An expanded toolkit for gene tagging based on MiMIC and scarless CRISPR tagging in *Drosophila*.

David Li-Kroeger[1,*], Oguz Kanca[1,*], Pei-Tseng Lee[1], Sierra Cowan[2], Michael Lee[2], Manish Jaiswal[1,3,$], Jose Luis Salazar[1,4], Yuchun He[1,3], and Hugo J. Bellen [1,3,4, 5,6,&]

1 Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA.
2 Department of Biochemistry and Cell Biology, Rice University Houston, TX 77030, USA
3 Howard Hughes Medical Institute, Baylor College of Medicine, Houston, TX 77030, USA
4 Program in Developmental Biology, Baylor College of Medicine, Houston, TX 77030, USA
5 Department of Neuroscience, Baylor College of Medicine, Houston, TX 77030, USA
6 Jan and Dan Duncan Neurological Research Institute, Texas Children's Hospital (TCH), Houston, TX 77030, USA

Key words: Double Header, gene editing, T2A-GAL4, RMCE, protein tagging, Trojan exons, swappable insertion cassette (SIC), NMNAT

[*]contributed equally to this work
[$] current affiliation: TIFR Centre for Interdisciplinary Sciences, Hyderabad, India
[&] corresponding author; hbellen@bcm.edu

Abstract

We generated new genetic tools to efficiently tag genes in *Drosophila*. Double Header (DH) utilizes intronic MiMIC/CRIMIC insertions to generate artificial exons for GFP mediated protein trapping or T2A-GAL4 gene trapping *in vivo* based on CRE recombinase to avoid embryo injections. DH significantly increases integration efficiency compared to previous strategies and faithfully reports the expression pattern of genes and proteins. The second technique targets genes lacking coding introns using a two-step cassette exchange. First, we replace the endogenous gene with an excisable compact dominant marker using CRISPR making a null allele. Second, the insertion is replaced with a protein::tag cassette.

This sequential manipulation allows the generation of numerous tagged alleles or insertion of other DNA fragments that facilitates multiple downstream applications. Both techniques allow precise gene manipulation and facilitate detection of gene expression, protein localization and assessment of protein function, as well as numerous other applications.

Introduction

Comprehensive gene annotation is a central challenge in the post-genomic era. *Drosophila melanogaster* offers more sophisticated genetic approaches and tools to assess gene function and expression than other multicellular model organisms (Bier et al., 2018; Cox et al., 2017; Germani et al.,2018; Heigwer et al., 2018; Kanca et al., 2017; Kaufman 2017; Simpson & Looger, 2018). A versatile tool used for functional gene annotation in *Drosophila* is MiMIC, a *minos*-based transposon that integrates a Swappable Integration Cassette (SIC) in the genome (Venken et al., 2011a). MiMIC SICs contain a cassette nested between two *attP* sites that can be exchanged with any DNA sequence flanked with *attB* sites through Recombinase Mediated Cassette Exchange (RMCE) by ΦC31 integrase. When a MiMIC is integrated in an intron of a gene flanked on both sides by coding exons (hereafter referred to as a coding intron), the SIC can easily be exchanged with an artificial exon that encodes *Splice Acceptor (SA)-$(GGS)_4$ linker-EGFP-FlAsH tag-StrepII tag-TEV protease cleavage site-3XFlag-$(GGS)_4$ linker-Splice Donor (SD)* (abbreviated as *GFP tag*) (Venken et al., 2011a; Nagarkar-Jaiswal et al., 2015a,b). The GFP tagged endogenous proteins report the subcellular localization of the gene product and are functional in ~75% of tested genes (Nagarkar-Jaiswal et al., 2015a). Importantly functional GFP protein traps can be used for multiple assays. These include chromatin Immunoprecipitation (ChIP) of transcription factors (Negre et al., 2011), Immunoprecipitation (IP)-Mass Spectroscopy (MS) (David-Morrison et al., 2016; Neumuller et al., 2012; Yoon et al., 2017), rapid conditional removal of gene products (Caussinus et al., 2011; Lee et al., 2018b; Nagarkar-Jaiswal et al., 2015a; Neumuller et al., 2012; Wissel et al., 2016) and sequestration of tagged proteins (Harmansa et al., 2017; Harmansa et al., 2015). Hence, tagging an endogenous gene with GFP enables numerous applications to dissect gene function.

The SIC in MiMICs can also be replaced by an artificial exon that encodes *SA-T2A-GAL4-polyA signal* (abbreviated as *T2A-GAL4*) (Diao et

al., 2015; Gnerer et al., 2015; Lee et al., 2018a). T2A-GAL4 creates a mutant allele by truncating the protein at the insertion site but also expresses GAL4 with the spatial-temporal dynamics of the targeted gene. Hence, T2A-GAL4 facilitates the replacement of the gene of interest with fly or human UAS-cDNAs (Bellen & Yamamoto, 2015; Senturk & Bellen, 2017; Wangler et al., 2017; Lee et al., 2018a), allowing us to assess putative disease associated variants and permitting structure function analysis of the protein of interest. Moreover, these gene-specific GAL4 stocks can be used to drive a variety of UAS constructs to further identify and probe the function of the cells expressing the gene using UAS-Fluorescent proteins or numerous other UAS constructs (Venken et al., 2011b). This is especially useful for genes that are not abundantly expressed, providing a means to amplify the signal, as GAL4 drives overexpression of the UAS transgenes (Diao et al., 2015; Lee et al., 2018a). In summary, MiMIC, applications allow the acquisition of valuable data about the function of the gene as well as the cells in which the gene is expressed.

Given the usefulness of MiMICs, the *Drosophila* Gene Disruption Project (GDP) (http://flypush.imgen.bcm.tmc.edu/pscreen/index.php) has generated and mapped 17,500 MiMIC insertion stocks (Nagarkar-Jaiswal et al., 2015a; Venken et al., 2011a). This collection includes insertions within introns for ~1860 genes, each of which can be converted to a *GFP*-tagged protein trap and/or a *T2A-GAL4* gene trap (Nagarkar-Jaiswal et al., 2015a,b; Lee et al., 2018a). However, we needed to develop a complementary strategy to generate resources for genes that do not have a MiMIC randomly inserted within a coding intron. To that end, we recently developed CRIMIC (CRISPR mediated Integration Cassette); a Cas9/CRISPR Homology Directed Repair (HDR) mediated approach that integrates a modified SIC *attP-FRT-SA-T2A-GAL4-polyA-FRT-attP* in a coding intron of choice. This approach greatly expands the number of genes that can be tagged using MiMIC-like technology from 1,860 to ~6,000 (Lee et al., 2018a) allowing about forty percent of *Drosophila* protein coding genes to be targeted with SICs.

RMCE cassettes can either be injected into embryos as part of a circular plasmid or can be circularized *in vivo* from an initial insertion locus in the genome through Cre/loxP or Flp/FRT mediated recombination (Diao et al., 2015; Nagarkar-Jaiswal et al., 2015b). Importantly, RMCE cassettes can replace a SIC in either orientation with equal probability due to inverted symmetric *attP* sequences. Therefore 50% of the insertions are inserted in the opposite orientation of transcription and will not be included in the

mature mRNA. Hence, only half of all successful exchange events will result in protein or gene trap lines.

Here, we show that by combining GFP protein traps and T2A-GAL4 gene traps in a single RMCE construct, named Double Header (DH), we significantly increased the number of productive RMCE events for MiMIC/CRIMIC containing genes to generate protein or gene trap alleles. Importantly, we expand the ability to target SICs into genes regardless of the presence of introns to allow access to virtually any gene in the fly genome based on CRISPR/Cas9 mediated HDR. This provides a means to create robust null alleles with simple screening, and to convert the SIC insertion using any DNA, creating scarless modifications to facilitate numerous downstream applications.

Results

Double Header (DH) improves the tagging rate of MiMIC containing genes

SICs in coding introns can be converted into GFP protein traps or T2A-GAL4 gene traps through RMCE. However, because RMCE of SICs in MiMICs and CRIMICs can occur in either orientation, only one out of two events produces a tag that is incorporated in the gene product. Moreover, each RMCE experiment generates only a protein trap or a gene trap, requiring two independent injections or crosses to generate both reagents. In order to reduce the effort to generate these genetic tools we engineered DH, a construct that combines the two key RMCE cassettes to replace intronic MiMICs and CRIMICs: 1) *SA-T2A-GAL4-polyA* ($DH^{T2A\text{-}GAL4}$) which generates a gene trap that expresses GAL4 in the expression domain of the gene while typically inactivating gene function, in one orientation and the 2) *SA-GFP-SD* (DH^{GFP}) which generates a GFP protein trap, in the opposite orientation. Hence, insertion of DH via RMCE in a MiMIC in either orientation should result in two valuable reagents. This compound RMCE cassette is flanked by two inverted ϕC31 *attB* sites in a vector backbone that contains other features including mini *white* as shown in Fig. 1A. The presence of *white* in the plasmid backbone allows for a counter selection against the integration of the whole plasmid when incorporated into a white$^-$ background, ensuring that only the DNA between the *attB* sites integrates.

The artificial exons that are integrated into MiMIC or CRIMIC sites need to be in frame with the preceding Splice Donor (SD) to create a functional tag. Because exon/intron boundaries can occur at any one of

three positions in a codon (phase 0, +1, +2), we generated three different DH plasmids (Sequences can be found in the Supp. info.). Each construct contains the same codon phase for the two modules. Given that the RMCE cassette is about 4.8 kB, or about 1.5 times the size of the single T2A-GAL4 cassette, we anticipated a lower integration rate. We tested the integration efficacy by injecting 30 strains carrying a coding intronic MiMIC insertion. On average, ~400 embryos were injected for each MiMIC line together with a plasmid that encodes the ΦC31 integrase. We screened for the loss of yellow+ in the progeny of the injected animals, as MiMICs carry the *yellow*+ marker (Supp. Fig1A). We did not observe the integration of *white* indicating that a single RMCE event leading to the integration of the entire plasmid is rare. For 16 out of 30 MiMICs, we were able to isolate yellow$^-$ flies that carry a DH integration (Fig 1B and Supp. Fig 1B). We determined the orientation of DH inserts through single fly PCR (Supp. Fig 2) and determined the orientation of the cassette in 47 out of 72 DH RMCE events. For 6 MiMICs we obtained both orientations and for the other 10 MiMICs we obtained one or the other orientation (Fig 1B, Supp. Fig 1B). Hence, we generated a total of 22 new reagents, increasing the overall rate of productive RMCE events by injection (Nagarkar-Jaiswal et al., 2015a). Note that the integration efficiency of DH construct by injection is lower than the smaller SICs: ~50% versus ~66% (Nagarkar-Jaiswal et al. 2015a). However, the number of productive events increases tagging efficacy somewhat as every successful event produces a useful line: 74% versus 66%.

In vivo Double Header mobilization based on transgenes

In an effort to avoid embryo injections and to increase integration efficiency of DH, we developed an *in vivo* RMCE strategy using genetic crosses similar to Trojan exons developed for *T2A-GAL4* (Diao et al., 2015). We integrated the same constructs as in Fig. 1A, one for each reading frame, in the genome through P-element mediated transformation. These insertions serve as jump starter constructs because the RMCE cassettes in these transgenes can be excised from their initial landing sites by expressing Cre recombinase in germ line cells (Fig 2). The crossing scheme is outlined in Supp. Fig. 3A. We generated jump starter insertions in second and third chromosomes for all 3 possible phases of DH and generated double balanced stocks for subsequent crosses. We tested the efficacy of integration by crosses for DH for third chromosome MiMICs. For 12 out of 13 MiMICs tested we obtained integration of DH. We determined the

orientation of DH for 48 out of 102 insertions by PCR. The inconclusive insertions either showed no PCR amplification in one end or both ends of the MiMIC (48/102) or in rare cases conflicting PCR amplification that indicates integration in both orientations (6/102). Interestingly 44/54 inconclusive inserts happened in only two of the MiMICs, indicating locus specific issues. For 6 out of 12 MiMICs we obtained both orientations and for 6 we obtained one or the other, resulting in 18 tagged genes (Supp. Fig. 3B). Hence, the genetic strategy is about twice as efficient (18 constructs for 13 crossed MiMICs versus 22 constructs for 30 injected MiMICs) as the injection strategy in generating RMCE events and requires significantly less effort.

Double Header reports the expression pattern of the tagged gene and protein

We proceeded to test if DH functions as expected. We determined the expression patterns of genes tagged in both orientations in third instar larval brain and adult brains for *MI01487* (*kibra*), *MI05208* [*5-HT2B(5-hydroxytryptamine receptor 2B)*], *MI06794* [*Lgr4(Leucine-rich repeat-containing G protein-coupled receptor 4)*], *MI06872* (*CG34383*), *MI08614 [Dgk (Diacyl glycerol kinase)*], *MI11741* (*CG12206*) and *MI15073* (*CG9132*) (Fig. 3). As we selected a few MiMICS that were previously tagged with *T2A-GAL4* by the Gene Disruption Project as positive controls (Lee et al., 2018a; Diao et al, 2015) (*MI01487* (*kibra*), *MI06794* (*Lgr4*), *MI06872* (*CG34383*), *MI08614* (*Dgk*), and *MI11741* (*CG12206*), we were able to compare expression patterns obtained by $DH^{T2A\text{-}GAL4}$ to expression patterns obtained by single *T2A-GAL4*. (http://flypush.imgen.bcm.tmc.edu/pscreen/rmce/). In all cases the expression pattern was very similar to what was previously reported (Fig. 3, (Lee et al., 2018a). In addition, tracheal expression of *CG12206* is consistent with a previous report (Chandran et al., 2014) and the $5HT2B^{T2A\text{-}GAL4}$ expression pattern in the adult brain matches an independently generated T2A-GAL4 (Gnerer et al., 2015) (Supp. Fig. 4). In all cases, the DH^{GFP} insertions show consistent patterns of expression in third instar larval brains, albeit at much lower levels than the $DH^{T2A\text{-}GAL4}$ insertions at the same MiMIC site (Fig. 3). Note that in adult brains almost no signal of DH^{GFP} was detected, in agreement with previous observations, with the

exception of *MI15073* which shows ubiquitous expression, (Diao et al., 2015; Lee et al., 2018a) (Supp. Fig. 4). These results indicate that neither the size nor the design of DH alters the functionality or expression patterns of the tagged genes and proteins.

A compact cassette to target a SIC to intron-less genes

To tag a gene containing a MiMIC/CRIMIC, the SIC should be integrated within a suitable coding intron, leaving about 50-60% of all *Drosophila* genes that encode proteins inaccessible. Targeting genes without introns by directly fusing tags in the proper reading frame has been very difficult because HDR is much less efficient than non-homologous end joining (NHEJ) (Gratz et al., 2014). Hence, to expand the range of targetable genes requires precise, seamless genome editing. Gene editing using HDR is well suited to modifying genes without introducing unwanted changes (Bier et al., 2018). HDR repairs double stranded DNA breaks using a donor template that contains two homology regions, each typically about 1,000 nucleotides, which flank the desired changes. Recombination on either side of the break replaces the regions with the template, precisely modifying the locus. We therefore developed a novel SIC compatible with HDR (Fig. 4 and Suppl. Fig. 5) that could be targeted to loci regardless of the presence of introns.

To make a SIC that is HDR compatible, three features are important: (1) a dominant marker for screening that is compact for ease of insertion via HDR (Li et al., 2014), (2) a method to insert the SIC in the desired location, and (3) a strategy to remove the SIC for replacement with the desired end product. To design a compact marker that is compatible with Golden Gate cloning we focused on the *yellow* gene which has well characterized enhancers (Geyer & Corces, 1987). We identified a 575 nucleotide regulatory region that when fused to the promoter and *yellow* coding sequence creates a 2.9 kilobase cassette that reliably drives expression only in the wing (Suppl. Fig. 6). We refer to this cassette as y^{wing2+}.

To enable targeting y^{wing2+} into precise locations in the genome, we first define a region (or gene) of interest (ROI) flanked by two Cas9 target sites comprised of a 20 nucleotide guide sequence and an NGG PAM

(Jinek et al., 2012; Sternberg et al., 2014) (Fig. 4A). We then design a HDR donor template with y^{wing2+} flanked by homology regions. As the donor template removes part of the Cas9 target sequences neither the donor cassette nor in the final product are cleaved upon HDR. Injecting the donor template along with sgRNA expression plasmids into embryos carrying a germline specific source of Cas9, followed by screening offspring for yellow+ wings provides a straightforward method to generate robust null alleles for the gene.

Lastly, to make the cassette removable, we flanked the SIC with the nucleotides “GG” and “CC” upstream and downstream of the y^{wing2+} marker, respectively. Upon insertion, this creates two novel Cas9 target sites that are not present in the endogenous sequence (box inset of Fig. 4A), which can be used to remove the inserted cassette for final replacement via a second round of HDR. Finally, to facilitate cloning, the y^{wing2+} cassette was made compatible with Golden Gate assembly (Engler et al., 2009). We also generated templates for designing replacement HDR constructs containing GFP and mCherry for protein tags or T2A-Gal4 that are compatible with Golden Gate assembly (Suppl. Fig. 5).

The y^{wing2+} SIC efficiently replaces genomic loci

We tested the efficacy of replacing the coding sequence of 10 loci with y^{wing2+} (Table 1). Nine out of ten injections led to successful integration of the cassette. We injected an average of ~500 embryos for each gene and recovered 1 to 6 independent founder lines for a total of 22 insertion events or ~2 founder animals/gene. Sanger sequencing confirmed the correct insertion of all but one HDR event, in which the entire plasmid backbone had been integrated. As shown in Table 1, seven insertions are homozygous lethal. To test whether the lethality was specific to the removal of the targeted gene, we attempted rescue of the lethality with a genomic duplication of the locus for 4 genes (*Nmnat (Nicotinamide mononucleotide adenylyltransferase), CG13390, Med27 (Mediator complex subunit 27), and CG11679,* and tested for failure to complement molecularly defined deletions for two genes (*ubiquilin* and *Nmnat*) (Zhai et al., 2006). In all cases, lethality mapped to the targeted locus showing that no second site

lethal mutation were induced in these lines. For the gene almondex (amx), the y^{wing2+} insertion produced flies that were female sterile. Female sterility was previously reported for *amx* and a genomic fragment previously reported to rescue female sterility also rescued female sterility in $amx^{\Delta cds,ywing2+}$(Jakobsdottir et al, 2016). For the gene (*Stub1(STIP1 homology and U-box containing protein 1))*, four positive lines were recovered; two are homozygous lethal while two are viable and fertile. Sanger sequencing confirmed the correct insertion of the cassette in all four lines, suggesting that the gene is not essential. Hence, the lethality is either caused by off-target cleavage events or the presence of a floating lethal mutation in the original strain. Thus while off-target cleavage may have occurred, this evidence suggests that it is not common, in agreement with what we have observed when we use CRIMIC (Lee et al., 2018a). In summary, we created null alleles for 9 genes and show that the y^{wing2+} knock-in cassette inserted precisely based on Sanger sequencing.

Step two: removal of y^{wing2} allows "scarless" modification of endogenous loci

The y^{wing2} cassette is designed to introduce two new gRNA target sites upon replacing the endogenous locus. These newly introduced gRNA target sites can now be used for the removal of the cassette via CRISPR/Cas9 mediated HDR and replacement with the desired DNA sequence. We attempted to incorporate protein tags for five genes (Table 2). We successfully incorporated tags for *Nmnat, Stub1, CG11679* and *Med27* but failed for *amx.* We tagged *Nmnat* and *Stub1* with GFP, and *CG11679* and *Med27* with Flag tags (see Mat. and Meth.). Internally GFP tagged *Nmnat* (*Nmnat::GFP::Nmnat)* and C-terminally Flag tagged *CG11679* (*CG11679::Flag)* reverted the lethality of the y^{wing2} knock-in allele and hence produced functional proteins. However, the C-terminal Flag tagged *Med27* (*Med27::Flag)* is recessive pupal lethal, similar to the y^{wing2} knock-in allele, suggesting that the C-terminal Flag tag disrupts protein function. Because the loss of *Stub1* (Table 1) does not result in an overt phenotype, we cannot determine if *Stub1::GFP* is functional but Sanger sequencing showed that the replacement of y^{wing2+} with *Stub1::GFP*

342 happened precisely. Taken together, the data indicate that Cas9 mediated
343 cassette replacement occurred correctly for four out of five genes.
344
345 **A case study: structure function analysis of *NMNAT***
346
347 To highlight the utility of the y^{wing2+} scarless replacement strategy, we
348 performed a structure function analysis of *Nmnat. Nmnat* is an enzyme
349 with Nicotinamide adenine dinucleotide (NAD) synthase activity that also
350 functions as a molecular chaperone (Zhai et al., 2006, 2008). Additionally,
351 the Nmnat family is highly conserved, required for neuronal survival and
352 protects neurons from a variety of neurodegenerative insults in flies and
353 humans (Ali et al., 2013; Brazill et al., 2017; Lau et al., 2009). A previous
354 report generated a *Nmnat* null allele and determined that its loss causes
355 lethality in first instar larva (Zhai et al., 2006). For our structure function
356 analysis, we first created a null allele of *Nmnat* by replacing the entire
357 coding sequence (CDS) with y^{wing2+}, $Nmnat^{\Delta CDS,ywing2+}$ (see Table 1). The
358 resulting flies are homozygous first instar larval lethal, consistent with a
359 known protein null, and can be rescued by a 3 kb Nmnat transgene known
360 to rescue the lethality associated with the loss of Nmnat (Zhai et al., 2006).
361 We then replaced the y^{wing2+} SIC with internally GFP tagged versions of
362 wild-type and three variants of *Nmnat* (Table 2 & Fig. 5A,B). These variants
363 of *Nmnat* are known or predicted to affect specific molecular functions of
364 Nmnat *in vitro* (Fig. 5B): (1) $Nmnat::GFP::Nmnat^{W129G}$ reduces NAD
365 synthase activity (Zhai et al., 2006) (2) $Nmnat::GFP::Nmnat^{\Delta 251\ldots 257}$ disrupts
366 the ATP binding motif required for chaperone function (Zhai et al., 2006; Ali
367 et al., 2016), and (3) $Nmnat::GFP::Nmnat^{C344S,\ C345S}$ is predicted to disrupt
368 critical palmitoylation sites that, in vertebrates, are required for membrane
369 association and protein turnover (Lau et al., 2010; Mayer et al., 2010). All
370 cassettes correctly replaced the y^{wing2+} SIC based on Sanger sequencing.
371 We chose three independent lines of wild-type and each variant for
372 further analysis. Homozygous $Nmnat::GFP::Nmnat^{WT}$ flies are viable and
373 fertile, suggesting that the internal GFP tag does not overtly affect Nmnat
374 function. In contrast, homozygous $Nmnat::GFP::Nmnat^{\Delta 251\ldots 257}$ or
375 $Nmnat::GFP::Nmnat^{\Delta 251\ldots 257}$ / $Nmnat^{\Delta 4790\text{-}1}$ animals die as 1st instar larvae
376 similar to $Nmnat^{\Delta CDS,ywing2}$, suggesting that $Nmnat::GFP::Nmnat^{\Delta 251\ldots 257}$

behaves as a null allele (Zhai et al., 2006). Homozygous *Nmnat::GFP::Nmnat*^{W129G} flies grow slowly and die prior to pupariation at late 3rd instar larval stage, suggesting that it is a hypomorphic allele. Finally, the *Nmnat::GFP::Nmnat*$^{C344S,\ C345S}$ flies that lack the putative palmitoylation sites are viable and fertile, but exhibit reduced lifespan relative to *Nmnat::GFP::Nmnat*^{WT} controls, indicating that it behaves as a weak hypomorph.

To determine protein levels and localization we stained the brains of heterozygous GFP-tagged animals (Fig. 5C). Antibody staining against GFP showed robust staining for wild-type *Nmnat::GFP::Nmnat*^{WT} (Fig. 5C). Most immunofluorescence signal is confined to the nucleus and cell body of neurons, but low levels of signal are observed in axons (Fig. 5C). In contrast, *Nmnat::GFP::Nmnat*$^{\Delta 251\ldots 257}$ surprisingly produced no detectable protein (Fig. 5C), suggesting that the removal of the ATP binding domain severely affects the stability of the mRNA or protein, consistent with the observation that is a genetic null allele. On the other hand, *Nmnat::GFP::Nmnat*^{W129G} showed a mildly reduced signal when compared to wild-type, and the GFP localization was also seen mainly in the nucleus and cell body. Finally, *Nmnat::GFP::Nmnat*$^{C344S,\ C345S}$ flies show an obvious increase in GFP levels consistent with the hypothesis that this site is required for protein degradation as previously shown in vertebrate NMNAT (Mayer et al., 2010; Milde et al., 2013; Lau et al., 2010,). In summary, our data document the ability to perform structure-function analyses in the endogenously GFP tagged locus.

Discussion

Here we describe two methods to facilitate endogenous tagging and functional annotation of genes in *Drosophila*. DH doubles the success rate of RMCE using readily available MiMIC/CRIMIC lines to generate GFP protein traps or T2A-GAL4 gene traps. On the other hand, the *y*$^{wing2+}$ SIC mediated two-step scarless gene tagging strategy offers a means to manipulate more than 6,000 genes that cannot be targeted with the artificial exon approaches. Together these technologies facilitate the use of MiMICs

and expand the capabilities of cassette swapping to include virtually all genes in *Drosophila*.

Recently two other compound RMCE cassettes that encode two different modules in opposing orientations have been reported (Fisher et al., 2017; Nagarkar-Jaiswal et al., 2017). Flip-Flop contains a protein trap that can be inverted by Flp-FRT to a mutant allele, conditionally inactivating the gene in mitotic and postmitotic cells. The mutagenic module encodes *SA-T2A-mcherry-polyA* (Nagarkar-Jaiswal et al., 2017). In FlpStop on the other hand the non-mutagenic orientation does not encode a protein, but inversion by Flp leads to a gene trap *SA-stop cassette-polyA* (Fisher et al. 2017). Hence, these methods create conditional alleles. In contrast, DH creates alleles that are final and cannot be inverted or altered by Flp expression. This allows the use of Flp/FRT for independent manipulations in the DH background.

Previously, Nagarkar-Jaiswal et al. (2015b) used a Flp-mediated *in vivo* mobilization strategy to create GFP protein traps for intronic MiMIC containing genes whereas Diao et al. (2015) used a Cre-mediated *in vivo* mobilization scheme to create T2A-GAL4 gene traps. Both the GFP tag and T2A-GAL4 provide complementary means to assess gene function and expression. By combining the two in a single vector, DH greatly improves the rate of RMCE, the breath of applications, and the amount of labor involved.

We observed that although T2A-GAL4 is highly successful in marking the gene expression domains in the adult brain, for most genes the GFP tagged protein signal is often weak, consistent with previous results (Lee et al., 2018a; Diao et al., 2015). Even when these GFP tagged proteins cannot be used to detect the gene product, they can still be very useful for biochemical applications or to knock down the gene through a variety of methods to create conditional alleles (Caussinus et al., 2011; Harmansa et al., 2017; Harmansa et al., 2015; Nagarkar-Jaiswal et al., 2015a; Neumuller et al., 2012; Lee et al., 2018b).

Interestingly for three out of 28 MiMICs, we detected numerous DH RMCE inserts, judged by loss of the *yellow* marker, but for many of these events we could not determine conclusively the orientation or presence of DH by PCR. However, these false positives are easily identified by single

fly PCR (Supp. Fig. 2). Moreover, we could identify positive events in the MiMICs where the false positive rate was high, showing that the high false positive rate for these MiMICs does not limit the technique.

Finally, the GAL4 >UAS system can also be used to assess the function of cells, particularly neurons. Multiple features of neurons, including electrophysiological properties, can be modulated or assessed using established UAS constructs (Venken et al., 2011b). These include the UAS-Tetanus Toxin, UAS-Kir2.1 or UAS-Shibire^{ts} to silence neurons (Sweeney et al., 1995; Baines et al., 2001; Kitamoto, 2001); UAS-TRPM8 or UAS-ChannelRhodopsin 2 (ChR2) to activate neurons (Peabody et al., 2009; Schroll et al., 2006); and UAS-GCaMP to assess changes in Calcium concentrations (Chen et al., 2013) or UAS-ASAP2 that acts as voltage sensor (Yang et al., 2016) to assess neuronal activity. Hence, the cells that express T2A-Gal4 associated with a specific gene can be manipulated in numerous ways.

Given that 50-60% of the protein coding genes do not contain suitable introns numerous genes are not amenable for tagging based on our approaches. Inserting tags in genes that lack large (>150 bp) introns creates two main challenges: 1) screening for a precise rare gene editing event is very time consuming and 2) inserting extraneous sequences for RMCE within or near coding regions often create mutations and indels which disrupt protein function. Scarless gene editing offers obvious advantages for manipulating these loci, however, few options currently exist. Scarless gene editing can be achieved through the use of single-stranded DNA donors (Gratz et al., 2014, Xue et al., 2014). However, they are limited in size to ~200 nucleotides by current synthesis methods (Korona et al., 2017). Accordingly, sgRNA sites must be close to the specific nucleotides to be edited and because they cannot carry visible markers they require laborious screening methods to find flies carrying the correct gene editing event. Two strategies have been proposed to integrate fluorescent markers in fly genes using double stranded DNA donor plasmids and remove them to perform scarless genome editing (reviewed in Bier et al., 2018). One method, the Scarless-dsRed system (http://flycrispr.molbio.wisc.edu/scarless), relies on the *piggybac* transposase to remove a dominant marker by precise excision after the

gene editing event has been confirmed. However, as of yet no data have been reported to determine its efficiency or efficacy. While we were developing and testing our approach, a similar method, pGEM-wingGFP-tan, was reported (Lamb et al., 2017). This method integrates two Cas9 target sequences on either end of a GFP marker driven by a wing promotor to replace a locus. Unlike y^{wing2+} which is readily visible in adults, the reported wing-GFP marker needs to be scored within a narrow developmental stage in pupae with a fluorescent microscope (Lamb et al., 2017). Moreover, Lamb et al. (2017) report a high rate of backbone insertion, where we observed only a single case out of 11 genes with our y^{wing2+} approach. Since the methodology was only applied to a single gene, we cannot compare our data with Lamb et al (2017).

For 9 out of 10 genes that we targeted with y^{wing2+}, we obtained at least one correctly inserted SIC from an injection of ~400 embryos. The y^{wing2+} marker is easy to score in adult flies and is compatible with Golden Gate and Gibson assembly (Engler et al., 2009; Gibson et al., 2009), greatly facilitating its application to virtually any locus within the genome. The first step creates a null allele which provides an essential reference point for all subsequent genetic and molecular manipulations. Given that most genes that lack introns are rather small, they are poor targets for chemical or transposon mediated mutagenesis, although they can be targeted with CRISPR based on NHEJ. The y^{wing2+} SIC offers an easy way to completely remove these small genes and is not labor intensive.

The major advantage of the y^{wing2+} SIC is that it creates a highly versatile line. Multiple manipulations within the region of interest can be performed in parallel in essentially an isogenic background using this SIC. Although we did not observe widespread off-target mutations, we suggest the use of multiple lines where possible. We have shown that the cassette swapping via the y^{wing2+} SIC occurs precisely with a high success rate and demonstrated its usefulness both for gene tagging and for structure function analysis.

In summary, the two methodologies and accompanying tool kits presented here complement and expand existing MiMIC and CRIMIC approaches. The combination of these methodologies should enable

endogenous tagging and manipulation of most fly genes, an invaluable resource for the fly research community.

Materials and Methods

Cloning of DH

Sequence of the DH plasmids can be found in Supp. file. Briefly, SA- EGFP-FlAsH-StrepII-TEVcs-3xFlag-SD cassette was PCR amplified from pBS-KS-attB1-2-PT-SA-SD-pX (corresponding to the codon phase)-EGFP-FlAsH-StrepII-TEV-3xFlag (DGRC # 1298, (Venken et al., 2011)) with tags_for_BsiWI, tags_rev_AvrII primers. This fragment is cloned in, a modified pTGEM plasmid of pX (corresponding to the codon phase) where the loxP site before 3XP3RFP is deleted and a BsiWI site is integrated after an AvrII site using BsiWI and AvrII restriction sites. Resulting vector is cut with XbaI-BsiWI and cloned into pC-(loxP2-attB2-SA(1)-T2A-Gal4-Hsp70) (Addgene # 62955, (Diao et al., 2015)) modified to include a BsiWI site after SD, generating DH pX.

Cloning of constructs to test *yellow* wing enhancer expression

A sub-region of the *yellow* dominant marker from P{EPgy2} (Bellen et al., 2011) that contains the promoter, coding sequence and UTRs was subcloned into the plasmid pattB (Accession # KC896839 (Bischof et al., 2013)) using oligos DLK0048 and DLK0049 (see suppl. file for table showing oligonucleotides sequences used) flanked by XhoI and XbaI to make the vector pBS II SK-attB yMP *w*+. Either the full sequence of the *yellow* enhancers (oligos DLK0054 & DLK0056) (Geyer & Corces, 1987) or sub-fragments (oligos DLK0054&DLK0057 or DLK0055&DLK0056 - see supp. Fig. 6) were then subcloned into pattB yMP *w*+ to make pattB expression constructs. Once expression of the markers were verified, *miniwhite*+ and the *loxP* sites were removed by digestion with ApaI and NotI, the ends blunted with DNA Polymerase I, Large (Klenow) Fragment (NEB) and ligated using T4 Ligase (NEB). As only the full body, *full wing* and *wing2* gave positive expression, *wing2* was chosen for use as the most compact construct and a multiple cloning site was then added using annealed oligos DLK0022&DLK0023 cut with SpeI and XbaI into XbaI digested pattB y^{wing2+}, and sequence verified for insert direction to make

functional plasmids for ϕC31 mediated transgenesis. The sequence of pattB pattB *y*^{wing2} can be found in the Supp. file.

Cloning *y*^{wing2} Golden Gate donor template

Three versions of p{*y*$^{wing2+}$} were cloned: for BsmBI, BsaI, and BbsI. First, annealed oligos (DLK320&DLK338, DLK322&DLK339, and DLK324&DLK340 for BsmBI, BbsI and BsaI, respectively) containing the appropriate SacI overhang and ends-in TypeIIS restriction sites with a short intervening random nucleotide spacer inserted into the pM14 plasmid backbone (Lee et al., 2018a) digested with enzymes SacI and EcoRV. The *y*$^{wing2+}$ dominant reporter was then subcloned by PCR using oligos DLK326&DLK327 from p-attB *y*^{wing2} into the pM14 + spacer backbone digested with BbsI, BsaI or BsmBI. The sequence of p{*y*$^{wing2+}$} can be found in the Supp. file.

Cloning the GFP, mCherry and T2A-Gal4 donor templates

Three versions of p{*GFP*} were cloned: for BsmBI, BsaI, and BbsI. Annealed oligos (DLK461&DLK462, DLK463&DLK464, and DLK465&DLK466 for BsmBI, BbsI and BsaI, respectively) containing the appropriate SacI overhang and ends-in TypeIIS restriction sites were inserted into the pM14 plasmid backbone (Lee et al., 2018a) digested with enzymes SacI and EcoRV to make p{spacer-L-L} where L denotes $(GGS)_4$ linker. The linker-GFP sequence was generated by PCR from PM14 (REF) and using oligos DLK225&DLK300 and subcloned into p{spacer-L-GFP-L}. The second linker was then added from annealed oligos DLK554&DLK555 to make the BbsI version of p{L--L}. Finally, the linker-GFP-linker was subcloned into versions of p{spacer-L -L} for BsaI and BsmBI to make plasmids compatible with each enzyme. Additional template vectors were produced for mCherry and T2A-Gal4 inserts which can also be found in Supp. file. mCherry was subcloned into p{spacer-L-L} from the Flip-Flop cassette which was designed with silent mutations to remove BsaI sites (Nagarkar-Jaiswal, et al., 2017) and T2A-Gal4 inserts were subcloned from pM14 (Lee et al., 2018) into p{spacer-L- L} . Sequences of can be found in the suppl. file.

Cloning HDR donor injection plasmids

Donor constructs were generated as previously described (Housden and Perrimon, 2016). Briefly, homology arms were PCR amplified from genomic DNA using Q5 polymerase (NEB), run on an agarose gel and

purified with the QIAquick Gel Extraction Kit (Qiagen). The homology arms, pBH donor vector and either p{y^{wing2}} or p{L-GFP-L} cassette were combined by Golden Gate assembly (Engler et al., 2009) using the appropriate type IIS restriction enzyme (BbsI, BsaI, or BsmBI). The resulting reaction products were transformed into Stbl2 Chemically Competent Cells (ThermoFisher), and plated overnight under kanamycin selection. Colonies were cultured for 24 hr at 30°C and DNA was prepared by miniprep (QIAGEN). The entire homology arm and partial sequences of the adjacent cassette were verified prior to injection. Additional cloning information and sequences for all HDR donor plasmids can be found in Supp. file.

Cloning sgRNA expression constructs

sgRNA expression constructs were cloned into the vector pCFD3-dU6:3gRNA (addgene plasmid #49410) using established protocols (http://www.crisprflydesign.org/wp-content/uploads/2014/05/Cloning-with-pCFD3.pdf). Sequences of sgRNAs can be found in the Supp. File.

Cloning Genomic rescue constructs

Genomic fragments were cloned into pattB (Ascension # KC896839 (Bischof et al., 2013) by PCR from wild-type genomic DNA and inserted BamHi/NotI for CG13390 (oligos DLK823&DLK824) and XhoI/AvrII (into the XbaI site of pattB) for Med27 (oligos DLK966&DLK969)

Fly lines

MiMIC stocks are obtained from Bloomington Drosophila Stock Center.

The fly lines to conduct RMCE of DH with crosses are:

y[1] M{vas-int.Dm}ZH-2A w[*], P{y[+mDint2]=Crey}1b; ; *TM3, Sb[1]/TM2*
y[1] M{vas-int.Dm}ZH-2A w[*], P{y[+mDint2]=Crey}1b; *Sco/CyO*
y[1] w[*]; P{DHp0-7 [w+]};TM2/TM6, Tb[1]
y[1] w[*]; P{DHp1-7 [w+]};TM2/TM6, Tb[1]
y[1] w[*]; P{DHp2-4 [w+]};TM2/TM6, Tb[1]
y[1] w[*]; Kr[If1] wg[Sp1]/CyO; P{DHp0-8 [w+]}
y[1] w[*]; Kr[If1] wg[Sp1]/CyO; P{DHp1-6 [w+]}
y[1] w[*]; Kr[If1] wg[Sp1]/CyO; P{DHp2-6A [w+]}

Cas9 stocks for CRISPR experiments carried an isogenic chromosome on either X, II or III derived from *y w* and were either y^1 *M{nos-Cas9.P}ZH-2A w** (from Bloomington stock center) or y^1 *(iso X)* w^**; +/+; attP2(y+){nos-Cas9(y+)}* (Ren et al., 2013) in which the *y+* marker was mutagenized by injecting sgRNA expression plasmids (pCFD3-y1 and pCFD3-y2) against the yellow coding sequence. The isogenic chromosomes (X, II, III) were sequenced using whole genome sequencing (Human Genome Sequencing Center at Baylor College of Medicine) and the sequence files are available upon request.

Generation of Double Header transgenics

RMCE to generate DH insertions by injections is depicted in Supp. Fig. 1A and in (Nagarkar-Jaiswal et al., 2017). Briefly, the DH plasmid of the correct phase (500 ng/ul final concentration) is mixed with ΦC31 integrase helper plasmid (400ng/ul final concentration) and injected in embryos of MiMIC stocks. The crossing scheme for generating DH insertions is depicted in Supp. Fig. 3A. 5-7 crosses with 10-15 virgins of *hs-Cre vas-dΦC31int*; ;*TM2/TM6* were set with 7-10 males from individual MiMIC lines. The vials are flipped every second day to prevent overcrowding. 5-10 crosses are set in the subsequent generations. The resulting individual y- flies are selected to set up stocks. Note that SICs are flanked by FRT sites in the newly generated CRIMIC alleles. Hence, the Flp/FRT mobilization schemes described by Nagarkar-Jaiswal et al. (2015 b) cannot be used to generate GFP protein traps in these CRIMICs, given that expression of the Flp would excise the SIC. For the same reason, Flp/FRT cannot be used in combination with the CRIMIC T2A-GAL4 alleles for experiments that require both T2A-GAL4 and Flp.

PCR determination of DH orientation

For each DH insertion stock four PCRs are set. Primer pairs are MiMIC_5'_for-GFP_DH_for, MiMIC_3'_rev-T2A_GAL4_rev, MiMIC_5'_for-T2A_GAL4_rev, and MiMIC_3'_rev-GFP_DH_for. Correct RMCE events result in 2 out of 4 successful amplicons (Supp. Fig. 2). We isolated a number of lines where only one out of two amplicons was detected or gave conflicting results. These lines are not included in the analysis in this manuscript.

CRISPR/Cas9 injections for scarless y^{wing2} gene editing

Embryos (designated G_0) at less than one hour post-egg laying carrying the appropriate Cas9 allele were injected with a mixture of 150-200 ng/ul of donor plasmid and 25 ng/ul of each sgRNA expression construct, transferred to standard media after 24-48 hours, and crossed to *yw* flies. For experiments to insert the y^{wing2+} cassette, offspring were screened for presence of $yellow^{+}$ wings several days after eclosion and crossed to appropriate balancers. For experiments to replace the y^{wing2+} cassette, G_0 flies were crossed to *y w* flies carrying appropriate balancers and F1 offspring screened for loss of yellow+ wings. Individual F1 founders were backcrossed to *y w* flies carrying appropriate balancers, allowed to establish larvae, and screened for presence of the insert via PCR (see Supp. file).

Confocal Imaging

Confocal imaging was conducted as in the previous study (Lee et al., 2018a). In brief, dissected adult brains were fixed in 4% paraformaldehyde/1xPBS overnight, then penetrated with 2% Triton X-100/1xPBS at 4°C overnight. The larval brains or other tissues were fixed in 4% paraformaldehyde/ 1xPBS at 4°C for at least 2 hours, transferred to 0.5% Triton X-100/1xPBS for overnight 4°C incubation. For adult brains, the samples were vacuumed for 1 hour at room temperature, and left overnight in the same solution for penetration at 4°C. For immunostaining of GFP, the samples were incubated with anti-GFP antibody conjugated with FITC (1:500) (Abcam, RRID: AB_305635) in 1xPBS with 0.5% Triton X-100 overnight. Samples were cleared and mounted in RapiClear® (SunJin Lab Co.) and imaged with a Zeiss LSM 880 Confocal Microscope under a 20x objective lens.

Figure 1. Double Header optimizes RMCE outcome of MiMICs. (A) Schematics of the Double Header construct and RMCE outcomes. Double Header constructs contain a Splice Acceptor (SA)- super folder GFP-FlAsH-StrepII-TEV-3xFlag (EGFP) – Splice Donor (SD) in one orientation and a SA-T2A-GAL4-polyA in the other orientation. Insertion in the GFP orientation results in GFP protein trap whereas insertion in the T2A-GAL4 orientation results in T2A-GAL4 gene trap. (B) Double Header injection statistics.

Figure 2. Double Header integration through crosses facilitates RMCE. (A) Schematics of the Double Header transgene mobilization *in vivo*. Double Header transgenes contain *loxP* sites that can be used to mobilize the RMCE cassette *in vivo*, without the need for injection. (B) Double header crossing statistics.

Figure 3. Examples of gene expression patterns obtained by Double Header. Each MiMIC, *MI01487*, *MI05208*, *MI06794*, MI06872, *MI08614*, *MI11741* and *MI15073*, was converted to either T2A-GAL4 protein traps or GFP protein traps by Double Header insertion. The expression in the larval CNS is shown with either *T2A-GAL4>UAS-mCD8::GFP* or *GFP-tag* (GFP and mCD8::GFP, green). The affected genes are labelled above. Scale bar: 50 µm.

Figure 4. Schematic of a two-step system for scarless gene editing. (A) In step 1, a cassette containing a dominant marker flanked by nucleotides GG and CC replaces an endogenous locus via Homology Directed Repair (HDR) following double strand breaks produced by Cas9 cleavage. The removal of the intervening sequence between the Cas9 cut sites alters the sgRNA target sequences preventing cleavage of the donor construct or the modified DNA. Screening for the dominant marker facilitates identification of CRISPR gene editing events while the flanking nucleotides GG (boxed inset) and CC create novel Cas9 target sites, allowing subsequent excision. (B) In step 2 the insert is removed and replaced with any DNA via a second round of HDR with new sgRNA sequences, facilitating the scarless insertion of any DNA sequence desirable.

Figure 5. y^{wing2+} cassette-swapping facilitates structure-function analyses. (A) Schematic of the *Nmnat::GFP::Nmnat* donor construct for replacing the inserted y^{wing2+} SIC at the *Nmnat* locus. (B) *Nmnat::GFP::Nmnat* variants used in the structure function experiment. Red * denotes approximate location of altered sequence(s) (C) images of adult brains of *Nmnat::GFP::Nmnat* ^{wt} (Top left) *Nmnat::GFP::Nmnat*^{W129G} (Bottom left) *Nmnat::GFP::Nmnat*$^{\Delta 251\ldots 257}$ (Top right) and *Nmnat:GFP:Nmnat*$^{C344S,\ C345S}$ (bottom right).

Table 1. Summary statistics for y^{wing2+} SIC insertion experiments

Table 2. Summary statistics for y^{wing2+} cassette-swap experiments

Supplementary Figure 1. Injection data for Double Header. (A) The crossing scheme to generate Double Header RMCE events through embryo injection. (B) PCR determination of the Double Header orientation. ? indicates the number of lines where the PCR pattern was ambiguous. Yellow MiMICs: Double Header insertions in both orientations, Green MiMICs: only GFP protein trap orientation, Blue MiMICs: only T2A-GAL4 gene trap orientation.

Supplementary Figure 2. PCR strategy to identify Double Header orientation. Primers are indicated on the Double Header construct that inserted in an intronic MiMIC. For each Double Header insertion 4 single fly PCR reactions were set with MiMIC inwards primers that bind outside the attP sites and are directed inwards (MiMIC_5'_for, MiMIC_3'_rev) and GFP and T2A-GAL4 specific outwards primers (DH_GFP_rev, DH_T2A-for). A correct insertion results in amplicons in two out of four PCRs.

Supplementary Figure 3. Crossing scheme for Double Header and data of integration. (A) Crossing scheme for mobilizing double header RMCE cassette and selecting RMCE events. (B) PCR determination of the Double Header orientation. ? indicates the number of lines where PCR pattern was ambiguous. Yellow MiMICs: Double Header insertions in both orientations, Green MiMICs: only GFP protein trap orientation, Blue MiMICs: only T2A-GAL4 gene trap orientation.

Supplementary Figure 4. Examples of gene expression patterns obtained by Double Header insertions in MiMICs in adult brain. Each MiMIC *MI01487*, *MI05208*, *MI06794*, *MI06872*, *MI08614*, *MI11741* and *MI15073*, was converted to either a T2A-GAL4 gene trap or GFP protein trap by Double Header insertion. The expression in the adult brain is shown with either *T2A-GAL4>UAS-mCD8::GFP* or *GFP-tag* (GFP and mCD8::GFP, green). The affected genes are labelled above. Arrow: Ellipsoid body (*MI05208-5HT2B*). Scale bar: 50 µm.

Supplementary Figure 5. Template vectors for cloning *yellow* expression constructs. (A) Vectors for ΦC31 mediated site specific integration with *yellow* wing and body expression constructs. (B**)** *yellow* wing and body expression constructs compatible with Golden Gate cloning

for SIC insertion via HDR. (C**)** Template donors compatible with Golden Gate cloning for yellow cassette swapping via HDR.

Supplementary Figure 6. Mapping cis-regulatory modules for the *yellow* gene. (A) Structure of the *yellow* dominant marker showing previously characterized enhancers for wing and body (Geyer & Corces, 1987). (B) regions tested for *yellow* expression (black lines) when fused to a minimal promoter and the *yellow* gene. The table at the bottom right indicates the presence/absence of dark pigmentation in either the body or the wing for each cis-regulatory module (CRM) tested.

Acknowledgements

We thank the Bloomington *Drosophila* Stock Center and the Developmental Studies Hybridoma Bank for antibodies. We thank Zelun Wang for technical help, and Megan Campbell and Shinya Yamamoto for reading the manuscript and providing helpful suggestions as well as Clair Hu for valuable information. Confocal microscopy was performed in the BCM IDDRC Neurovisualization Core, supported by the NICHD (U54HD083092). This research was supported by NIH grants R01GM067858, Robert A and Renee E Belfer Family Foundation, and the Huffington Foundation to H.J.B. D.L.K. was supported by the Brain Disorders training grant NIH T32NS043124. HJB is an investigator of the Howard Hughes Medical Institute.

Bibliography

Ali, Y. O., H. M. Allen, L. Yu, D. Li-Kroeger, D. Bakhshizadehmahmoudi, A. Hatcher, C. McCabe, J. Xu, N. Bjorklund, G. Taglialatela, D. A. Bennett, P. L. De Jager, J. M. Shulman, H. J. Bellen and H. C. Lu (2016). "NMNAT2:HSP90 Complex Mediates Proteostasis in Proteinopathies." PLoS Biol 14(6): e1002472.

Ali, Y. O., D. Li-Kroeger, H. J. Bellen, R. G. Zhai and H. C. Lu (2013). "NMNATs, evolutionarily conserved neuronal maintenance factors." Trends Neurosci **36**(11): 632-640.

Baines, R. A., J. P. Uhler, A. Thompson, S. T. Sweeney and M. Bate (2001). "Altered electrical properties in Drosophila neurons developing without synaptic transmission." J Neurosci **21**(5): 1523-1531.

Bellen, H. J., R. W. Levis, Y. He, J. W. Carlson, M. Evans-Holm, E. Bae, J. Kim, A. Metaxakis, C. Savakis, K. L. Schulze, R. A. Hoskins and A. C. Spradling (2011). "The Drosophila gene disruption project: progress using transposons with distinctive site specificities." Genetics **188**(3): 731-743.

Bellen, H. J. and S. Yamamoto (2015). "Morgan's legacy: fruit flies and the functional annotation of conserved genes." Cell **163**(1): 12-14.

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): 1-18.

Bischof, J., M. Bjorklund, E. Furger, C. Schertel, J. Taipale and K. Basler (2013). "A versatile platform for creating a comprehensive UAS-ORFeome library in Drosophila." Development **140**(11): 2434-2442.

Brazill, J. M., C. Li, Y. Zhu and R. G. Zhai (2017). "NMNAT: It's an NAD(+) synthase... It's a chaperone... It's a neuroprotector." Curr Opin Genet Dev **44**: 156-162.

Brown, J. B., N. Boley, R. Eisman, G. E. May, M. H. Stoiber, M. O. Duff, B. W. Booth, J. Wen, S. Park, A. M. Suzuki, K. H. Wan, C. Yu, D. Zhang, J. W. Carlson, L. Cherbas, B. D. Eads, D. Miller, K. Mockaitis, J. Roberts, C. A. Davis, E. Frise, A. S. Hammonds, S. Olson, S. Shenker, D. Sturgill, A. A. Samsonova, R. Weiszmann, G. Robinson, J. Hernandez, J. Andrews, P. J. Bickel, P. Carninci, P. Cherbas, T. R. Gingeras, R. A. Hoskins, T. C. Kaufman, E. C. Lai, B. Oliver, N. Perrimon, B. R. Graveley and S. E. Celniker (2014). "Diversity and dynamics of the Drosophila transcriptome." Nature **512**(7515): 393-399.

Caussinus, E., O. Kanca and M. Affolter (2011). "Fluorescent fusion protein knockout mediated by anti-GFP nanobody." Nat Struct Mol Biol **19**(1): 117-121.

Chandran, R. R., E. Iordanou, C. Ajja, M. Wille and L. Jiang (2014). "Gene expression profiling of Drosophila tracheal fusion cells." Gene Expr Patterns **15**(2): 112-123.

Chen, T. W., T. J. Wardill, Y. Sun, S. R. Pulver, S. L. Renninger, A. Baohan, E. R. Schreiter, R. A. Kerr, M. B. Orger, V. Jayaraman, L. L. Looger, K. Svoboda and D. S. Kim (2013). "Ultrasensitive fluorescent proteins for imaging neuronal activity." Nature **499**(7458): 295-300.

Cox, J. E., C. S. Thummel and J. M. Tennessen (2017). "Metabolomic Studies in Drosophila." Genetics **206**(3): 1169-1185.

David-Morrison, G., Z. Xu, Y. N. Rui, W. L. Charng, M. Jaiswal, S. Yamamoto, B. Xiong, K. Zhang, H. Sandoval, L. Duraine, Z. Zuo, S. Zhang and H. J. Bellen (2016). "WAC Regulates mTOR Activity by Acting as an Adaptor for the TTT and Pontin/Reptin Complexes." Dev Cell **36**(2): 139-151.

Diao, F., H. Ironfield, H. Luan, F. Diao, W. C. Shropshire, J. Ewer, E. Marr, C. J. Potter, M. Landgraf and B. H. White (2015). "Plug-and-play genetic access to drosophila cell types using exchangeable exon cassettes." Cell Rep **10**(8): 1410-1421.

Engler, C., R. Gruetzner, R. Kandzia and S. Marillonnet (2009). "Golden gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes." PLoS One **4**(5): e5553.

Fisher, Y. E., H. H. Yang, J. Isaacman-Beck, M. Xie, D. M. Gohl and T. R. Clandinin (2017). "FlpStop, a tool for conditional gene control in Drosophila." Elife **6**.

Germani, F., C. Bergantinos and L. A. Johnston (2018). "Mosaic Analysis in Drosophila." Genetics **208**(2): 473-490.

Geyer, P. K. and V. G. Corces (1987). "Separate regulatory elements are responsible for the complex pattern of tissue-specific and developmental transcription of the yellow locus in Drosophila melanogaster." Genes Dev **1**(9): 996-1004.

Gibson, D. G., L. Young, R. Y. Chuang, J. C. Venter, C. A. Hutchison, 3rd and H. O. Smith (2009). "Enzymatic assembly of DNA molecules up to several hundred kilobases." Nat Methods **6**(5): 343-345.

Gnerer, J. P., K. J. Venken and H. A. Dierick (2015). "Gene-specific cell labeling using MiMIC transposons." Nucleic Acids Res **43**(8): e56.

Gratz, S. J., F. P. Ukken, C. D. Rubinstein, G. Thiede, L. K. Donohue, A. M. Cummings and K. M. O'Connor-Giles (2014). "Highly specific and efficient CRISPR/Cas9-catalyzed homology-directed repair in Drosophila." Genetics **196**(4): 961-971.

Harmansa, S., I. Alborelli, D. Bieli, E. Caussinus and M. Affolter (2017). "A nanobody-based toolset to investigate the role of protein localization and dispersal in Drosophila." Elife **6**.

Harmansa, S., F. Hamaratoglu, M. Affolter and E. Caussinus (2015). "Dpp spreading is required for medial but not for lateral wing disc growth." Nature **527**(7578): 317-322.

Heigwer, F., F. Port and M. Boutros (2018). "RNA Interference (RNAi) Screening in Drosophila." Genetics **208**(3): 853-874.

Jakobsdottir, J., S. J. van der Lee, J. C. Bis, V. Chouraki, D. Li-Kroeger, S. Yamamoto, M. L. Grove, A. Naj, M. Vronskaya, J. L. Salazar, A. L. DeStefano, J. A. Brody, A. V. Smith, N. Amin, R. Sims, C. A. Ibrahim-Verbaas, S. H. Choi, C. L. Satizabal, O. L. Lopez, A. Beiser, M. A. Ikram, M. E. Garcia, C. Hayward, T. V. Varga, S. Ripatti, P. W. Franks, G. Hallmans, O. Rolandsson, J. H. Jansson, D. J. Porteous, V. Salomaa, G. Eiriksdottir, K. M. Rice, H. J. Bellen, D. Levy, A. G. Uitterlinden, V. Emilsson, J. I. Rotter, T. Aspelund, H. Cohorts for, c. Aging Research in Genomic Epidemiology, C. Alzheimer's Disease Genetic, Genetic, c. Environmental Risk in Alzheimer's Disease, C. J. O'Donnell, A. L. Fitzpatrick, L. J. Launer, A. Hofman, L. S. Wang, J. Williams, G. D. Schellenberg, E. Boerwinkle, B. M. Psaty, S. Seshadri, J. M. Shulman, V. Gudnason and C. M. van Duijn (2016). "Rare Functional Variant in TM2D3 is Associated with Late-Onset Alzheimer's Disease." PLoS Genet **12**(10): e1006327.

Jinek, M., K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna and E. Charpentier (2012). "A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity." Science 337(6096): 816-821.

Kanca, O., H. J. Bellen and F. Schnorrer (2017). "Gene Tagging Strategies To Assess Protein Expression, Localization, and Function in Drosophila." Genetics **207**(2): 389-412.

Kaufman, T. C. (2017). "A Short History and Description of Drosophila melanogaster Classical Genetics: Chromosome Aberrations, Forward Genetic Screens, and the Nature of Mutations." Genetics **206**(2): 665-689.

Kitamoto, T. (2001). "Conditional modification of behavior in Drosophila by targeted expression of a temperature-sensitive shibire allele in defined neurons." J Neurobiol **47**(2): 81-92.

Korona, D., S. A. Koestler and S. Russell (2017). "Engineering the Drosophila Genome for Developmental Biology." J Dev Biol 5(4).

Lamb, A. M., E. A. Walker and P. J. Wittkopp (2017). "Tools and strategies for scarless allele replacement in Drosophila using CRISPR/Cas9." Fly (Austin) **11**(1): 53-64.

Lau, C., C. Dolle, T. I. Gossmann, L. Agledal, M. Niere and M. Ziegler (2010). "Isoform-specific targeting and interaction domains in human nicotinamide mononucleotide adenylyltransferases." J Biol Chem **285**(24): 18868-18876.

Lau, C., M. Niere and M. Ziegler (2009). "The NMN/NaMN adenylyltransferase (NMNAT) protein family." Front Biosci (Landmark Ed) **14**: 410-431.

Lee, P. T., J. Zirin, O. Kanca, W. W. Lin, K. L. Schulze, D. Li-Kroeger, R. Tao, C. Devereaux, Y. Hu, V. Chung, Y. Fang, Y. He, H. Pan, M. Ge, Z. Zuo, B. E. Housden, S. E. Mohr, S. Yamamoto, R. W. Levis, A. C. Spradling, N. Perrimon and H. J. Bellen (2018a). "A gene-specific T2A-GAL4 library for *Drosophila*." Elife **7**.

Lee, P. T., G. Lin, W. W. Lin, F. Diao, B. H. White and H. J. Bellen (2018b). "A kinase-dependent feedforward loop affects CREBB stability and long term memory formation." Elife 7.

Li, K., G. Wang, T. Andersen, P. Zhou and W. T. Pu (2014). "Optimization of genome engineering approaches with the CRISPR/Cas9 system." PLoS One 9(8): e105779.

Mayer, P. R., N. Huang, C. M. Dewey, D. R. Dries, H. Zhang and G. Yu (2010). "Expression, localization, and biochemical characterization of nicotinamide mononucleotide adenylyltransferase 2." J Biol Chem **285**(51): 40387-40396.

Nagarkar-Jaiswal, S., S. Z. DeLuca, P. T. Lee, W. W. Lin, H. Pan, Z. Zuo, J. Lv, A. C. Spradling and H. J. Bellen (2015). "A genetic toolkit for tagging intronic MiMIC containing genes." Elife **4**.

Nagarkar-Jaiswal, S., P. T. Lee, M. E. Campbell, K. Chen, S. Anguiano-Zarate, M. C. Gutierrez, T. Busby, W. W. Lin, Y. He, K. L. Schulze, B. W. Booth, M. Evans-Holm, K. J. Venken, R. W. Levis, A. C. Spradling, R.

A. Hoskins and H. J. Bellen (2015). "A library of MiMICs allows tagging of genes and reversible, spatial and temporal knockdown of proteins in Drosophila." Elife **4**.

Nagarkar-Jaiswal, S., S. N. Manivannan, Z. Zuo and H. J. Bellen (2017). "A cell cycle-independent, conditional gene inactivation strategy for differentially tagging wild-type and mutant cells." Elife **6**.

Negre, N., C. D. Brown, L. Ma, C. A. Bristow, S. W. Miller, U. Wagner, P. Kheradpour, M. L. Eaton, P. Loriaux, R. Sealfon, Z. Li, H. Ishii, R. F. Spokony, J. Chen, L. Hwang, C. Cheng, R. P. Auburn, M. B. Davis, M. Domanus, P. K. Shah, C. A. Morrison, J. Zieba, S. Suchy, L. Senderowicz, A. Victorsen, N. A. Bild, A. J. Grundstad, D. Hanley, D. M. MacAlpine, M. Mannervik, K. Venken, H. Bellen, R. White, M. Gerstein, S. Russell, R. L. Grossman, B. Ren, J. W. Posakony, M. Kellis and K. P. White (2011). "A cis-regulatory map of the Drosophila genome." Nature **471**(7339): 527-531.

Neumuller, R. A., F. Wirtz-Peitz, S. Lee, Y. Kwon, M. Buckner, R. A. Hoskins, K. J. Venken, H. J. Bellen, S. E. Mohr and N. Perrimon (2012). "Stringent analysis of gene function and protein-protein interactions using fluorescently tagged genes." Genetics **190**(3): 931-940.

Peabody, N. C., J. B. Pohl, F. Diao, A. P. Vreede, D. J. Sandstrom, H. Wang, P. K. Zelensky and B. H. White (2009). "Characterization of the decision network for wing expansion in Drosophila using targeted expression of the TRPM8 channel." J Neurosci **29**(11): 3343-3353.

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(29): E2967-2976.

Ren, X., J. Sun, B. E. Housden, Y. Hu, C. Roesel, S. Lin, L. P. Liu, Z. Yang, D. Mao, L. Sun, Q. Wu, J. Y. Ji, J. Xi, S. E. Mohr, J. Xu, N. Perrimon and J. Q. Ni (2013). "Optimized gene editing technology for Drosophila melanogaster using germ line-specific Cas9." Proc Natl Acad Sci U S A **110**(47): 19012-19017.

Schroll, C., T. Riemensperger, D. Bucher, J. Ehmer, T. Voller, K. Erbguth, B. Gerber, T. Hendel, G. Nagel, E. Buchner and A. Fiala (2006). "Light-induced activation of distinct modulatory neurons triggers appetitive or aversive learning in Drosophila larvae." Curr Biol **16**(17): 1741-1747.

Senturk, M. and H. J. Bellen (2017). "Genetic strategies to tackle neurological diseases in fruit flies." Curr Opin Neurobiol **50**: 24-32.

Simpson, J. H. and L. L. Looger (2018). "Functional Imaging and Optogenetics in Drosophila." Genetics **208**(4): 1291-1309.

Sternberg, S. H., S. Redding, M. Jinek, E. C. Greene and J. A. Doudna (2014). "DNA interrogation by the CRISPR RNA-guided endonuclease Cas9." Nature 507(7490): 62-67.

Sweeney, S. T., K. Broadie, J. Keane, H. Niemann and C. J. O'Kane (1995). "Targeted expression of tetanus toxin light chain in Drosophila specifically eliminates synaptic transmission and causes behavioral defects." Neuron **14**(2): 341-351.

Venken, K. J., K. L. Schulze, N. A. Haelterman, H. Pan, Y. He, M. Evans-Holm, J. W. Carlson, R. W. Levis, A. C. Spradling, R. A. Hoskins and H. J. Bellen (2011). "MiMIC: a highly versatile transposon insertion resource for engineering Drosophila melanogaster genes." Nat Methods **8**(9): 737-743.

Venken, K. J., J. H. Simpson and H. J. Bellen (2011). "Genetic manipulation of genes and cells in the nervous system of the fruit fly." Neuron **72**(2): 202-230.

Wangler, M. F., Y. Hu and J. M. Shulman (2017). "Drosophila and genome-wide association studies: a review and resource for the functional dissection of human complex traits." Dis Model Mech **10**(2): 77-88.

Wissel, S., A. Kieser, T. Yasugi, P. Duchek, E. Roitinger, J. Gokcezade, V. Steinmann, U. Gaul, K. Mechtler, K. Forstemann, J. A. Knoblich and R. A. Neumuller (2016). "A Combination of CRISPR/Cas9 and Standardized RNAi as a Versatile Platform for the Characterization of Gene Function." G3 (Bethesda) **6**(8): 2467-2478.

Xue, Z., M. Ren, M. Wu, J. Dai, Y. S. Rong and G. Gao (2014). "Efficient gene knock-out and knock-in with transgenic Cas9 in Drosophila." G3 (Bethesda) **4**(5): 925-929.

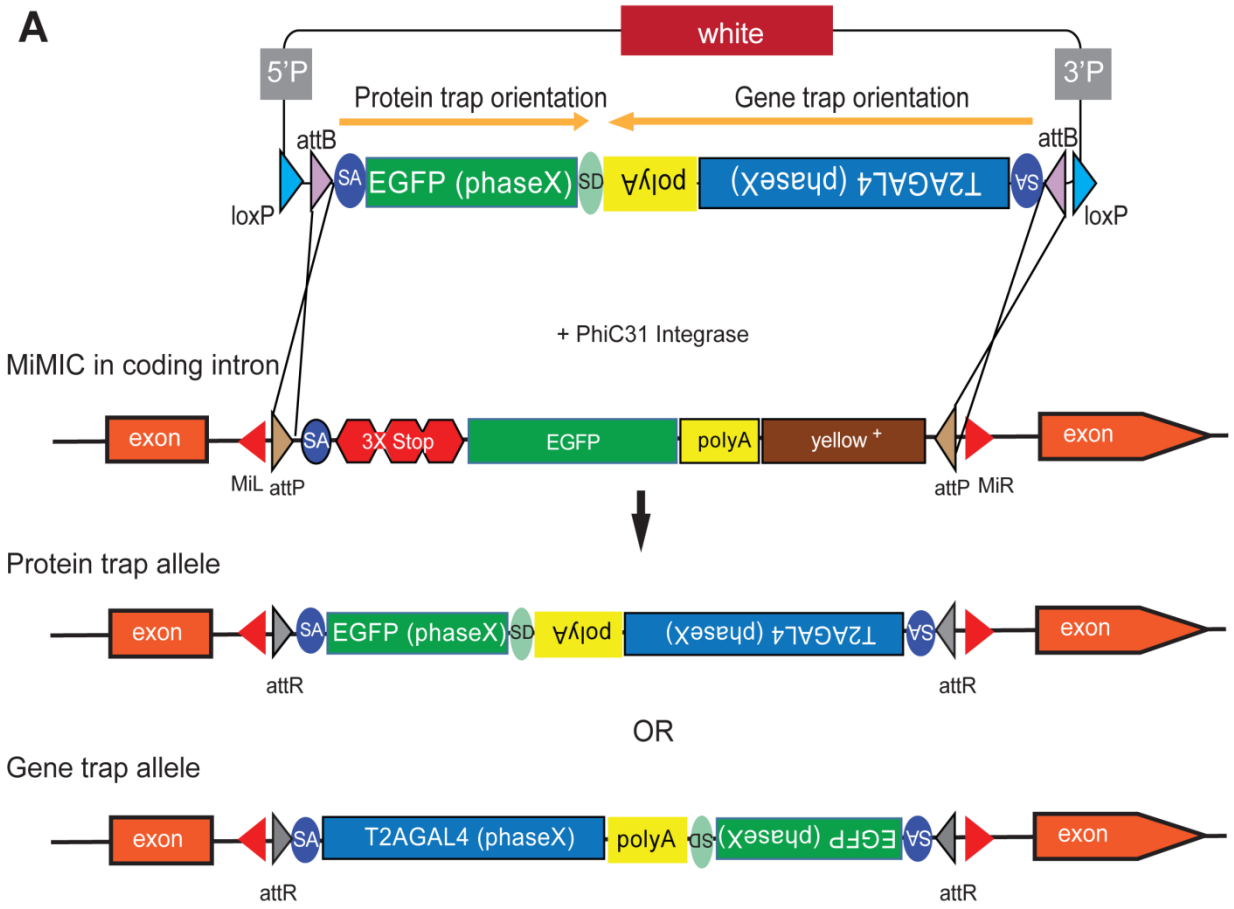
966 Yang, H. H., F. St-Pierre, X. Sun, X. Ding, M. Z. Lin and T. R. Clandinin (2016). "Subcellular Imaging of
967 Voltage and Calcium Signals Reveals Neural Processing In Vivo." Cell **166**(1): 245-257.
968 Yoon, W. H., H. Sandoval, S. Nagarkar-Jaiswal, M. Jaiswal, S. Yamamoto, N. A. Haelterman, N. Putluri, V.
969 Putluri, A. Sreekumar, T. Tos, A. Aksoy, T. Donti, B. H. Graham, M. Ohno, E. Nishi, J. Hunter, D. M.
970 Muzny, J. Carmichael, J. Shen, V. A. Arboleda, S. F. Nelson, M. F. Wangler, E. Karaca, J. R. Lupski and
971 H. J. Bellen (2017). "Loss of Nardilysin, a Mitochondrial Co-chaperone for alpha-Ketoglutarate
972 Dehydrogenase, Promotes mTORC1 Activation and Neurodegeneration." Neuron **93**(1): 115-131.
973 Zhai, R. G., Y. Cao, P. R. Hiesinger, Y. Zhou, S. Q. Mehta, K. L. Schulze, P. Verstreken and H. J. Bellen
974 (2006). "Drosophila NMNAT maintains neural integrity independent of its NAD synthesis activity."
975 PLoS Biol **4**(12): e416.
976 Zhai, R. G., F. Zhang, P. R. Hiesinger, Y. Cao, C. M. Haueter and H. J. Bellen (2008). "NAD synthase
977 NMNAT acts as a chaperone to protect against neurodegeneration." Nature **452**(7189): 887-891.
978

979

980

981 # Figures

Li-Kroeger, Kanca et al. Figure 1. Double header optimizes RMCE outcome for MiMICs.

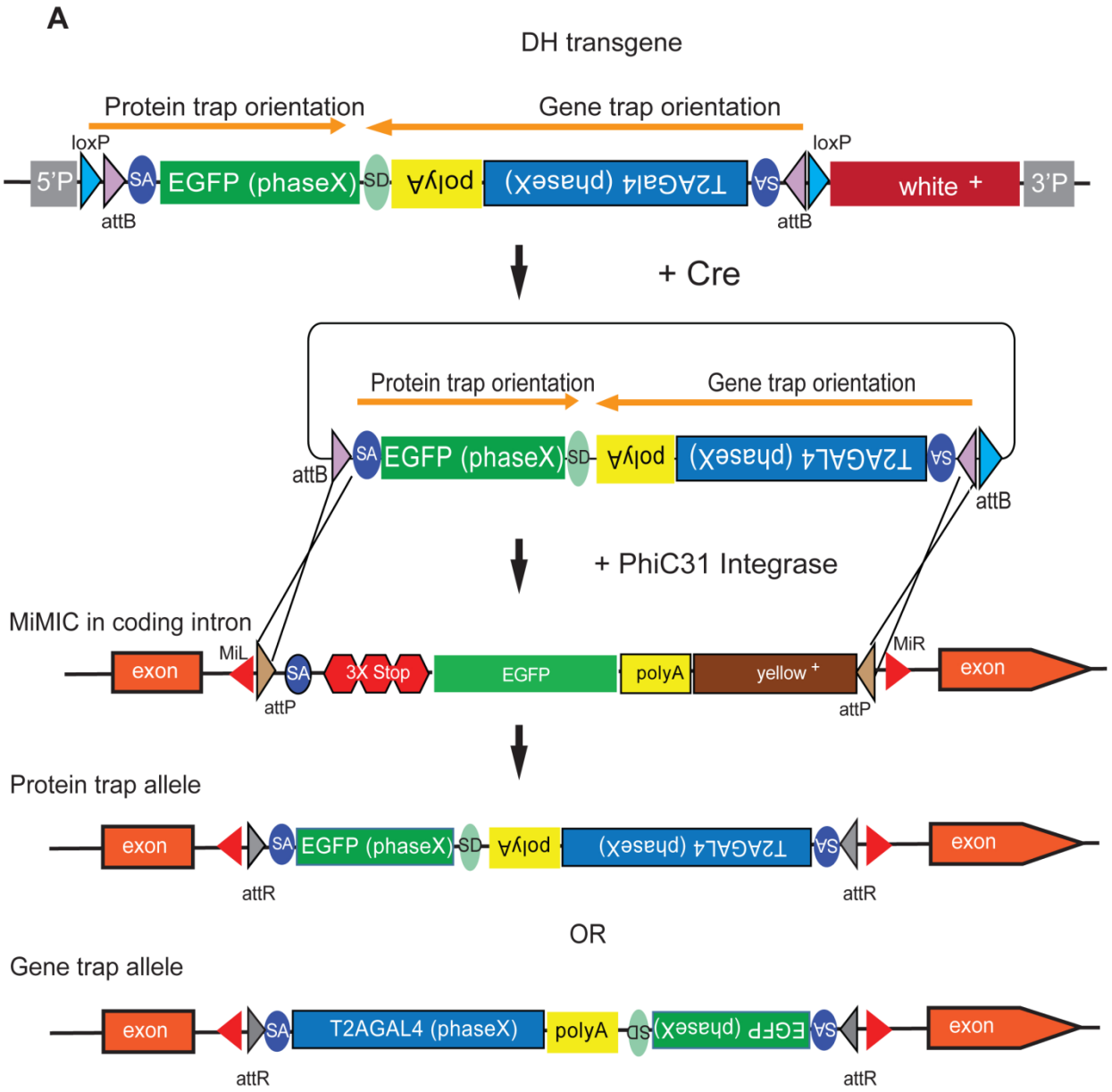


B

	Total	Both orientations	GFP only	T2A-GAL4 only
Injected MiMICS	30	6	8	2

982

Li-Kroeger, Kanca et al. Figure 2. Double header integration through crosses facilitates RMCE

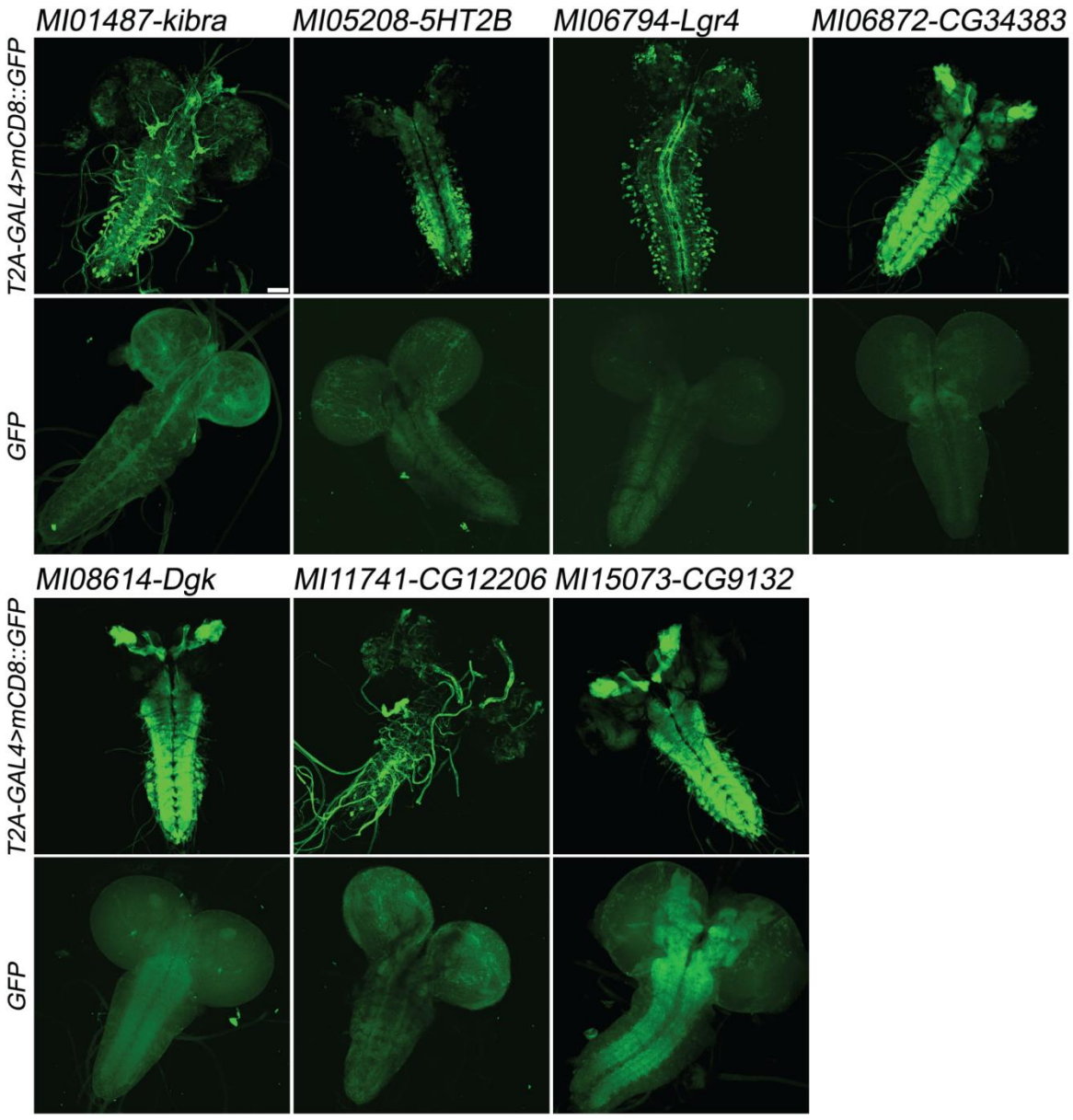


B

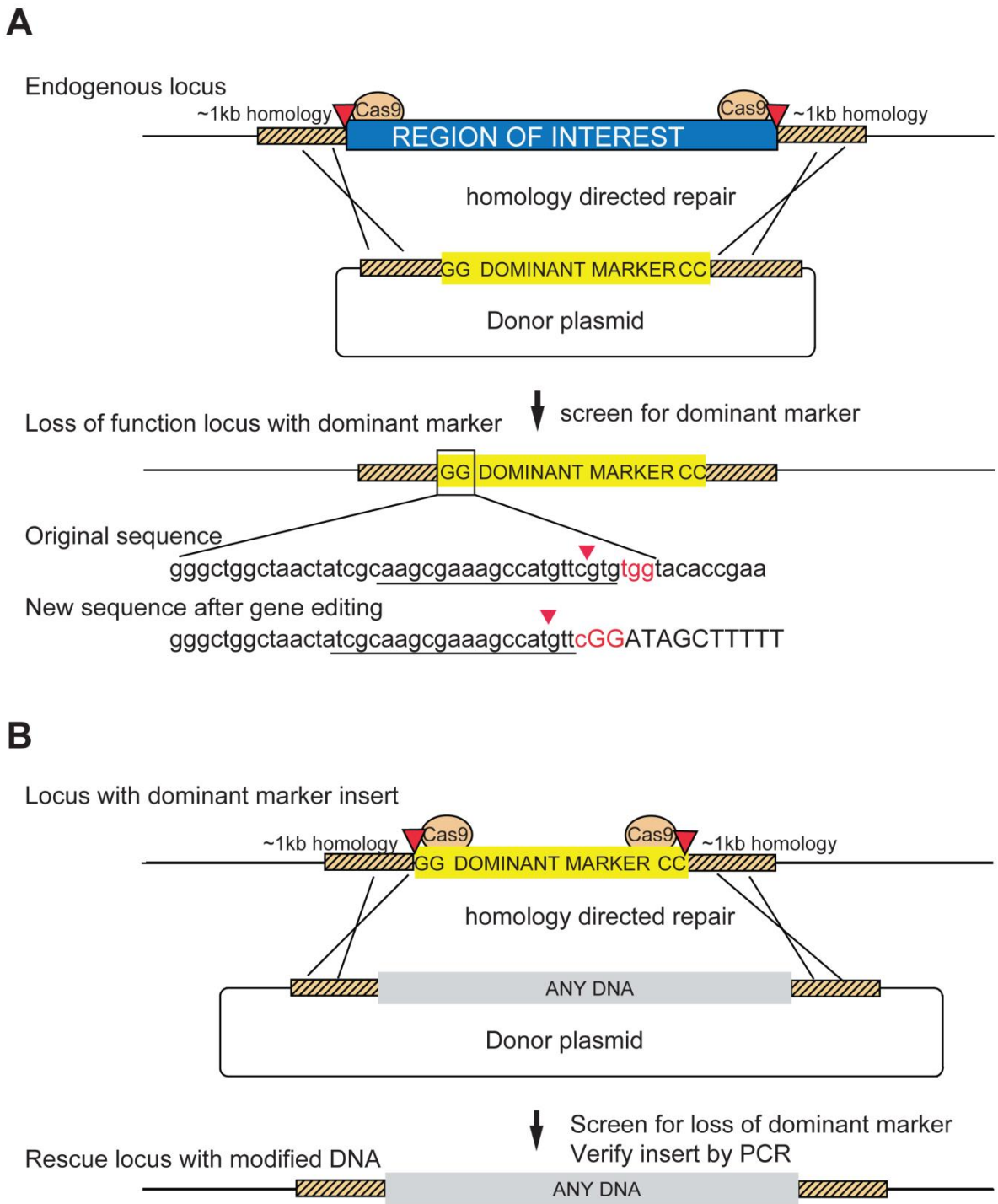
	Total	Both orientations	GFP only	T2A-GAL4 only
Crossed MiMICS	13	6	3	3

983

Li-Kroeger, Kanca et al. Figure 3. Expression patterns obtained by Double Header insertions in MiMICs in 3rd instar larval CNS.



Li-Kroeger, Kanca et al. Figure 4. Outline of two step gene editing strategy



987

988

989

Table 1. Summary statistics for cassette knock-in experiments

Construct	genotype injected :	No. independent positive lines obtained	Lethality	Rescue of lethality/failure to complement
y^{wing2+} Δ*Nmnat*	y^1 *M{nos-Cas9.P}ZH-2A w**	6	lethal	Genomic Fragment (Zhai et al 2006)/NmnatΔ4790-1
y^{wing2+} Δ*Stub1*	y^1 *M{nos-Cas9.P}ZH-2A w**	4	Viable/Fertile*	ND
y^{wing2+} Δ*Ubqn*	*yw iso#6; +/+; attP2(y-){nos-Cas9}*	2	lethal	Fails to complement Ubqn1
y^{wing2+} Δ*ltp-r83A*	y^1 *M{nos-Cas9.P}ZH-2A w**	1	lethal	ND
y^{wing2+} Δ*CG18769*	y^1 *M{nos-Cas9.P}ZH-2A w**	2	lethal	ND
y^{wing2+} Δ*CG13390*	y^1 *M{nos-Cas9.P}ZH-2A w**	2	lethal	Rescued by Genomic Fragment (this study)
y^{wing2+} Δ*Med27*	*yw iso#6; +/+; attP2(y-){nos-Cas9}*	1	lethal	Rescued by Genomic Fragment (this study)
y^{wing2+} Δ*CG11679*	*yw iso#6; +/+; attP2(y-){nos-Cas9}*	2	lethal	Rescued by genomic duplication BSC Dp(1:3) 304
Y^{wing2+} Δ*rho*	y^1 *M{nos-Cas9.P}ZH-2A w**	0	NA	NA
y^{wing2+} Δ*amx*	*yw iso#6; +/+; attP2(y-){nos-Cas9}*	2	Female sterile	Sterility Rescued by Genomic Fragment (Jakobsdottir et al., 2016)

* two of four lines

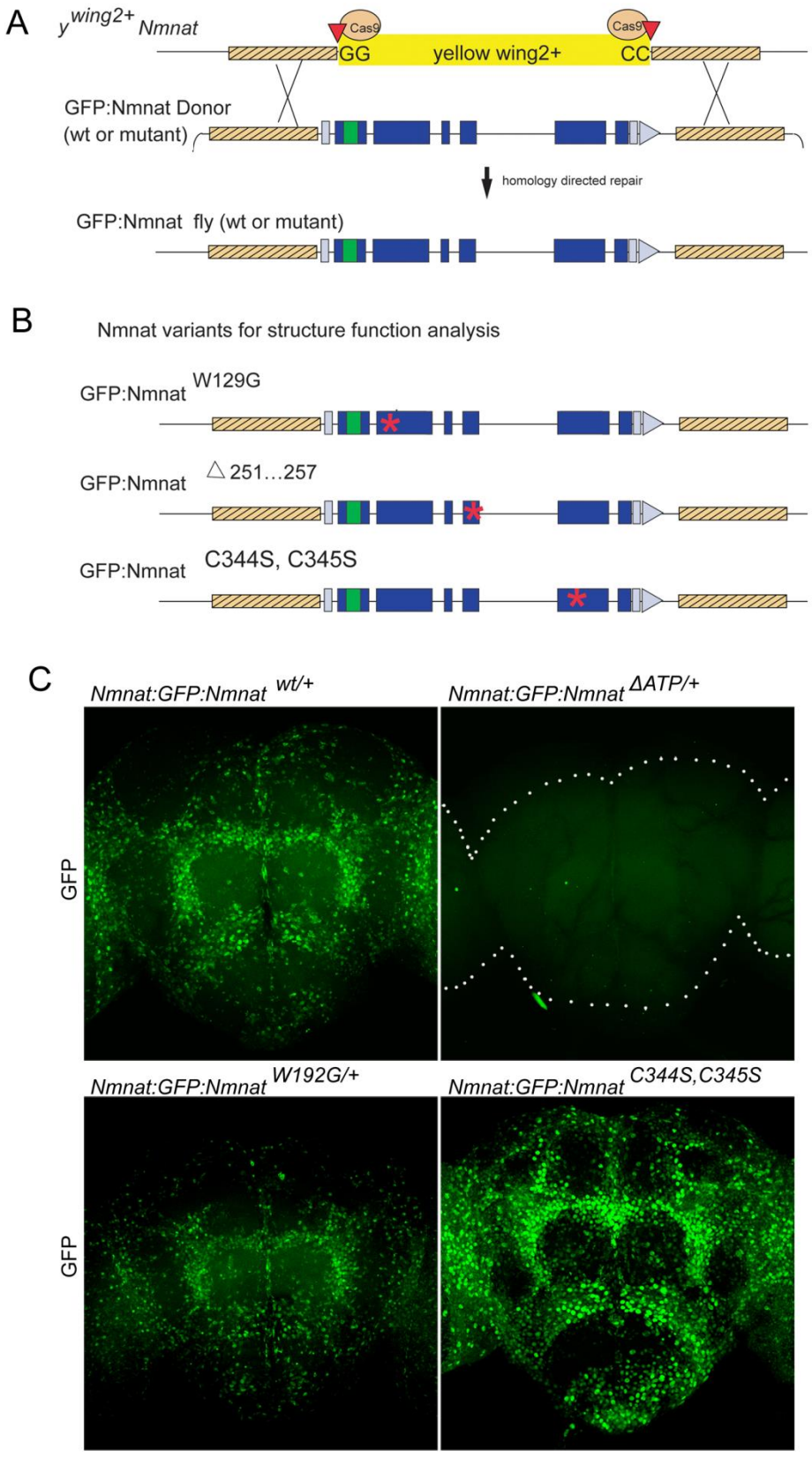
0

Table 2. Summary statistics for cassette swapping experiments

Construct	injected genotype:	No. embryos injected	No. fertile adults	no vials with y- flies	% of y- flies confirmed positive
Nmnat:GFP:Nmnat ^{wt} #1	*y1 M{nos-Cas9.P}ZH-2A w*;+;* y^{wing2+} Δ*Nmnat /TM6B*	514	7	4	6%
Nmnat:GFP:Nmnat ^{wt} #2	*y1 M{nos-Cas9.P}ZH-2A w*;+;* y^{wing2+} Δ*Nmnat /TM6B*	607	16	5	21%
Nmnat:GFP:Nmnat^{W129G} #1	*y1 M{nos-Cas9.P}ZH-2A w*;+;* y^{wing2+} Δ*Nmnat /TM6B*	653	0	-	-
Nmnat:GFP:Nmnat^{W129G} #2	*y1 M{nos-Cas9.P}ZH-2A w*;*3KB NMNAT GRC; y^{wing2+} Δ*Nmnat*	418	31	3	55%
Nmnat:GFP:Nmnat$^{\Delta 251...257}$	*y1 M{nos-Cas9.P}ZH-2A w*;*3KB NMNAT GRC; y^{wing2+} Δ*Nmnat*	496	29	11	24%
Nmnat:GFP:Nmnat$^{C344S,\ C345S}$	*y1 M{nos-Cas9.P}ZH-2A w*;*3KB NMNAT GRC; y^{wing2+} Δ*Nmnat*	386	30	12	14%
Stub1:GFP	*y1 M{nos-Cas9.P}ZH-2A w**; y^{wing2+} Δ*Stub1*	235	62	2	66%
CG11679:Flag	y^{wing2+} Δ*CG11679*/FM7 Kgal4,UAS GFP;+/+; *attP2(y-){nos-Cas9*	976	12*	3	33%
Med27:flag	*y1 M{nos-Cas9.P}ZH-2A w*;+;* y^{wing2+} Δ*Med27*	833	17	3	29%
Amx:GFP	y^{wing2+} Δ*amx;+/+; attP2(y-){nos-Cas9*	648	34	0	-

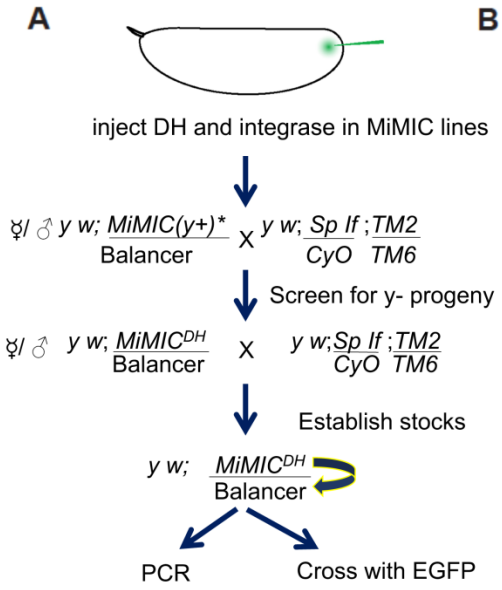
* excluding FM7 homozygotes and hemizygotes

Li-Kroeger, Kanca et al. Fig. 5:Nmnat structure affects its expression



995

Li-Kroeger, Kanca et al. Supp. Figure 1. Injection data for Double header.

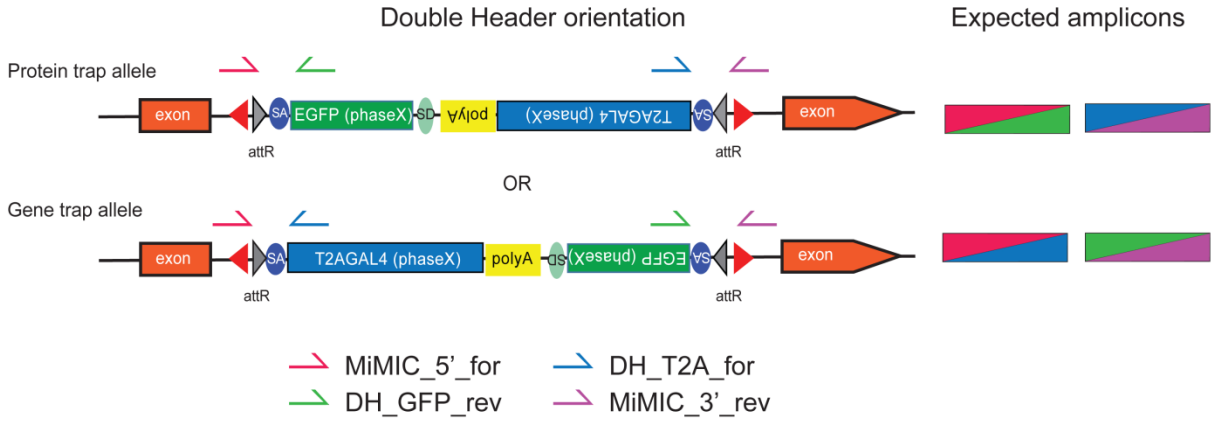


B

MiMIC line	# of survivors/ crosses	# of steriles	Transformants	GFP	T2A-GAL4	?
MI03136	105	13	10	5	5	0
MI09900	29	7	1	1	0	0
MI08614	18	1	2	1	1	0
MI01646	18	4	1	1	0	0
MI04010	24	9	3	3	0	0
MI10071	98	9	17	4	3	10
MI06794	31	0	6	1	2	3
MI13728	78	1	5	5	0	0
MI11741	32	0	5	1	2	2
MI00986	24	2	3	0	2	1
MI05741	32	1	5	0	2	3
MI09222	18	0	2	1	0	1
MI10889	36	2	1	1	0	0
MI15073	22	0	4	1	2	1
MI02926	30	1	1	1	0	0
MI15214	16	0	2	1	0	1
MI09222	28	2	2	1	0	1
MI12839	11	0	2	0	0	2

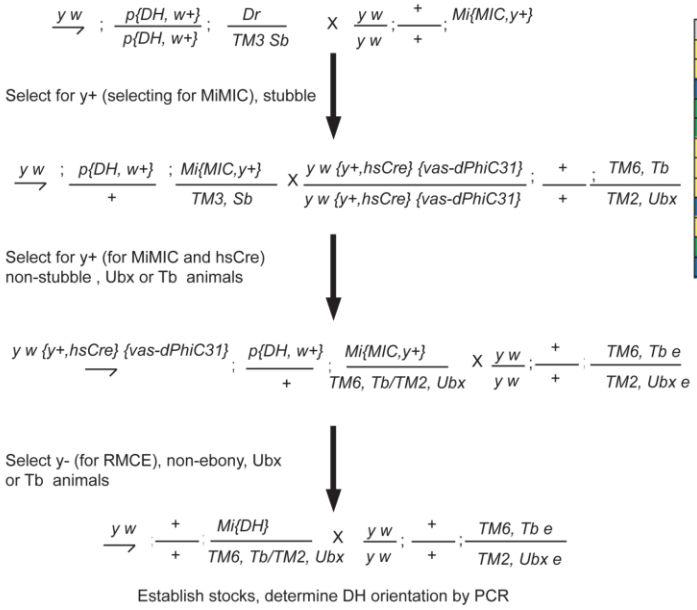
T2A-GAL4 only
GFP only
Both

Li-Kroeger, Kanca et al. Supp. Figure 2. PCR strategy to determine Double Header orientation.



Li-Kroeger, Kanca et al. Supp. Figure 3 Crossing scheme for RMCE with Double Header and data of integration.

A



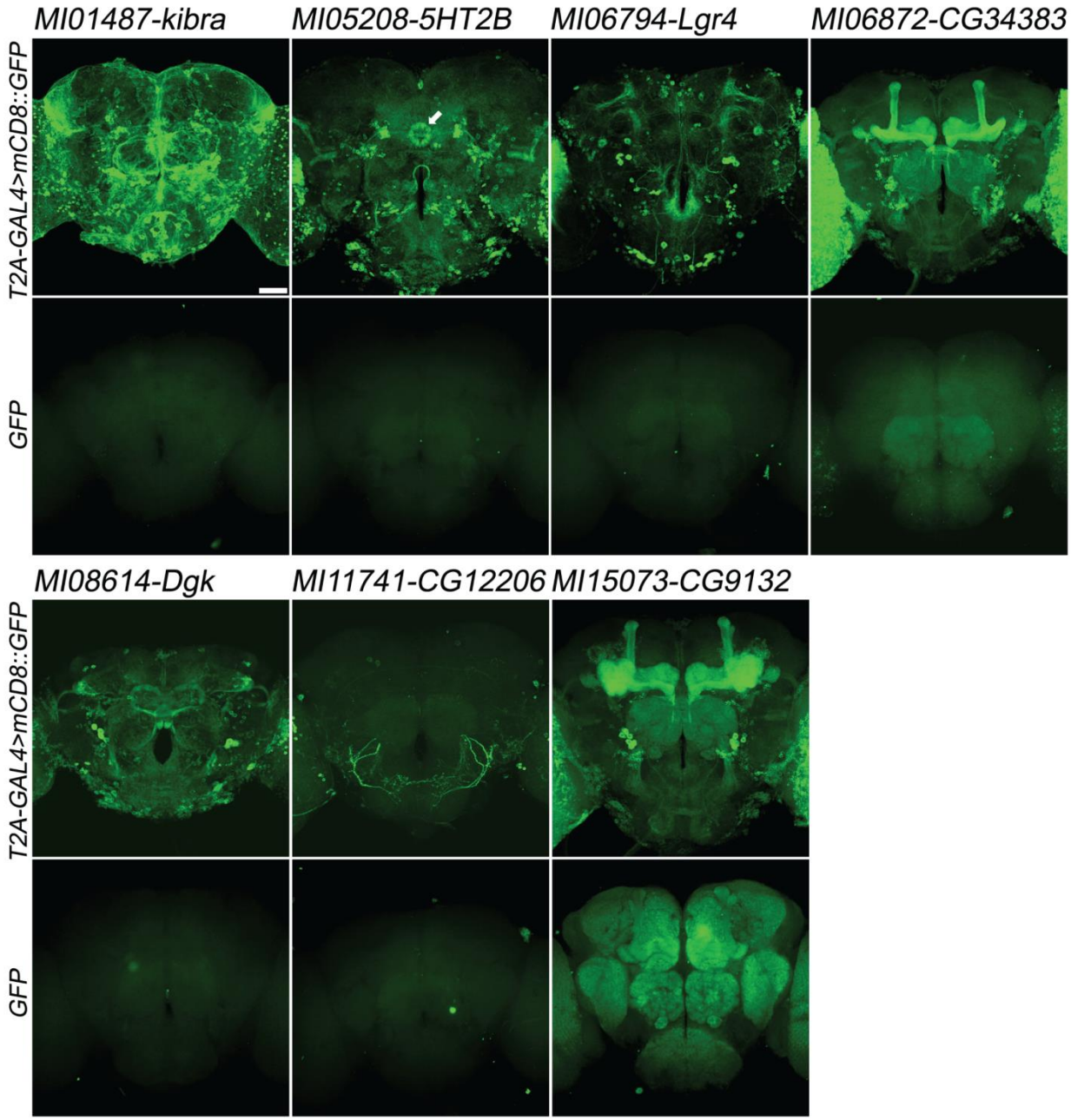
B

MiMIC line	yellow -	GFP	T2A-GAL4	?
MI06179	34	2	1	31
MI01487	3	1	2	0
MI00445	8	0	5	3
MI00494	2	2	0	0
MI12643	8	5	0	3
MI00805	12	2	5	5
MI05208	6	3	3	0
MI06872	16	2	3	11
MI8818	1	0	1	0
MI06873	3	1	1	1
MI02915	1	1	0	0
MI05871	8	0	8	0

1000

1001

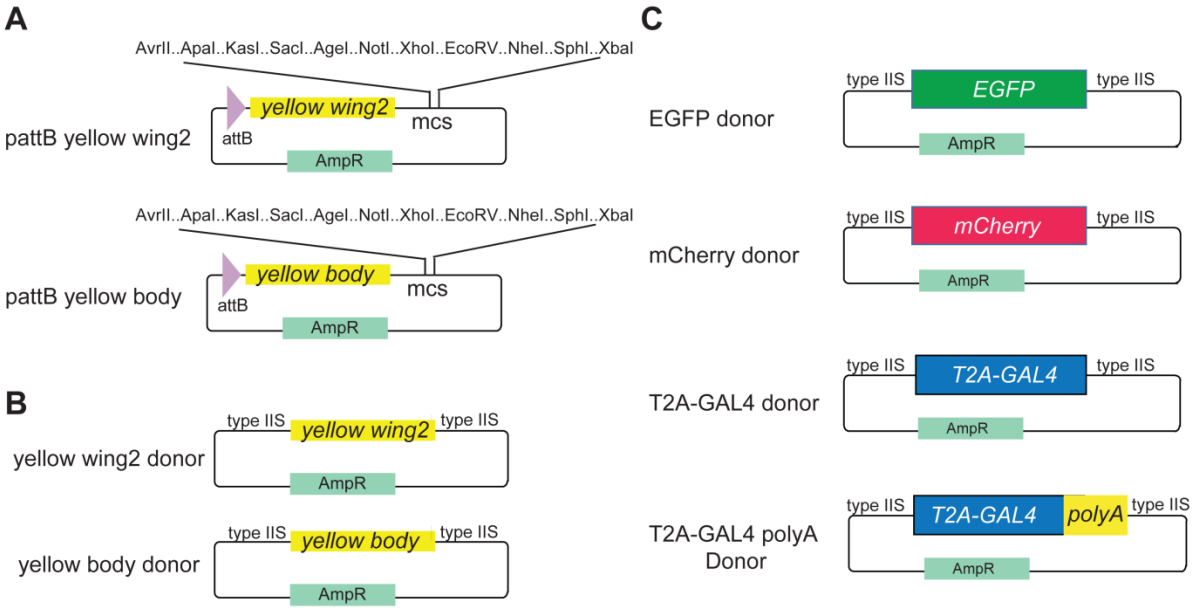
Li-Kroeger, Kanca et al. Supp. Figure 4. Expression patterns obtained by Double Header insertions in MiMICs in adult brain.



1002

1003

Li-Kroeger, Kanca et al. Supp. Figure 5. Vectors used for two step gene editing



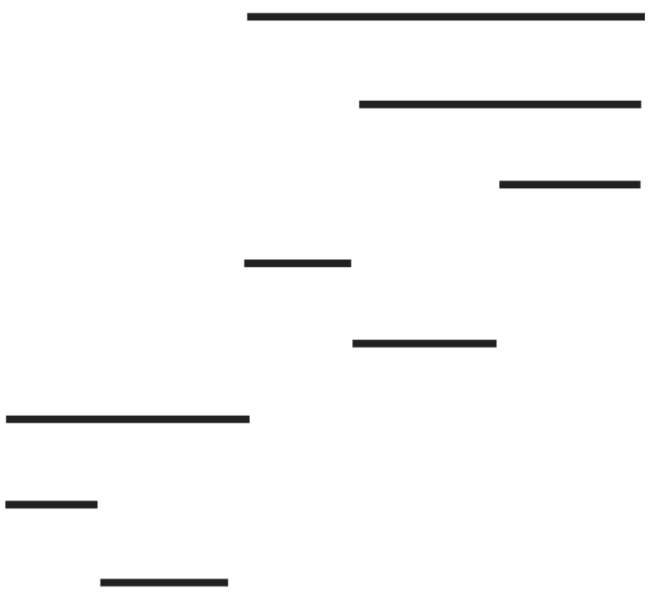
1004

1005

Li-Kroeger, Kanca et al. Supp. Fig 6: Mapping Cis-Regulatory Modules of the *yellow* gene

A

B



CRM	expression	
	body	wing
full body	+	-
body1	-	-
body2	-	-
body3	-	-
body4	-	-
full wing	-	+
wing1	-	-
wing2	-	+
min. prom.	-	-

1006

1007

1008 # Supplementary information

oligo name	sequence
DLK0048	ctagcaCTCGAGaaaaaaaccttcatataaaacgcggccgac
DLK0049	ctagcaTCTAGAcctgcaggtcaacggatcctac
DLK0022	See MCS of pattB ywing2+
DLK0023	See MCS of pattB ywing2+ (reverse complement)
DLK0054	ctgacGCGGCCGCatgattatcgcccgattaccacattgagtg
DLK0055	ctgacGCGGCCGCctctattcaatcgggacagtggaaattgac
DLK0056	atcgGTCGACtccacaaatcagatctctttggccataattg
DLK0057	atcgGTCGACgtcaatttccactgtcccgattgaatagag
DLK0225	CCGCGGTGGGAAGACttGGAGGTTCCGGTGGAA GCGGAGGTAGCGGCggatccGTGTCCAAGGG
DLK300	GAGAGTCTAGAAGAGAGGTCTCGCTTGTACAGCTCATCCATGCCCAG
DLK320	ggTGGCGTCTCtGGATgattaactagtgTTCCCGAGACGGAT
DLK322	ggTGGGAAGACttGGATgattaactagtgTTCCCcGTCTTCGAT
DLK324	ggTGGGGTCTCtGGATgattaactagtgTTCCCGAGACCGAT
DLK326	GAGATCGGTCTCGGGAACCTGCAGGTCAACGGATCCTAC
DLK327	ggTGGGGTCTCtGGATctctattcaatcgggacagtggaaattgac
DLK338	ATCGAAGACGGGGAACACTAGTTAATCATCCAAGTCTTCCCACCGC
DLK339	ATCGGTCTCGGGAACACTAGTTAATCATCCAGAGACCCCACCGC
DLK340	ATCCGTCTCGGGAACACTAGTTAATCATCCAGAGACGCCACCGC
DLK461	CGGTCTCtTCCGTCTAGATTCTcGAGACCGAT
DLK462	ATCGGTCTCgAGAATCTAGACGGAaGAGACCGAGCT
DLK463	CCGTCTCtTCCGTCTAGATTCTcGAGACGGAT
DLK464	ATCCGTCTCgAGAATCTAGACGGAaGAGACGGAGCT
DLK465	CGAAGACttTCCGTCTAGATTCTccGTCTTCGAT
DLK466	ATCGAAGACggAGAATCTAGACGGAaaGTCTTCGAGCT
DLK554	ggcgGATCCGGAGGTAGCGGTGGAAGCGGAGGTTCTcGAGACCtctct
DLK555	ctagagagaGGTCTCgAGAACCTCCGCTTCCACCGCTACCTCCGGATC
DLK823	gcgcGGATCCggcaatgattcccgcttgc
DLK824	acacacGCGGCCGCcagctggttcgatagctttcg
DLK966	ctctctctcgaggcgatttaacagtccgtcaag
DLK969	acacaccctagggttggatcacaagttttcattgttgc
MiMIC_5'_for	gctaccttaatctcaagaagagcaaaacaaaagc
MiMIC_3'_rev	cgcggcgtaatgtgatttactatcatac
GFP_DH_for	gcaccacgccggtgaacag
T2A_GAL4_DH_rev	ctgtagcggcactcccagttgttc

1009
1010
1011
1012

1013 sgRNAs used in study

Nmnat ywing2+ KI upstream	AAGCGAAAGCAATGTTCGTG
Nmnat ywing2+ KI downstream	CACGTTCGTTACCAAGTTGA
Stub1 ywing2+ KI upstream	ATTCTGTCGGCTCTTAAATT
Stub1 ywing2+ KI downstream	CGATATCGATAGAGAGTTAG
Ubqn ywing2+ KI upstream	CCGCTATCGATTCATCGAT
Ubqn ywing2+ KI downstream	ATAGTAGAGGGTAAAAGACG
Itp-r83A ywing2+ KI upstream	AGGTCCAACTAATTCACTAA
Itp-r83A ywing2+ KI downstream	GGTGGAGTTCTACGCTCGAT
CG18769 ywing2+ KI upstream	TTTGGGGAAGTGAAGGGAGT
CG18769 ywing2+ KI downstream	GCACACACAATATTTGCCTT
CG13390 ywing2+ KI upstream	AGCCCCAGGAATCCAGTTAG
CG13390 ywing2+ KI downstream	TTGCTGAAGTCTTGTGACCA
CG11679 ywing2+ KI upstream	AATGGTATAGATCTACGTTG
CG11679 ywing2+ KI downstream	TGGCTATATAATATAGTAGC
Med27 ywing2+ KI upstream	TGATGAGACGAAATTACTGT
Med27 ywing2+ KI downstream	TATGTACTTTGAATGCATGC
rho ywing2+ KI upstream	TCCAAATTCCGAGCGGTCTc
rho ywing2+ KI downstream	GTTTTCTGCGTCTGACTCGC
amx ywing2+ KI upstream	GAAGATCTTGCTATTCCTAA
amx ywing2+ KI downstream	TCCATTTAAGTTGTGACCAT
Nmnat ywing2+ replacement upstream	TCGCAAGCGAAAGCAATGTT
Nmnat ywing2+ replacement downstream	CAAAAGCATGGGCAGCCGTC
Stub1 ywing2+ replacement upstream	ATTGATTCTGTCGGCTCTTA
Stub1 ywing2+ replacement downstream	ATTGTTTTCACACTACCGCT
CG11679 ywing2+ replacement upstream	cacagcgttggccatccttt
CG11679 ywing2+ replacement downstream	AAAGGCAATCGGAGGCCGGC
Med27 ywing2+ replacement upstream	CGATTGAATAGAGATCCtgt
Med27 ywing2+ replacement downstream	GTTGACCTGCAGGTTCCtgc
amx ywing2+ replacement upstream	TTGAGAAGATCTTGCTATTC
amx ywing2+ replacement downstream	CTACGACTAAAGCGGCCAAT

Sequence of Double Header constructs (Reverse orientated elements are marked with the same color but with underline)

GGCCAGACCCACGTAGTCCAGCGGCAGATCGGCGGCGGAGAAGTTAAGCGTCTCCAGGA
TGACCTTGCCCGAACTGGGGCACGTGGTGTTCGACGATGTGCAGCTAATTTCGCCCGGCT
CCACGTCCGCCCATTGGTTAATCAGCAGACCCTCGTTGGCGTAACGGAACCATGAGAGGT
ACGACAACCATTTGAGGTATACTGGCACCGAGCCCGAGTTCAAGAAGAAGGCGTTTTTCCA
TAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA
CCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCC
TGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGC
GCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTG
GGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGT
CTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGG
ATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTAC
GGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAA
AAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGT
TTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTA
CGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATC
AAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTAT
ATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCG
ATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACG
GGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGG
CTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTG
CAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTC
GCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCG
TCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCC
CCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTT
GGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCA
TCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTAT
GCGGCGACCGAGTTGCTCTTGCCCGGCGTCAACACGGGATAATACCGCGCCACATAGCA
GAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTA
CCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTT
TTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGG
AATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCA
TTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAA
TAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTAT
CATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTCTCGCGCGTTTCGGT
GATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAA
GCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTC
GGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGT
GTGAAATACCGCACCGAATCGCGCGGAACTAACGACAGTCGCTCCAAGGTCGTCGAACAA
AAGGTGAATGTGTTGCGGAGAGCGGGTGGGAGACAGCGAAAGAGCAACTACGAAACGTG

1058 GTGTGGTGGAGGTGAATTATGAAGAGGGCGCGCGATTTGAAAAGTATGTATATAAAAAATA
1059 TATCCCGGTGTTTTATGTAGCGATAAACGAGTTTTTGATGTAAGGTATGCAGGTGTGTAAGT
1060 CTTTTGGTTAGAAGACAAATCCAAAGTCTACTTGTGGGGATGTTCGAAGGGGAAATACTTG
1061 TATTCTATAGGTCATATCTTGTTTTTATTGGCACAAATATAATTACATTAGCTTTTTGAGGGG
1062 GCAATAAACAGTAAACACGATGGTAATAATGGTAAAAAAAAAAACAAGCAGTTATTTCGGAT
1063 ATATGTCGGCTACTCCTTGCGTCGGGCCCGAAGTCTTAGAGCCAGATATGCGAGCACCCG
1064 GAAGCTCACGATGAGAATGGCCAGACCATGATGAAATAACATAAGGTGGTCCCGTCGGCA
1065 AGAGACATCCACTTAACGTATGCTTGCAATAAGTGCGAGTGAAAGGAATAGTATTCTGAGT
1066 GTCGTATTGAGTCTGAGTGAGACAGCGATATGATTGTTGATTAACCCTTAGCATGTCCGTG
1067 GGGTTTGAATTAACTCATAATATTAATTAGACGAAATTATTTTTAAAGTTTTATTTTTAATAAT
1068 TTGCGAGTACGCAAAGCTTGGCTGCATCCAACGCGCAATTAACCCTCACTAAAGGGAACAA
1069 AAGCTGGGTACCGGGCCCCCCCTCGAGATAACTTCGTATAATGTATGCTATACGAAGTTAT
1070 CAATTGGCTTATTGGGATTTCAATTTAAGATGCTGCTAAATAAAGTTAGTCACTTACGGACT
1071 GGGGAGGCACCCCGTCATCACTCGCCAGAAATGCAATTGATGTAGGTCACGGTCTCGAAG
1072 CCGCGGTGCGGGTGCCAGGGCGTGCCCTTGGGCTCCCCGGGCGCGTACTCCACCTCAC
1073 CCATCTGGTCCATCATGATGGTCGACTCTAGAAGTCGATCCAACATGGCGACTTGTCCCAT
1074 CCCCGGCATGTTTAAATATACTAATTATTCTTGAACTAATTTTAATCAACCGATTTATCTCTC
1075 TTCCGCAGGTGGGAGGTTCCGGTGGAAGCGGAGGTAGCGGCGGATCCGAGGGCCGCGG
1076 CAGCCTGCTGACCTGCGGCGATGTGGAGGAGAACCCCGGGCCCATGAAGCTGCTGAGCA
1077 GCATCGAGCAGGCCTGCGATATCTGCCGCCTGAAGAAGCTGAAGTGCAGCAAGGAGAAG
1078 CCCAAGTGCGCCAAGTGCCTGAAGAACAACTGGGAGTGCCGCTACAGCCCCAAGACCAA
1079 GCGCAGCCCCCTGACCCGCGCCCACCTGACCGAGGTGGAGAGCCGCCTGGAGCGCCTG
1080 GAGCAGCTGTTCCTGCTGATCTTCCCCCGCGAGGATCTGGATATGATCCTGAAGATGGATA
1081 GCCTGCAGGATATCAAGGCCCTGCTGACCGGCCTGTTCGTGCAGGATAACGTGAACAAGG
1082 ATGCCGTGACCGATCGCCTGGCCAGCGTGGAGACCGATATGCCCCTGACCCTGCGCCAG
1083 CACCGCATCAGCGCCACCAGCAGCAGCGAGGAGAGCAGCAACAAGGGCCAGCGCCAGCT
1084 GACCGTGAGCATCGATAGCGCCGCCCACCACGATAACAGCACCATCCCCCTGGATTTCAT
1085 GCCCCGCGATGCCCTGCACGGCTTCGATTGGAGCGAGGAGGATGATATGAGCGATGGCC
1086 TGCCCTTCCTGAAGACCGATCCCAACAACAACGGCTTCTTCGGCGATGGCAGCCTGCTGT
1087 GCATCCTGCGCAGCATCGGCTTCAAGCCCGAGAACTACACCAACAGCAACGTGAACCGCC
1088 TGCCCACCATGATCACCGATCGCTACACCCTGGCCAGCCGCAGCACCACCAGCCGCCTG
1089 CTGCAGAGCTACCTGAACAACTTCCACCCCTACTGCCCCATCGTGCACAGCCCCACCCTG
1090 ATGATGCTGTACAACAACCAGATCGAGATCGCCAGCAAGGATCAGTGGCAGATCCTGTTCA
1091 ACTGCATCCTGGCCATCGGCGCCTGGTGCATCGAGGGCGAGAGCACCGATATCGATGTGT
1092 TCTACTACCAGAACGCCAAGAGCCACCTGACCAGCAAGGTGTTCGAGAGCGGCAGCATCA
1093 TCCTGGTGACCGCCCTGCACCTGCTGAGCCGCTACACCCAGTGGCGCCAGAAGACCAAC
1094 ACCAGCTACAACTTCCACAGCTTCAGCATCCGCATGGCCATCAGCCTGGGCCTGAACCGC
1095 GATCTGCCCAGCAGCTTCAGCGATAGCAGCATCCTGGAGCAGCGCCGCCGCATCTGGTG
1096 GAGCGTGTACAGCTGGGAGATCCAGCTGAGCCTGCTGTACGGCCGCAGCATCCAGCTGA
1097 GCCAGAACACCATCAGCTTCCCCAGCAGCGTGGATGATGTGCAGCGCACCACCACCGGC
1098 CCCACCATCTACCACGGCATCATCGAGACCGCCCGCCTGCTGCAGGTGTTCACCAAGATC
1099 TACGAGCTGGATAAGACCGTGACCGCCGAGAAGAGCCCCATCTGCGCCAAGAAGTGCCT
1100 GATGATCTGCAACGAGATCGAGGAGGTGAGCCGCCAGGCCCCCAAGTTCCTGCAGATGG
1101 ATATCAGCACCACCGCCCTGACCAACCTGCTGAAGGAGCACCCCTGGCTGAGCTTCACCC

1102 GCTTCGAGCTGAAGTGGAAGCAGCTGAGCCTGATCATCTACGTGCTGCGCGATTTCTTCA
1103 CCAACTTCACCCAGAAGAAGAGCCAGCTGGAGCAGGATCAGAACGATCACCAGAGCTACG
1104 AGGTGAAGCGCTGCAGCATCATGCTGAGCGATGCCGCCCAGCGCACCGTGATGAGCGTG
1105 AGCAGCTACATGGATAACCACAACGTGACCCCCTACTTCGCCTGGAACTGCAGCTACTACC
1106 TGTTCAACGCCGTGCTGGTGCCCATCAAGACCCTGCTGAGCAACAGCAAGAGCAACGCCG
1107 AGAACAACGAGACCGCCCAGCTGCTGCAGCAGATCAACACCGTGCTGATGCTGCTGAAGA
1108 AGCTGGCCACCTTCAAGATCCAGACCTGCGAGAAGTACATCCAGGTGCTGGAGGAGGTGT
1109 GCGCCCCCTTCCTGCTGAGCCAGTGCGCCATCCCCCTGCCCCACATCAGCTACAACAACA
1110 GCAACGGCAGCGCCATCAAGAACATCGTGGGCAGCGCCACCATCGCCCAGTACCCCACC
1111 CTGCCCGAGGAGAACGTGAACAACATCAGCGTGAAGTACGTGAGCCCCGGCAGCGTGGG
1112 ACCCAGCCCCGTGCCCCTGAAGAGCGGCGCCAGCTTCAGCGATCTGGTGAAGCTGCTGA
1113 GCAACCGCCCCCCCAGCCGCAACAGCCCCGTGACCATCCCCCGCAGCACCCCCAGCCAC
1114 CGCAGCGTGACCCCCTTCCTGGGCCAGCAGCAGCAGCTGCAGAGCCTGGTGCCCCTGAC
1115 CCCCAGCGCCCTGTTCGGCGGCGCCAACTTCAACCAGAGCGGCAACATCGCCGATAGCA
1116 GCCTGAGCTTCACCTTCACCAACAGCAGCAACGGCCCCAACCTGATCACCACCCAGACCA
1117 ACAGCCAGGCCCTGAGCCAGCCCATCGCCAGCAGCAACGTGCACGATAACTTCATGAACA
1118 ACGAGATCACCGCCAGCAAGATCGATGATGGCAACAACAGCAAGCCCCTGAGCCCCGGCT
1119 GGACCGATCAGACCGCCTACAACGCCTTCGGCATCACCACCGGCATGTTCAACACCACCA
1120 CCATGGATGATGTGTACAACTACCTGTTCGATGATGAGGATACCCCCCCCAACCCCAAGAA
1121 GGAGTAAAATGAATCGTAGATACTGAAAAACCCCGCAAGTTCACTTCAACTGTGCATCGTG
1122 CACCATCTCAATTTCTTTCATTTATACATCGTTTTGCCTTCTTTTATGTAACTATACTCCTCTA
1123 AGTTTCAATCTTGGCCATGTAACCTCTGATCTATAGAATTTTTTAAATGACTAGAATTAATGC
1124 CCATCTTTTTTTTGGACCTAAATTCTTCATGAAAATATATTACGAGGGCTTATTCAGAAGCTT
1125 ATCGATACCGTCGACTAAAGCCAAATAGAAATTATTCAGTTCTGGCTTAAGTTTTTAAAAGT
1126 GATATTATTTATTTGGTTGTAACCAACCAAAAGAATGTAAATAACTAATACATAATTATGTTA
1127 GTTTTAAGTTAGCAACAAATTGATTTTAGCTATATTAGCTACTTGGTTAATAAATAGAATATAT
1128 TTATTTAAAGATAATTCGTTTTTATTGTCAGGGAGTGAGTTTGCTTAAAAACTCGTTTGGATC
1129 CGACGTCTAATCCTAGGAGAAGTTCAAATGGGCTTTCATATTTACTTTTTTCTCTGAGTGGG
1130 TAAGGTCATTGTGGGGTCTTAAGGGCTAAAGGGTAGTGATGATGAAGGCATTTGATGCCAT
1131 TGTTCAATAACTTACCTCGCCGGAACCTCCGCTTCCACCGCTACCTCCGGATCCGCCCTTG
1132 TCGTCATCGTCCTTGTAATCAATATCGTGGTCCTTGTAGTCGCCGTCGTGATCCTTGTAATC
1133 GCTCTGGAAGTACAGGTTTTCCTTCTCGAACTGGGGGTGGGACCAGTTGCCGCTTCCCGG
1134 TTCCATGCAGCATCCTGGGCAGCAGTTCAGGAAGCTGCCGCTGCTCTCCTTGTACAGCTC
1135 ATCCATGCCCAGGGTGATGCCGGCGGCGGTCACGAACTCCAGCAGCACCATGTGATCGC
1136 GCTTCTCGTTGGGGTCCTTGGACAGCACGCTCTGGGTGCTCAGGTAGTGGTTATCGGGCA
1137 GCAGCACTGGGCCGTCGCCGATGGGGGTGTTCTGCTGGTAGTGATCGGCCAGCTGCACG
1138 GAGCCATCCTCCACATTGTGGCGGATCTTGAAGTTGGCCTTGATGCCGTTCTTCTGCTTAT
1139 CGGCGGTGATGTACACGTTGTGGCTGTTGAAGTTGTACTCCAGCTTGTGGCCCAGGATGT
1140 TGCCATCCTCCTTGAAATCGATGCCCTTCAGCTCGATGCGGTTCACCAGGGTATCGCCCTC
1141 GAACTTCACCTCGGCGCGGGTCTTGTAGGTGCCGTCATCCTTGAAGCTGATGGTGCGCTC
1142 CTGCACGTAGCCCTCGGGCATGGCGCTCTTGAAGAAATCGTGCTGCTTCATGTGATCGGG
1143 GTAGCGGCTGAAGCACTGCACGCCGTAGGTCAGGGTGGTCACCAGGGTGGGCCAGGGCA
1144 CGGGCAGCTTGCCGGTGGTGCAGATGAACTTCAGGGTCAGCTTGCCGTTGGTGGCGTCG
1145 CCCTCGCCCTCGCCGCGCACGCTGAACTTGTGGCCGTTCACGTCGCCATCCAGCTCCACC

1146 AGGATGGGCACCACGCCGGTGAACAGCTCCTCGCCCTTGGACACGGATCCGCCGCTACC
1147 TCCGCTTCCACCGGAACCTCCCACCTGCGGAAGAGAGATAAATCGGTTGATTAAAATTAGT
1148 TCAAGAATAATTAGTATATTTAAACATGCCGGGGATGGGACAAGTCGCCATGTTGGATCGA
1149 CTTCCGTACGCTTGTCGACCATCATGATGGACCAGATGGGTGAGGTGGAGTACGCGCCCG
1150 GGGAGCCCAAGGGCACGCCCTGGCACCCGCACCGCGGCTTCGAGACCGTGACCTACATC
1151 ATACAATAAATATTCACACCCCTAAAAAACGAAGCTGCATTTGCAAATTGCATACAATTATG
1152 CCAGAGTAAAAATTATAATTGTTATTATTTAAATTTGAATAACTTCGTATAATGTATGCTATAC
1153 GAAGTTATAGATCCACTAGAAGGCCTAATTCTAGTATGTATGTAAGTTAATAAAACCCATTTT
1154 TGCGGAAAGTAGATAAAAAAAACATTTTTTTTTTTTACTGCACTGGATATCATTGAACTTATC
1155 TGATCAGTTTTAAATTTACTTCGATCCAAGGGTATTTGATGTACCAGGTTCTTTCGATTACCT
1156 CTCACTCAAAATGACATTCCACTCAAAGTCAGCGCTGTTTGCCTCCTTCTCTGTCCACAGAA
1157 ATATCGCCGTCTCTTTCGCCGCTGCGTCCGCTATCTCTTTCGCCACCGTTTGTAGCGTTAC
1158 CTAGCGTCAATGTCCGCCTTCAGTTGCACTTTGTCAGCGGTTTCGTGACGAAGCTCCAAGC
1159 GGTTTACGCCATCAATTAAACACAAAGTGCTGTGCCAAAACTCCTCTCGCTTCTTATTTTTG
1160 TTTGTTTTTTGAGTGATTGGGGTGGTGATTGGTTTTGGGTGGGTAAGCAGGGGAAAGTGTG
1161 AAAAATCCCGGCAATGGGCCAAGAGGATCAGGAGCTATTAATTCGCGGAGGCAGCAAACA
1162 CCCATCTGCCGAGCATCTGAACAATGTGAGTAGTACATGTGCATACATCTTAAGTTCACTTG
1163 ATCTATAGGAACTGCGATTGCAACATCAAATTGTCTGCGGCGTGAGAACTGCGACCCACAA
1164 AAATCCCAAACCGCAATCGCACAAACAAATAGTGACACGAAACAGATTATTCTGGTAGCTG
1165 TGCTCGCTATATAAGACAATTTTTAAGATCATATCATGATCAAGACATCTAAAGGCATTCATT
1166 TTCGACTACATTCTTTTTTACAAAAAATATAACAACCAGATATTTTAAGCTGATCCTAGATGC
1167 ACAAAAAATAAATAAAAGTATAAACCTACTTCGTAGGATACTTCGTTTTGTTCGGGGTTAGA
1168 TGAGCATAACGCTTGTAGTTGATATTTGAGATCCCCTATCATTGCAGGGTGACAGCGGAGC
1169 GGCTTCGCAGAGCTGCATTAACCAGGGCTTCGGGCAGGCCAAAAACTACGGCACGCTCCT
1170 GCCACCCAGTCCGCCGGAGGACTCCGGTTCAGGGAGCGGCCAACTAGCCGAGAACCTCA
1171 CCTATGCCTGGCACAATATGGACATCTTTGGGGCGGTCAATCAGCCGGGCTCCGGATGGC
1172 GGCAGCTGGTCAACCGGACACGCGGACTATTCTGCAACGAGCGACACATACCGGCGCCC
1173 AGGAAACATTTGCTCAAGAACGGTGAGTTTCTATTCGCAGTCGGCTGATCTGTGTGAAATC
1174 TTAATAAAGGGTCCAATTACCAATTTGAAACTCAGTTTGCGGCGTGGCCTATCCGGGCGAA
1175 CTTTTGGCCGTGATGGGCAGTTCCGGTGCCGGAAAGACGACCCTGCTGAATGCCCTTGCC
1176 TTTCGATCGCCGCAGGGCATCCAAGTATCGCCATCCGGGATGCGACTGCTCAATGGCCAA
1177 CCTGTGGACGCCAAGGAGATGCAGGCCAGGTGCGCCTATGTCCAGCAGGATGACCTCTTT
1178 ATCGGCTCCCTAACGGCCAGGGAACACCTGATTTTCCAGGCCATGGTGCGGATGCCACGA
1179 CATCTGACCTATCGGCAGCGAGTGGCCCGCGTGGATCAGGTGATCCAGGAGCTTTCGCTC
1180 AGCAAATGTCAGCACACGATCATCGGTGTGCCCGGCAGGGTGAAAGGTCTGTCCGGCGG
1181 AGAAAGGAAGCGTCTGGCATTCGCCTCCGAGGCACTAACCGATCCGCCGCTTCTGATCTG
1182 CGATGAGCCCACCTCCGGACTGGACTCATTTACCGCCCACAGCGTCGTCCAGGTGCTGAA
1183 GAAGCTGTCGCAGAAGGGCAAGACCGTCATCCTGACCATTCATCAGCCGTCTTCCGAGCT
1184 GTTTGAGCTCTTTGACAAGATCCTTCTGATGGCCGAGGGCAGGGTAGCTTTCTTGGGCACT
1185 CCCAGCGAAGCCGTCGACTTCTTTTCCTAGTGAGTTCGATGTGTTTATTAAGGGTATCTAG
1186 CATTACATTACATCTCAACTCCTATCCAGCGTGGGTGCCCAGTGTCCTACCAACTACAATC
1187 CGGCGGACTTTTACGTACAGGTGTTGGCCGTTGTGCCCGGACGGGAGATCGAGTCCCGT
1188 GATCGGATCGCCAAGATATGCGACAATTTTGCTATTAGCAAAGTAGCCCGGGATATGGAGC
1189 AGTTGTTGGCCACCAAAAATTTGGAGAAGCCACTGGAGCAGCCGGAGAATGGGTACACCT

ACAAGGCCACCTGGTTCATGCAGTTCCGGGCGGTCCTGTGGCGATCCTGGCTGTCGGTG
CTCAAGGAACCACTCCTCGTAAAAGTGCGACTTATTCAGACAACGGTGAGTGGTTCCAGTG
GAAACAAATGATATAACGCTTACAATTCTTGGAAACAAATTCGCTAGATTTTAGTTAGAATTG
CCTGATTCCACACCCTTCTTAGTTTTTTTCAATGAGATGTATAGTTTATAGTTTTGCAGAAAA
TAAATAAATTTCATTTAACTCGCGAACATGTTGAAGATATGAATATTAATGAGATGCGAGTAA
CATTTTAATTTGCAGATGGTTGCCATCTTGATTGGCCTCATCTTTTTGGGCCAACAACTCAC
GCAAGTGGGCGTGATGAATATCAACGGAGCCATCTTCCTCTTCCTGACCAACATGACCTTT
CAAAACGTCTTTGCCACGATAAATGTAAGTCTTGTTTAGAATACATTTGCATATTAATAATTT
ACTAACTTTCTAATGAATCGATTCGATTTAGGTGTTCACCTCAGAGCTGCCAGTTTTTATGA
GGGAGGCCCGAAGTCGACTTTATCGCTGTGACACATACTTTCTGGGCAAAACGATTGCCG
AATTACCGCTTTTTCTCACAGTGCCACTGGTCTTCACGGCGATTGCCTATCCGATGATCGG
ACTGCGGGCCGGAGTGCTGCACTTCTTCAACTGCCTGGCGCTGGTCACTCTGGTGGCCAA
TGTGTCAACGTCCTTCGGATATCTAATATCCTGCGCCAGCTCCTCGACCTCGATGGCGCTG
TCTGTGGGTCCGCCGGTTATCATACCATTCCTGCTCTTTGGCGGCTTCTTCTTGAACTCGG
GCTCGGTGCCAGTATACCTCAAATGGTTGTCGTACCTCTCATGGTTCCGTTACGCCAACGA
GGGTCTGCTGATTAACCAATGGGCGGACGTGGAGCCGGGCGAAATTAGCTGCACATCGTC
GAACACCACGTGCCCCAGTTCGGGCAAGGTCATCCTGGAGACGCTTAACTTCTCCGCCGC
CGATCTGCCGCTGGACTACGTGGGTCTGGCCATTCTCATCGTGAGCTTCCGGGTGCTCGC
ATATCTGGCTCTAAGACTTCGGGCCCGACGCAAGGAGTAGCCGACATATATCCGAAATAAC
TGCTTGTTTTTTTTTTTACCATTATTACCATCGTGTTTACTGTTTATTGCCCCCTCAAAAAGCT
AATGTAATTATATTTGTGCCAATAAAAACAAGATATGACCTATAGAATACAAGTATTTCCCCT
TCGAACATCCCCACAAGTAGACTTTGGATTTGTCTTCTAACCAAAAGACTTACACACCTGCA
TACCTTACATCAAAAACTCGTTTATCGCTACATAAAACACCGGGATATATTTTTTATATACAT
ACTTTTCAAATCGCGCGCCCTCTTCATAATTCACCTCCACCACACCACGTTTCGTAGTTGCT
CTTTCGCTGTCTCCCACCCGCTCTCCGCAACACATTCACCTTTTGTTCGACGACCTTGGAG
CGACTGTCGTTAGTTCCGCGCGATTCGGTTCGCTCAAATGGTTCCGAGTGGTTCATTTCGT
CTCAATAGAAATTAGTAATAAATATTTGTATGTACAATTTATTTGCTCCAATATATTTGTATAT
ATTTCCCTCACAGCTATATTTATTCTAATTTAATATTATGACTTTTTAAGGTAATTTTTTGTGA
CCTGTTCGGAGTGATTAGCGTTACAATTTGAACTGAAAGTGACATCCAGTGTTTGTTCCTTG
TGTAGATGCATCTCAAAAAAATGGTGGGCATAATAGTGTTGTTTATATATATCAAAAATAACA
ACTATAATAATAAGAATACATTTAATTTAGAAAATGCTTGGATTTCACTGGAACTAGAATTAA
TTCGGCTGCTGCTCTAAACGACGCATTTCGTACTCCAAAGTACGAATTTTTTCCCTCAAGCT
CTTATTTTCATTAAACAATGAACAGGACCTAACGCACAGTCACGTTATTGTTTACATAAATGA
TTTTTTTTACTATTCAAACTTACTCTGTTTGTGTACTCCCACTGGTATAGCCTTCTTTTATCTT
TTCTGGTTCAGGCTCTATCACTTTACTAGGTACGGCATCTGCGTTGAGTCGCCTCCTTTTAA
ATGTCTGACCTTTTGCAGGTGCAGCCTTCCACTGCGAATCATTAAAGTGGGTATCACAAATT
TGGGAGTTTTCACCAAGGCTGCACCCAAGGCTCTGCTCCCACAATTTTCTCTTAATAGCAC
ACTTCGGCACGTGAATTAATTTTACTCCAGTCACAGCTTTGCAGCAAAATTTGCAATATTTC
ATTTTTTTTTATTCCACGTAAGGGTTAATGTTTTCAAAAAAAAATTCGTCCGCACACAACCTT
TCCTCTCAACAAGCAAACGTGCACTGAATTTAAGTGTATACTTCGGTAAGCTTCGGCTATCG
ACGGGACCACCTTATGTTATTTCATCATG

NNNN Origin of replication
NNNN P-Element inverted repeats

NNNN LoxP
NNNN attB
NNNN MHC splice acceptor
NNNN $(GSS)_4$ linker*
NNNN Amp resistance
NNNN T2A-GAL4_polyA
NNNN Splice donor
NNNN EGFP-FIAsH tag-StrepII tag-TEV protease cleavage site-3XFlag
NNNN white

*Linker for p1 GTGGCGGAGGTTCCGGTGGAAGCGGAGGTAGCGGCGGATCC
for p2 GTCGGGAGGTTCCGGTGGAAGCGGAGGTAGCGGCGGATCC
(sense orientations)

Sequence of y^{wing2+} donor template vector- BbsI version

CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTT
TTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAG
GGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGT
CAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAATC
AAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCG
ATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGA
AAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACA
CCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCGCAA
CTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGG
GATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTA
AAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGAGCTCCACC
GCggTGGGAAGACttGGATCTCTATTCAATCGGGACAGTGGAAATTGACTATTTTATTTATATT
AATGAACTTATTTTTAATTTGGCTTAAGTTACTAAGGGGTACTAATAGTTTGAGCGCAGTGC
ATGTCATGGGGACATGTGCAATTGTGTGTAAGCGGGAAGTGATCGCGGCCTTCCGAATTT
GGCCATGCCAAATAATCCCAGCTCGAAAGGAGGGGACCCGGCGGTCAGGGCCATGGACA
TTGAACTTGAAAAAAAAAAACACAAAAATATATAACACAAAACGGAAAATGCTGTGTACCGC
TTATGTTAGAGAAGTTGAGCAACGGGTTTTTCGTTTTGCAGTCACGATGGATTTCCAAATTA
GTGTAGGAGGGGGGAGGGGAGGGAGGGAGATAATGTCCAGGCTGCCATAAGTGGGGAAT
AAGGAAAATAAAACATGAAACACGGGTCGGGCAATGTCATGCGGTATTCGGCTTTGCTTTC
CGCCCAAGTTGAAGTGATCCTGTGTGTAAATAATGTCGAATGTTGCCGGTCGGTTGCATAA
GCGTTAGTcaattatggccaaagagatctgatttgtggaGTCGAGAAAAAAACCTTCATATAAAACGCGGC
CGACATATTATGGCCACCAGTCGTTACCGCGCCACGGTCCACAGAAGAGGATTAAAAAAAT
ATCACACAGCCGAAGGCTAGAGAAGAACCCCCTATAGCTGAACATATATAAACAAATATATT
TTTTTTTATTGCCAACACACTTTGGCTTAAGTGTTAAGAGTGATTGTCAGCTTAGAGCTAAG
TGCAATGTTCCAGGACAAAGGGTGGATCCTTGTGACCCTGATCACCTTGGTGACGCCGTC
TTGGGCTGCTTACAAACTTCAGGAGCGATATAGTTGGAGCCAGCTGGACTTTGCTTTCCCG
AATACCCGACTAAAGGACCAAGCTCTGGCTAGTGGAGATTATATTCCGCAAAATGCTCTAC

1278 CTGTTGGAGTCGAACACTTTGGCAATCGGTTATTCGTCACTGTTCCCCGCTGGCGTGATGG
1279 GATTCCGGCCACTCTGACCTATATAAACATGGACCGCAGTTTGACGGGTTCACCGGAGCTA
1280 ATTCCGTATCCAGATTGGCGCTCAAATACAGCTGGAGATTGCGCCAACAGTATTACCACTG
1281 CCTACCGCATTAAAGTGGATGAGTGTGGTCGGCTGTGGGTTTTGGACACTGGAACCGTGG
1282 GCATCGGCAATACCACCACTAATCCGTGCTCCTATGCGGTAAATGTCTTTGACTTGACCAC
1283 GGATACGCGAATTCGGAGATACGAGCTACCTGGCGTGGACACAAATCCAAATACTTTCATA
1284 GCTAACATTGCCGTGGATATAGGCAAAAATTGCGATGATGCATATGCCTATTTTGCCGATG
1285 AATTGGGATACGGCTTGATTGCTTACTCCTGGGAACTGAACAAGTCCTGGAGATTCTCGGC
1286 ACATTCGTATTTTTTCCCCGATCCATTGAGGGGCGATTTCAATGTCGCTGGTATTAACTTCC
1287 AATGGGGCGAGGAGGGTATATTTGGTATGTCCCTTTCGCCCATTCGATCGGATGGTTATCG
1288 TACCCTGTACTTTAGTCCGTTAGCAAGTCATCGACAATTTGCCGTATCCACGAGGATTTTGA
1289 GGGATGAAACCAGGACGGAAGATAGCTATCATGACTTTGTTGCCTTAGATGAACGGGGTC
1290 CAAACTCCCATACCACTTCACGTGTGATGAGCGATGATGGAATTGAGCTGTTCAATTTAATA
1291 GATCAAAATGCAGTGGGTTGCTGGCACTCATCAATGCCGTACTCACCGCAATTTCATGGCA
1292 TTGTGGATCGCGATGACGTTGGCTTAGTTTTTCCGGCCGATGTGAAAATTGATGAGAACAA
1293 AAACGTTTGGGTTCTATCCGATAGGATGCCCGTTTTCTTGCTGTCTGACTTGGATTATTCAG
1294 ATACTAATTTCCGAATTTACACGGCTCCCTTGGCCACTTTAATTGAGAATACTGTGTGTGAT
1295 TTGAGGAATAACGCCTATGGGCCGCCAAATACCGTTTCAATACCAAAACAAGCCGTTTTGC
1296 CAATGGGTCCACCGTTATATACGAAACAATATCGTCCTGTCTTGCCACAGAAACCTCAGAC
1297 CAGCTGGGCTTCCTCGCCGCCTCCTCCAAGTCGCACTTATTTGCCCGCCAATTCAGGCAAT
1298 GTAGTCTCCAGTATTAGTGTCTCTACAAATTCTGTGGGTCCTGCAGGAGTGGAGGTGCCAA
1299 AGGCCTATATTTTCAACCAGCACAACGGCATAAATTACGAGACAAGTGGTCCCCATCTATTT
1300 CCCACCCATCAACCCGCCCAACCGGGTGGCCAGGATGGTGGGTTAAAAACTTATGTGAAT
1301 GCCCGCCAATCTGGGTGGTGGCATCATCAGCATCAAGGTTAACATAATCCTACACACGGTA
1302 CTTGGGTATATTCTCACACACTCGATTGATGTAAAGAATATTTAAAGACAACAACATAGGGC
1303 AACAGCGGTTAAAAAAACCACATGACGTATGAGCAAGTGGCAAATCAATACTTTATCTAGTT
1304 ATGTTAAGCAAAAAATAACAATAAATCAACTTTTTTTTGAAGGTTAAGAGTTTACGCAATTTT
1305 CTTGAGCGGAAAAAGCGGAAAAAATGTAAGTATGCATAAATTCTAAATATATCAACAACTGT
1306 ACATTTTCTGGAGTACTACTACCAGGCAAGAAAGTAGGTTGATAAAGCTATGCACAAGATCT
1307 TGTTTGGGTGCAGGGAAAGTTCAACTTAATCGCTCAATTTGAGATCGCCTGGTCGCTTGAG
1308 ATTCGACTGTAATTGAAATTTTTGCTTTTGATCGGAGCCAGACTTCAGACGGGGCAAACAAA
1309 AAGACTTTGTTGGTGGTAGGGTAGGATCCGTTGACCTGCAGGTTCCccGTCTTCGATATCA
1310 AGCTTATCGATACCGTCGACCTCGAGGGGGGGCCCGGTACCCAGCTTTTGTTCCCTTTAG
1311 TGAGGGTTAATTGCGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTT
1312 ATCCGCTCACAATTCCACACAACATACGAGCCGGGAGCATAAAGTGTAAAGCCTGGGGTG
1313 CCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGG
1314 AAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGC
1315 GTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGC
1316 GGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATA
1317 ACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCC
1318 GCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGC
1319 TCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGA
1320 AGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTT
1321 CTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGT

AGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGC
GCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGG
CAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCT
TGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCT
GAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCT
GGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAG
AAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGG
GATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAG
TTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAG
TGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTC
GTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCG
CGtGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCC
GAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGG
AAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGG
CATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCA
AGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGA
TCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAA
TTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGT
CATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATA
ATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCG
AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCC
AACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGC
AAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCT
TTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATG
TATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCAC

NNNN y^{wing2+}

Sequence of the pATTBy^{wing2+} vector

CTGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTG
ACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCG
CCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGAT
TTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGG
GCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGT
GGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATA
AGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACG
CGAATTTTAACAAAATATTAACGCTTACAATTTCCATTCGCCATTCAGGCTGCGCAACTGTT
GGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGT
GCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACG
ACGGCCAGTGAATTGTAATACGACTCACTATAGGGCGAATTGGGTACGTACCGGCCGCCT
CTATTCAATCGGGACAGTGGAAATTGACTATTTTATTTATATTAATGAACTTATTTTTAATTTG

1366 GCTTAAGTTACTAAGGGGTACTAATAGTTTGAGCGCAGTGCATGTCATGGGGACATGTGCA
1367 ATTGTGTGTAAGCGGGAAGTGATCGCGGCCTTCCGAATTTGGCCATGCCAAATAATCCCAG
1368 CTCGAAAGGAGGGGACCCGGCGGTCAGGGCCATGGACATTGAACTTGAAAAAAAAAAACA
1369 CAAAAATATATAACACAAAACGGAAAATGCTGTGTACCGCTTATGTTAGAGAAGTTGAGCAA
1370 CGGGTTTTTCGTTTTGCAGTCACGATGGATTTCCAAATTAGTGTAGGAGGGGGGAGGGGA
1371 GGGAGGGAGATAATGTCCAGGCTGCCATAAGTGGGGAATAAGGAAAATAAAACATGAAAC
1372 ACGGGTCGGGCAATGTCATGCGGTATTCGGCTTTGCTTTCCGCCCAAGTTGAAGTGATCCT
1373 GTGTGTAAATAATGTCGAATGTTGCCGGTCGGTTGCATAAGCGTTAGTcaattatggccaaagagat
1374 ctgatttgtggaGTCGAGAAAAAAACCTTCATATAAAACGCGGCCGACATATTATGGCCACCAGT
1375 CGTTACCGCGCCACGGTCCACAGAAGAGGATTAAAAAAATATCACACAGCCGAAGGCTAG
1376 AGAAGAACCCCCTATAGCTGAACATATATAAACAAATATATTTTTTTTTTATTGCCAACACACT
1377 TTGGCTTAAGTGTTAAGAGTGATTGTCAGCTTAGAGCTAAGTGCAATGTTCCAGGACAAAG
1378 GGTGGATCCTTGTGACCCTGATCACCTTGGTGACGCCGTCTTGGGCTGCTTACAAACTTCA
1379 GGAGCGATATAGTTGGAGCCAGCTGGACTTTGCTTTCCCGAATACCCGACTAAAGGACCA
1380 AGCTCTGGCTAGTGGAGATTATATTCCGCAAAATGCTCTACCTGTTGGAGTCGAACACTTT
1381 GGCAATCGGTTATTCGTCACTGTTCCCCGCTGGCGTGATGGGATTCCGGCCACTCTGACC
1382 TATATAAACATGGACCGCAGTTTGACGGGTTCACCGGAGCTAATTCCGTATCCAGATTGGC
1383 GCTCAAATACAGCTGGAGATTGCGCCAACAGTATTACCACTGCCTACCGCATTAAAGTGGA
1384 TGAGTGTGGTCGGCTGTGGGTTTTGGACACTGGAACCGTGGGCATCGGCAATACCACCAC
1385 TAATCCGTGCTCCTATGCGGTAAATGTCTTTGACTTGACCACGGATACGCGAATTCGGAGA
1386 TACGAGCTACCTGGCGTGGACACAAATCCAAATACTTTCATAGCTAACATTGCCGTGGATA
1387 TAGGCAAAAATTGCGATGATGCATATGCCTATTTTGCCGATGAATTGGGATACGGCTTGATT
1388 GCTTACTCCTGGGAACTGAACAAGTCCTGGAGATTCTCGGCACATTCGTATTTTTTCCCCG
1389 ATCCATTGAGGGGCGATTTCAATGTCGCTGGTATTAACTTCCAATGGGGCGAGGAGGGTAT
1390 ATTTGGTATGTCCCTTTCGCCCATTCGATCGGATGGTTATCGTACCCTGTACTTTAGTCCGT
1391 TAGCAAGTCATCGACAATTTGCCGTATCCACGAGGATTTTGAGGGATGAAACCAGGACGGA
1392 AGATAGCTATCATGACTTTGTTGCCTTAGATGAACGGGGTCCAAACTCCCATACCACTTCA
1393 CGTGTGATGAGCGATGATGGAATTGAGCTGTTCAATTTAATAGATCAAAATGCAGTGGGTT
1394 GCTGGCACTCATCAATGCCGTACTCACCGCAATTTCATGGCATTGTGGATCGCGATGACGT
1395 TGGCTTAGTTTTTCCGGCCGATGTGAAAATTGATGAGAACAAAAACGTTTGGGTTCTATCC
1396 GATAGGATGCCCGTTTTCTTGCTGTCTGACTTGGATTATTCAGATACTAATTTCCGAATTTA
1397 CACGGCTCCCTTGGCCACTTTAATTGAGAATACTGTGTGTGATTTGAGGAATAACGCCTAT
1398 GGGCCGCCAAATACCGTTTCAATACCAAAACAAGCCGTTTTGCCAATGGGTCCACCGTTAT
1399 ATACGAAACAATATCGTCCTGTCTTGCCACAGAAACCTCAGACCAGCTGGGCTTCCTCGCC
1400 GCCTCCTCCAAGTCGCACTTATTTGCCCGCCAATTCAGGCAATGTAGTCTCCAGTATTAGT
1401 GTCTCTACAAATTCTGTGGGTCCTGCAGGAGTGGAGGTGCCAAAGGCCTATATTTTCAACC
1402 AGCACAACGGCATAAATTACGAGACAAGTGGTCCCCATCTATTTCCCACCCATCAACCCGC
1403 CCAACCGGGTGGCCAGGATGGTGGGTTAAAAACTTATGTGAATGCCCGCCAATCTGGGTG
1404 GTGGCATCATCAGCATCAAGGTTAACATAATCCTACACACGGTACTTGGGTATATTCTCACA
1405 CACTCGATTGATGTAAAGAATATTTAAAGACAACAACATAGGGCAACAGCGGTTAAAAAAAC
1406 CACATGACGTATGAGCAAGTGGCAAATCAATACTTTATCTAGTTATGTTAAGCAAAAAATAA
1407 CAATAAATCAACTTTTTTTTGAAGGTTAAGAGTTTACGCAATTTTCTTGAGCGGAAAAAGCG
1408 GAAAAAATGTAAGTATGCATAAATTCTAAATATATCAACAACTGTACATTTTCTGGAGTACTA
1409 CTACCAGGCAAGAAAGTAGGTTGATAAAGCTATGCACAAGATCTTGTTTGGGTGCAGGGAA

1410 AGTTCAACTTAATCGCTCAATTTGAGATCGCCTGGTCGCTTGAGATTCGACTGTAATTGAAA
1411 TTTTTGCTTTTGATCGGAGCCAGACTTCAGACGGGGCAAACAAAAAGACTTTGTTGGTGGT
1412 AGGGTAGGATCCGTTGACCTGCAGGTCTAGTCCTAGGTAGAGGGCCCTTAAGGCGCCTAT
1413 GGAGCTCTCAGACCGGTTGAGGCGGCCGCTACGCTCGAGGATGGATATCTATGCTAGCGA
1414 GATCGATACGTGCATGCGATTCTAGAGATCCACTAGTGTCGACGATGTAGGTCACGGTCTC
1415 GAAGCCGCGGTGCGGGTGCCAGGGCGTGCCCTTGGGCTCCCCGGGCGCGTACTCCACC
1416 TCACCCATCTGGTCCATCATGATGAACGGGTCGAGGTGGCGGTAGTTGATCCCGGCGAAC
1417 GCGCGGCGCACCGGGAAGCCCTCGCCCTCGAAACCGCTGGGCGCGGTGGTCACGGTGA
1418 GCACGGGACGTGCGACGGCGTCGGCGGGTGCGGATACGCGGGGCAGCGTCAGCGGGTT
1419 CTCGACGGTCACGGCGGGCATGTCGACACTAGTTCTAGCCAGCTTTTGTTCCCTTTAGTGA
1420 GGGTTAATTTCGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCC
1421 GCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTA
1422 ATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAAC
1423 CTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATT
1424 GGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCG
1425 AGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGC
1426 AGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTT
1427 GCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAG
1428 TCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTC
1429 CCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCT
1430 TCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTC
1431 GTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTA
1432 TCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCA
1433 GCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAG
1434 TGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGC
1435 CAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAG
1436 CGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGAT
1437 CCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTT
1438 TGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTA
1439 AATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAG
1440 GCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTA
1441 GATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGA
1442 CCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCG
1443 CAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCT
1444 AGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCG
1445 TGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCG
1446 AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTT
1447 GTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTC
1448 TTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTC
1449 TGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACC
1450 GCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAAC
1451 TCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTG
1452 ATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAAT
1453 GCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTC

AATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTT
AGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCAC
y^{wing2+}
y min promoter+CDS+3'UTR (bp1680-4014)
MCS: 4014-4130

Sequence of the linker_EGFP_linker donor template

CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTT
TTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAG
GGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGT
CAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAATC
AAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCG
ATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGA
AAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACA
CCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCGCAA
CTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGG
GATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTA
AAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGAGCTCGAAG
ACttTCCGGAGGTTCCGGTGGAAGCGGAGGTAGCGGCggatccGTGTCCAAGGGCGAGGAG
CTGTTCACCGGCGTGGTGCCCATCCTGGTGGAGCTGGATGGCGACGTGAACGGCCACAA
GTTCAGCGTGCGCGGCGAGGGCGAGGGCGACGCCACCAACGGCAAGCTGACCCTGAAG
TTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTGGTGACCACCCTGACC
TACGGCGTGCAGTGCTTCAGCCGCTACCCCGATCACATGAAGCAGCACGATTTCTTCAAG
AGCGCCATGCCCGAGGGCTACGTGCAGGAGCGCACCATCAGCTTCAAGGATGACGGCAC
CTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGATACCCTGGTGAACCGCATCGAGC
TGAAGGGCATCGATTTCAAGGAGGATGGCAACATCCTGGGCCACAAGCTGGAGTACAACT
TCAACAGCCACAACGTGTACATCACCGCCGATAAGCAGAAGAACGGCATCAAGGCCAACT
TCAAGATCCGCCACAATGTGGAGGATGGCTCCGTGCAGCTGGCCGATCACTACCAGCAGA
ACACCCCCATCGGCGACGGCCCAGTGCTGCTGCCCGATAACCACTACCTGAGCACCCAGA
GCGTGCTGTCCAAGGACCCCAACGAGAAGCGCGATCACATGGTGCTGCTGGAGTTCGTGA
CCGCCGCCGGCATCACCCTGGGCATGGATGAGCTGTACAAGGGCGGATCCGGAGGTAGC
GGTGGAAGCGGAGGTTCTccGTCTTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGG
GGGGCCCGGTACCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTTGGCGTAAT
CATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGA
GCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTG
CGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAAT
CGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCA
CTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCG
GTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGC
CAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGC
CCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGA
CTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCC

TGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAG
CTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCA
CGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAA
CCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGC
GAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAG
AAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGT
AGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGC
AGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGA
CGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATC
TTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAA
CTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTT
CGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTAC
CATCTGGCCCCAGTGCTGCAATGATACCGCGtGACCCACGCTCACCGGCTCCAGATTTATC
AGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGC
CTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTT
TGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGG
CTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAA
AAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTA
TCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTT
TTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGT
TGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGC
TCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATC
CAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCG
TTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACAC
GGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATT
GTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGC
ACATTTCCCCGAAAAGTGCCAC

Sequence of the Linker_mCherry_Linker donor template

CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTT
TTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAG
GGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGT
CAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAATC
AAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCG
ATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGA
AAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACA
CCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCGCAA
CTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGG
GATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTA
AAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGAGCTCCACC
gcggTGGGAAGACttTCCGGAGGTTCCGGTGGAAGCGGAGGTAGCGGCggatccGTGAGCAAG
GGCGAGGAGGATAACATGGCCATCATCAAGGAGTTCATGCGCTTCAAGGTGCACATGGAG
GGCTCCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCGCCCCTACG

1542 AGGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGGTGGCCCCCTGCCCTTCGCCTGG
1543 GACATCCTGTCCCCTCAGTTCATGTACGGCTCCAAGGCCTACGTGAAGCACCCCGCCGAC
1544 ATCCCCGACTACTTGAAGCTGTCCTTCCCCGAGGGCTTCAAGTGGGAGCGCGTGATGAAC
1545 TTCGAGGACGGCGGCGTGGTGACCGTGACCCAGGACTCCTCCCTGCAGGACGGCGAGTT
1546 CATCTACAAGGTGAAGCTGCGCGGCACCAACTTCCCCTCCGACGGCCCCGTAATGCAGAA
1547 GAAAACCATGGGCTGGGAGGCCTCCTCCGAGCGGATGTACCCCGAGGACGGCGCCCTGA
1548 AGGGCGAGATCAAGCAGAGGCTGAAGCTGAAGGACGGCGGCCACTACGACGCTGAGGTC
1549 AAGACCACCTACAAGGCCAAGAAGCCCGTGCAGCTGCCCGGCGCCTACAACGTCAACATC
1550 AAGTTGGACATCACCTCCCACAACGAGGACTACACCATCGTGGAACAGTACGAACGCGCC
1551 GAGGGCCGCCACTCCACCGGCGGCATGGACGAGCTGTACAAGggcgGATCCGGAGGTAGC
1552 GGTGGAAGCGGAGGTTCTccGTCTTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGG
1553 GGGGCCCGGTACCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTTGGCGTAAT
1554 CATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGA
1555 GCCGGGAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTG
1556 CGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAAT
1557 CGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCA
1558 CTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCG
1559 GTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGC
1560 CAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGC
1561 CCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGA
1562 CTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCC
1563 TGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAG
1564 CTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCA
1565 CGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAA
1566 CCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGC
1567 GAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAG
1568 AAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGT
1569 AGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGC
1570 AGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGA
1571 CGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATC
1572 TTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAA
1573 CTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTT
1574 CGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTAC
1575 CATCTGGCCCCAGTGCTGCAATGATACCGCGtGACCCACGCTCACCGGCTCCAGATTTATC
1576 AGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGC
1577 CTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTT
1578 TGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGG
1579 CTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAA
1580 AAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTA
1581 TCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTT
1582 TTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGT
1583 TGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGC
1584 TCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATC
1585 CAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCG

TTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACAC
GGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATT
GTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGC
ACATTTCCCCGAAAAGTGCCAC

Sequence of the T2A_Gal4 (Not ATG, NoStop codon) donor template

CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTT
TTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAG
GGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGT
CAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAATC
AAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCG
ATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGA
AAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACA
CCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCGCAA
CTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGG
GATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTA
AAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGAGCTCGAAG
ACttGAGGGCCGCGGCAGCCTGCTGACCTGCGGCGATGTGGAGGAGAACCCCGGGCCCAT
GAAGCTGCTGAGCAGCATCGAGCAGGCCTGCGATATCTGCCGCCTGAAGAAGCTGAAGTG
CAGCAAGGAGAAGCCCAAGTGCGCCAAGTGCCTGAAGAACAACTGGGAGTGCCGCTACA
GCCCCAAGACCAAGCGCAGCCCCCTGACCCGCGCCCACCTGACCGAGGTGGAGAGCCG
CCTGGAGCGCCTGGAGCAGCTGTTCCTGCTGATCTTCCCCCGCGAGGATCTGGATATGAT
CCTGAAGATGGATAGCCTGCAGGATATCAAGGCCCTGCTGACCGGCCTGTTCGTGCAGGA
TAACGTGAACAAGGATGCCGTGACCGATCGCCTGGCCAGCGTGGAAACCGATATGCCCCT
GACCCTGCGCCAGCACCGCATCAGCGCCACCAGCAGCAGCGAGGAGAGCAGCAACAAGG
GCCAGCGCCAGCTGACCGTGAGCATCGATAGCGCCGCCCACCACGATAACAGCACCATC
CCCCTGGATTTCATGCCCCGCGATGCCCTGCACGGCTTCGATTGGAGCGAGGAGGATGAT
ATGAGCGATGGCCTGCCCTTCCTGAAAACCGATCCCAACAACAACGGCTTCTTCGGCGAT
GGCAGCCTGCTGTGCATCCTGCGCAGCATCGGCTTCAAGCCCGAGAACTACACCAACAGC
AACGTGAACCGCCTGCCCACCATGATCACCGATCGCTACACCCTGGCCAGCCGCAGCACC
ACCAGCCGCCTGCTGCAGAGCTACCTGAACAACTTCCACCCCTACTGCCCCATCGTGCAC
AGCCCCACCCTGATGATGCTGTACAACAACCAGATCGAGATCGCCAGCAAGGATCAGTGG
CAGATCCTGTTCAACTGCATCCTGGCCATCGGCGCCTGGTGCATCGAGGGCGAGAGCACC
GATATCGATGTGTTCTACTACCAGAACGCCAAGAGCCACCTGACCAGCAAGGTGTTCGAG
AGCGGCAGCATCATCCTGGTGACCGCCCTGCACCTGCTGAGCCGCTACACCCAGTGGCG
CCAAAAGACCAACACCAGCTACAACTTCCACAGCTTCAGCATCCGCATGGCCATCAGCCTG
GGCCTGAACCGCGATCTGCCCAGCAGCTTCAGCGATAGCAGCATCCTGGAGCAGCGCCG
CCGCATCTGGTGGAGCGTGTACAGCTGGGAGATCCAGCTGAGCCTGCTGTACGGCCGCA
GCATCCAGCTGAGCCAGAACACCATCAGCTTCCCCAGCAGCGTGGATGATGTGCAGCGCA
CCACCACCGGCCCCACCATCTACCACGGCATCATCGAGACTGCCCGCCTGCTGCAGGTGT
TCACCAAGATCTACGAGCTGGATAAGACCGTGACCGCCGAGAAGAGCCCCATCTGCGCCA
AGAAGTGCCTGATGATCTGCAACGAGATCGAGGAGGTGAGCCGCCAGGCCCCCAAGTTC

1630 CTGCAGATGGATATCAGCACCACCGCCCTGACCAACCTGCTGAAGGAGCACCCCTGGCTG
1631 AGCTTCACCCGCTTCGAGCTGAAGTGGAAGCAGCTGAGCCTGATCATCTACGTGCTGCGC
1632 GATTTCTTCACCAACTTCACCCAGAAGAAGAGCCAGCTGGAGCAGGATCAGAACGATCAC
1633 CAGAGCTACGAGGTGAAGCGCTGCAGCATCATGCTGAGCGATGCCGCCCAGCGCACCGT
1634 GATGAGCGTGAGCAGCTACATGGATAACCACAACGTGACCCCCTACTTCGCCTGGAACTG
1635 CAGCTACTACCTGTTCAACGCCGTGCTGGTGCCCATCAAGACCCTGCTGAGCAACAGCAA
1636 GAGCAACGCCGAGAACAACGAAACGGCCCAGCTGCTGCAGCAGATCAACACCGTGCTGAT
1637 GCTGCTGAAGAAGCTGGCCACCTTCAAGATCCAGACCTGCGAGAAGTACATCCAGGTGCT
1638 GGAGGAGGTGTGCGCCCCCTTCCTGCTGAGCCAGTGCGCCATCCCCCTGCCCCACATCA
1639 GCTACAACAACAGCAACGGCAGCGCCATCAAGAACATCGTGGGCAGCGCCACCATCGCC
1640 CAGTACCCCACCCTGCCCGAGGAGAACGTGAACAACATCAGCGTGAAGTACGTGAGCCCC
1641 GGCAGCGTGGGACCCAGCCCCGTGCCCCTGAAGAGCGGCGCCAGCTTCAGCGATCTGGT
1642 GAAGCTGCTGAGCAACCGCCCCCCCAGCCGCAACAGCCCCGTGACCATCCCCCGCAGCA
1643 CCCCCAGCCACCGCAGCGTGACCCCCTTCCTGGGCCAGCAGCAGCAGCTGCAGAGCCTG
1644 GTGCCCCTGACCCCCAGCGCCCTGTTCGGCGGCGCCAACTTCAACCAGAGCGGCAACAT
1645 CGCCGATAGCAGCCTGAGCTTCACCTTCACCAACAGCAGCAACGGCCCCAACCTGATCAC
1646 CACCCAGACCAACAGCCAGGCCCTGAGCCAGCCCATCGCCAGCAGCAACGTGCACGATA
1647 ACTTCATGAACAACGAGATCACCGCCAGCAAGATCGATGATGGCAACAACAGCAAGCCCC
1648 TGAGCCCCGGCTGGACCGATCAGACCGCCTACAACGCCTTCGGCATCACCACCGGCATGT
1649 TCAACACCACCACCATGGATGATGTGTACAACTACCTGTTCGATGATGAGGATACCCCCCC
1650 CAACCCCAAGAAGGAGccGTCTTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGG
1651 GGCCCGGTACCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTTGGCGTAATCAT
1652 GGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCC
1653 GGGAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGT
1654 TGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCG
1655 GCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACT
1656 GACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGT
1657 AATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCA
1658 GCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCC
1659 CCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACT
1660 ATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCT
1661 GCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAG
1662 CTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCA
1663 CGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAA
1664 CCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGC
1665 GAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAG
1666 AAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGT
1667 AGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGC
1668 AGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGA
1669 CGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATC
1670 TTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAA
1671 CTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTT
1672 CGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTAC
1673 CATCTGGCCCCAGTGCTGCAATGATACCGCGtGACCCACGCTCACCGGCTCCAGATTTATC

AGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGC
CTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTT
TGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGG
CTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAA
AAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTA
TCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTT
TTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGT
TGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGC
TCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATC
CAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCG
TTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACAC
GGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATT
GTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGC
ACATTTCCCCGAAAAGTGCCAC

SEQUENCE OF THE T2A_GAL4_POLYA DONOR TEMPLATE

CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTT
TTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAG
GGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGT
CAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAATC
AAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCG
ATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGA
AAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACA
CCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCGCAA
CTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGG
GATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTA
AAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGAGCTCCACC
GCGGTGGCGTCTCTAAACGGTCTCTAAACGAAGACTTAAACATGCATGTACTGACGGACAC
ACCGAAGCCCCGGCGGCAACCCTCAGCGGATGCCCCGGGGCTTCACGTTTTCCCAGGTC
AGAAGCGGTTTTCGGGAGTAGTGCCCCAACTGGGGTAACCTTTGAGTTCTCTCAGTTGGG
GGCGTAGGGCCGCCGACATGACACAAGGGGTTGTGACCGGGGTGGACACGTACGCGGG
TGCTTACGACCGTCAGTCGCGCGAGCGCGAGAAGTTCCTATTCTCTAGAAAGTATAGGAAC
TTCGAATTGAGTCGATCCAACATGGCGACTTGTCCCATCCCCGGCATGTTTAAATATACTAA
TTATTCTTGAACTAATTTTAATCAACCGATTTATCTCTCTTCCGCAGGTGGGAGGTTCCGGT
GGAAGCGGAGGTAGCGGCACCGGTGCATGCGAGGGCCGCGGCAGCCTGCTGACCTGCG
GCGATGTGGAGGAGAACCCCGGGCCCATGAAGCTGCTGAGCAGCATCGAGCAGGCCTGC
GATATCTGCCGCCTGAAGAAGCTGAAGTGCAGCAAGGAGAAGCCCAAGTGCGCCAAGTGC
CTGAAGAACAACTGGGAGTGCCGCTACAGCCCCAAGACCAAGCGCAGCCCCCTGACCCG
CGCCCACCTGACCGAGGTGGAGAGCCGCCTGGAGCGCCTGGAGCAGCTGTTCCTGCTGA
TCTTCCCCCGCGAGGATCTGGATATGATCCTGAAGATGGATAGCCTGCAGGATATCAAGG
CCCTGCTGACCGGCCTGTTCGTGCAGGATAACGTGAACAAGGATGCCGTGACCGATCGCC
TGGCCAGCGTGGAAACCGATATGCCCCTGACCCTGCGCCAGCACCGCATCAGCGCCACC
AGCAGCAGCGAGGAGAGCAGCAACAAGGGCCAGCGCCAGCTGACCGTGAGCATCGATAG
CGCCGCCCACCACGATAACAGCACCATCCCCCTGGATTTCATGCCCCGCGATGCCCTGCA

1718 CGGCTTCGATTGGAGCGAGGAGGATGATATGAGCGATGGCCTGCCCTTCCTGAAAACCGA
1719 TCCCAACAACAACGGCTTCTTCGGCGATGGCAGCCTGCTGTGCATCCTGCGCAGCATCGG
1720 CTTCAAGCCCGAGAACTACACCAACAGCAACGTGAACCGCCTGCCCACCATGATCACCGA
1721 TCGCTACACCCTGGCCAGCCGCAGCACCACCAGCCGCCTGCTGCAGAGCTACCTGAACAA
1722 CTTCCACCCCTACTGCCCCATCGTGCACAGCCCCACCCTGATGATGCTGTACAACAACCA
1723 GATCGAGATCGCCAGCAAGGATCAGTGGCAGATCCTGTTCAACTGCATCCTGGCCATCGG
1724 CGCCTGGTGCATCGAGGGCGAGAGCACCGATATCGATGTGTTCTACTACCAGAACGCCAA
1725 GAGCCACCTGACCAGCAAGGTGTTCGAGAGCGGCAGCATCATCCTGGTGACCGCCCTGC
1726 ACCTGCTGAGCCGCTACACCCAGTGGCGCCAAAAGACCAACACCAGCTACAACTTCCACA
1727 GCTTCAGCATCCGCATGGCCATCAGCCTGGGCCTGAACCGCGATCTGCCCAGCAGCTTCA
1728 GCGATAGCAGCATCCTGGAGCAGCGCCGCCGCATCTGGTGGAGCGTGTACAGCTGGGAG
1729 ATCCAGCTGAGCCTGCTGTACGGCCGCAGCATCCAGCTGAGCCAGAACACCATCAGCTTC
1730 CCCAGCAGCGTGGATGATGTGCAGCGCACCACCACCGGCCCCACCATCTACCACGGCAT
1731 CATCGAGACTGCCCGCCTGCTGCAGGTGTTCACCAAGATCTACGAGCTGGATAAGACCGT
1732 GACCGCCGAGAAGAGCCCCATCTGCGCCAAGAAGTGCCTGATGATCTGCAACGAGATCGA
1733 GGAGGTGAGCCGCCAGGCCCCCAAGTTCCTGCAGATGGATATCAGCACCACCGCCCTGA
1734 CCAACCTGCTGAAGGAGCACCCCTGGCTGAGCTTCACCCGCTTCGAGCTGAAGTGGAAGC
1735 AGCTGAGCCTGATCATCTACGTGCTGCGCGATTTCTTCACCAACTTCACCCAGAAGAAGAG
1736 CCAGCTGGAGCAGGATCAGAACGATCACCAGAGCTACGAGGTGAAGCGCTGCAGCATCAT
1737 GCTGAGCGATGCCGCCCAGCGCACCGTGATGAGCGTGAGCAGCTACATGGATAACCACA
1738 ACGTGACCCCCTACTTCGCCTGGAACTGCAGCTACTACCTGTTCAACGCCGTGCTGGTGC
1739 CCATCAAGACCCTGCTGAGCAACAGCAAGAGCAACGCCGAGAACAACGAAACGGCCCAG
1740 CTGCTGCAGCAGATCAACACCGTGCTGATGCTGCTGAAGAAGCTGGCCACCTTCAAGATC
1741 CAGACCTGCGAGAAGTACATCCAGGTGCTGGAGGAGGTGTGCGCCCCCTTCCTGCTGAG
1742 CCAGTGCGCCATCCCCCTGCCCCACATCAGCTACAACAACAGCAACGGCAGCGCCATCAA
1743 GAACATCGTGGGCAGCGCCACCATCGCCCAGTACCCCACCCTGCCCGAGGAGAACGTGA
1744 ACAACATCAGCGTGAAGTACGTGAGCCCCGGCAGCGTGGGACCCAGCCCCGTGCCCCTG
1745 AAGAGCGGCGCCAGCTTCAGCGATCTGGTGAAGCTGCTGAGCAACCGCCCCCCCAGCCG
1746 CAACAGCCCCGTGACCATCCCCCGCAGCACCCCCAGCCACCGCAGCGTGACCCCCTTCC
1747 TGGGCCAGCAGCAGCAGCTGCAGAGCCTGGTGCCCCTGACCCCCAGCGCCCTGTTCGGC
1748 GGCGCCAACTTCAACCAGAGCGGCAACATCGCCGATAGCAGCCTGAGCTTCACCTTCACC
1749 AACAGCAGCAACGGCCCCAACCTGATCACCACCCAGACCAACAGCCAGGCCCTGAGCCA
1750 GCCCATCGCCAGCAGCAACGTGCACGATAACTTCATGAACAACGAGATCACCGCCAGCAA
1751 GATCGATGATGGCAACAACAGCAAGCCCCTGAGCCCCGGCTGGACCGATCAGACCGCCT
1752 ACAACGCCTTCGGCATCACCACCGGCATGTTCAACACCACCACCATGGATGATGTGTACAA
1753 CTACCTGTTCGATGATGAGGATACCCCCCCCAACCCCAAGAAGGAGTAATAAGGCGCGCC
1754 CAAAGATCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAATGCAG
1755 TGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGC
1756 TGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGGT
1757 GTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGCTGATTATGATCATA
1758 ATTCGAGCTCGCCCGGGGATCTAATTCAATTAGAGACTAATTCAATTAGAGCTAATTCAATT
1759 AGGATCCAAGCTTATCGATTTCGAACCCTCGACCGCCGGAGTATAAATAGAGGCGCTTCGT
1760 CTACGGAGCGACAATTCAATTCAAACAAGCAAAGTGAACACGTCGCTAAGCGAAAGCTAAG
1761 CAAATAAACAAGCGCAGCTGAACAAGCTAAACAATCGGACTAGAGCCGGTCGCCGGCCGG

1762 CCACCATGGTGTCCAAGGGCGAGGAGCTGTTCACCGGCGTGGTGCCAATTCTGGTGGAG
1763 CTGGATGGCGACGTGAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGACG
1764 CCACCTATGGAAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCAT
1765 GGCCAACCCTCGTGACCACCCTGACCTATGGCGTGCAGTGCTTCAGCCGCTACCCCGATC
1766 ACATGAAGCAGCACGATTTCTTCAAGAGCGCCATGCCCGAGGGCTACGTGCAGGAGCGCA
1767 CCATCTTTTTCAAGGATGACGGCAACTACAAGACCCGCGCCGAAGTGAAGTTCGAGGGCG
1768 ATACCCTCGTGAACCGCATCGAGCTGAAGGGCATCGATTTCAAGGAGGATGGAAACATCC
1769 TGGGCCACAAGCTGGAGTACAACTACAACAGCCACAACGTGTACATCATGGCCGACAAGC
1770 AGAAGAACGGCATCAAGGCCAACTTCAAGATCCGCCACAACATCGAGGATGGCGGCGTGC
1771 AGCTGGCCGATCACTACCAGCAGAACACCCCAATCGGCGACGGCCCAGTGCTGCTGCCC
1772 GATAACCATTACCTGAGCACCCAGAGCGCCCTGAGCAAGGATCCCAACGAGAAGCGCGAC
1773 CACATGGTGCTGCTGGAGTTTGTGACCGCCGCCGGCATTACCCTGGGCATGGATGAGCTG
1774 TACAAGTAGGATCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAAT
1775 GCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTAT
1776 AAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGG
1777 AGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGCTGATTATGAT
1778 CAGAAGTTCCTATTCTCTAGAAAGTATAGGAACTTCTCGCGCTCGCGCGACTGACGGTCGT
1779 AAGCACCCGCGTACGTGTCCACCCCGGTCACAACCCCTTGTGTCATGTCGGCGGCCCTAC
1780 GCCCCCAACTGAGAGAACTCAAAGGTTACCCCAGTTGGGGCACTACTCCCGAAAACCGCT
1781 TCTGACCTGGGAAAACGTGAAGCCCCGGGGCATCCGCTGAGGGTTGCCGCCGGGGCTTC
1782 GGTGTGTCCGTCAGTACGCATGCCGCGTTGTCTTCCGCGTGAGACCCGCGTGAGACGGAT
1783 ATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCCGGTACCCAGCTTTTGTTCCCT
1784 TTAGTGAGGGTTAATTGCGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAAT
1785 TGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGG
1786 GTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTC
1787 GGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTT
1788 TGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGC
1789 TGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGG
1790 ATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGG
1791 CCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGAC
1792 GCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTG
1793 GAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCT
1794 TTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGT
1795 GTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCT
1796 GCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACT
1797 GGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTT
1798 CTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTG
1799 CTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCG
1800 CTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCA
1801 AGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAA
1802 GGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGA
1803 AGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATC
1804 AGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCG
1805 TCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATAC

CGCGTGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGG
CCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCG
GGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACA
GGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGAT
CAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCC
GATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCAT
AATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAA
GTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGA
TAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGG
CGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCAC
CCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG
GCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTC
CTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAA
TGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCAC

Protocol for designing donor templates for HDR using the y^{wing2+} dominant marker constructs

Overview:
Homology arms are cloned to span between the donor vector backbone and the yellow marker inserts to allow recombination to replace the locus of interest with the screenable marker. We will use the golden gate reaction to assemble the product:

Vector---upstream homology arm----GG..yellow dominant marker…CC-downstream homology arm---vector.

The GG and CC will produce new directional PAM sites allowing the original wild type sequence to be reinserted into the resultant fly strain without being cut by the new guide RNA + Cas9 combination. Internal sequences can thus be modified and reinserted to modify the locus of interest in any way desired.

For the reaction, directional restriction enzyme cuts will produce 4bp sticky ends with 5' overhangs.

Put together in the following order by the golden gate reaction the result will be:

pBH…GACC…upstrmHMA…GGAT…yellow marker…TTCC…dwnstrmHMA…TATA…pBH

Where the color shaded sequences represent the overhangs between the respective pieces.

PROTOCOL:

I. define region (or gene) of interest (ROI - note that subsequent modifications are confined to the ROI)
   a. Use a suitable program to find Cas9 target sites flanking the ROI

b. optional – it is helpful to build a sequence file as you design (see blank template below)
c. recommended: verify sequence of actual fly strain to be injected at all sgRNA+PAM sites and all primer sites for PCR
d. Choose enzyme for Golden Gate assembly (BbsI, BsaI or BsmBI) by examining the sequence ~1kb surrounding your region of interest for the presence of the enzyme. If all three enzymes are in the sequence, use Gibson of NEB HiFi assembly instead
    i. Choose upstream boundary of ROI by presence of suitable sgRNA+PAM sequences
        1. locate suitable PAM site flanking ROI upstream, note cut site (ie 3nts prior to NGG PAM relative to sense strand)
        2. Verify that nucleotides GG are not immediately downstream of cut site on sense orientation (this will interfere with cassette removal)
        3. Recommended: check and minimize predicted off-targets on same chromosome as ROI
        4. Recommended: check predicted efficiency of sgRNA >5.5 (http://www.flyrnai.org/evaluateCrispr/)
    II. Predict novel sgRNA formed upon insertion (if cassette removal is desired. If only creating null mutant skip this step)
        1. Note sequence of 21 nucleotides immediately upstream of cut site
        2. Add “GG” to end of 21 nucleotides immediately upstream of cut site = novel upstream sgRNA+PAM
        3. Recommended: check and minimize predicted off-targets on same chromosome as ROI
        4. Recommended: check predicted efficiency of sgRNA >5.5 (http://www.flyrnai.org/evaluateCrispr/)
        5. Optional: add up to 3 nucleotides between 21 nucleotide sequence and “GG” to increase specificity and/or predicted efficiency (these must be added to reverse primer for upstream homology arm – see below.
        6. If novel upstream sgRNA is not suitable, return to step i. 1 and choose another sgRNA+PAM sequence
        7. If novel upstream sgRNA is suitable (no precited off targets, efficiency >5.5), note the sequence of 20 nucleotides preceeding the NGG created (be sure to include optional 1-3 nucleotides if applicaple from step ii. 5.) and proceed to cloning upstream sgRNA from step i. and novel upstream sgRNA from step ii. (if cassette removal is desired) using appropriate protocol for sgRNA expression construct
    ii. Clone upstream homology arm
        1. Design reverse primer
            a. Choose the annealing sequence for PCR= 18-30 nucleotides immediately upstream of the cutsite

i. Should end exactly at the cut site
ii. Should have sufficient Tm for PCR
iii. Should have little secondary structure and 40%-60% GC content, depending on Taq used

b. Add the 1-3 nucleotides from step ii. 5. If applicable
c. Add the overhang GGAT followed by the reverse orientation of the typeIIs restriction enzyme binding site chosen in step I. D
d. Add 4-6 random nucleotids so the enzyme is not at the end of the sequence
e. Reverse primer for upstream homology arm is **the reverse complement** of: a+b+c+d

2. Design forward primer
   a. Add 4-6 random nucleotids so the enzyme is not at the end of the sequence
   b. Insert forward orientation of the typeIIs restriction enzyme binding site chosen in step I. D
   c. Add the overhang GACC
   d. Choose the annealing sequence for PCR= 18-30 nucleotides approximately 500-1000bp upstream of the cutsite
      i. Should end exactly at the cut site
      ii. Should have sufficient Tm for PCR
      iii. Should have little secondary structure and 40%-60% GC content, depending on Taq used
   e. Forward primer for upstream homology arm is a+b+c+d
3. PCR from wild-type genomic DNA from fly strain to be injected

iii. choose downstream boundary of ROI by presence of suitable sgRNA+PAM sequences
   1. locate suitable PAM site flanking ROI downstream, note cut site (ie 3nts prior to NGG PAM relative to sense strand)
   2. Verify that nucleotides CC are not immediately upstream of cut site on sense orientation (this will interfere with cassette removal)
   3. Recommended: check and minimize predicted off-targets on same chromosome as ROI (http://tools.flycrispr.molbio.wisc.edu/targetFinder/)
   4. Recommended: check predicted efficiency of sgRNA >5.5 (http://www.flyrnai.org/evaluateCrispr/)

iv. Predict novel sgRNA formed upon insertion (if cassette removal is desired. If only creating null mutant skip this step)
   1. Note sequence of 21 nucleotides immediately downstream of cut site
   2. Add “CC” to beginning of 21 nucleotides immediately downstream of cut site = novel downstream sgRNA+PAM

3. Recommended: check and minimize predicted off-targets on same chromosome as ROI (http://tools.flycrispr.molbio.wisc.edu/targetFinder/)
4. Recommended: check predicted efficiency of sgRNA >5.5 (http://www.flyrnai.org/evaluateCrispr/)
5. Optional: add up to 3 nucleotides between 21 nucleotide sequence and "CC" to increase specificity and/or predicted efficiency (these must be added to reverse primer for upstream homology arm – see below.
6. If novel downstream sgRNA is not suitable, return to step iii. 1 and choose another sgRNA+PAM sequence
7. If novel downstream sgRNA is suitable (no predicted off targets, efficiency >5.5), note the sequence of 20 nucleotides preceding the NGG created (be sure to include optional 1-3 nucleotides if applicable from step ii. 5.) and proceed to cloning downstream sgRNA from step iii. and novel downstream sgRNA from step iv. (if cassette removal is desired) using appropriate protocol for sgRNA expression construct

v. Clone downstream homology arm
   1. Design forward primer
      a. Add 4-6 random nucleotides so the enzyme is not at the end of the sequence
      b. Insert forward orientation of the typeIIs restriction enzyme binding site chosen in step I. D followed by the overhang TTCC
      c. Add the 1-3 nucleotides from step iv. 5. If applicable
      d. Choose the annealing sequence for PCR= 18-30 nucleotides immediately downstream of the cutsite
         i. Should start exactly after the cut site
         ii. Should have sufficient Tm for PCR
         iii. Should have little secondary structure and 40%-60% GC content, depending on Taq used
      e. Forward primer for upstream homology arm is a+b+c+d
   2. Design reverse primer
      a. Choose the annealing sequence for PCR= 18-30 nucleotides approximately 500-1000 nucleotides downstream of the cutsite
         i. Should start right after at the cut site
         ii. Should have sufficient Tm for PCR
         iii. Should have little secondary structure and 40%-60% GC content, depending on Taq used
      b. Add the 1-3 nucleotides from step iv. 5. If applicable

c. Add the overhang TATA followed by the reverse orientation of the typeIIs restriction enzyme binding site chosen in step I. D
d. Add 4-6 random nucleotides so the enzyme is not at the end of the sequence
e. Reverse primer for downstream homology arm is **the reverse complement** of: a+b+c+d

3. PCR from wild-type genomic DNA from fly strain to be injected

vi. Proceed to Golden Gate Assembly of HDR donor template plasmid

Golden gate reaction using pBH as the destination vector

Originally by Ben Housden – modified by David Li-Kroeger. 10 or (15ul rx)

1. PCR amplify upstream and downstream homology arms from genomic DNA prepared from the Cas9 line to be used for injection. Arms should be 500-1000bp each and should not contain cut sites for the restriction enzyme to be used for golden gate cloning.

2. Gel purify homology arms and normalize concentrations to 10ng/ul. (note: equimolar ratios of components is important)

3. Set up a golden gate reaction including the following:

a. 15 (20ng) pBH-destination plasmid (2386bp)
b. Equimolar amounts upstream homology arm
c. Equimolar amounts downstream homology arm
d. Equimolar amounts insert donor plasmid (ie y^{wing2+})
e. 1 (1.5) ul buffer: cutsmart for BsaI; Buffer 2.1 for BbsI and for **BsmBI use T4 ligase buffer** (All buffers are NEB)
f. 1 (1.5) ul 10mM ATP
g. 0.5 (0.75) ul **high concentration** T4 ligase (NEB)
h. 0.5 (0.75) ul restriction enzyme (BsaI, BbsI or BsmBI (NEB))
i. water to 10 (15ul) total volume

4. Run golden gate program in a PCR machine:

a. 37C – 3 mins

b. 20C – 2 mins

c. Goto a. 10-20X

d. 37C – 5 mins

e. Remove promptly and put on ice

For BsmBI use the following conditions (remember to use buffer for T4 ligase)

5. 1. 42C for 5mins
6. 2. 16C for 5mins
7. 3. Cycle 1-2 for 20times
8. 4. 60C for 10mins
9. 5. 80C for 10min
10. 6. 4C forever.
11. Transform in Stable2 cells and plate with kanamycin selection

Xform into stbl2 cells and Grow o/n at 30C

Select colonies and digest by restriction enzymes to give appropriate banding pattern to evaluate presence of both homology arms and insert
Sequence verify homology arms and flanking sequences of insert and backbone

Overhangs for typeIIs restriction enzymes

BsaI

UPSTREAM OVERHANG DOWNSTREAM OVERHANG

5’GGTCTCt 3’ 5’ NNNNcGAGACC 3’
3’CCAGAGaNNNN 5’ 3’ gCTCTGG 5’

BbsI

UPSTREAM OVERHANG DOWNSTREAM OVERHANG

5’GAAGACtt 3’ 5’ NNNNccGTCTTC 3’
3’CTTCTGaaNNNN 5’ 3’ ggCAGAAG 5’

BsmBI

UPSTREAM OVERHANG DOWNSTREAM OVERHANG

5’CGTCTCt 3’ 5’ NNNNcGAGACG 3’
3’GCAGAGaNNNN 5’ 3’ gCTCTGC 5’

pBH_y^{wing2+} blank template

insert sequence of Downstream homology arm here… (note that the ends may vary due to enzyme used - delete sequences in red to the appropriate TATA)

TATACGTCTCtTATAGAAGACttTATAttctagaGCCGTCCCGTCAAGTCAGCGTAATGCTCTGCC
AGTGTTACAACCAATTAACCAATTCTGATTAGAAAAACTCATCGAGCATCAAATGAAACTGC
AATTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGTTTCTGTAATGAAGGA
GAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCCG
ACTCGTCCAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAATAAGGTTATCAAGTGA
GAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGCTTATGCATTTCTTTC
CAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGCATCAACCAAAC
CGTTATTCATTCGTGATTGCGCCTGAGCGAGtCGAAATACGCGATCGCTGTTAAAAGGACA
ATTACAAACAGGAATCGAATGCAACCGGCGCAGGAACACTGCCAGCGCATCAACAATATTT
TCACCTGAgTCAGGATATTCTTCTAATACCTGGAATGCTGTTTTCCCGGGGATCGCAGTGG
TGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGTCGGAAGAGGCATAAA
TTCCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACATCATTGGCAACGCTACCTTTGC
CATGTTTCAGAAACAACTCTGGCGCATCGGGCTTCCCATACAATCGATAGATTGTCGCACC
TGATTGCCCGACATTATCGCGAGCCCATTTATACCCATATAAATCAGCATCCATGTTGGAAT
TTAATCGCGGCCTGGAGCAAGACGTTTCCCGTTGAATATGGCTCATAACACCCCTTGTATT
ACTGTTTATGTAAGCAGACAGTTTTATTGTTCATGATGATATATTTTTATCTTGTGCAATGTA
ACATCAGAGATTTTGAGACATGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTG
AGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCG
GTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCA
GAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAA
CTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGT
GGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAG
CGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACAC
CGAACTGAGATACCTACAGCGTGAGCtaTGAGAAAGCGCCACGCTTCCCGAAGGGAGAAA
GGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTT
CCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAG
CGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGTC
GActcgagGACCtGAGACCGACCtGAGACGGACCttGTCTTC

Insert sequence of the Upstream homology arm (note that the ends may vary due to enzyme used remove to the appropriate GACC)

GGATCTCTATTCAATCGGGACAGTGGAAATTGACTATTTTATTTATATTAATGAACTTATTTT
TAATTTGGCTTAAGTTACTAAGGGGTACTAATAGTTTGAGCGCAGTGCATGTCATGGGGAC
ATGTGCAATTGTGTGTAAGCGGGAAGTGATCGCGGCCTTCCGAATTTGGCCATGCCAAATA
ATCCCAGCTCGAAAGGAGGGGACCCGGCGGTCAGGGCCATGGACATTGAACTTGAAAAAA
AAAAACACAAAAATATATAACACAAAACGGAAAATGCTGTGTACCGCTTATGTTAGAGAAGT
TGAGCAACGGGTTTTTCGTTTTGCAGTCACGATGGATTTCCAAATTAGTGTAGGAGGGGGG
AGGGGAGGGAGGGAGATAATGTCCAGGCTGCCATAAGTGGGGAATAAGGAAAATAAAACA
TGAAACACGGGTCGGGCAATGTCATGCGGTATTCGGCTTTGCTTTCCGCCCAAGTTGAAGT
GATCCTGTGTGTAAATAATGTCGAATGTTGCCGGTCGGTTGCATAAGCGTTAGTCAATTAT

2095 GGCCAAAGAGATCTGATTTGTGGAGTCGAGAAAAAAACCTTCATATAAAACGCGGCCGAC
2096 ATATTATGGCCACCAGTCGTTACCGCGCCACGGTCCACAGAAGAGGATTAAAAAAATATCA
2097 CACAGCCGAAGGCTAGAGAAGAACCCCCTATAGCTGAACATATATAAACAAATATATTTTTT
2098 TTTATTGCCAACACACTTTGGCTTAAGTGTTAAGAGTGATTGTCAGCTTAGAGCTAAGTGCA
2099 ATGTTCCAGGACAAAGGGTGGATCCTTGTGACCCTGATCACCTTGGTGACGCCGTCTTGG
2100 GCTGCTTACAAACTTCAGGAGCGATATAGTTGGAGCCAGCTGGACTTTGCTTTCCCGAATA
2101 CCCGACTAAAGGACCAAGCTCTGGCTAGTGGAGATTATATTCCGCAAAATGCTCTACCTGT
2102 TGGAGTCGAACACTTTGGCAATCGGTTATTCGTCACTGTTCCCCGCTGGCGTGATGGGATT
2103 CCGGCCACTCTGACCTATATAAACATGGACCGCAGTTTGACGGGTTCACCGGAGCTAATTC
2104 CGTATCCAGATTGGCGCTCAAATACAGCTGGAGATTGCGCCAACAGTATTACCACTGCCTA
2105 CCGCATTAAAGTGGATGAGTGTGGTCGGCTGTGGGTTTTGGACACTGGAACCGTGGGCAT
2106 CGGCAATACCACCACTAATCCGTGCTCCTATGCGGTAAATGTCTTTGACTTGACCACGGAT
2107 ACGCGAATTCGGAGATACGAGCTACCTGGCGTGGACACAAATCCAAATACTTTCATAGCTA
2108 ACATTGCCGTGGATATAGGCAAAAATTGCGATGATGCATATGCCTATTTTGCCGATGAATTG
2109 GGATACGGCTTGATTGCTTACTCCTGGGAACTGAACAAGTCCTGGAGATTCTCGGCACATT
2110 CGTATTTTTTCCCCGATCCATTGAGGGGCGATTTCAATGTCGCTGGTATTAACTTCCAATGG
2111 GGCGAGGAGGGTATATTTGGTATGTCCCTTTCGCCCATTCGATCGGATGGTTATCGTACCC
2112 TGTACTTTAGTCCGTTAGCAAGTCATCGACAATTTGCCGTATCCACGAGGATTTTGAGGGA
2113 TGAAACCAGGACGGAAGATAGCTATCATGACTTTGTTGCCTTAGATGAACGGGGTCCAAAC
2114 TCCCATACCACTTCACGTGTGATGAGCGATGATGGAATTGAGCTGTTCAATTTAATAGATCA
2115 AAATGCAGTGGGTTGCTGGCACTCATCAATGCCGTACTCACCGCAATTTCATGGCATTGTG
2116 GATCGCGATGACGTTGGCTTAGTTTTTCCGGCCGATGTGAAAATTGATGAGAACAAAAACG
2117 TTTGGGTTCTATCCGATAGGATGCCCGTTTTCTTGCTGTCTGACTTGGATTATTCAGATACT
2118 AATTTCCGAATTTACACGGCTCCCTTGGCCACTTTAATTGAGAATACTGTGTGTGATTTGAG
2119 GAATAACGCCTATGGGCCGCCAAATACCGTTTCAATACCAAAACAAGCCGTTTTGCCAATG
2120 GGTCCACCGTTATATACGAAACAATATCGTCCTGTCTTGCCACAGAAACCTCAGACCAGCT
2121 GGGCTTCCTCGCCGCCTCCTCCAAGTCGCACTTATTTGCCCGCCAATTCAGGCAATGTAGT
2122 CTCCAGTATTAGTGTCTCTACAAATTCTGTGGGTCCTGCAGGAGTGGAGGTGCCAAAGGC
2123 CTATATTTTCAACCAGCACAACGGCATAAATTACGAGACAAGTGGTCCCCATCTATTTCCCA
2124 CCCATCAACCCGCCCAACCGGGTGGCCAGGATGGTGGGTTAAAAACTTATGTGAATGCCC
2125 GCCAATCTGGGTGGTGGCATCATCAGCATCAAGGTTAACATAATCCTACACACGGTACTTG
2126 GGTATATTCTCACACACTCGATTGATGTAAAGAATATTTAAAGACAACAACATAGGGCAACA
2127 GCGGTTAAAAAAACCACATGACGTATGAGCAAGTGGCAAATCAATACTTTATCTAGTTATGT
2128 TAAGCAAAAAATAACAATAAATCAACTTTTTTTTGAAGGTTAAGAGTTTACGCAATTTTCTTG
2129 AGCGGAAAAAGCGGAAAAAATGTAAGTATGCATAAATTCTAAATATATCAACAACTGTACAT
2130 TTTCTGGAGTACTACTACCAGGCAAGAAAGTAGGTTGATAAAGCTATGCACAAGATCTTGTT
2131 TGGGTGCAGGGAAAGTTCAACTTAATCGCTCAATTTGAGATCGCCTGGTCGCTTGAGATTC
2132 GACTGTAATTGAAATTTTTGCTTTTGATCGGAGCCAGACTTCAGACGGGGCAAACAAAAAG
2133 ACTTTGTTGGTGGTAGGGTAGGATCCGTTGACCTGCAGGTTCC
2134
2135
2136
2137 oligos
2138 sequencing from pBH forward DLK349 GCCACCTCTGACTTGAGCGTCG

sequencing from pBH reverse	DLK356		
GAGTTTTCTCCTTCATTACAGAAACGGC			
sequencing from ywing2+ reverse	DLK406	GCACTGCGCTCAAACTATTAGTACCCC	
sequencing from ywing2+ reverse	DLK407	CGGAGCCAGACTTCAGACGGG	