days (PCR, Gel image, Sequence)	Cross with UAS-GFP, 5-14 days

Table 1 – Knock-in timeline comparison in S2R+ cells and in vivo for CRISPaint and HDR

Supplemental Figure 1

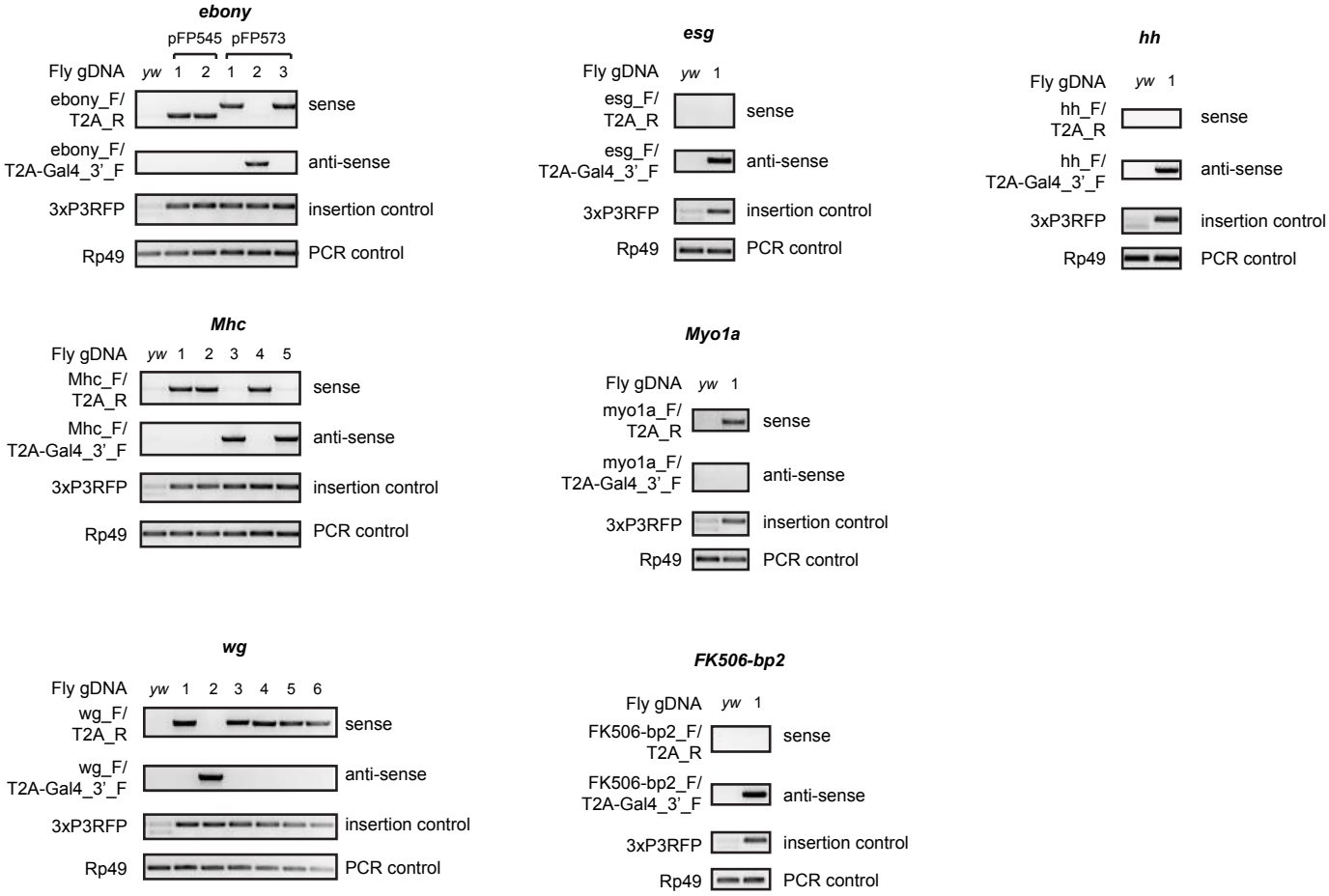


Supplemental Figure 1 – Next-generation sequencing analysis of mNeonGreen insertion sites in transfected S2R+ cells using CRISPresso2

Supplemental Figure 2

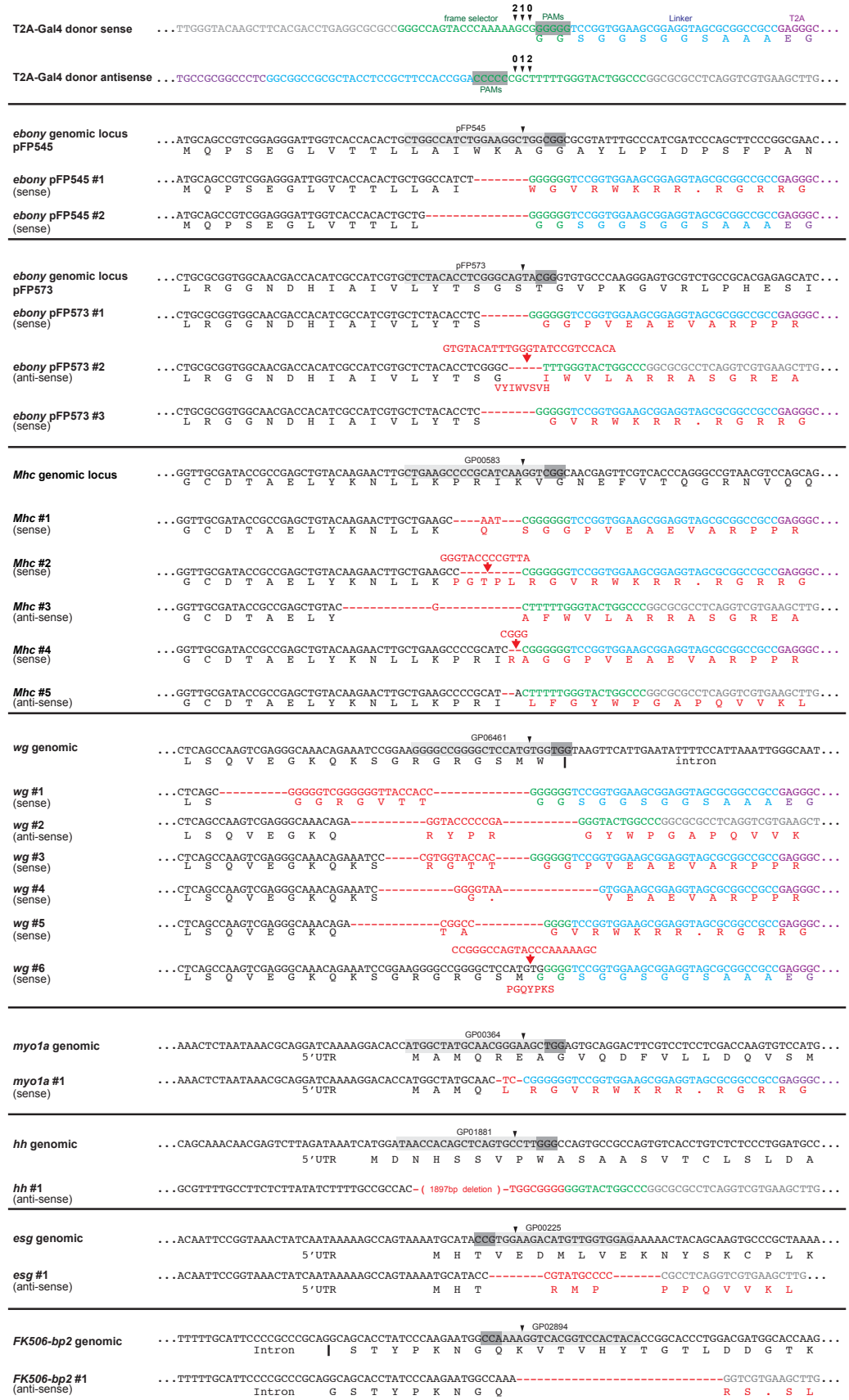
Sense insertion

Antisense insertion



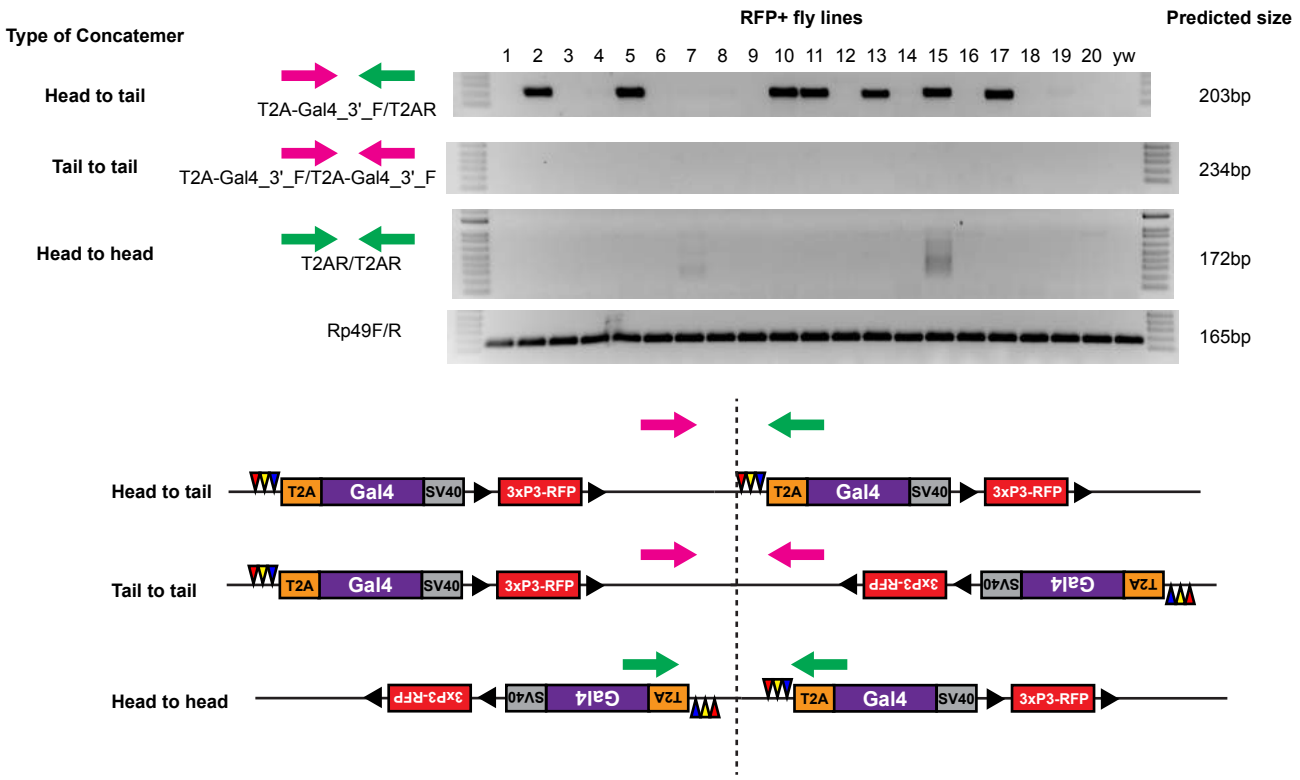
Supplemental Figure 2 – Diagnostic PCR of *CRISPaint-T2A-Gal4* insertions in RFP+ fly lines to confirm their insertion site and orientation.

Supplemental Figure 3



Supplemental Figure 3 – Sequence structure of *CRISPaint-T2A-Gal4* insertions in RFP+ fly lines

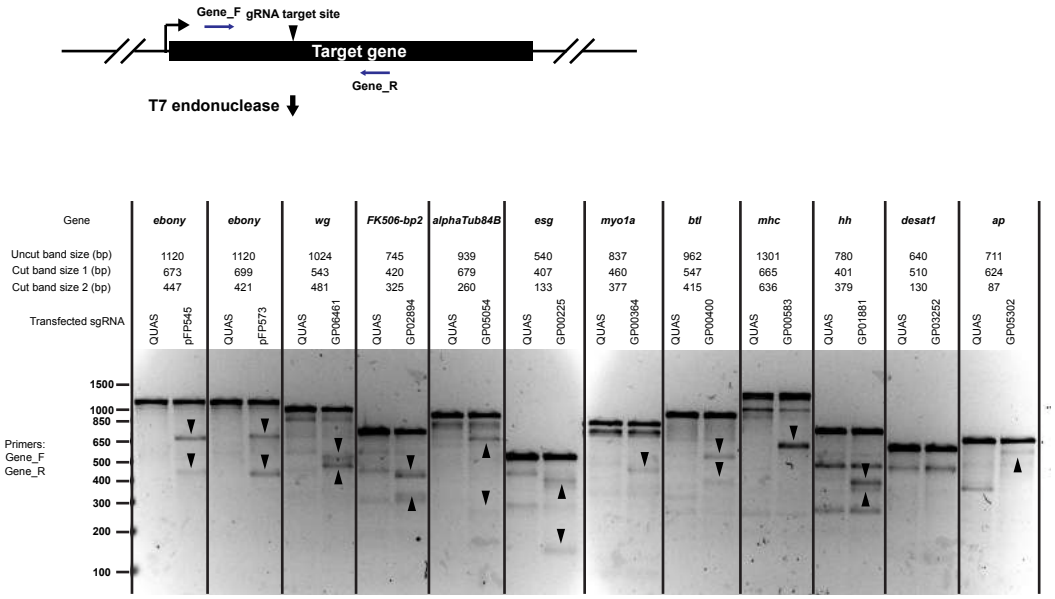
Supplemental Figure 4



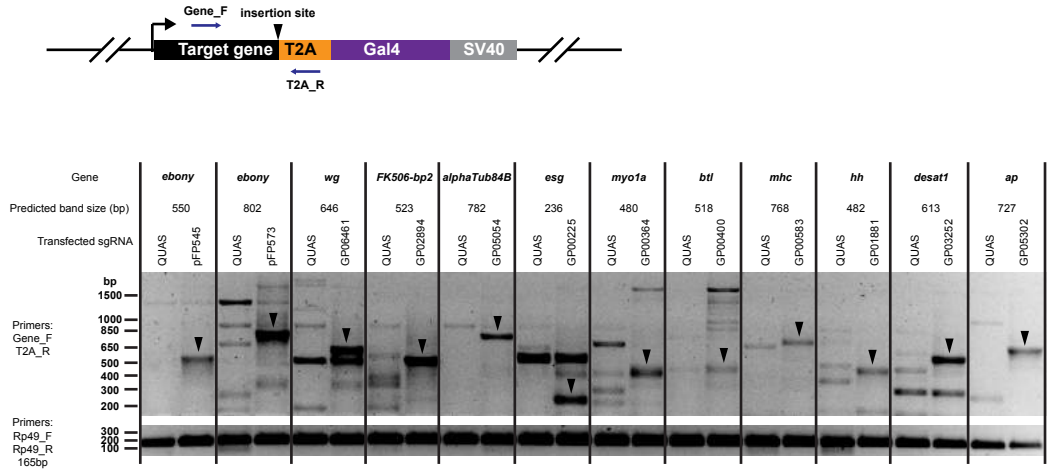
Supplemental Figure 4 – Diagnostic PCR of *CRISPaint-T2A-Gal4* insertions in RFP+ fly lines to detect concatemers

Supplemental Figure 5

A

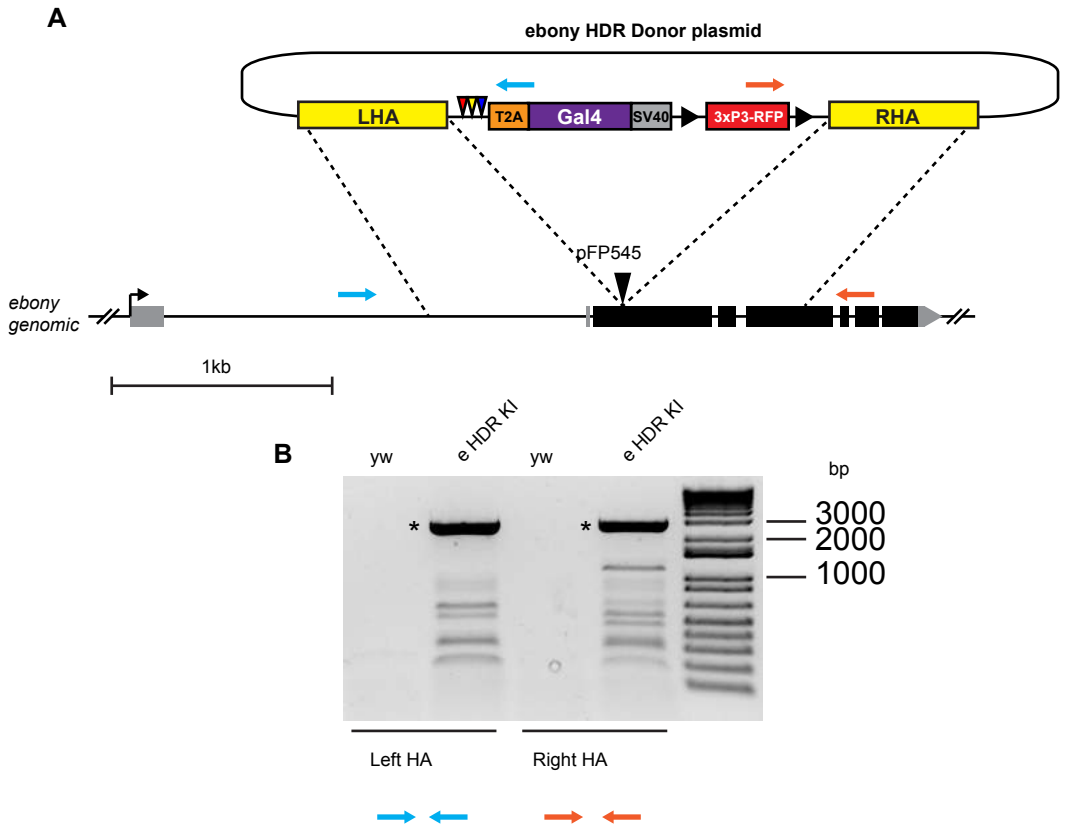


B



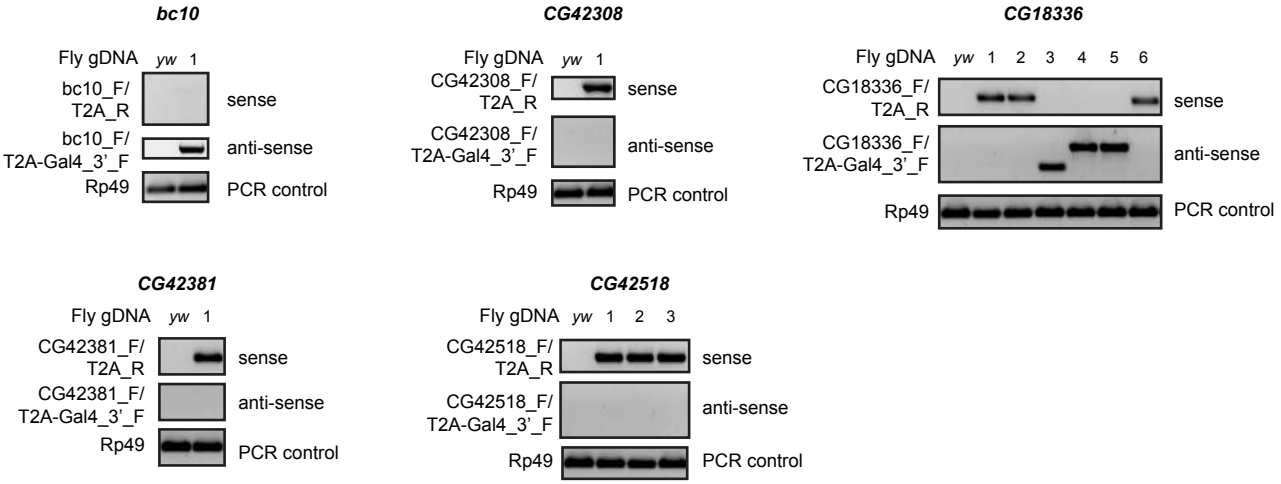
Supplemental Figure 5 – Cutting efficiency of 12 sgRNAs in transfected S2R+ cells. (**A**) T7 endonuclease assay. (**B**) Diagnostic PCR to detect for presence of sense orientation *CRISPaint-T2A-Gal4* insertion events.

Supplemental Figure 6



Supplemental Figure 6 – Construction of *ebony-Gal4* HDR insertion. (**A**) Targeting strategy of *ebony* HDR donor plasmid and *ebony* genomic locus. (**B**) Diagnostic PCR of *ebony-Gal4* HDR insertion. Asterisks indicate correct size band.

Supplemental Figure 7



Supplemental Figure 7 – Diagnostic PCR of CRISPaint-T2A-Gal4 insertions in smORF genes

	Gene and sgRNA information			
Gene Name	**Protein localization**	**sgRNA plasmid (TRiP)**	**sgRNA target site with PAM**	**frame selector sgRNA**
Act5c	actin	GP07595	GCGGTGCACAATGGAGGGGCCGG	2
His2Av	chromatin	GP07596	GGTGCAGGATCCGCAGCGGAAGG	2
αTub84B	microtubules	GP07609	CGGCGAGGGTGAGGGCGCTGAGG	2
Lamin	nuclear envelope	GP07612	CGCAGCGTGACCGCCGTGGACGG	1
Untransfected				

	CRISPresso2 output			
Gene Name	**seamless**	**in-frame in/del**	**frameshift in/del**	**% in-frame**
Act5c	5543	1343	13068	34.50937155
His2Av	2384	2620	8279	37.6722126
αTub84B	726	3034	6901	35.26873652
Lamin	971	801	1757	50.21252479
Untransfected				

	FACs Transfected			FACs Puro-selected		
Gene Name	**mNeonGreen+**	**Single cells total**	**% mNeonGreen+**	**mNeonGreen+**	**Single cells total**	**% mNeonGreen**
Act5c	1963	80589	2.435816302	45355	50883	89.1358607
His2Av	1554	80291	1.935459765	32780	46298	70.80219448
αTub84B	150	78752	0.190471353	16682	52643	31.6889235
Lamin	385	81688	0.471305455	34826	49234	70.73567047
Untransfected	2	83869	0.002384671	10	55656	0.017967515

	Cell counting transfected (n=6)			Cell counting Puro-selected (n=3)		
Gene Name	**mNeonGreen+**	**Total cells**	**Average % mNeonGreen+**	**mNeonGreen+**	**Total cells**	**Average % mNeonGreen+**
Act5c	17/17/26/18/22/25	571/679/577/599/574/613	3.5	342/280/318	397/398/377	80.3
His2Av	26/21/12/19/13/19	634/634/568/525/669/612	3	314/327/315	397/402/400	79.7
αTub84B				412/400/403	412/400/403	100
Lamin				270/282/259	382/388/374	70.9
Untransfected						

Supplemental Table 1 – Quantification of CRISPaint-mNeonGreen insertion events in transfected and puro-selected S2R+ cells by CRISPresso2, FACs, and cell counting of confocal images.

					endogenous gene			
Clone name	**Gene-mNeonGreen**	**predicted mNeonGreen fusion (kDa)**	**Observed Localization**	**mNeon insertion site**	**allele 1**	**allele 2**	**allele 3**	**allele 4**
A3	Act5c	71.4	very infrequent small rods	seamless	1bp deletion	1bp insertion		
A5	Act5c	71.4	very frequent long rods	seamless	numerous in/dels			
A19	Act5c	71.4	no rods, cyto and nuclear	N/A	numerous in/dels			
B1	His2Av	44.5	cyto and nuclear	N/A	4bp deletion	19bp deletion		
B11	His2Av	44.5	nuclear	seamless	3bp deletion	2bp deletion	1bp deletion	27bp deletion
B14	His2Av	44.5	nuclear	3 bp deletion, 1bp change	19bp deletion	3bp deletion	1bp deletion	
C2	alphatub84B	79.5	mitotic spindles	seamless	wt	6bp deletion	1bp deletion	
C6	alphatub84B	79.5	mitotic spindles	seamless	insertion of 1482bp from alphatub84D			
C13	alphatub84B	79.5	mitotic spindles	3bp deletion	wt	6bp insertion		
D1	Lam	100.9	cytoplasmic, nuclear	N/A	8bp deletion	4bp deletion	3bp insertion	
D6	Lam	100.9	nuclear lamina	seamless	3bp deletion			
D9	Lam	100.9	nuclear lamina, polarized	seamless	N/A			

Supplemental Table 2 – Molecular characterization of single cell cloned mNeonGreen-expressing S2R+ lines.

Gene	Chr.	target gene sgRNA	target site (no PAM)	Frame selector sgRNA	# G0 crosses	# G0 founders	% founders/total	# G1 progeny screened	# G1 RFP+ progeny	% RFP G1/total G1
Mhc	2	GP00583	CTGAAGCCCCGCATCAAGGT	1	29	6	21	2086	21	1.01
ebony	3	pFP545	TGGCCATCTGGAAGGCTGG	1	16	3	19	1298	11	0.85
FK506-bp2	2	GP02894	TGTAGTGGACCGTGACCTTT	1	13	2	15	901	2	0.22
wg	2	GP06461	GGGGCCGGGGCTCCATGTGG	0	40	6	15	4423	26	0.59
ebony	3	pFP573	TCTACACCTCGGGCAGTAC	1	23	3	13	2121	27	1.27
esg	2	GP00225	CTCCACCAACATGTCTTCCA	2	15	1	7	772	1	0.13
Myo1a	2	GP00364	ATGGCTATGCAACGGGAAGC	1	17	1	6	860	1	0.12
hh	3	GP01881	TAACCACAGCTCAGTGCCTT	2	39	2	5	1975	3	0.15
alphaTub84B	3	GP05054	GCTGACGGTAGGTTCCGGTA	1	6	0	0	506	0	0.00
btl	2	GP00400	GGCAAAAGTGCCGATCACGC	2	26	0	0	2085	0	0.00
Desat1	3	GP03252	AAGCTGCAGGAGGACTCCAC	1	7	0	0	386	0	0.00
ap	2	GP05302	CACCACCTGTAGCACATCAA	0	29	0	0	2928	0	0.00

Supplemental Table 3 – Germ line knock-in efficiency of *CRISPaint-T2A-Gal4*

#	Gene targeted	sgRNA	G0 cross #	Head-tail insertion by PCR?	Splinkerette 5' (PCR Purification)	Splinkerette 5' (Gel Purification)	Splinkerette 3' (PCR Purification)	Splinkerette 3' (Gel Purification)	Insert orientation	Insert frame	Insertion site sequence	Gal4 expression	Homozygous phenotype	Complementation test
1	ebony	pFP545	1	no	ebony 5' insertion site	no bands	ebony 3' insertion site	ebony 3' insertion site	Sense	Out	8bp del	none	dark cuticle pigment	dark cuticle pigment over e1 on TM3
2	ebony	pFP545	2	yes	ebony 3' insertion site	Band 1: ebony 3' insertion site Band 2: 3' end of donor plasmid	5' end of donor plasmid	5' end of donor plasmid	Sense	In	15bp del	anterior and posterior trachae in larvae, whole body cuticle in pupae and adult	lethal	dark cuticle pigment over e1 on TM3
3	ebony	pFP573	1	no	internal donor plasmid	internal donor plasmid	vermillion	vermillion	Sense	Out	7bp del	none	dark cuticle pigment	dark cuticle pigment over e1 on TM3
4	ebony	pFP573	2	no	ebony 3' insertion site	ebony 3' insertion site	vermillion	vermilllion	Antisense	N/A	5bp del, 25bp ins	none	dark cuticle pigment	dark cuticle pigment over e1 on TM3
5	ebony	pFP573	3	yes	3' end of donor plasmid	3' end of donor plasmid	vermillion	bad sequence data	Sense	Out	8bp deletion	none	dark cuticle pigment	dark cuticle pigment over e1 on TM3
6	hh	GP01881	1	no	bad sequence data	no bands	vermillion	Band 1: hh 5' insertion site Band 2: hh 3' insertion site Band 3: vermillion	Antisense	N/A	1897bp del, 8bp ins	none	lethal	fail to complement hh[AC] and Df(3R)ED5296
7	Mhc	GP00583	1	no	Mhc 5' insertion site	no bands	vermillion	Band 1: Mhc 5' insertion site Band 2: vermillion	Sense	Out	10bp del, 3bp ins	muscle	lethal	fail to complement P{lacW}Mhck10423 and Df(2L)H20
8	Mhc	GP00583	2	no	Mhc 5' insertion site	no bands	bad sequence data	bad sequence data	Sense	Out	9bp del, 13bp ins	muscle	lethal	fail to complement P{lacW}Mhck10423 and Df(2L)H20
9	Mhc	GP00583	3	no	Mhc 3' insertion site	Mhc 3' insertion site	bad sequence data	bad sequence data	Antisense	N/A	27bp del, 1bp ins	none	lethal	fail to complement P{lacW}Mhck10423 and Df(2L)H20
10	Mhc	GP00583	4	yes	Mhc 3' insertion site	Band 1: 3' end of donor plasmid Band 2: 3' end of donor plasmid	bad sequence data	Band 1: 5' end of donor plasmid Band 2: internal donor plasmid	Sense	Out	2bp del, 4bp ins	none	lethal	fail to complement P{lacW}Mhck10423 and Df(2L)H20
11	Mhc	GP00583	5	yes	Mhc 3' insertion site	Band 1: Mhc 3' insertion site Band 2: 3' end of donor plasmid	5' end of donor plasmid	5' end of donor plasmid	Antisense	N/A	2bp deletion	none	lethal	fail to complement P{lacW}Mhck10423 and Df(2L)H20
12	Myo1a	GP00364	1	no	Myo1a 5' insertion site	Myo1a, 5' insertion site	internal donor plasmid	internal donor plasmid	Sense	Out	4bp del, 2bp ins	gut	viable	N/A
13	esg	GP00225	1	yes	esg, 3' insertion site	Band 1: esg, 3' insertion site Band 2: bad sequence data	vermillion	Band 1: 5' end of donor plasmid Band 2: vermillion	Antisense	N/A	25bp del, 10bp	none	lethal	fail to complement P{enG}esg[G66]
14	FK506-bp2	GP02894	1	no	FK506-bp2 3' insertion site	FK506-bp2 3' insertion site	FK506-bp2 5' insertion site	Band 1: FK506-bp2 5' insertion site Band 2: vermillion	Antisense	N/A	30bp del	none	viable	N/A
15	wg	GP06461	1	yes	wg 3' insertion site	Band 1: wg 3' insertion site Band 2: 3' end of donor plasmid	vermillion	Band 1: wg 5' insertion site Band 2: 5' end of donor plasmid	Sense	In	45bp del, 21bp ins	wingless, strong expression	lethal	fail to complement wg[l-17], wg[l-8], and Df(2L)BSC291
16	wg	GP06461	2	no	3' end of donor plasmid	3' plasmid	wg 5' insertion site	5' end of donor plasmid	Antisense	N/A	33bp del, 11bp ins	none	lethal	fail to complement wg[l-17], wg[l-8], and Df(2L)BSC291
17	wg	GP06461	3	yes	wg 5' insertion site	Band 1: wg 5' insertion site Band 2: wg 3' insertion site Band 3: 3' end of donor plasmid	vermillion	wg 3' insertion site	Sense	Out	21bp del, 11bp ins	none	lethal	fail to complement wg[l-17], wg[l-8], and Df(2L)BSC291
18	wg	GP06461	4	no	wg 5' insertion site	Band 1: wg 5' insertion site Band 2: 3' end of donor plasmid	wg 3' insertion site	Band 1: wg 3' insertion site Band 2: attB sequence?	Sense	Out	32bp del, 7bp ins	wingless, weak expression	lethal	fail to complement wg[l-17], wg[l-8], and Df(2L)BSC291
19	wg	GP06461	5	no	wg 5' insertion site	wg 5' insertion site	wg 3' insertion site	vermillion	Sense	Out	28bp del, 5bp ins	none	lethal	fail to complement wg[l-17], wg[l-8], and Df(2L)BSC291
20	wg	GP06461	6	no	wg 5' insertion site	Band 1: wg 5' insertion site Band 2: wg 3' insertion site	vermillion	no band	Sense	In	21bp insertion	none	lethal	fail to complement wg[l-17], wg[l-8], and Df(2L)BSC291

Supplemental Table 4 – Molecular and phenotypic characterization of 20 RFP+ fly strains, each carrying a distinct *CRISPaint-T2A-Gal4* insertion.

gene	sgRNA name	sgRNA sequence	DRSC efficiency score
ebony	pFP545	gTGGCCATCTGGAAGGCTGG	5.56968
ebony	pFP573	gTCTACACCTCGGGCAGTAC	5.08877
wg	GP06461	GGGGCCGGGGCTCCATGTGG	5.61857
FK506-bp2	GP02894	TGTAGTGGACCGTGACCTTT	6.5446
alphaTub84B	GP05054	GCTGACGGTAGGTTCCGGTA	6.20359
esg	GP00225	CTCCACCAACATGTCTTCCA	9.35069
Myo1a	GP00364	ATGGCTATGCAACGGGAAGC	7.98579
btl	GP00400	GGCAAAAGTGCCGATCACGC	3.52487
Mhc	GP00583	CTGAAGCCCCGCATCAAGGT	8.00559
hh	GP01881	TAACCACAGCTCAGTGCCTT	6.99912
Desat1	GP03252	AAGCTGCAGGAGGACTCCAC	6.60024
ap	GP05302	CACCACCTGTAGCACATCAA	5.73411

Supplemental Table 5 – Efficiency scores for 12 sgRNAs used in germ line knock-ins

Gene targeted	FBgn	Size (aa)	Human homolgoue	sgRNA	sgRNA target site	G0 cross #	Insert orientation
bc10	FBgn0040239	79	BLCAP	GP01409	GGGTTTCGGTATCAGTAGTA	1	Antisense
CG42308	FBgn0259204	61	C12orf73	GP01629	ATGCCCGCCGGAGTTTCCTG	1	Sense
CG18336	FBgn0040763	94	FAM166A	GP02853	GGGTAGCACCATCCCTCATC	1	Sense
						2	Sense
						3	Antisense
						4	Antisense
						5	Antisense
						6	Sense
CG42381	FBgn0259727	76	CMC4	GP01437	TGGTCAGACAGGCTTGTATG	1	Sense
CG42518	FBgn0260392	85	NCBP2-AS2	GP06479	GCAGCTCATCCAGCGCATGG	1	Sense
						2	Sense
						3	Sense

Supplemental Table 6 – smORF genes targeted for knock-in with pCRISPaint-T2A-Gal4-3xP3-RFP

Name	Sequence	How used
JB880_frame+0_gRNA_top	GTCGgccagtacccaaaaagcggg	for cloning pCFD3-CRISPaint_frameselector
JB881_frame+0_gRNA_bot	AAACcccgcttttttgggtactggc	for cloning pCFD3-CRISPaint_frameselector
JB882_frame+1_gRNA_top	GTCGggccagtacccaaaaagcgg	for cloning pCFD3-CRISPaint_frameselector
JB883_frame+1_gRNA_bot	AAACccgcttttttgggtactggcc	for cloning pCFD3-CRISPaint_frameselector
JB884_frame+2_gRNA_top	GTCGgggccagtacccaaaaagcg	for cloning pCFD3-CRISPaint_frameselector
JB885_frame+2_gRNA_bot	AAACcgcttttttgggtactggccc	for cloning pCFD3-CRISPaint_frameselector
JB886_target-T2A-Gal4overlap_gBlock	ttcacgacctgaggcgcgccgggccagtacccaaaaagcgggggtccggtggaagcggaggtagcgcggccgccgagg gccgcggcagcctgctgacctgcggcgatgtggaggagaaccccgggcccGCTAGCatgaagctactgtcttctatcga acaagcatgcgatatttgccgacttaaaaagctcaagtgctccaaagaaaaaccgaagtgcgccaagtgtct	gBlock containing CRISPaint site and T2A and overlap sequence
JB877_Gal4_F	atgaagctactgtcttctatcgaaca	to amplify Gal4-3xP3RFP
JB878_target-T2A-Gal4_P3RFP_R	gggaacaaaagctggagctcataacttcgtatagcatacattatacgaagttatCGTATGGGCCTTCGCTGCTTACAG	to amplify Gal4-3xP3RFP
JB915_sfGFP_HIKI_T2A_F	tgtggaggagaaccccgggcccgctagcGTGTCCAAGGGCGAGGAG	for cloning into JAB290 cut with NheI/KpnI
JB916_sfGFP_HIKI_R	caaagatcctctagaggtaccCTACTTGTACAGCTCATCCATGC	for cloning into JAB290 cut with NheI/KpnI
JB1000_LexGAD_HIKI_T2A_F	tgtggaggagaaccccgggcccgctagcATGCCACCCAAGAAGAAGC	for cloning into JAB290 cut with NheI/KpnI
JB1001_LexGAD_HIKI_R	caaagatcctctagaggtaccCTACTCCTTCTTTGGGTTCGG	for cloning into JAB290 cut with NheI/KpnI
JB998_QF2_HIKI_T2A_F	tgtggaggagaaccccgggcccgctagcATGCCACCCAAGCGCAAA	for cloning into JAB290 cut with NheI/KpnI
JB999_QF2_HIKI_R	caaagatcctctagaggtaccTCACTGTTCGTATGTATTAATGTCG	for cloning into JAB290 cut with NheI/KpnI
JB1055_Cas9-T2A-EGFP_HIKI_T2A_F	tgtggaggagaaccccgggcccgctagcATGGATTACAAGGATCACGATG	for cloning into JAB290 cut with NheI/KpnI
JB1056_Cas9-T2A-EGFP_HIKI_R	caaagatcctctagaggtaccTTAGTACAGCTCGTCCATGCC	for cloning into JAB290 cut with NheI/KpnI
JB1002_FLPo_HIKI_T2A_F	tgtggaggagaaccccgggcccgctagcatgagccagttcgacatcct	for cloning into JAB290 cut with NheI/KpnI
JB1003_FLPo_HIKI_R	caaagatcctctagaggtacctcagatccgcctgttgatgtagc	for cloning into JAB290 cut with NheI/KpnI
JB1004_Gal80_HIKI_T2A_F	tgtggaggagaaccccgggcccgctagcATGGACTACAACAAGAGATCTTCG	for cloning into JAB290 cut with NheI/KpnI
JB1005_Gal80_HIKI_R	caaagatcctctagaggtaccTTATAAACTATAATGCGAGATATTGCTAA	for cloning into JAB290 cut with NheI/KpnI
JB1008_Nluc_HIKI_T2A_F	tgtggaggagaaccccgggcccgctagcATGGTCTTCACACTCGAAGATTT	for cloning into JAB290 cut with NheI/KpnI
JB1009_Nluc_HIKI_R	caaagatcctctagaggtaccTTACGCCAGAATGCGTTC	for cloning into JAB290 cut with NheI/KpnI
JB1032_Gal4DBD_HIKI_T2A_F	tgtggaggagaaccccgggcccgctagcATGCTGGAGATCCGCGCC	for cloning into JAB290 cut with NheI/KpnI
JB1033_Gal4DBD_HIKI_R	caaagatcctctagaggtaccTTACGATACCGTCAGTTGCC	for cloning into JAB290 cut with NheI/KpnI
JB1030_p65_HIKI_T2A_F	tgtggaggagaaccccgggcccgctagcATGGATAAAGCGGAATTAATTCC	for cloning into JAB290 cut with NheI/KpnI
JB1031_p65_HIKI_R	caaagatcctctagaggtaccTTACTTGCCGCCGCCCAG	for cloning into JAB290 cut with NheI/KpnI
JB917_sfGFP_HIKI_F	ggtggaagcggaggtagcgcggccgccGTGTCCAAGGGCGAGGAG	for cloning into JAB290 cut with NotI/KpnI
JB916_sfGFP_HIKI_R	caaagatcctctagaggtaccCTACTTGTACAGCTCATCCATGC	for cloning into JAB290 cut with NotI/KpnI
JB1355_FK506-bp2_geno_1F	acgcgccaaaatacaaaaac	amplification of endogenous target gene and knock-in
JB1356_FK506-bp2_geno_1R	GCCTATTCGACCTTGAGCAG	amplification of endogenous target gene
JB1357_alphaTub84B_geno_1F	tttgtgtgggcaaaattcaa	amplification of endogenous target gene and knock-in
JB1358_alphaTub84B_geno_1R	GCTTGGACTTCTTGCCGTAG	amplification of endogenous target gene
JB1359_esg_geno_1F	CGTTTGGTATTTGTGCATCG	amplification of endogenous target gene and knock-in
JB1360_esg_geno_1R	GTAGGGCGACATGTGGAAGT	amplification of endogenous target gene
JB1361_btl_geno_1F	aactaagggaggggcaaaaa	amplification of endogenous target gene and knock-in
JB1362_btl_geno_1R	CGTCCACCAAGGATTTGAGT	amplification of endogenous target gene
JB1363_Desat1_geno_1F	aatccacctggtgcttgttc	amplification of endogenous target gene and knock-in
JB1364_Desat1_geno_1R	GTAACCGAAGGCGATGATGT	amplification of endogenous target gene
JB1365_ap_geno_1F	ttgcaaatctgtcaggaacg	amplification of endogenous target gene and knock-in
JB1366_ap_geno_1R	ATCTGGACACGAGGATGAGG	amplification of endogenous target gene
JB959_ebony_geno_F	gcattagcctgcattgcata	amplification of endogenous target gene and knock-in
JB960_ebony_geno_R	CACGCCCTCATCGAAATAGT	amplification of endogenous target gene
JB961_wg_geno_F	CAGTTAAGCGTTGGCACTGA	amplification of endogenous target gene and knock-in
JB962_wg_geno_R	ttgttgcatctctgcggtag	amplification of endogenous target gene
JB963_Myo1a_geno_F	TCGTCGTCATCAACAGAAGC	amplification of endogenous target gene and knock-in
JB964_Myo1a_geno_R	tctggagtggaaccgaaaac	amplification of endogenous target gene
JB965_Mhc_geno_F	cggctaaagactgacccaaa	amplification of endogenous target gene and knock-in
JB966_Mhc_geno_R	CTCTTGCTCCATGACGAACA	amplification of endogenous target gene
JB967_hh_geno_F	TCGTACTCGCACTCGAACAC	amplification of endogenous target gene and knock-in
JB1388_hh_geno_6F	aaatcaaagctggaccaaatc	amplification of endogenous target gene and knock-in, used with hh #1 insertion
JB968_hh_geno_R	GTTGTAGTTGGGCACGAGGT	amplification of endogenous target gene
JB900_T2A_R	cggggttctcctccacat	reverse primer for amplifying T2A-Gal4 insert in sense orientation
JB958_T2A-Gal4_3'_F	gttttcccagtcacgacgtt	reverse primer for amplifying T2A-Gal4 insert in antisense orientation
JB659_3P3dsred_seq1F	ACTCCAAGCTGGACATCACC	to amplify 3xP3-dsred as a control
JB660_3P3dsred_seq1R	CGAGGGTTCGAAATCGATAA	to amplify 3xP3-dsred as a control
JB713_Rp49_F	ATCGGTTACGGATCGAACAA	to amplify endogenous Rp49 gene as a control
JB714_Rp49_R	GACAATCTCCTTGCGCTTCT	to amplify endogenous Rp49 gene as a control
JB1051_alphaTub84B_C-term_F	CCTTCGTCCACTGGTACGTT	amplification of endogenous target gene and knock-in
JB1052_Act5c_C-term_F	CGTCGACCATGAAGATCAAG	amplification of endogenous target gene and knock-in
JB1053_His2Av_C-term_F	CTCCTCGCCACTTACAGCTC	amplification of endogenous target gene and knock-in
JB1054_Lam_C-term_F	GCCGACAACACTAGGACGAT	amplification of endogenous target gene and knock-in
JB1050_mNeonGreen_R	GGGAGAGAGGCGTTATCCTC	reverse primer for amplifying gene-mNeonGreen insert
JB1120_act5c_seqnextgenR	CGACTTCTCCTCCTCCTCCT	amplification of endogenous target gene
JB1121_His2av_seqnextgenR	TCGTCGGTGTTTTAGCTTGTC	amplification of endogenous target gene
JB1122_alphatub_seqnexgenR	GCGATTGGAAGCGTAAACAC	amplification of endogenous target gene
JB1123_Lam_seqnextgenR	GTGTTGTGCTGCGTTTGATT	amplification of endogenous target gene
JB1192_Lam_2F	GCGGCTAATCAACGAGAAAG	amplification of non-knock-in Lamin gene
JB1193_Lam_2R	TCTGTTGTCAGGAGCGTTTG	amplification of non-knock-in Lamin gene
JB1194_Lam_3F	ACGAGGAGCAGATGCAGATT	amplification of non-knock-in Lamin gene
JB1195_Lam_3R	GGTCTAAACCGGGAGAAAGC	amplification of non-knock-in Lamin gene
JB1196_Lam_4F	AGCTGCAGAACCTGAACGAT	amplification of non-knock-in Lamin gene
JB1197_Lam_4R	ACTAGCCGAACCCAGGATTT	amplification of non-knock-in Lamin gene
JB1198_Lam_5F	cccattacaagcgacgattt	amplification of non-knock-in Lamin gene
JB1199_Lam_5R	GAACAGCTCCACTCCTCCAG	amplification of non-knock-in Lamin gene
JB1200_Lam_6F	GTCTCGGTCTGCTCCTCATC	amplification of non-knock-in Lamin gene
JB1201_Lam_6R	GAATGGCATAGCCACCACTT	amplification of non-knock-in Lamin gene
JB1202_Lam_7F	CATTTGCAAGATGGTGGTTG	amplification of non-knock-in Lamin gene
JB1203_Lam_7R	ccaattaggccaacactgct	amplification of non-knock-in Lamin gene
JB1050_mNeonGreen_R	GGGAGAGAGGCGTTATCCTC	amplification and sequencing primer, 5' end of mNeonGreen
JB1423_Not1-MCS-BamHI_gBlock	agggttttcccagtcacgacgttgtaaaacgacggccagtgaattgtaatacgactcactatagggcgaattgggtaca agcttcacgacctgaggcgcgccgggccagtacccaaaaagcgggggtccggtggaagcggaggtagcgcggccgcac GCTAGCtgCTCGAGacGGTACCtaGAATTCacAGATCTataacttcgtataatgtatgctatacgaagttatcgtacgG GATCTAATTCAATTAGAGACTAATTCAATTAGAGCTAATTCAATTAGGATCCAAGCTTATCGATTTCGAACCCTCGACC GCCGGAGTATAAATAGAGGCGCTTCGTCTACGGAGCGACAATTCAATTCAAACAA	gBlock containing multiple cloning site to construct pCRISPaint-3xP3-RFP
JB1426_pCRISPaint_3xP3RFP_backbone_F	gggccgcgactctagatcataatc	to amplify pCRISPaint-3xP3-RFP backbone to swap out RFP sequence
JB1427_pCRISPaint_3xP3RFP_backbone_R	ggtggcgaccggcttcga	to amplify pCRISPaint-3xP3-RFP backbone to swap out RFP sequence
JB1428_GFP_overlapCRISPaint3xP3RFP_F	gctcgaagccggtcgccaccatggtgtccaagggcgag	to amplify EGFP to Gibson clone into pCRISPaint-3xP3-RFP
JB1429_GFP_overlapCRISPaint3xP3RFP_R	atgatctagagtcgcggcccctacttgtacagctcatccatgc	to amplify EGFP to Gibson clone into pCRISPaint-3xP3-RFP
JB939_PCR1fwrd_frame0_1_2_pCFD5	GCGGCCCGGGTTCGATTCCCGGCCGATGCAgccagtacccaaaaagcgggGTTTTAGAGCTAGAAATAGCAAG	to construct pCFD5-frame_selector_0,1,2
JB940_PCR1rev_frame0_1_2_pCFD5	ccgctttttgggtactggccTGCACCAGCCGGGAATCGAACCC	to construct pCFD5-frame_selector_0,1,2
JB941_PCR2fwrd_frame0_1_2_pCFD5	ggccagtacccaaaaagcggGTTTTAGAGCTAGAAATAGCAAG	to construct pCFD5-frame_selector_0,1,2
JB942_PCR2rev_frame0_1_2_pCFD5	ATTTTAACTTGCTATTTCTAGCTCTAAAACcgctttttgggtactggcccTGCACCAGCCGGGAATCGAACCC	to construct pCFD5-frame_selector_0,1,2
JB1294_ebony_LHA_T2A-Gal4_2F	tgggtacaagcttcacgacctgagggcacgggtaattgtcgttctt	to amplify ebony left homology arm
JB1295_ebony_LHA_T2A-Gal4_2R	ccccgcttttgggtactggcccggtGCCTTCCAGATGGCCAGC	to amplify ebony left homology arm
JB975_ebony_RHA_T2A-Gal4_F	atgtatgctatacgaagttatgagctTGGCGGCGCGTATTTGCC	to amplify ebony right homology arm
JB976_ebony_RHA_T2A-Gal4_R	accctcactaaagggaacaaaagctggagctCCTCAGGAACCTGCGAGTAG	to amplify ebony right homology arm
JB1484_ebony_HDRgeno_F	gggcaaacatagtgcaggtt	use with to verify ebony-T2A-Gal4 HDR knock-in
JB1485_ebony_HDRgeno_R	AGACCCTTGGGCAGGTAGTT	to verify ebony-T2A-Gal4 HDR knock-in
JB1422_bc10_CRISPaint_F	TAAACGACCCGAAAATCGAG	amplification of endogenous target gene and knock-in
JB1413_CG42308_CRISPaint_F	AAACACGGCACCTTTGAATC	amplification of endogenous target gene and knock-in
JB1414_CG18336_CRISPaint_F	TCTCCTGTCAGGCCAGCTAT	amplification of endogenous target gene and knock-in
JB1416_CG42381_CRISPaint_F	CAACCGCCAAGAAGAAGAAG	amplification of endogenous target gene and knock-in
JB1417_CG42518_CRISPaint_F	CCACAAGgtgagtccttggt	amplification of endogenous target gene and knock-in
JB1045_SPLNK-GATC-TOP	GATCCCACTAGTGTCGACACCAGTCTCTAATTTTTTTTTTCAAAAAAA	splinkerette oligos for annealing
JB1046_SPLNK-BOT	CGAAGAGTAACCGTTGCTAGGAGAGACCGTGGCTGAATGAGACTGGTGTCGACACTAGTGG	splinkerette oligos for annealing
JB1047_SPLNK#1	CGAAGAGTAACCGTTGCTAGGAGAGACC	splinkerette round 1 primer
JB1048_SPLNK#2	GTGGCTGAATGAGACTGGTGTCGAC	splinkerette round 2 primer
JB1057_T2A-Gal4_5'#1	agcggagacctttggtttt	5' splinkerette round 1 primer
JB1058_T2A-Gal4_5'#2	tcggtttttctttggagcac	5' splinkerette round 2 primer
JB1059_T2A-Gal4_5'Seq	cggggttctcctccacat	5' splinkerette sequencing primer
JB1060_T2A-Gal4_3'#1	gtgctgcaaggcgattaagt	3' splinkerette round 1 primer
JB1061_T2A-Gal4_3'#2	gttttcccagtcacgacgtt	3' splinkerette round 2 primer
JB1062_T2A-Gal4_3'Seq	GTAAAACGACGGCCAG	3' splinkerette sequencing primer

Supplemental Table 7 – Oligo and dsDNA sequences