Gene Knock-Ins in *Drosophila* Using Homology-Independent Insertion of Universal Donor Plasmids

Justin A. Bosch,* Ryan Colbeth,* Jonathan Zirin,* and Norbert Perrimon*,†,1
*Department of Genetics, Blavatnik Institute, Harvard Medical School and †Howard Hughes Medical Institute, Boston, Massachusetts 02115
ORCID IDs: 0000-0001-8499-1566 (J.A.B.); 0000-0001-7542-472X (N.P.)

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 [clustered regularly interspaced short palindromic repeat (CRISPR)-assisted insertion tagging] homology-independent knock-in method for *Drosophila melanogaster*, which uses CRISPR/Cas9 and nonhomologous end joining to insert "universal" donor plasmids into the genome. 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-red fluorescent protein* 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, insertions/deletions 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. Therefore, homology-independent insertion in *Drosophila* is a fast and simple alternative to HDR that will enable researchers to dissect gene function.

KEYWORDS CRISPR/Cas9; knock-in; homology-independent; gene function; CRISPaint

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 with which 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).

Copyright © 2020 by the Genetics Society of America
doi: https://doi.org/10.1534/genetics.119.302819
Manuscript received September 13, 2019; accepted for publication October 25, 2019; published Early Online November 4, 2019.
Supplemental material available at figshare: https://doi.org/10.25386/genetics.10136588.
[1]Corresponding author: Department of Genetics, Blavatnik Institute, Harvard Medical School, 77 Ave. Louis Pasteur, NRB 336, Boston, MA 02115. E-mail: perrimon@genetics.med.harvard.edu

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 ($\leq$ 10 kb) 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 ($\sim$50–100 bp each) (Bier *et al.* 2018) or long ssDNAs of $\leq$ 2 kb (Quadros *et al.* 2017; Kanca *et al.* 2019) 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 nonhomologous 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/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 has been shown to be a fast, simple, and effective alternative to HDR 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), *Caenorhabditis elegans* (Katic *et al.* 2015), and *Daphnia* (Kumagai *et al.* 2017). However, this method has not yet been utilized in *Drosophila*.

Here, we use the CRISPaint [clustered regularly interspaced short palindromic repeat (CRISPR)-assisted insertion tagging] strategy to show that homology-independent insertion functions effectively in *Drosophila*. First, we inserted a mNeonGreen universal donor plasmid into four target genes in S2R+ cells, and used puromycin resistance (PuroR) to efficiently select for cell lines with fluorescently tagged endogenous proteins. Second, we inserted a *T2A-Gal4 3xP3-RFP* (red fluorescent protein) 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 sequences, 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.

Materials and Methods

Plasmid cloning

pCFD3-frame_selector_0,1,2 plasmids [#127553-127555; Addgene, #1482-1484 Drosophila Genomics Resource Center (DGRC)] were cloned by ligating annealed oligos encoding single guide RNAs (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 (Transgenic RNAi Project: https://fgr.hms.harvard.edu/) or obtained from Filip Port (Port *et al.* 2015). sgRNA plasmids targeting coding sequences close to the stop codon were *GP07595* (*Act5c*) (#130278; Addgene, #1492; DGRC), *GP07596* (*His2Av*), *GP07609* (*αTub84B*), and *GP07612* (*Lam*). sgRNA plasmids targeting coding sequences close to the start codon were *GP06461* (*wg*), *GP02894* (*FK506-bp2*), *GP05054* (*αTub84B*), *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 (#127556; Addgene, #1481; DGRC) was constructed using Gibson assembly (E2611; New England Biolabs, Beverly, MA) of three DNA fragments: (1) *Gal4-SV40-3xP3-RFP* was PCR amplified from *pHD-Gal4-DsRed* (unpublished results; Gratz *et al.* 2014); (2) a linear plasmid backbone generated by digesting *pWalium10-roe* (Perkins *et al.* 2015) with *Asc*I/*Sac*I; and (3) a synthesized double-stranded DNA (dsDNA) fragment (gBlock; Integrated DNA Technologies) encoding the CRISPaint target site, linker sequence, Thosea asigna virus 2A (T2A) self-cleaving peptide, and ends that overlap the other two fragments.

pCRISPaint-3xP3-RFP (# 130276; Addgene, #1490; DGRC) was constructed by digesting *pCRISPaint-T2A-Gal4-3xP3-RFP* with *Not*1/*Bam*HI to remove *T2A-Gal4* sequences and using Gibson to add a gBlock with a multiple cloning site for future manipulation.

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

pCRISPaint-T2A-ORF-3xP3-RFP donor plasmids (#127557–127565; Addgene, #1485–1489, 1493–1497; DGRC) were cloned by PCR amplifying the ORFs and Gibson cloning into *CRISPaint-T2A-Gal4-3xP3-RFP* cut with *Nhe*I/*Kpn*I. ORF sequences were amplified from templates as follows: super folder GFP (*sfGFP*) [amplified from *pUAS* (*upstream activating sequence*)-*TransTimer* (He *et al.* 2019)], *LexGAD* [amplified from p*CoinFLP-LexGAD/Gal4* (*Bosch et al.* 2015)], *QF2* (amplified from #80274; Addgene), *Cas9-T2A-GFP* (amplified from a template kindly provided by Raghuvir Viswanatha, Cas9 from *Streptococcus pyogenes*), *FLPo* (amplified from #24357; Addgene), *Gal80* (amplified from #17748; Addgene), *Nluc* (amplified from #62057; Addgene), *Gal4DBD* (amplified from #26233; Addgene), and *p65* (amplified from #26234; Addgene).

pCRISPaint-sfGFP-3xP3-RFP (#127566; Addgene, DGRC; #1486) was cloned by PCR amplifying the *sfGFP* coding sequence and Gibson cloning into *CRISPaint-T2A-Gal4-3xP3-RFP* cut with *Not*I/*Kpn*I.

pCFD5-frame_selectors_0,1,2 (#131152; Addgene) was cloned by PCR amplifying two fragments encoding the three sgRNAs that cut the CRISPaint target site and Gibson cloning into *pCFD5* (Port and Bullock 2016). This plasmid allows expression of sgRNAs that are separated by transfer RNA sequences, and thus *pCFD5-frame_selectors_0,1,2* expresses all three frame-selector sgRNAs simultaneously.

pLHA-T2A-Gal4-3xP3-RFP-RHA_ebony was cloned by digesting *pCRISPaint-T2A-Gal4-3xP3-RFP* with *Asc*I/*Sac*I, and purifying insert and backbone DNA together (QIAQuick; QIAGEN, Valencia, CA). Homology arms were amplified from *Drosophila* genomic DNA (*yw*, single fly) using Phusion polymerase (Left homology arm (LHA): 1592 bp and Right homology arm: 1544 bp). Insert, backbone, and homology arms were assembled by Gibson cloning.

See Supplemental Material, Table S6 for oligo and dsDNA sequences, and Addgene and the DGRC for plasmid sequences.

Cell culture

Drosophila S2R+ cells stably expressing Cas9 and an mCherry protein trap in *Clic* (S2R+-MT::Cas9, stock #268, DGRC) (Viswanatha *et al.* 2018) were cultured at 25°, using Schneider's medium (21720-024; Thermo Fisher Scientific) with 10% fetal bovine serum (A3912; Sigma [Sigma Chemical], St. Louis, MO) and 50 U/ml penicillin/streptomycin (15070-063; Thermo Fisher Scientific). S2R+ cells were transfected using Effectene (301427; QIAGEN) following the manufacturer's instructions. Plasmid mixes were composed of sgRNA-expressing plasmids (see above) and *pCRISPaint-mNeonGreen-PuroR* (Schmid-Burgk *et al.* 2016). Cells were transfected with plasmid mixes in six-well dishes at 1.8×10^6 cells/ml, split at a dilution of 1:6 after 3–4 days, and incubated with 2 μg/ml puromycin (540411; Calbiochem, San Diego, CA). Every 3–5 days, the media was replaced with fresh puromycin until the cultures became confluent (~12–16 days). For single-cell cloning experiments, cultures were split 1:3 2 days before sorting. Cells were resuspended in fresh media, triturated to break up cell clumps, and pipetted into a cell-straining FACS tube (352235; Corning). Single cells expressing mNeonGreen were sorted into single wells of a 96-well plate containing 50% conditioned media and 50% fresh media using an Aria-594 instrument at the Harvard Medical School Division of Immunology's Flow Cytometry Facility. Once colonies were visible by eye (3–4 weeks), they were expanded and screened for mNeonGreen fluorescence.

Fly genetics and embryo injections

Flies were maintained on standard fly food at 25°. Fly stocks were obtained from the Perrimon laboratory collection or Bloomington *Drosophila* Stock Center (BDSC; indicated with BL#). Stocks used in this study are as follows: *yw* (Perrimon laboratory), *yw/Y hs-hid* (BL8846), *yw*; *nos-Cas9attP40/CyO* (derived from BL78781), *yw*;; *nos-Cas9attP2* (derived from BL78782), *yw*; *Sp hs-hid/CyO* (derived from BL7757), *yw*;; *Dr hs-hid/TM3,Sb* (derived from BL7758), *UAS-2xGFP* (BL6874), *wg1-17/CyO* (BL2980), *wg1-8/CyO* (BL5351), *Df(2L)BSC291/CyO* (BL23676), *Mhc[k10423]/CyO* (BL10995), Df(2L)H20/CyO (BL3180), *Df(2L)ED8142/SM6a* (BL24135), *hh[AC]/TM3 Sb* (BL1749), *Df(3R)ED5296/TM3, Sb* (BL9338), *esgG66/CyO UAS-GFP* (BL67748), *Df(2R)Exel6069/CyO* (BL7551), ywCre; D/TM3, Sb (BL851), Dp(2;1)Sco[rv23]; 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 $NaPO_4$ and 5 mM KCl), and adjusted to 200 ng/μl. Plasmids were mixed equally by volume, and mixes were injected into *Drosophila* embryos using standard procedures. For targeting genes on chromosome 2, plasmid mixes were injected into *yw*;; *nos-Cas9attP2* embryos. For targeting genes on chromosome 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 the collection of large numbers of virgin flies by incubating larvae and pupae at 37° for 1 hr. G1 flies were screened for RFP expression in the adult eye on a Zeiss ([Carl Zeiss], Thornwood, NY) 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° for 2 hr 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-bottomed 384-well plate (6007558; Perkin-Elmer [Perkin Elmer-Cetus], Norwalk, CT). For fixed-cell images, cells were incubated with 4% paraformaldehyde for 30 min, washed with PBS with 0.1% Triton X-100 (PBT) three times for 5 min each, stained with 1:1000 DAPI (D1306; Thermo Fisher Scientific) and 1:1000 phalloidin-tetramethylrhodamine (TRITC) (P1951; Sigma), and washed with PBS. Plates were imaged on an IN Cell Analyzer 6000 (GE Healthcare) using a 20× or 60× objective. Images were processed using Fiji software.

Wing imaginal discs from third instar larvae were dissected in PBS, fixed in 4% paraformaldehyde, and permeabilized in PBT. For Wg staining, carcasses were blocked for 1 hr in 5% normal goat serum (S-1000; Vector Laboratories, Burlingame, CA) at room temperature, and incubated with 1:50 mouse anti-wg (4D4; DSHB) primary antibody and 1:500 anti-mouse 488 (A-21202; Molecular Probes, Eugene, OR) secondary antibody. Primary and secondary antibody incubations were performed at 4° overnight. All carcasses were stained with DAPI and phalloidin-TRITC, and mounted on glass slides with vectashield (H-1000; Vector Laboratories) 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. Preselection cultures were obtained by collecting 500 μl of culture 3–4 days after transfection. Postselection cultures were obtained after at least 2 weeks of puromycin incubation. Nontransfected cells were used as a negative control. For each sample, 100,000 cells were counted, and FlowJo software was used to analyze and graph the data. Forward scatter-A (FSC-A) *vs*. Green Fluorescent Protein-A (GFP-A) was plotted and we defined mNeonGreen+ cells by setting a signal intensity threshold where <0.02% of negative controls were 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 (model). For transfected cells, six fields containing ≥ 200 cells were quantified (*i.e.*, $n = 6$). For puromycin-selected cells, three fields containing ≥ 200 cells were quantified (*i.e.*, $n = 3$).

Western blotting

Single-cell cloned cell lines were grown until confluent and 1 ml of resuspended cells was centrifuged at 250 × *g* for 10 min. The cell pellet was resuspended in 1 ml ice-cold PBS, recentrifuged, and the pellet was lysed in 250 μl 2× SDS-sample buffer and boiled for 5 min. Next, 10 μl was loaded on a 4–20% Mini-Protean TGX SDS-Page gel (4561096; Bio-Rad, Hercules, CA), transferred to a PVDF membrane (IPFL00010; Millipore, Bedford, MA), blocked in 5% nonfat dry milk, then subjected to primary blotting using mouse anti-mNeonGreen (1:1000; Chromtek 32F6) or human Fragment Antigen-Binding (hFAB) Rhodamine anti-Actin (12004164; Bio-Rad), and secondary blotting using 1:3000 anti-mouse HRP (NXA93; Amersham, Piscataway, NJ), imaging using ECL (34580; Thermo Fisher Scientific) on a ChemiDoc MP Imaging System (Bio-Rad).

PCR, sequencing, and sgRNA cutting assays

S2R+ cell genomic DNA was isolated using QuickExtract (QE09050; Lucigen). Fly genomic DNA was isolated by grinding a single fly in 50 μl squishing buffer (10 mM Tris-Cl pH 8.2, 1 mM EDTA, and 25 mM NaCl) with 200 μg/ml proteinase K (3115879001; Roche), and incubating at 37° for 30 min and 95° for 2 min. PCR was performed using Taq polymerase (TAKR001C; Clontech) when running DNA fragments on a gel and Phusion polymerase (M-0530; New England Biolabs) was used when DNA fragments were sequenced. DNA fragments corresponding to *mNeonGreen* or *T2A-Gal4* insertion sites were amplified using primer pairs where one primer binds to the genomic sequence and the other primer binds to the insert. To amplify nonknock-in sites, we used primers that flank the sgRNA target site. Primer pairs used for gel analysis and/or Sanger sequencing were designed to produce DNA fragments <1 kb. Primer pairs used for next-generation sequencing of the insertion site were designed to produce DNA 200–280 bp fragments. DNA fragments were run on a 1% agarose gel for imaging or purified on QIAquick columns (28115; QIAGEN) for sequencing analysis. See Table S6 for oligo sequences.

Sanger sequencing was performed at the Dana-Faber/Harvard Cancer Center DNA Resource Core facility and chromatograms were analyzed using Lasergene 13 software (DNASTAR). Next-generation sequencing was performed at the Massachusetts General Hospital Center for Computational and Integrative Biology DNA Core. Fastq files were analyzed using CRISPresso2 (Clement *et al.* 2019) by entering the PCR fragment sequence into the exon specification window and setting the window size to 10 bases. Quantification of insertion types (seamless, in-frame indel, and frameshift indel) was taken from the allele plot and frameshift analysis outputs of CRISPresso2. The small proportion of "unmodified" reads that were not called by frameshift analysis were not included in the quantification.

T7 endonuclease assays (M0302L; New England Biolabs) were performed following the manufacturer's instructions.

Splinkerette sequencing was performed as previously described (Potter and Luo 2010). Briefly, genomic DNA was isolated by single-fly squishing (described above) and digested using enzyme *Bst*YI (R0523S; New England Biolabs) at 60° overnight, and heat inactivated for 20 min at 80°. Digested DNA was ligated with annealed splinkerette oligonucleotides overnight at 16°. PCR annealing temperatures were 60° (round 1) and 64° (round 2). Round 2 PCR products were either run on an agarose gel and gel fragments purified using a QIAquick Gel Extraction Kit (28704; QIAGEN), or purified using ExoSAP-IT PCR clean-up reagent (78200.200.UL; Applied Biosystems, Foster City, CA). Oligo sequences used for splinkerette sequencing are listed in Table S6. The 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 are available from Addgene and the DGRC. Gal4-expressing fly strains are available at the BDSC (#83624–83627). Two mNeonGreen-expressing S2R+ cell lines are available from the DGRC [His2Av-mNeonGreen clone B11 (DGRC #295) and Lamin-mNeonGreen clone D6 (DGRC #296)]. Oligo and dsDNA sequences are listed in Table S6. The remaining flies, cells, and sequence data are available from the Perrimon laboratory on request. Supplemental material available at figshare: https://doi.org/10.25386/genetics.10136588.

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, which is known to function efficiently in cultured *Drosophila* cells (Bottcher *et al.* 2014; Viswanatha *et al.* 2018) and the germ line (Gratz *et al.* 2013; 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, and luciferase) are seemingly compatible for expression in *Drosophila*.

The CRISPaint system works by introducing three components into Cas9-expressing cells: (1) an 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 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 are predicted not to 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-termini. 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 the 3′ coding sequences of *Drosophila Actin 5C Actin5c*, *His2Av*, *αTub84B*, and *Lamin* because these genes are expressed in S2R+ cells (Hu *et al.* 2017) and encode proteins with known subcellular localizations (actin filaments, chromatin, microtubules, and the nuclear envelope, respectively). For donor plasmid, we used *pCRISPaint-mNeonGreen-T2A-PuroR* (Schmid-Burgk *et al.* 2016), which contains a frame-selector sgRNA target site upstream of coding sequence for the fluorescent mNeonGreen protein and PuroR protein linked by a cleavable T2A peptide sequence. Only integration of the donor plasmid in-frame with the target coding sequence should result in the translation of mNeonGreen-T2A-PuroR (Figure 1A).

We transfected Cas9-expressing S2R+ cells (Viswanatha *et al.* 2018) with a mix of three plasmids: *pCRISPaint-mNeonGreen-T2A-PuroR* donor, target-gene sgRNA, and the appropriate frame-selector sgRNA (Figure 1A and Table S1). As an initial method to detect knock-in events, we used PCR to amplify the predicted insertion sites from transfected cells. Using primers that are specific to the target gene and *mNeonGreen* sequence, we successfully amplified sense-orientation *gene-mNeonGreen* DNA fragments for all four genes (Figure 1B). Furthermore, next-generation sequencing of these amplified fragments revealed that 34–50% of sense-orientation insertions were in-frame with the target gene (Figure 1B, and Figure S1 and Table S1); 7–27% of sense-orientation insertions were seamless, which is slightly lower than previously observed (Schmid-Burgk *et al.* 2016).

Next, we quantified in-frame knock-in frequency by measuring mNeonGreen fluorescence in transfected S2R+ cells. Flow cytometry-based cell counting of transfected cells revealed that the numbers of mNeonGreen+ cells ranged from 0.19 to 2.4% (Figure 1C and Table S1), in agreement with human cultured cells (Schmid-Burgk *et al.* 2016). These results were confirmed by confocal analysis of transfected cells, which showed mNeonGreen fluorescence in a small subset of cells (Figure 1C); 3.2% (*Act5c*) and 2.4% (*His2Av*) of transfected cells expressed mNeonGreen (Figure 1C and Table S1), which roughly agreed with flow cytometry cell counting. Finally, mNeonGreen localized to the expected subcellular compartments, most obviously observed by His2Av-mNeonGreen and Lam-mNeonGreen colocalization with the nucleus, and Act5C-mNeonGreen and *αTub84B*-mNeonGreen exclusion from the nucleus (Figure 1C). These results suggest that a significant number of transfected S2R+ cells received in-frame insertion of mNeonGreen at their C-terminus using the CRISPaint homology-independent insertion method.

For knock-in cells to be useful in experiments, it is important to derive cultures where most cells, if not all, carry the insertion. Therefore, we enriched for in-frame insertion events using puromycin selection (Figure 1D). After a 2-week incubation of transfected S2R+ cells with puromycin, flow cytometry and confocal analysis revealed that 31.7–89.1% of cells expressed mNeonGreen and exhibited correct subcellular localization (Figure 1E and Table S1). For *αTub84B*, cell counting by flow cytometry greatly underestimated the number of mNeonGreen+ cells counted by confocal analysis, likely because mNeonGreen expression level was so low. These results demonstrate that puromycin selection is a fast and efficient method of selecting for mNeonGreen expressing knock-in cells.

A subset of cells in puromycin-selected cultures had no mNeonGreen expression or unexpected localization (Figure 1E). Since each culture is composed of different cells with independent insertion events, we used FACS to derive single-cell cloned lines expressing mNeonGreen for further characterization (Figure 2A). At least 14 single-cell cloned lines were isolated for each target gene and imaged by confocal microscopy. Within a given clonal culture, each cell exhibited the same mNeonGreen localization (Figure 2B), confirming our single-cell cloning approach and demonstrating that the insertion was genetically stable over many cell divisions.

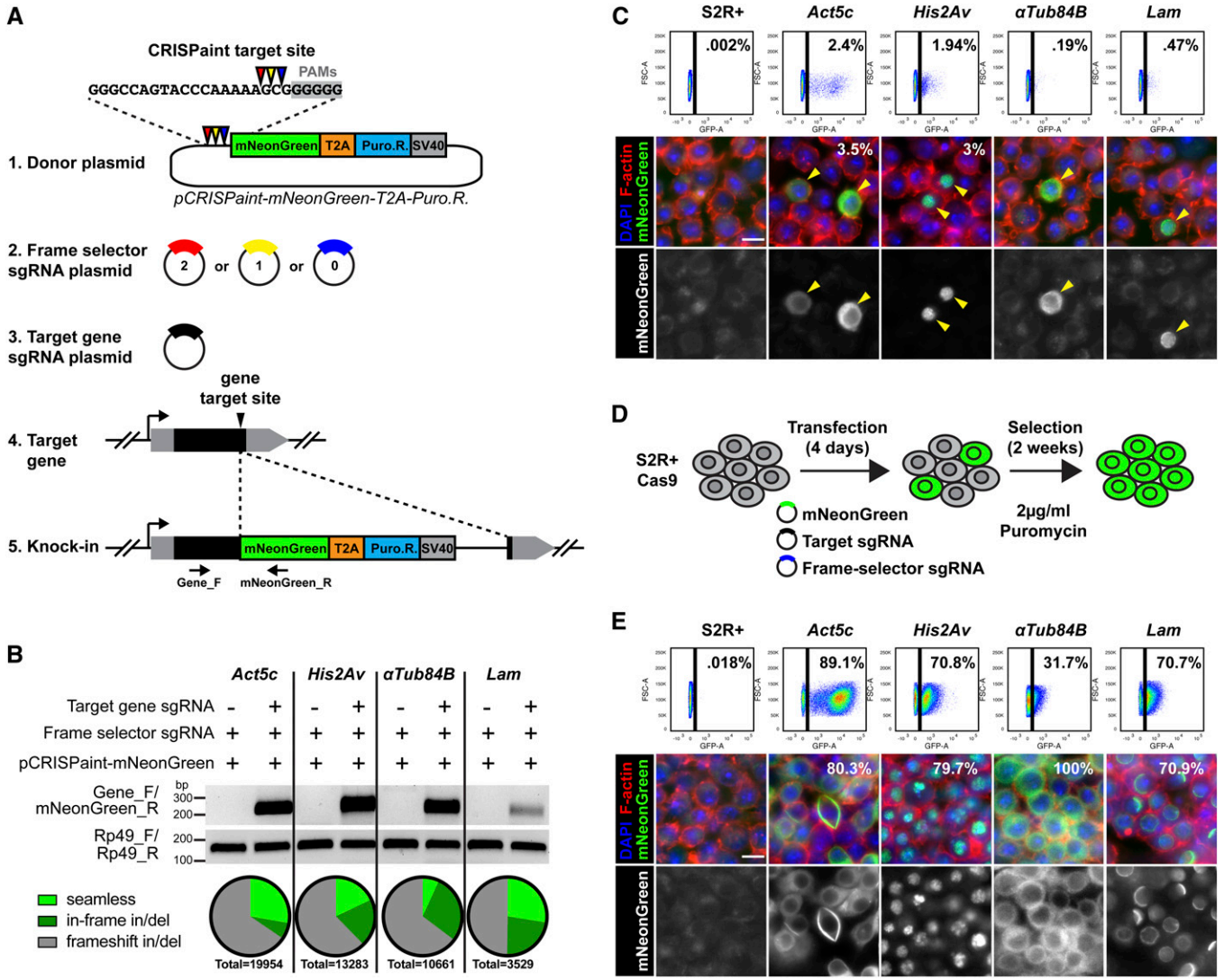

Figure 1 Knock-in of *mNeonGreen-T2A-PuroR* into *Drosophila* S2R+ cells using homology-independent insertion. (A) Schematic of CRISPaint knock-in approach. *mNeonGreen-T2A-PuroR* is inserted into 3′ coding sequence. (B) Analysis of knock-in efficiency of transfected cells by diagnostic PCR (DNA gel image) and next-generation sequencing (pie charts). (C) Analysis of knock-in efficiency of transfected cells by FACs and confocal microscopy. Numbers indicate percentages of cells with fluorescence. F-actin stained using phalloidin-TRITC (red), nuclei labeled with DAPI (blue), mNeonGreen signal is in green. Bar, 10 μm. (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. Bar, 10 μm. CRISPaint, clustered regularly interspaced short palindromic repeat-assisted insertion tagging; Puro.R, puromycin resistance; sgRNA, single guide RNA; T2A, Thosea asigna virus 2A self-cleaving peptide; TRITC, tetramethylrhodamine.

Importantly, while many clones exhibited the predicted mNeonGreen localization, a subset of the clonal cell lines displayed an unusual localization pattern (Figure 2B). For example, 3/21 Act5c-mNeonGreen clones had localization in prominent rod structures and 12/14 Lamin-mNeonGreen clones had asymmetric localization in the nuclear envelope (Figure 2B). In addition, some clones had diffuse mNeonGreen localization in the cytoplasm and nucleus (Figure 2B).

To better characterize the insertions in single-cell cloned lines, we further analyzed three clones per gene (12 total), selecting different classes when possible (correct localization, unusual localization, and diffuse localization) (Figure 2C and Table S2). Using PCR amplification of the predicted insertion site (Figure 1A and Figure 2D) and sequencing of amplified fragments (Figure 2E and Table S2), we determined that all clones with correct or unusual mNeonGreen localization contained an in-frame insertion of mNeonGreen with the target gene. In contrast, we were unable to amplify DNA fragments from the expected insertion site in clones with diffuse mNeonGreen localization (Figure 2D). Western blotting of cell lysates confirmed that only clones with in-frame mNeonGreen insertion express fusion proteins that match the predicted molecular weights (Figure 2F). All together, these results suggest that clones with correct mNeonGreen localization are likely to contain an in-frame insert in the correct target gene.

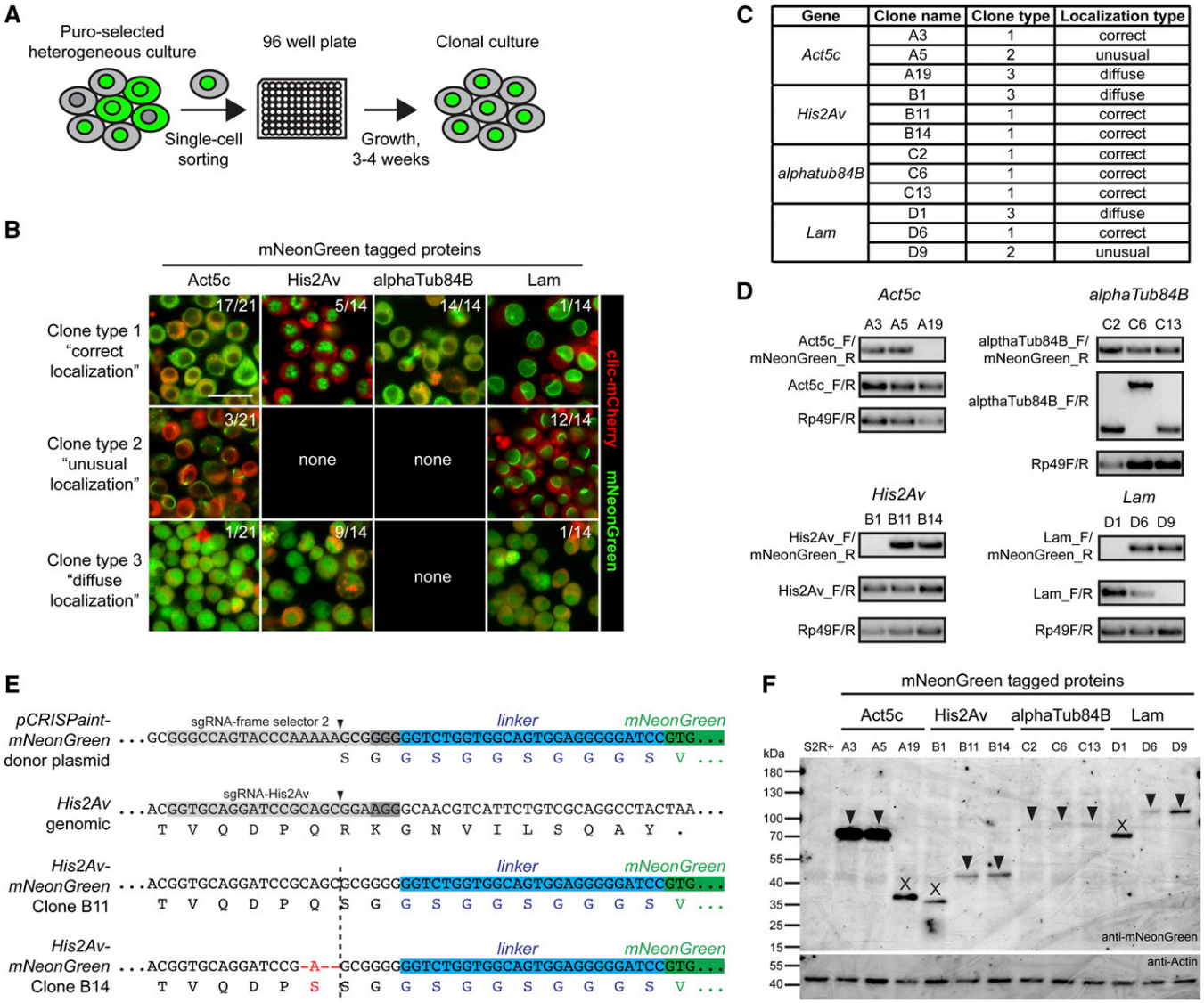

Figure 2 Analysis of S2R+ mNeonGreen-expressing single-cell cloned lines. (A) Schematic of FACS isolation of single-cell clones expressing mNeonGreen. (B) Confocal images of live mNeonGreen-expressing cell lines, categorized into three clone types. Numbers indicate the frequency of each clone type for each gene targeted. Images show fluorescence from Clic-mCherry (red) and mNeonGreen (green). Bar, 25 μm. (C) Single-cell cloned lines retained for further analysis. (D) Agarose gel with PCR fragments amplified from knock-in (Gene_F/mNeonGreen_R) and nonknock-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 gray 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. CRISPaint, clustered regularly interspaced short palindromic repeat-assisted insertion tagging; PAM, protospacer adjacent motif; sgRNA, single guide RNA.

S2R+ cells are polyploid (Lee *et al.* 2014), and clones expressing mNeonGreen could bear one or more insertions. Furthermore, indels induced at the noninsertion locus could disrupt protein function. To explore these possibilities, we amplified the noninsertion locus in our single-cell cloned lines, and used Sanger and next-generation sequencing to analyze the DNA fragments (Figure 2D and Table S2). For each gene, we could find indels occurring at the noninsertion sgRNA cut site. For example, we could distinguish four distinct alleles in clone B11: a 3-bp deletion, a 2-bp deletion, a 1-bp deletion, and a 27-bp deletion. In addition, we identified an unusual mutation in clone C6, where a 1482-bp DNA fragment inserted at the sgRNA cut site, which corresponds to a region from *αTub84D*. We assume that this large insertion was caused by homologous recombination, since *αTub84D* and *αTub84B* share 92% genomic sequence identity (FlyBase, http://flybase.org/). For *Act5c-mNeonGreen* clones A5 and A19, numerous indel sequences were found, suggesting that this region has an abnormal number of gene copies. We were unable to amplify a DNA fragment of the non-insertion allele from Lam-*mNeonGreen* D9, despite follow-up PCRs using primers that bind genomic sequences further away from the insertion site (data not shown).

Loss-of-function knock-in fly lines by homology-independent insertion in the germ line

We next explored whether homology-independent insertion could function in the *Drosophila* germ line for the purpose of generating knock-in fly strains. As opposed to antibiotic selection in cultured cells, we wanted to identify transgenic flies using a visible body marker that was not dependent on target gene expression. In addition, we wanted to target insertions to 5′ coding sequences to create loss-of-function alleles by premature protein truncation. Finally, we wanted to test if insertion of a reporter could capture the expression of the target gene. To accomplish these goals, we constructed a new universal donor plasmid called *pCRISPaint-T2A-Gal4-3xP3-RFP* (Figure 3A). This donor contains a frame-selector sgRNA target site upstream of *T2A-Gal4*, which encodes a self-cleaving form of the transcription factor (Diao and White 2012). This donor plasmid also contains the transgenesis marker *3xP3-RFP*, which expresses red fluorescence in larval tissues and the adult eye (Berghammer *et al.* 1999) (Figure 3A).

To test if homology-independent insertion functions in the germ line and determine its approximate efficiency, we targeted 11 genes using *pCRISPaint-T2A-Gal4-3xP3-RFP* (Table S3). These genes were selected based on their known loss-of-function phenotypes, expression patterns, and availability of sgRNA plasmids from the TRiP (https://fgr.hms.harvard.edu/). Plasmid mixes of donor, frame-selector, and target-gene sgRNA were injected into *nos-Cas9* embryos, and the resulting G0 progeny were crossed to *yw*. G1 progeny were screened for RFP fluorescence in adult eyes, and each RFP+ founder fly was crossed to balancer flies to establish a stable stock (Figure 3B). Figure 3C and Table S3 show the integration efficiency results for each gene, and Table S4 has information on 20 RFP+ lines, each derived from a different G0 founder. For injections that produced RFP+ animals, the frequency of G0 crosses yielding RFP+ G1 progeny ranged from 5 to 21% (Figure 3C and Table S3). For example, when targeting *ebony* with sgRNA *pFP545*, three out of 16 G0 crosses produced ≥ 1 RFP+ G1 flies. In 1 month following embryo injection, we obtained balanced RFP+ lines for 7 out of 11 genes targeted (64%) (Figure 3C and Table S3).

Next, we used PCR and sequencing to confirm the insertion sites in our RFP+ lines. For each target site, we used two primer pairs to detect the presence of an on-target insertion as well as its orientation (Figure S2). Gel images of PCR-amplified DNA showed that all 20 of our RFP+ lines contained an insertion in the correct target site (Figure S2 and Table S4), where 13 insertions were in the sense orientation and 7 were antisense (Figure 3D). Sequence analysis of PCR fragments confirmed that the insertions were present at the target site (Figure S3 and Table S4). Unexpectedly, all 20 lines contained genomic sequence indels at the insertion site (Figure S3), unlike the frequent seamless insertions observed in cultured cells. Germ line indels were also predominantly longer than indels in cultured cells. In particular, a remarkably long 1896-bp genomic deletion was found at the *hh* insertion site. Regardless, these results demonstrate that all 20 independently derived RFP+ lines contained *pCRISPaint-T2A-Gal4-3xP3-RFP* at the target site.

Next, we analyzed insertion lines for loss-of-function phenotypes. Flies with insertions in *wg*, *Mhc*, *hh*, and *esg* were homozygous lethal (Table S4), which is consistent with characterized mutations in these genes (FlyBase). Similarly, complementation tests revealed that insertions in *wg*, *Mhc*, *hh*, and *esg* were lethal *in trans* with loss-of-function alleles or genomic deletions spanning the target gene (Table S4). In addition, homozygous insertions in *hh* and *wg* caused a "lawn of denticles" phenotype in embryos (Bejsovec and Martinez Arias 1991) (Figure 3F) and flies with heterozygous insertions in *Mhc* were flightless (Mogami *et al.* 1986) (data not shown), all reminiscent of previously characterized mutants. Four insertions in *ebony* produced flies with dark cuticle pigment when homozygous (Figure 3E and Table S4) or *in trans* with $ebony^1$ (data not shown). In one case (*ebony pFP545 #2)*, the insertion was homozygous lethal but was viable over $ebony^1$, and *trans*-het flies exhibited dark cuticle pigment. An insertion in *myo1a* (also known as *Myo31DF*) was viable and flies exhibited genital rotation phenotypes observed in previously reported loss-of-function alleles (Spéder *et al.* 2006) (Figure 3G). An insertion in *FK506-bp2* was homozygous viable, though this gene has not been well characterized in *Drosophila*. In all, these results indicate that insertion of *pCRISPaint-T2A-Gal4-3xP3-RFP* into 5′ coding sequence can disrupt gene function.

An important consideration when using genome editing is off-target effects. For example, in addition to on-target insertions, our RFP+ fly lines could contain off-target insertions or other mutations that disrupt nontarget genes. We noticed that RFP fluorescence always cosegregated with the target chromosome when balancing insertions (data not shown), suggesting that multiple insertions may be uncommon. Furthermore, we did not detect evidence of off-target insertions by splinkerette PCR (Potter and Luo 2010), which largely reconfirmed the on-target sites (Table S4). However, this analysis was problematic because some sequence traces aligned to the donor plasmid or to genomic sequence on the opposite expected side of the target site, suggesting that insertions could be concatemers. Others have observed concatemers when performing NHEJ knock-ins (Auer *et al.* 2014; Kumagai *et al.* 2017), and we found that seven of our RFP+ insertions were head-to-tail concatemers as determined by PCR (Figure S4). We could rescue the lethality of our *esg* insertion using a duplication chromosome (stock # BL6230), suggesting that there were no other second-site lethal mutations on this chromosome. Finally, one of five *ebony* insertions (*ebony pFP545 #2*) was homozygous lethal, suggesting it contains a second-site lethal mutation.

We did not obtain insertions when targeting *ap*, *αTub84B*, *btl*, or *Desat1*. Therefore, we investigated whether the sgRNAs targeting these genes were functional. Four sgRNAs used for germ line knock-ins had an acceptable efficiency score of > 5, with the exception being the sgRNA targeting *btl* (Table S5). Performing a T7 endonuclease assay from

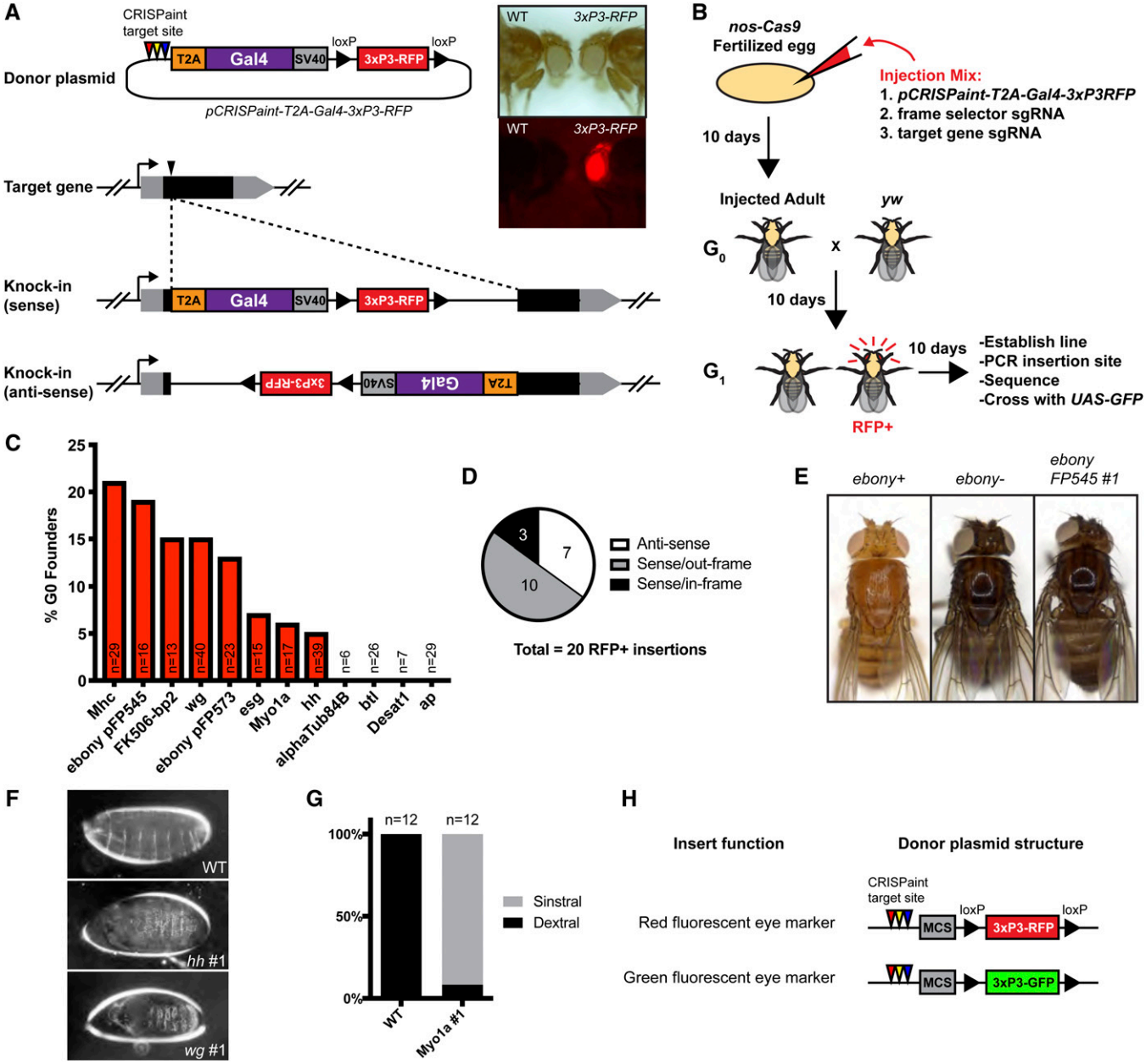

Figure 3 Germ line knock-ins using homology-independent insertion. (A) Schematic of knock-in approach. *pCRISPaint-T2A-Gal4-3xP3-RFP* is inserted into 5′ coding sequence. Inset shows images of adult flies with 3xP3-RFP fluorescence in the eye. Top panel is brightfield, bottom panel is fluorescence. (B) 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 WT (*yw*) flies 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. CRISPaint, clustered regularly interspaced short palindromic repeat-assisted insertion tagging; RFP, red fluorescent protein; sgRNA, single guide RNA; UAS, upstream activating sequence; WT, wild-type; T2A, Thosea asigna virus 2A self-cleaving peptide; MCS, Multiple Cloning Site; sfGFP, super folder GFP.

transfected cells revealed that sgRNAs targeting *ap*, *αTub84B*, and *btl* can cut at the target site, whereas the results with *Desat1* were inconclusive (Figure S5A). As an alternative functional test, we used PCR to detect knock-in events in S2R+ cells transfected with the *pCRISPaint-T2A-Gal4-3xP3-RFP* donor plasmid. This showed that sgRNAs targeting *ap*, *αTub84B*, *btl*, and *Desat1* can lead to successful knock-in of *pCRISPaint-T2A-Gal4-3xP3-RFP* and suggests that the sgRNAs are functional (Figure S5B). Finally, we sequenced the sgRNA target sites in the *nos-Cas9* fly strains and found a SNP in the *btl* sgRNA-binding site (data not shown), whereas all 10 remaining sgRNAs had no SNPs in the target site. In summary, we conclude that the sgRNAs targeting *ap*, *αTub84B*, *btl*, and *Desat1* are able to induce

cleavage at their target site in S2R+ cells, but that the sgRNA targeting *btl* may not function in the germ line using our *nos-Cas9* strains.

Collectively, our results demonstrate that the *pCRISPaint-T2A-Gal4-3xP3-RFP* universal donor plasmid inserts into target loci in the fly germ line and can be used to generate loss-of-function fly lines. Therefore, we generated two simplified universal donor plasmids that were more tailored to this purpose: *pCRISPaint-3xP3-RFP* and *pCRISPaint-3xP3-GFP* (Figure 3H). These donor plasmids contain green or red fluorescent markers, and a useful multiple cloning site for adding additional insert sequence.

Knock-ins that express Gal4 under the control of the target gene

We next determined if any of our 20 RFP+ lines expressed Gal4 from the target gene. Since the *T2A-Gal4* reporter gene in *pCRISPaint-T2A-Gal4-3xP3-RFP* is promoterless, we reasoned that only insertions in the sense orientation and in-frame with the target gene should express Gal4. We took an unbiased approach by crossing all 20 RFP+ lines to *UAS-GFP* and screening progeny for GFP fluorescence throughout development. From this effort, we identified six Gal4-expressing lines (Figure 4A, and Table S4). *wg-T2A-Gal4* (#1 and 4), *Mhc-T2A-Gal4* (#1 and 2), and *Myo1a-T2A-Gal4* #1 insertions were expressed in the imaginal disc, larval muscle, and larval gut (Figure 4B and Table S4), respectively, which matched the known expression patterns for these genes. *wg-T2A-Gal4 #1* and *#4* insertions were expressed in a distinctive Wg pattern in the wing disc pouch (Figure 4C). *ebony-T2A-Gal4 pFP545 #2* was expressed in the larval brain and throughout the pupal body (Figure 4D), consistent with a previous study (Hovemann *et al.* 1998). Interestingly, it was also expressed in the larval trachea, with particularly strong expression in the anterior and posterior spiracles. Indeed, classical *ebony* mutations are known to cause dark pigment in larval spiracles (Brehme 1941).

Next, we compared the Gal4 expression results with our sequence analysis of the insertion sites. As expected, all lines that expressed Gal4 had insertions that were in the sense orientation (Figures S2 and S3, and Table S4). For example, *ebony-T2A-Gal4 pFP545 #2* contains a 15-bp genomic deletion that is predicted to keep *T2A-Gal4* in-frame with *ebony* coding sequence. Similarly, *wg-T2A-Gal4* #1 contains an in-frame 45-bp deletion and 21-bp insertion. Remarkably, *wg-T2A-Gal4* #4 contains a frameshift indel (Figure S3), yet still expresses Gal4 in the Wg pattern, albeit at significantly lower levels than *wg-T2A-Gal4* #1 (Figure 4C). In addition, *Mhc-T2A-Gal4* lines #1 and #2, and *Myo1a-T2A-Gal4* #1 each have indels that put *T2A-Gal4* out-of-frame with the target-gene coding sequence. These findings confirm that our Gal4-expressing lines have *T2A-Gal4* inserted in the sense orientation, but that in-frame insertion with the target-gene coding sequence is not necessarily a requirement for Gal4 expression.

Artifacts due to homology-independent insertion, such as indels at the insertion site, could conceivably interfere with Gal4 expression. For example, *ebony* expression has not been well studied, so we did not know if *ebony-T2A-Gal4 pFP545 #2* ("*ebony-Gal4* CRISPaint") was an accurate reporter. To address this issue, we generated a precise in-frame HDR insertion in *ebony* ("*ebony-Gal4* HDR") using the same insert sequence (*T2A-Gal4-3xP3-RFP*) and target-gene sgRNA (*pFP545*). *ebony-Gal4* HDR knock-in fly lines were validated by PCR (Figure S6), and *ebony-Gal4* HDR homozygous flies were viable and had dark cuticle pigment (data not shown), suggesting on-target knock-in. By crossing CRISPaint and HDR knock-in alleles with *UAS-GFP*, we found that their expression patterns were similar at larval, pupal, and adult stages (Figure 4, D and E). However, *ebony-Gal4* CRISPaint had higher levels of RFP fluorescence, and expressed Gal4 in the larval anal pad and gut, which is coincident with expression of *3xP3-RFP*. Since this insertion is a concatemer (Figure S4 and Table S4), we speculated that multiple copies of *3xP3-RFP* lead to increased RFP fluorescence and ectopic expression of Gal4. Unfortunately, we could not test this because our efforts to remove *3xP3-RFP* using Cre/*loxP* excision were unsuccessful; progeny expressing Cre and *ebony-Gal4* CRISPaint were apparently lethal, whereas we easily generated RFP-free derivatives of *ebony-Gal4* HDR using the same crossing scheme. In contrast, *Mhc-T2A-Gal4* lines #1 and #2, and *Myo1a-T2A-Gal4* #1 do not express in the anal pad, and gut and anal pad expression in *wg-T2A-Gal4* #1 likely corresponds to normal *wg* expression in these tissues (Takashima and Murakami 2001) (Figure 4B). In conclusion, these results suggest that Gal4-expressing CRISPaint insertions can capture the endogenous expression pattern of the target gene, albeit with caveats that we demonstrated (also see *Discussion*).

To facilitate the insertion of other sequences into the germ line to generate expression reporters, we created 10 additional universal donor plasmids (Figure 4F). These include T2A-containing donors with sequence encoding the alternative binary reporters LexGAD, QF2, and split-Gal4, as well as Cas9 nuclease, FLP recombinase, Gal80 repression protein, NanoLuc luminescence reporter, and superfolder GFP. In addition, we generated *pCRISPaint-sfGFP-3xP3-RFP*, which can be used to insert into 3′ coding sequence, generating a C-terminal GFP fusion protein. Finally, we created a single plasmid expressing all three frame-selector sgRNAs (*pCFD5-frame-selectors_0,1,2*). Since indels at the insertion site in the germ line effectively randomize the coding frame, we reasoned that simultaneous expression of all three sgRNAs would maximize cutting and linearization of a CRISPaint universal donor plasmid.

Discussion

Inserting large DNA elements into the genome by HDR requires a great deal of expertise and labor for the design, and construction, of donor plasmids. Some groups have developed strategies to improve the efficiency and scale at which homology arms are cloned into donor plasmids (Housden *et al.* 2014; Gratz *et al.* 2015; Kanca *et al.* 2019), but the root problem still remains. In this study, we

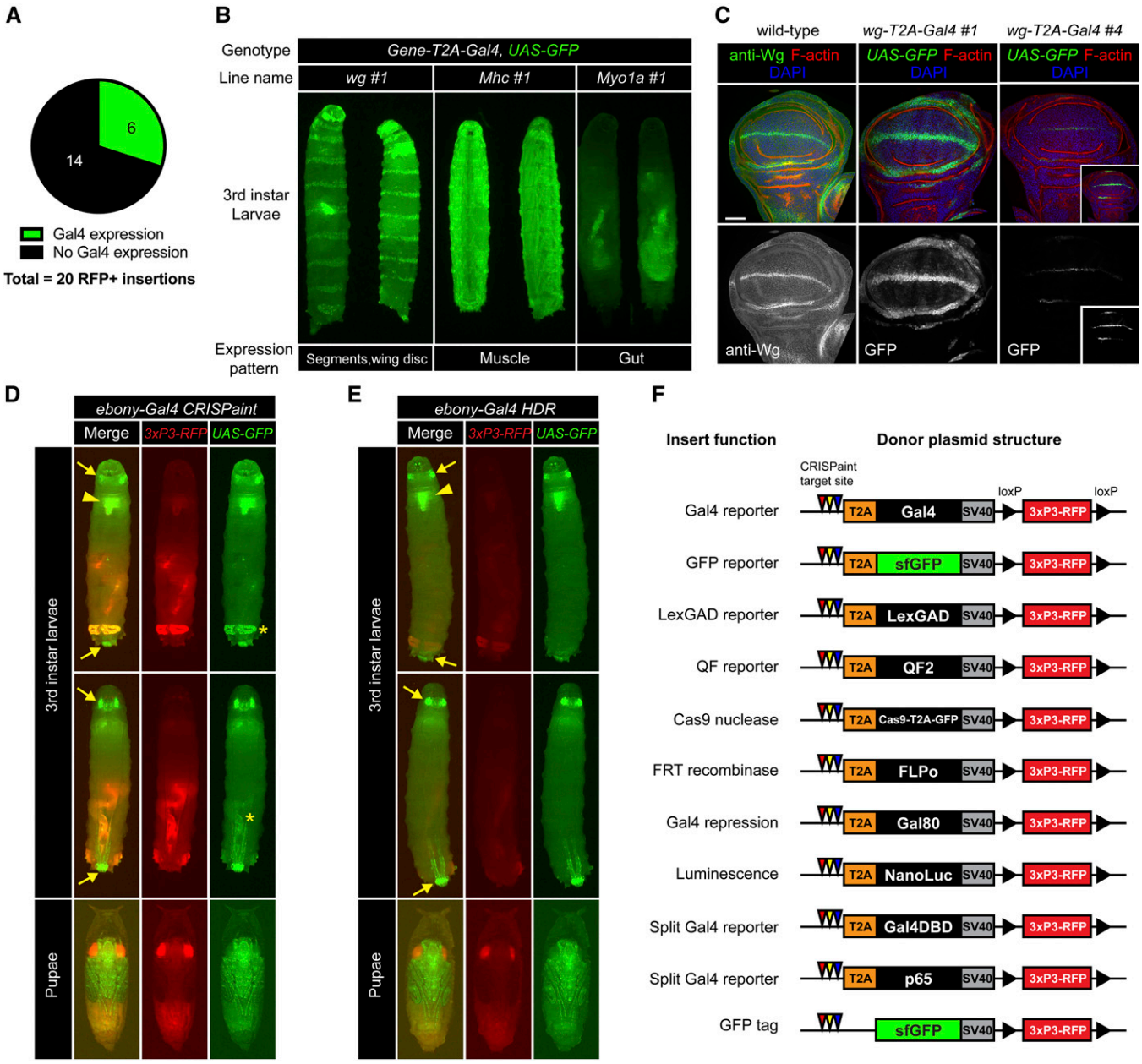

Figure 4 Germ line insertions that express *T2A-Gal4* under the control of the target gene. (A) Proportion of 20 RFP+ knock-in lines that express Gal4. (B) Fluorescence images of third-instar larvae with indicated genotypes. Two larvae of the same genotype are shown as rotated at different angles. Expression 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. Bar, 50 μm. (D and E) Fluorescence images of third-instar larvae and pupae. Fluorescence is from *3xP3-RFP* (red) and Gal4 expression with *UAS-GFP* (green). (D) Results with *ebony-T2A-Gal4 pFP545* #2 (*ebony-Gal4* CRISPaint) and (E) shows results with *ebony-Gal4* HDR. (F) Additional universal donor plasmids. CRISPaint, clustered regularly interspaced short palindromic repeat-assisted insertion tagging; HDR, homology-directed repair; RFP, red fluorescent protein; UAS, upstream activating sequence; T2A, Thosea asigna virus 2A self-cleaving peptide.

showed that homology-independent insertion can be used in *Drosophila* as an alternative to HDR for certain gene-targeting strategies in cultured cells and *in vivo*. With no donor design and construction, knock-in lines are simple and fast to obtain (see Table 1).

To perform homology-independent insertion in *Drosophila*, we adapted the mammalian CRISPaint system (Schmid-Burgk *et al.* 2016) because of its user-friendly design. CRISPaint donor plasmids are universal because they lack homology arms, and publicly available collections enable researchers to "mix and match" insert sequences and target genes. Indeed, we showed that *pCRISPaint-mNeonGreen-T2A-PuroR*, originally used in human cells, functions in *Drosophila* S2R+ cells (Figure 1). To linearize donor plasmids in *Drosophila* in three

Table 1 Timeline comparison for CRISPaint *vs.* HDR in S2R+ cells and *in vivo*

	Donor design and construction	sgRNA cloning	Timescale to generate single-cell cloned cell lines or transgenic fly stocks	Molecular characterization	Screening expression
S2R+ gene tagging (*e.g.*, mNeonGreen)					
HDR knock-in	Variable, 1–3 wk	3 D	2 mo	1–2 D (PCR, gel image)	1 D, confocal microscopy
CRISPaint knock-in	NA	3 D	2 mo	3 D (PCR, gel image, sequence)	1 D, confocal microscopy
In vivo KO					
HDR knock-in	Variable, 1–3 wk	3 D	1 mo	1–2 D (PCR, gel image)	—
In/del frameshift	NA	3 D	1 mo	3–7 D (PCR, gel image, sequence)	—
CRISPaint knock-in	NA	3 D	1 mo	3 D (PCR, gel image, sequence)	
In vivo gene tagging (*e.g.*, T2A-Gal4)					
HDR knock-in	Variable, 1–3 wk	3 D	1 mo	1–2 D (PCR, gel image)	Optional though recommended: cross with UAS-GFP, 5–14 D
CRISPaint knock-in	NA	3 D	1 mo	3 D (PCR, gel image, sequence)	Cross with UAS-GFP, 5–14 D

CRISPaint, clustered regularly interspaced short palindromic repeat-assisted insertion tagging; HDR, homology-directed repair; In/del, insertion/deletion; KO, knockout; UAS, upstream activating sequence; T2A, Thosea asigna virus 2A self-cleaving peptide.

coding frames, we generated plasmids expressing the three CRISPaint frame-selector sgRNAs under the *Drosophila U6* promoter. Similarly, we also generated a plasmid to express all three sgRNAs simultaneously for maximum cutting efficiency. Therefore, to perform a knock-in experiment, only an sgRNA that targets a gene of interest is needed, which is simple to produce by molecular cloning or ordering. In fact, all target-gene sgRNA plasmids used in this study were obtained from the TRiP facility where an ever-growing number of sgRNAs are being generated.

Unlike HDR, which can seamlessly insert any DNA into a target genomic location, homology-independent insertion is restricted to specific targeting scenarios. For example, since the entire *pCRISPaint* donor plasmid integrates into a genomic target site, it cannot be used to replace genes, make amino acid changes, or tag proteins at their N-termini. In addition, potential concatemerization of the insertion (Auer *et al.* 2014; Kumagai *et al.* 2017) prevents its use for strategies where a single copy of the insertion is required, such as artificial exon traps that exploit target gene splicing (Buszczak *et al.* 2007). Despite these limitations, homology-independent insertion was previously shown to be useful for generating loss-of-function alleles by premature truncation (Cristea *et al.* 2013; Auer *et al.* 2014; Katic *et al.* 2015; Katoh *et al.* 2017; Zhang *et al.* 2018) as well as C-terminal gene tagging (Lackner *et al.* 2015; Schmid-Burgk *et al.* 2016; Gao *et al.* 2019).

Using CRISPaint in the fly germ line, we rapidly generated loss-of-function alleles by inserting the selectable marker *3xP3-RFP* into target-gene coding sequence. This method has several advantages over isolating frameshift alleles using CRISPR-induced indels (Bier *et al.* 2018). First, identifying founder flies with RFP fluorescence is easier and faster than PCR genotyping/sequencing for frameshift indels. Second, insertion into 5′ coding sequence is more likely to disrupt the target gene, whereas frameshift alleles can sometimes retain protein activity (Tuladhar *et al.* 2019); however, insertions could also have unintended consequences, such as disrupting neighboring gene expression. Third, genetic crosses are simplified by tracking insertion alleles with a fluorescent marker, which could be especially useful in species other than *D. melanogaster* that do not have balancer chromosomes. Indeed, the combined use of *3xP3-RFP* and *3xP3-GFP* universal donor plasmids (Figure 3H) could facilitate the generation of double-mutant lines. While we did not find evidence for off-target insertions, knock-in lines should be vetted using traditional techniques, including molecular verification of the insertion site, the comparing independent insertions, outcrossing to wild-type, and performing complementation tests.

Using homology-independent insertion for gene tagging requires screening for properly expressed inserts, due to the unpredictable and error-prone nature of NHEJ. For example, donor plasmids can integrate in two orientations and indels at the insertion site can shift the coding frame. In cultured S2R+ cells, we used antibiotic resistance to efficiently select for mNeonGreen protein fusions (Figure 1D). However, such an assay is not feasible in flies, and so we screened germ line *T2A-Gal4* insertions for expression by crossing with *UAS-GFP* (Figure 3B and Figure 4). This yielded Gal4-expressing lines for *ebony*, *wg*, *Mhc*, and *Myo1a*. In contrast, we obtained no insertions for *αTub84B*, *btl*, *Desat1*, *ap*, and one non-Gal4-expressing insertion each for *FK506-bp2*, *esg*, and *hh*. This highlights the importance of obtaining multiple independently derived germ line insertions to screen for insert expression. Additional steps could be taken to obtain more

insertions per gene, such as increasing the number of injected embryos or reattempting with a different sgRNA. Since our knock-in efficiency is roughly similar to HDR [5–21% (Figure 3C) *vs.* 5–22% (Gratz *et al.* 2014), 46–88% (Port *et al.* 2015), and 7–42% (Gratz *et al.* 2015)], the limited number of independent germ line insertions may simply be a constraint of embryo injection-based transgenesis.

Sequence analysis of insertion sites in cultured cells and *in vivo* revealed expected and unexpected results. Single-cell cloned S2R+ lines expressing mNeonGreen fusion proteins had insertion sites that were seamless (seven out of nine) or caused an in-frame 3-bp deletion (two out of nine), which was consistent with human cell data (Schmid-Burgk *et al.* 2016). However, all sequenced germ line insertions (20) contained indels at the genomic insertion site, and indels were larger than in S2R+ cell experiments (Figure S3). Germ cells are known to differ in their NHEJ mechanisms compared to somatic cells (Preston *et al.* 2006; Ahmed *et al.* 2015), but it is not clear why this would prevent seamless insertions. In addition, four out of six Gal4-expressing insertions were out-of-frame relative to the target gene (Figure S3). We speculate that this could result from ribosome frameshifting (Ketteler 2012), an internal ribosome entry site (Komar and Hatzoglou 2005), alternative ORFs (Mouilleron *et al.* 2016), or Gal4 translation initiation from its own start codon. Ultimately, we consider this a fortuitous effect as long as Gal4 is expressed in the correct pattern. Conversely, *wg* #6 is in-frame but does not express Gal4, perhaps due to mutation of Gal4 sequences. Indeed, splinkerette sequencing results from 3/20 of our RFP+ insertions suggested that portions of the donor plasmid may be deleted during insertion (Table S4). These findings again highlight the importance of screening insertions for insert expression.

For gene-tagged cell lines and fly lines generated using homology-independent insertion, potential artifacts could impact faithful protein fusion or reporter expression. First, indels at the insertion site could introduce cryptic splice sites, delete an important sequence region, or cause nonoptimal codons that impact translation of the fusion protein. Second, bacterial sequences in the inserted plasmid may cause transgene silencing or impact neighboring gene expression (Chen *et al.* 2004; Suzuki *et al.* 2016); however, we note that thousands of transgenic fly lines contain bacterial sequences from φC31 integration (Perkins *et al.* 2015) with no reports of ill effects. Third, concatemer insertions could affect gene expression in unpredictable ways. Indeed, insertion *ebony-T2A-Gal4 pFP545* #2 was a head-to-tail concatemer (Figure S4) and had ectopic Gal4 expression from the *3xP3* enhancer (Figure 4, D and E). Finally, as noted above, the insertion frame may be important for expression, since the *wg-T2A-Gal4* #4 line was expressed weakly and in a nonfaithful patchy pattern in the wing disc. While these are important considerations, we note that 9/9 of our mNeonGreen-tagged S2R+ lines were localized properly, and that 5/6 Gal4-expressing lines appeared to be faithful reporters of their target genes, with the above exceptions noted.

Some of our S2R+ mNeonGreen protein fusion lines exhibited unusual or unexpected protein localization. In clones D6 and D9, Lamin-mNeonGreen localized to the nuclear envelope, but in D9 this localization was enriched asymmetrically in the direction of the previous plane of cell division. Since both clones had seamless mNeonGreen insertions, we believe the localization difference is caused by the lack of a nonknock-in locus in D9, whereas D6 contained an in-frame 3-bp deletion at the nonknock-in locus, likely retaining wild-type function. We saw a similar pattern for clones A3 and A5, both of which had seamless insertions in *Act5C*, but clone A5 exhibited distinct rod structures. αTub84B-mNeonGreen fluorescence and protein levels were low in all cell lines, despite *αTub84B* being highly expressed in S2R+ cells (Hu *et al.* 2017). We speculate that the αTub84B-mNeonGreen fusion protein is unstable and previous studies in other organisms have highlighted problems with C-terminal tagging of α-Tubulin (Carminati and Stearns 1997). Similarly, C-terminal tags can disrupt Lamin and Actin function (Davies *et al.* 2009; Nagasaki *et al.* 2017). Therefore, it is important to consider the existing knowledge of the protein when generating C-terminal protein fusions, and to examine multiple single-cell cloned lines. Importantly, this level of scrutiny is also required when using HDR to produce knock-ins. Further validation steps could be done, such as comparing tagged protein localization to antibody staining of the untagged wild-type protein.

Additional applications of homology-independent insertion in the *Drosophila* germ line can be imagined. For example, new insert sequences could be added to *pCRISPaint-3xP3-RFP* or *pCRISPaint-3xP3-GFP* starting plasmids. In addition, a modified donor plasmid could be constructed that does not integrate the bacterial plasmid backbone, such as providing donor plasmids as mini-circles (Schmid-Burgk *et al.* 2016; Suzuki *et al.* 2016), cutting the donor plasmid twice to liberate the insert fragment (Lackner *et al.* 2015; Suzuki *et al.* 2016; Gao *et al.* 2019), or using PCR-amplified inserts (Manna *et al.* 2019). Protein tagging could be performed *in vivo*, but the large indels observed in the germ line may limit its use to proteins that have C-terminal tails that can be deleted without altering protein function. Regardless, we generated a donor plasmid for *in vivo* sfGFP C-terminal fusions (Figure 4F). Finally, we note that most of our germ line donor plasmids contain enzyme restriction sites that can be used to insert genomic homology arms by molecular cloning, making them "dual-use" reagents for HDR and homology-independent insertion. Indeed, we used this approach to generate the *ebony-Gal4* HDR allele (see *Materials and Methods*).

In summary, we demonstrated that homology-independent insertion can be used as an alternative to HDR in *Drosophila*, enabling researchers to rapidly obtain knock-ins without donor design and construction. The most practical application of this approach is to perform C-terminal protein tagging in cultured cells and gene knockout by insertion *in vivo*. While we obtained *in vivo T2A-Gal4* gene-tagged insertions, it is low efficiency and thus less appealing compared

to HDR. However, the techniques required for screening insertions for *T2A-Gal4* expression are less specialized than those for constructing donor plasmids, making this trade off potentially attractive for those with less molecular biology expertise or who previously failed using HDR. Finally, our *in vivo* donor plasmids are immediately useable in other arthropod species because of the *3xP3*-fluorescent marker (Berghammer *et al.* 1999), a testament to the modularity of this knock-in system, and the benefit of a community of researchers creating and sharing universal donor plasmids.

Acknowledgments

We thank Jonathan Schmid-Burgk for advice and the *CRISPaint-mNeonGreen* donor plasmid; Claire Hu and the Transgenic RNAi Project for sgRNA design and construction; Ben Ewen-Campen for valuable comments on the manuscript; Stephanie Mohr, Oguz Kanca, and Hugo Bellen for helpful discussions; Raghuvir Viswanatha for the *Cas9-T2A-EGFP* template sequence; Rich Binari for help with mounting embryos on slides; Cathryn Murphy for general assistance; and the Harvard Medical School (HMS) Microscopy Resources on the North Quad Core and the HMS Department of Immunology's Flow Cytometry Facility for technical assistance. J.A.B. was supported by the Damon Runyon Foundation. This work was supported by National Institutes of Health grants R01 GM-084947, R01 GM-067761, and R24 OD-019847. N.P. is an investigator of the Howard Hughes Medical Institute.

Literature Cited

Ahmed, E. A., H. Scherthan, and D. G. de Rooij, 2015 DNA double strand break response and limited repair capacity in mouse elongated spermatids. Int. J. Mol. Sci. 16: 29923–29935. https://doi.org/10.3390/ijms161226214

Auer, T. O., and F. Del Bene, 2014 CRISPR/Cas9 and TALEN-mediated knock-in approaches in zebrafish. Methods 69: 142–150. https://doi.org/10.1016/j.ymeth.2014.03.027

Auer, T. O., K. Duroure, A. De Cian, J. P. Concordet, and F. Del Bene, 2014 Highly efficient CRISPR/Cas9-mediated knock-in in zebrafish by homology-independent DNA repair. Genome Res. 24: 142–153. https://doi.org/10.1101/gr.161638.113

Bassett, A. R., C. Tibbit, C. P. Ponting, and J. L. Liu, 2014 Mutagenesis and homologous recombination in Drosophila cell lines using CRISPR/Cas9. Biol. Open 3: 42–49. https://doi.org/10.1242/bio.20137120

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. https://doi.org/10.1534/genetics.111.126995

Berghammer, A. J., M. Klingler, and E. A. Wimmer, 1999 A universal marker for transgenic insects. Nature 402: 370–371. https://doi.org/10.1038/46463

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. https://doi.org/10.1534/genetics.117.1113

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. https://doi.org/10.1242/dev.114603

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. https://doi.org/10.1093/nar/gku289

Brehme, K. S., 1941 The effect of adult body color mutations upon the larva of Drosophila melanogaster. Proc. Natl. Acad. Sci. USA 27: 254–261. https://doi.org/10.1073/pnas.27.6.254

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. https://doi.org/10.1534/genetics.106.065961

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. https://doi.org/10.1083/jcb.138.3.629

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. https://doi.org/10.1038/sj.gt.3302231

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. https://doi.org/10.1038/s41587-019-0032-3

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. https://doi.org/10.1002/bit.24733

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. https://doi.org/10.1146/annurev-genom-082908-150150

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. https://doi.org/10.1534/genetics.111.136291

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 103: 583–597.e8. https://doi.org/10.1016/j.neuron.2019.05.047

Gratz, S. J., A. M. Cummings, J. N. Nguyen, D. C. Hamm, L. K. Donohue *et al.*, 2013 Genome engineering of Drosophila with the CRISPR RNA-guided Cas9 nuclease. Genetics 194: 1029–1035. https://doi.org/10.1534/genetics.113.152710

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. https://doi.org/10.1534/genetics.113.160713

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: 31.2.1–31.2.20. https://doi.org/10.1002/0471142727.mb3102s111

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: e46181. https://doi.org/10.7554/eLife.46181

Housden, B. E., S. Lin, and N. Perrimon, 2014 Cas9-based genome editing in Drosophila. Methods Enzymol. 546: 415–439. https://doi.org/10.1016/B978-0-12-801185-0.00019-2

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. https://doi.org/10.1016/S0378-1119(98)00440-5

Hu, Y., A. Comjean, N. Perrimon, and S. E. Mohr, 2017 The Drosophila gene expression tool (DGET) for expression analyses. BMC Bioinformatics 18: 98. https://doi.org/10.1186/s12859-017-1509-z

Kanca, O., J. Zirin, J. Garcia-Marques, S. M. Knight, D. Yang-Zhou *et al.*, 2019 An efficient CRISPR-based strategy to insert small and large fragments of DNA using short homology arms. Elife 8: e51539. https://doi.org/10.7554/eLife.51539

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. https://doi.org/10.1534/g3.115.019273

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. https://doi.org/10.1091/mbc.e17-01-0051

Ketteler, R., 2012 On programmed ribosomal frameshifting: the alternative proteomes. Front. Genet. 3: 242. https://doi.org/10.3389/fgene.2012.00242

Komar, A. A., and M. Hatzoglou, 2005 Internal ribosome entry sites in cellular mRNAs: mystery of their existence. J. Biol. Chem. 280: 23425–23428. https://doi.org/10.1074/jbc.R400041200

Kondo, S., and R. Ueda, 2013 Highly improved gene targeting by germline-specific Cas9 expression in Drosophila. Genetics 195: 715–721. https://doi.org/10.1534/genetics.113.156737

Korona, D., S. A. Koestler, and S. Russell, 2017 Engineering the Drosophila genome for developmental biology. J. Dev. Biol. 5: E16. https://doi.org/10.3390/jdb5040016

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. https://doi.org/10.1371/journal.pone.0186112

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. https://doi.org/10.1038/ncomms10237

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 [corrigenda: Genome Biol. 20: 53 (2019)]. https://doi.org/10.1186/gb-2014-15-8-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: e35574. https://doi.org/10.7554/eLife.35574

Manna, P. T., L. J. Davis and M. S. Robinson, 2019 Fast and cloning-free CRISPR/Cas9-mediated genomic editing in mammalian cells. Traffic. https://doi.org/10.1111/tra.12696

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. https://doi.org/10.1101/gr.145441.112

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. USA 83: 1393–1397. https://doi.org/10.1073/pnas.83.5.1393

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. https://doi.org/10.1093/nar/gkv1218

Nagasaki, A., S. T. Kijima, 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. https://doi.org/10.1247/csf.17016

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. https://doi.org/10.1534/genetics.115.180208

Port, F., and S. L. Bullock, 2016 Augmenting CRISPR applications in Drosophila with tRNA-flanked sgRNAs. Nat. Methods 13: 852–854. https://doi.org/10.1038/nmeth.3972

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. USA 111: E2967–E2976. https://doi.org/10.1073/pnas.1405500111

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. https://doi.org/10.1534/g3.115.019083

Potter, C. J., and L. Luo, 2010 Splinkerette PCR for mapping transposable elements in Drosophila. PLoS One 5: e10168. https://doi.org/10.1371/journal.pone.0010168

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. https://doi.org/10.1534/genetics.105.050138

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. https://doi.org/10.1186/s13059-017-1220-4

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. USA 110: 19012–19017. https://doi.org/10.1073/pnas.1318481110

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. https://doi.org/10.1038/ncomms12338

Singh, P., J. C. Schimenti, and E. Bolcun-Filas, 2015 A mouse geneticist's practical guide to CRISPR applications. Genetics 199: 1–15. https://doi.org/10.1534/genetics.114.169771

Spéder, P., G. Adám, and S. Noselli, 2006 Type ID unconventional myosin controls left-right asymmetry in Drosophila. Nature 440: 803–807. https://doi.org/10.1038/nature04623

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. https://doi.org/10.1038/nature20565

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. https://doi.org/10.1016/S0925-4773(00)00555-4

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. https://doi.org/10.1038/s41467-019-12028-5

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. https://doi.org/10.1002/wdev.214

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: e36333. https://doi.org/10.7554/eLife.36333

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. https://doi.org/10.1534/genetics.113.153825

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. https://doi.org/10.1186/s12915-018-0616-2

Communicating editor: K. O'Connor-Giles