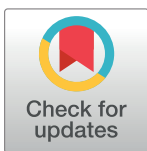# PLOS GENETICS

METHODS

High-efficiency CRISPR gene editing in *C. elegans* using Cas9 integrated into the genome

Matthew L. Schwartz, M. Wayne Davis, Matthew S. Rich[ORCID]**, Erik M. Jorgensen**[ORCID]*****

Howard Hughes Medical Institute and School of Biological Sciences, University of Utah, Salt Lake City, Utah, United States of America

* jorgensen@biology.utah.edu

Abstract

Gene editing in *C. elegans* using plasmid-based CRISPR reagents requires microinjection of many animals to produce a single edit. Germline silencing of plasmid-borne Cas9 is a major cause of inefficient editing. Here, we present a set of *C. elegans* strains that constitutively express Cas9 in the germline from an integrated transgene. These strains markedly improve the success rate for plasmid-based CRISPR edits. For simple, short homology arm GFP insertions, 50–100% of injected animals typically produce edited progeny, depending on the target locus. Template-guided editing from an extrachromosomal array is maintained over multiple generations. We have built strains with the Cas9 transgene on multiple chromosomes. Additionally, each Cas9 locus also contains a heatshock-driven Cre recombinase for selectable marker removal and a bright fluorescence marker for easy outcrossing. These integrated Cas9 strains greatly reduce the workload for producing individual genome edits.

Author summary

Germlines have evolved specialized mechanisms to protect themselves from invasions by transposons and viruses, which create barriers to genome editing techniques. For example, transgenes are silenced in the germline of the nematode *C. elegans*, thereby creating a barrier to CRISPR editing by Cas9. To facilitate gene editing, we built a collection of *C. elegans* strains in which Cas9 is never silenced. CRISPR is significantly more efficient in these animals, decreasing the effort researchers need to expend to get edited animals. The strains are available in multiple genetic backgrounds, and contain accessory transgenes to simplify downstream genetics. Together, these strains enable efficient, low-cost genome editing in *C. elegans*.

This is a *PLOS Genetics* Methods paper.

Introduction

CRISPR has revolutionized genetics by providing a simple method for targeted mutation, deletion or tagging of genes [1]. Since its initial demonstration, CRISPR has been rapidly adapted to facilitate gene editing in diverse species. Although the same core set of molecules are used to perform CRISPR editing in different species, methods must be optimized for individual species and applications. Efficient production pipelines are needed to minimize the workload to produce individual edits, and to facilitate systematic production of genome-wide resources.

In the nematode *C. elegans*, the Cas9 endonuclease is introduced into the germline via microinjection of either protein or DNA. In the protein method, Cas9 protein complexed *in vitro* with guide RNA is directly injected into the gonad along with a DNA repair template [2,3]. The injected ribonucleoprotein complex acts immediately in the gonad, producing targeted double-strand breaks within the meiotic oocyte progenitor cells. Co-injected repair templates can be used for homology-directed break repair, yielding edited F_1 progeny directly from the injected mother. In the DNA method, plasmids expressing Cas9 and the guide RNA are injected into the gonad along with a plasmid DNA repair template [4–6]. The injected plasmids are assembled into heritable extrachromosomal arrays in the injected gonad. Although cutting is occasionally observed directly in the injected gonad, templated repair events almost always occur in the germline of array-bearing F_1 animals. Edited progeny are typically recovered among the F_2–F_3 generations.

Direct injection of ribonucleoprotein complexes is fast and efficient–a single injected P_0 can yield dozens of independently edited F_1 progeny [3,7,8]. However, scalability is limited by two factors: (1) the reagents are expensive, and (2) edits must be identified by molecular or fluorescence screening of individual progeny because editing efficiency drops precipitously as repair templates exceed ~ 1.5 kb [9], precluding the use of selectable markers. Comparatively, plasmid reagents are inexpensive and can accommodate large repair templates, including selectable markers [4,5,10–12], but array-based edits appear a generation later, editing efficiency is low, and the plasmid-borne transgenes are rapidly silenced. In regard to efficiency, only 5–40% of injected P_0s yield a single edited progeny; most injected animals yield no modified progeny. Although selectable markers allow the rare edits to be identified without molecular screening, the effort required to injecting large numbers of animals limits throughput. Toward developing a pipeline for high-throughput editing in *C. elegans*, we sought to improve the efficiency of plasmid-based CRISPR editing by improving Cas9 expression in worms.

The bane of transgene-driven genome editing in *C. elegans* is silencing of transgenes in the germline. The worm germline detects and limits expression of foreign genetic elements using both sequence-specific and sequence independent mechanisms [13,14]. Germline silencing of array-borne transgenes is generally incomplete in the F_1 generation, but strengthens to completely eliminate expression in subsequent generations [15]. As the mechanisms of germline transgene silencing have been elucidated, effective strategies to mitigate silencing have been developed. These include reducing transgene copy number [16], increasing array complexity [15], removing piRNA complementarity [17], and including Periodic A/T Clusters (PATCs) within introns of the transgene [13,18]. We have leveraged this knowledge to produce Cas9 transgenes capable of constitutive germline expression.

Description of method

Here, we present a set of *C. elegans* strains engineered for high-efficiency CRISPR editing using plasmid reagents. The strains harbor integrated, single-copy insertions of germline-licensed Cas9 transgenes. To streamline post-CRISPR strain processing, each Cas9 locus also contains a heatshock-activated Cre recombinase to remove any selectable marker from the

edited locus. In addition, each locus is decorated with a fluorescent marker orthogonal to both GFP and mCherry to facilitate removal of the transgene by subsequent outcrossing. These strains greatly improve the efficiency of CRISPR modification using inexpensive plasmid reagents, providing a valuable addition to the *C. elegans* genome engineering toolkit.

Verification and comparison

Integrated Cas9 transgenes improve editing efficiency. Plasmid-based CRISPR editing strategies yield poor editing efficiency; typically only ~20% of successfully injected animals yield edited loci. Therefore to successfully edit with plasmids, many animals must be injected, and broods screened to identify rare events [10–12]. We found that Cas9 tagged with mCherry was not detectably expressed in the germline in most injected animals, consistent with previous observations [19]. These observations suggest that a lack of germline Cas9 expression is a limiting factor in plasmid-mediated editing. We therefore sought to improve germline Cas9 expression.

We tested whether an integrated single-copy transgene could drive constitutive germline Cas9 expression and generate guided double-strand breaks. We inserted Cas9 transgenes into the ttTi5605 Mos1 transposon insert site on chromosome II using MosSCI [16]. This locus is known to be licensed for germline expression [16]. To drive germline-specific expression, Cas9 transgenes were built with the *mex-5* promoter and *tbb-2* 3'UTR. We tested Cas9 efficiency in two strains. In one strain the Cas9 transgene was a cDNA interrupted by a single synthetic intron (derived from Addgene plasmid #46168) [6]. In another strain the Cas9 transgene was optimized for germline expression by including 4 large PATC-rich introns from the gene *smu-2* (Fig 1A)[13,18].

Each Cas9 expression strain was tested for basal activity by testing for double-strand break formation using the high-efficiency *dpy-10* sgRNA. Animals were injected with a plasmid expressing a *dpy-10* single-guide RNA (sgRNA). We monitored array-positive progeny for dumpy (Dpy) or rolling (Rol) phenotypes consistent with mutation of *dpy-10*. In the strain bearing the germline-optimized Cas9 construct (EG9615), ~25% of array-bearing F_1 animals (18/69) exhibited a Dpy or Rol phenotype. To determine if these were heritable mutations, we singled F_1 animals. Roughly half of these (8/15 singled F_1s) were caused by germline mutations, that is, the singled animals produced either 100% Dpy or one quarter Dpy progeny, consistent with germline mutation. By contrast, in strains expressing the Cas9 cDNA construct, we failed to detect *dpy-10* mutant animals among similar numbers of array positive progeny; this strain was not characterized further. Together, these data indicate that a single-copy transgene can express sufficient Cas9 to introduce cuts in the genome, but it must be licensed for germline expression.

As a more demanding test of Cas9 activity, we knocked a GFP tag into the N-terminus of SNB-1 (Fig 1B). GFP insertions were selected by co-insertion of an *unc-119(+)* transgene nested in an intron of GFP [12]. Editing efficiency was compared to plasmid injection of the germline-optimized Cas9 transgene, as well as injecting plasmid expressed Cas9 cDNA [12]. To normalize for variable injection quality, genome edits were only quantified from injected animals that generated heritable extrachromosomal arrays and were therefore verified as 'successfully injected'. Edits were observed in 100% of successfully injected P_0s using the integrated germline-optimized Cas9 transgene (Fig 1C), which is a profound improvement over our results using plasmid-injected cDNA (20% P_0s). Injecting the germline-optimized Cas9 transgene as a plasmid improved injections marginally (40% P_0s) and was silenced thereafter; this construct was not characterized further. The integrated transgene was then tested on a variety of genes to determine if it consistently outperformed plasmid injection of the cDNA.
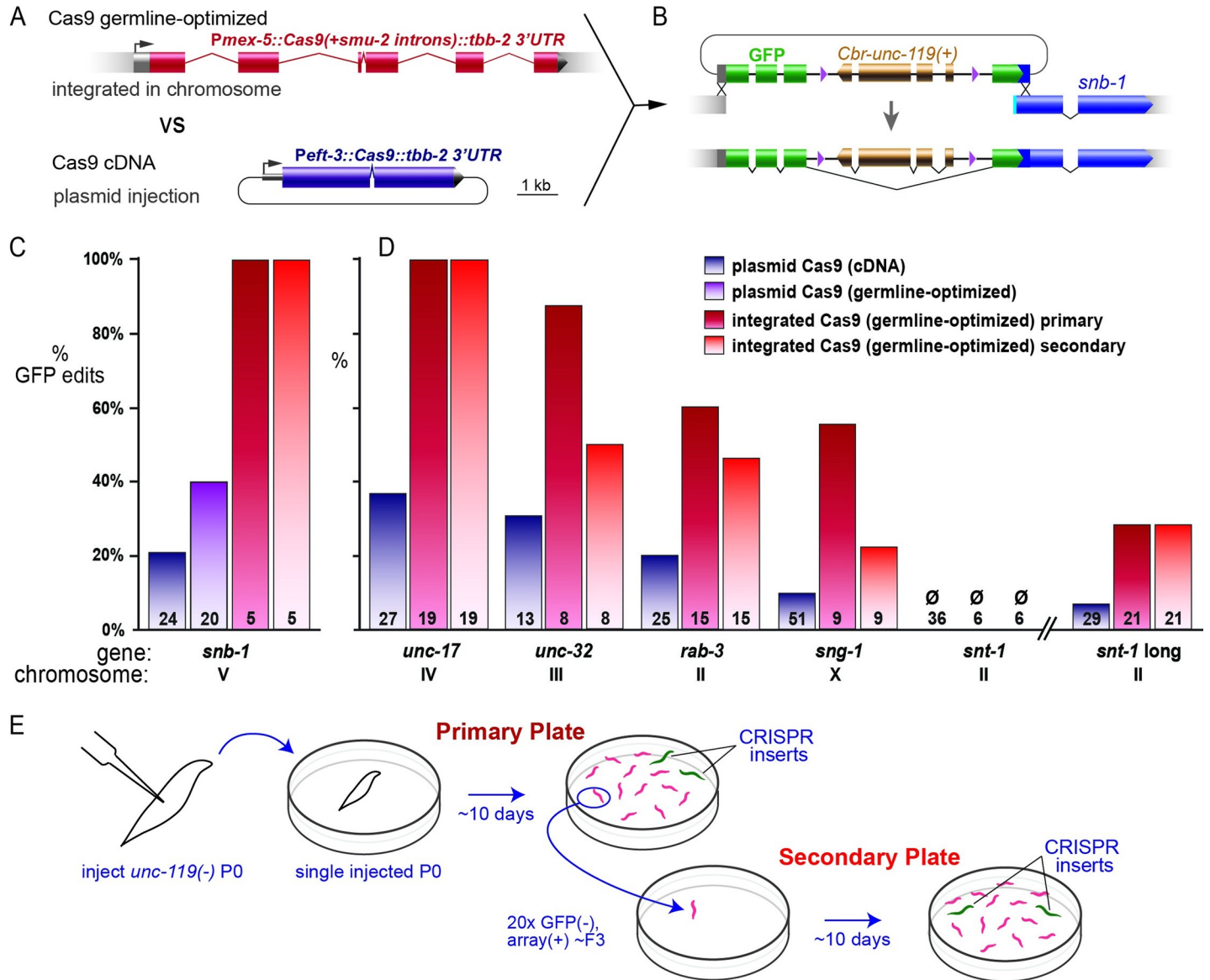
Fig 1. Germline-optimized Cas9 improves editing efficiency. A. Gene diagrams representing the germline optimized P*mex-5*::*Cas9* MosSCI construct and the standard plasmid P*eft-3*::*Cas9* expression construct. **B.** Diagram of GFP tagging of *snb-1* using the intron-nested *Cbr-unc-119* selectable marker. **C.** Comparison of CRISPR tagging efficiency at *snb-1* using array-borne P*eft-3*::*Cas9* containing a synthetic intron (dark blue) or germline-optimized introns (light purple) vs integrated, germline optimized P*mex-5*::*Cas9* (dark red and light red). Insertion frequency is the fraction of successfully injected P_0s which gave rise to a CRISPR tag among their progeny on the primary or secondary plate (dark red and light red, respectively). Numbers on bars indicate number of successfully injected P_0s. For germline-optimized Cas9, the tag was always GFP and inserts were confirmed by visual confirmation of GFP expression. For plasmid Cas9, tagging data is reproduced from a previous study with permission [12]. For plasmid Cas9 (dark blue and light purple), populations were only incubated until starvation of the primary plate. **D.** GFP insertion efficiency at multiple loci. Editing frequencies were calculated as in (C). **E.** Diagram of multigenerational editing experiment. For integrated Cas9 (dark red and light red), injected P_0 animals were singled on 'Primary' plates. Upon starvation, 'Primary plates' were scored for the presence of GFP(+) animals, and 20 array(+), GFP(-) animals were moved to a 'Secondary plate'. Upon starvation, 'Secondary plates' were inspected for GFP(+) animals that arose from new CRISPR events on the secondary plate. Each injected P_0 was assigned a unique identifier so that editing outcomes for individual lineages in both phases of the experiment could be collated (S1 Table).

We injected targeting plasmids specifying GFP insertion at five other synaptic genes previously used to assay insertion frequency using array-borne Cas9 [12]. For each locus, the targeting reagents comprise a repair template plasmid and a guide RNA expression plasmid. All genes were initially targeted with "short-arm" repair templates ($\leq$ 57 bp) assembled by SapTrap [12]. The *snt-1* gene was additionally targeted by a second "long-arm" construct (150 and

450 bp). Across all loci, GFP insertion per successfully injected P0 was 2.5- to 7-fold more frequent in the integrated Cas9 strain than in strains expressing Cas9 from extrachromosomal arrays (Fig 1D). The short-arm *snt-1* targeting strategy uses a guide RNA predicted to exhibit low activity due to a "C" base immediately preceding the PAM site (CNGG) [3,20,21]. The enhanced activity from integrated Cas9 was unable to overcome this guide RNA deficiency; whereas successful integration was observed for an upstream cut site requiring a long-arm template. For all other short-arm constructs, between 50% and 100% of injected P0's yielded tagged inserts. Thus, the integrated Cas9 transgene facilitates more efficient CRISPR editing than array-borne Cas9.

Integrated Cas9 transgenes facilitate multigenerational editing. We next explored whether the single-copy Cas9 transgene could support editing over multiple generations. In general, array-based genome editing in *C. elegans* is limited to a few generations after the initial injection, even though the editing arrays can transmit indefinitely. This is likely due to germline silencing of the Cas9 transgenes on the extrachromosomal arrays. The integrated and germline-licensed Cas9 is not silenced; all of the GFP targeting constructs (Fig 1) were injected in an integrated Cas9 strain more than ten generations after the initial integration of Cas9. Although the integration provides continuous Cas9 expression, it was not known if germline silencing mechanisms would limit guide RNA expression or repair template availability after the first few generations. We therefore monitored our GFP-tagging injections for new editing events beyond the F_2 generation, when germline silencing is usually established [15].

To detect new CRISPR events in later generations, we propagated unedited, array-bearing worms past the F_3 generation (that is, beyond starvation of the primary plate). Successful edits are easily scored by bright GFP expression in the synapse-rich nerve ring, which allowed us to reliably pick unedited individuals using a fluorescence dissecting microscope. Upon starvation of each plate founded by a single injected P_0 animal, we singled 20 animals that carried the array (*unc-119*+) but lacked an edited target locus (GFP-). When these 'secondary' plates starved (approximately 3 generations), the plates were inspected for GFP(+) worms (Fig 1E). For all except the defective short-arm *snt-1* target, new GFP insertions occurred on the secondary plates, albeit usually at a slightly lower rate than on the primary plates (Fig 1D). In most cases, lineages that generated inserts on the secondary plate had produced CRISPR insertions on primary plate (S1 Table). Interestingly, four lineages that failed to generate CRISPR inserts on the primary plate nevertheless yielded GFP inserts on the secondary plate. Thus, in the integrated Cas9 strain, *de novo* editing events routinely occur in generations after germline silencing of transgenes is normally established. Importantly, guide RNA expression and repair template accessibility are less sensitive to germline silencing mechanisms than Cas9 expression.

Building multiple Cas9 expression strains optimized for genome editing pipelines. Since integrated Cas9 transgenes delivered significantly higher success rates in GFP tagging experiments, we developed optimized strains to streamline the editing pipeline for the *C. elegans* community. These Cas9 strains include:

1. A fully outcrossed and uniform genetic background for *unc-119* rescue.

2. Cas9 integrations on different chromosomes.

3. A linked, inducible Cre recombinase to excise the selection marker in the edited locus.

4. A fluorescent marker in the Cas9 locus to facilitate out-crossing.

5. At least one Cas9 integration tightly linked to *unc-119*, so that the selectable marker is outcrossed simultaneously with the Cas9 transgene.

Inserting Cas9 transgenes in a clean genetic background. To generate a uniform genetic background, we thawed the standard N2 strain and generated a *de novo unc-119* mutation. We injected plasmids encoding Cas9 and a guide RNA targeting the region near the *ed3* allele of *unc-119*, and relied on error-prone repair to generate a new loss-of-function allele. We isolated a single *unc-119(-)* strain and designated the new allele *unc-119(ox819)*. Direct sequencing of *unc-119* revealed *ox819* is an 11 bp deletion flanking the targeted cut site (S1 Fig). *ox819* causes a frameshift after V110 (UNC-119a) and then appends 29 out-of-frame amino acids before a stop codon. *unc-119(ox819)* animals are phenotypically identical to *unc-119(ed3)* animals and can be rescued by expression of the smaller *unc-119* gene from the related nematode *C. briggsae* (*Cbr-unc-119*). The strain carrying *unc-119(ox819)* represents a relatively clean genetic background for gene editing since it was generated by targeted rather than random mutagenesis to limit secondary mutations, and the strain was immediately frozen to minimize the emergence of suppressor mutations during propagation of the strain in laboratory stocks.

To insert Cas9 expression cassettes at a variety of loci in the genome, we randomly integrated the germline optimized Cas9 expression cassette into the *unc-119(ox819)* background using a miniMos transposon (Fig 2). The transposon carried a Cas9 expression cassette, P*mex-5*::*Cas9* (+*smu-2* introns)::*tbb-2* 3'UTR, and a lox-flanked *unc-119(+)* locus for selection. We isolated 14 inserts and mapped the miniMos insertion site by inverse PCR (S2 Table). We first assessed germline Cas9 activity in the original *Cbr-unc-119+* isolates by testing the ability to cut the *dpy-10* locus. We injected the plasmid expressing the *dpy-10* sgRNA and assayed for the presence of Dpy or Rol animals. Second, after removing the floxed *Cbr-unc-119* rescue cassette through Cre expression, we assayed for efficient CRISPR-mediated GFP insertion into the *unc-32* gene (S2 Table). This screen identified four strains in which (1) Cas9 activity was high enough for robust targeted GFP insertion (> 50% of injected P0's gave rise to *unc-32*::*GFP* progeny), and (2) the Cas9 expression cassettes were distributed broadly across autosomes.

In three of the four strains, the Cas9 expression cassettes were inserted in unnamed genes without exacerbating phenotypes. Remarkably, in the fourth strain, the transposon (*oxTi1083*) landed in the *unc-119(ox819)* locus on chromosome III, satisfying the goal of a Cas9 integrant tightly linked to *unc-119*. The screen identified at least one strain (inserted into *Y69H2.3* on chromosome V) capable of targeting *dpy-10* but not capable of supporting homology-dependent GFP tagging of *unc-32* (S2 Table). Generating mutations in *dpy-10* by error-prone repair will be more efficient than inserting GFP through homology-dependent repair. The *Y69H2.3* insertion may express Cas9 weakly, which permits efficient untemplated mutagenesis but fails to support less efficient insertion events. Together, our screen results are consistent with previous observations that the local genomic environment can influence the overall level of germline expression [18], and that different levels of germline Cas9 expression can lead to varying editing efficiency.

Decorating Cas9 loci with inducible Cre and a fluorescent marker. We next incorporated a Cre locus and fluorescent marker to each locus to simplify genetic manipulations. To modify the miniMos insertions, we built a tagging construct that targets insertions into the 5' arm of the miniMos element (Fig 2, step 4). The targeting vector contains a lox2272-flanked *Cbr-unc-119* cassette to select for the modified miniMos insertion. After insertion, heat-shock activation of Cre removes the *unc-119(+)* transgene.

To each of the Cas9 loci we added a heatshock-driven Cre transgene (P*hsp16.41*::Cre::*tbb-2* 3'UTR). For tagging genes, we use a selection strategy in which a floxed *Cbr-unc-119* is nested in an intron of the inserted tag, thereby rescuing the *unc-119(ox819)* mutation at the genomic locus. The *Cbr-unc-119* gene is excised from the edit, leaving only a loxP site within an intron in GFP. Integrated heat-shock Cre eliminates the need for a secondary injection of the Cre-
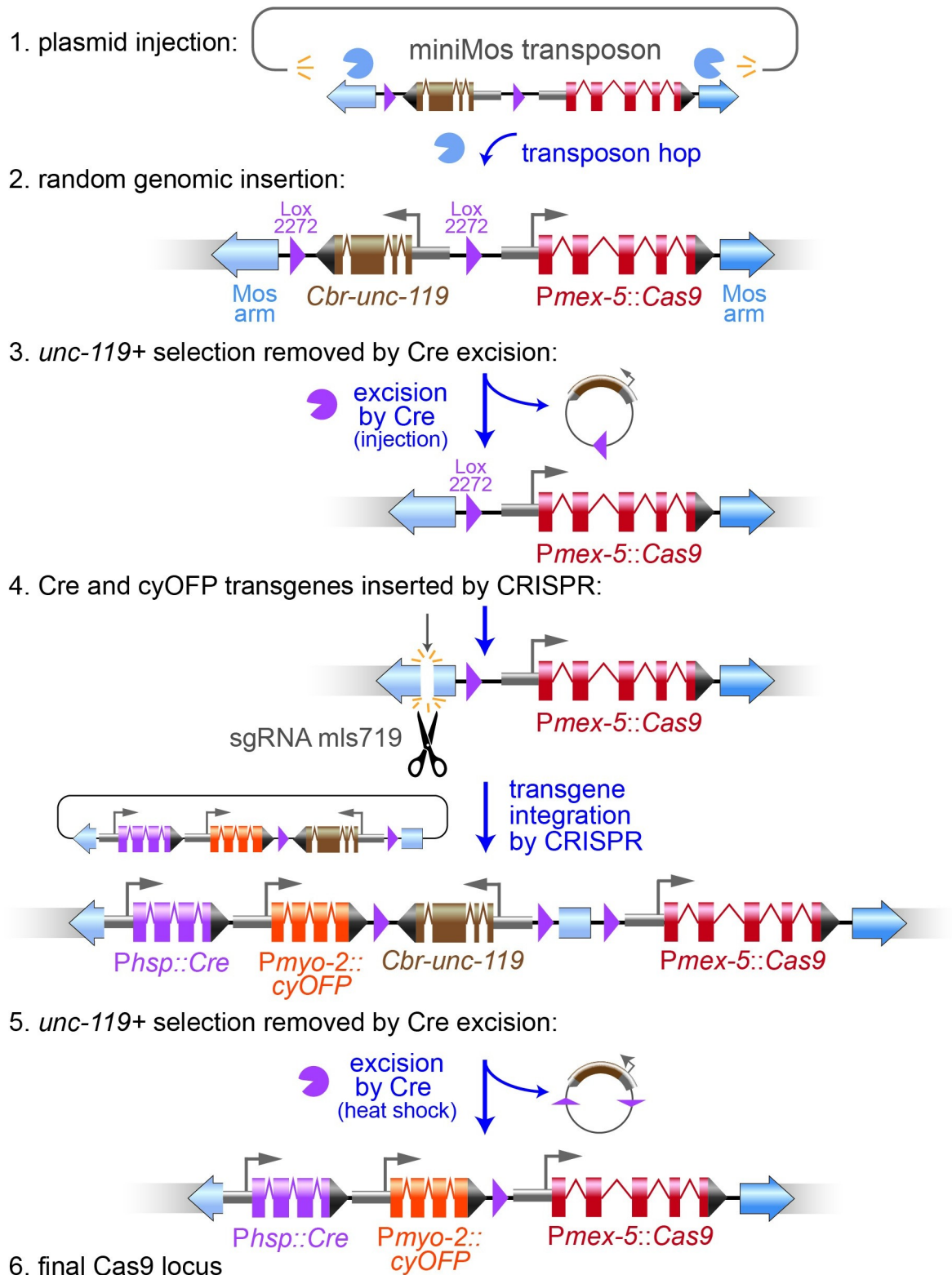
Fig 2. Building marked Cas9-Cre loci. (1) Cas9 transgenes were generated by injecting a miniMos transposon containing a germline-optimized Cas9 and a floxed *Cbr-unc-119* selectable marker into *unc-119(ox819)* mutants. (2) MiniMos elements inserted into the genome randomly and injection sites were mapped by inverse PCR and confirmed by directed PCR. (3) The *Cbr-unc-119* cassette was removed by injecting a plasmid expressing Cre. (4) The loci were then modified by CRISPR to include a Cre transgene and an Orange Fluorescent

expressing plasmid to excise the selectable marker and drives more efficient marker excision (Dickinson et al., 2015).

To simplify outcrossing of the Cas9 locus, we generated additional strains that incorporated a fluorescent marker. In these cases the CRISPR template contained transgenes for both Cre and an orange fluorescent protein (P*myo-2*::*2xNLS-cyOFP*::*let-858* 3'UTR*)*. cyOFP (cyan-excitable Orange Fluorescent Protein) is a long-Stokes-shift fluorescent protein that is spectrally separable from common green and red fluorophores. The nuclear localized, P*myo-2*-driven transgene produces a bright fluorescence pattern that is limited to the pharynx (S2 Fig); the marker is brightest in late larvae (L3 and L4) but dimmer in adults, providing a marker that is easy to track through crosses but that minimally obscures weak fluorescence signals from tagged target genes.

These modifications highlight the use of plasmid reagents to introduce single-copy transgenes into the genome by CRISPR (Fig 2, step 4). Although these single-gene constructs are larger than a simple GFP tag, the efficiency of transgene insertion was grossly consistent with GFP tagging. Specifically, for each transgene insertion, we recovered multiple independent insertions from ~10–20 injected P_0s. These observations suggest that integrated Cas9 can provide an alternative to MosSCI for single-copy gene insertion, particularly given that site selection is not limited to the existing characterized Mos1 insertion sites.

We quantified the activity of both Cre and Cas9 in each of the final strains. We assayed Cre activity by monitoring the efficiency of excision of the *Cbr-unc-119*+ marker from the locus (Fig 2, step 5). Heatshock-driven Cre acts efficiently from integration sites in the self-excising cassette (SEC) platform [10]. We found that heatshock-driven Cre excised *Cbr-unc-119*+ in ~50% of the progeny of heat-shocked animals (Fig 3A). We observed a low amount of Cre activity from the heat-shock transgene even in the absence of a *bona fide* heat-shock treatment. This 'leaky' Cre expression leads to marker excision events when animals are reared at 20–25˚C but is mostly eliminated by rearing animals at 15˚C.

After excision of the *Cbr-unc-119* marker from the Cas9-Cre locus, we quantified Cas9 activity in each strain by measuring the insertion frequency of *unc-32*::*gfp* (Fig 3B). The Cas9 insertions on chromosomes I and III maintained moderate-to-high levels of Cas9 activity. Unexpectedly, the Cas9 insertion on chromosome V (*W03F9.11*) was silenced by the addition of the cyOFP transgene (P*myo-2*::*2xNLS-cyOFP*::*let-858* 3'UTR). Multiple independent insertions of the cyOFP marker into the chromosome V site silenced Cas9 activity. Additionally, insertion of the fluorescent marker up to 0.3 map units away from the Cas9 transgene also resulted in silencing of Cas9 (see Methods). Thus, we were only able to produce fluorescently marked alleles for the chromosome I and chromosome III insertions. However, Cas9 expression remains robust in a simple version with just the Cre transgene inserted in the chromosome V Cas9-Cre locus.

Finally, to facilitate use with selectable markers other than *Cbr-unc-119*, we outcrossed the fluorescently marked chromosome I (*W01A8.6*) and chromosome III (*F53A2.9*) Cas9-Cre transgenes to a wild-type background. Because one of the Chr III inserts disrupts *unc-119* directly, this allele could not be segregated from an *unc-119* mutant background. Together, we have produced a suite of Cas9 and Cre expressing strains capable of supporting efficient plasmid and array-based CRISPR editing in *C. elegans*.

A Cre efficiency at Cas9 integration sites
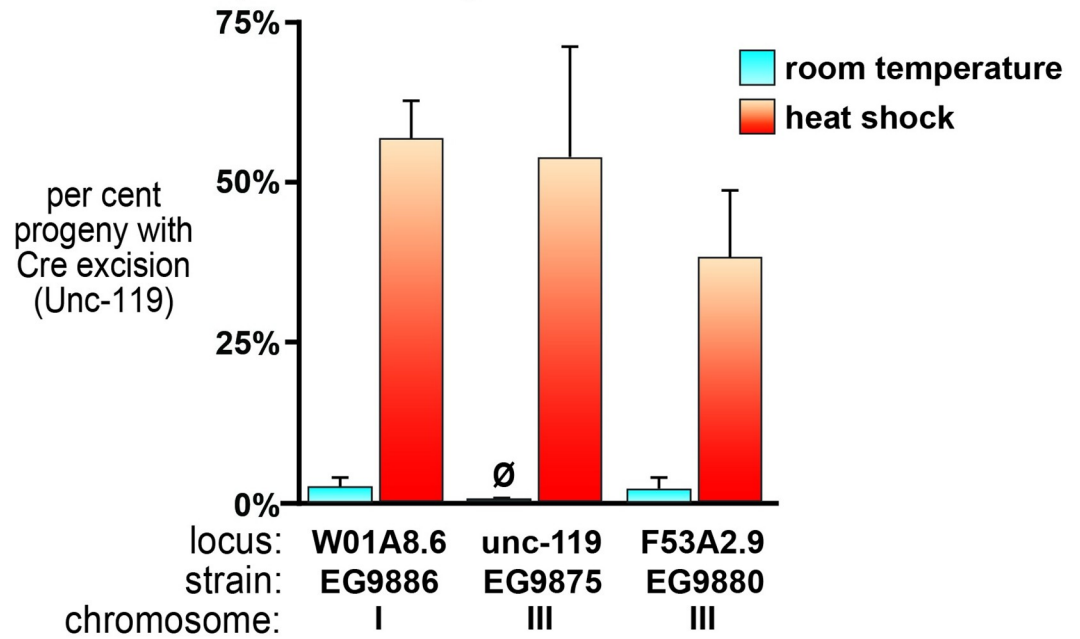
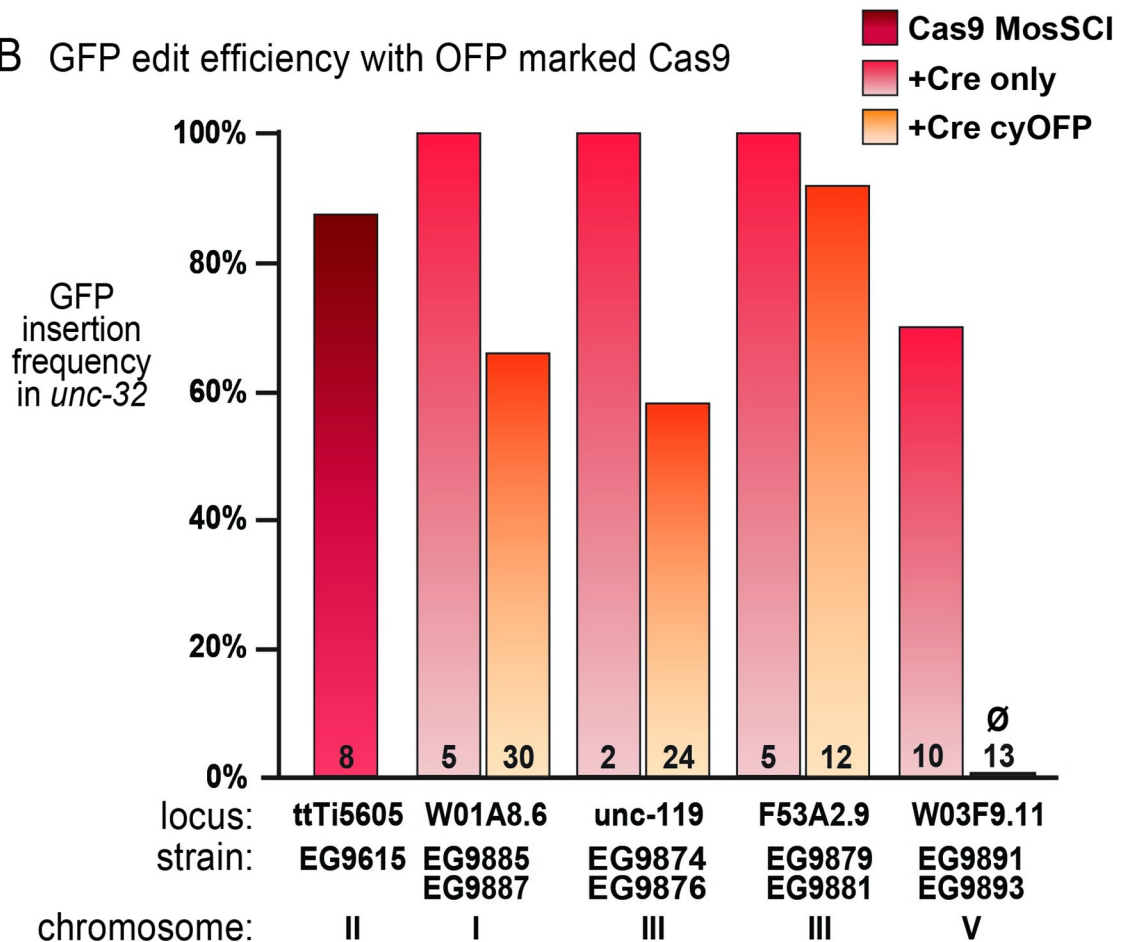## B GFP edit efficiency with OFP marked Cas9

Fig 3. Cre and Cas9 activity from integrated Cas9 loci. (A) *unc-119(+)* marker excision using integrated heat-shock Cre transgene. Cohorts of ~20 homozygous Unc-119(+) young larvae (L1 –L2) were either incubated at room temperature or at 32˚C for 2 hours ("heat shock"). After reaching adulthood, several hundred eggs (n = 223–1025) were collected from each cohort, and each hatched animal was scored for the presence of Unc-119(-) phenotype. Bars represent standard deviation for 3 independent experiments. **(B)** Insertion of GFP into the *unc-32* gene using CRISPR. The *unc-32* guide RNA was expressed from an injected plasmid; Cas9 was expressed from a transgene integrated at the indicated locus. A *Cbr-unc-119(+)* transgene is contained within an intron of the GFP tag. The loci either contained the Cre transgene alone (light red bars, top strain name) or the Cre and the Orange Fluorescent Protein transgenes (orange bars, bottom strain name). The first strain (EG9615) contains just the Cas9 transgene at the MosSCI insertion site (ttTi5605). Configurations of the transgene loci in these strains is illustrated in S4 Fig. Insertion frequency is the fraction of successfully injected P_0s which gave rise to a CRISPR tag insertion among their progeny. Numbers on bars indicate injected P_0s, chromosome indicates the location of the insertion containing the Cas9-Cre locus.

https://doi.org/10.1371/journal.pgen.1009755.g003

Discussion

In summary, we have generated specialized strains for efficient genome editing by CRISPR. These strains express the Cas9 enzyme in the germline constitutively, and can insert tags in a variety of genes at a success rate of 50%-100% of injected worms.

There are four important features to these strains that permit high expression of Cas9. First, to avoid germline silencing of the construct, Cas9 is expressed from an integrated transgene rather than from an injected plasmid. In *C. elegans* injected DNA forms extrachromosomal arrays, which are usually silenced in the germline [15,22,23]. Second, we inserted introns containing periodic A/T clusters (PATCs) into the Cas9 transgene. These clusters are simply a few adjacent A/T base pairs present with a 10 bp periodicity on a single face of the DNA molecule [24]. Inclusion of introns containing PATCs facilitates expression of genes in the *C. elegans* germline [13,18]. This specialization may protect the germline from invasion by transposons or viruses. Because foreign DNA is unlikely to contain these specialized introns, they will not be expressed in the germline. Third, because not all regions of chromosomes are permissive for germline expression, we tested 15 chromosomal integration sites, and 5 of these insertions expressed sufficient levels of Cas9 for efficient gene tagging with GFP. Fourth, we labelled 4 of these sites with a fluorescent marker and retested tagging efficiency; 3 of the 4 sites retained high Cas9 expression with the presence of the fluorescent marker.

The immediate advantage of the Cas9 expression strains is that they greatly reduce the injection burden for producing CRISPR edits with plasmid reagents. Previously, plasmid-based protocols required ~five times more injection effort than direct injection of the Cas9 protein. The significant improvement in efficiency afforded by the Cas9 expressing strains brings the injection burden for plasmid reagents on par with protein injections [8]. However, Cas9 ribonucleoprotein complex injections are likely to remain the preferred method for short edits including the introduction of point mutants. On the other hand, for gene tagging, transgene insertions, and gene-deletions that require long repair templates, the enhanced efficiency of Cas9 expressing worms coupled with selection within the tag will make plasmid-based editing an important alternative.

In previous DNA protocols Cas9 was expressed from a plasmid; edits largely occurred in the F1 germline only, apparently because the transgene in the extrachromosomal array is silenced thereafter. When Cas9 is expressed from an integrated transgene, editing is not silenced in subsequent generations, and unedited strains that carry the array will continue to yield GFP insertions in later generations. Continued editing capability suggests that the guide RNA plasmid expressed by a U6 RNA polymerase III promoter is still active on the extrachromosomal array, and the DNA repair template with the GFP tag is still a substrate for double-strand break repair of the cut site.

We have cultivated Cas9 expressing strains for over a year of continuous culture without loss of Cas9 activity, and we have performed multiple crosses with Cas9 transgenes and have

never observed silencing in progeny strains. Despite the observed robustness of Cas9 expression, we recommend maintaining a frozen stock of these strains and periodically replacing old cultures with fresh thaws. This practice will maintain Cas9 activity and limit potential off-target mutations from Cas9, which have not been characterized.

Nevertheless, there is variability in editing efficiency among our standard set of test loci for gene tagging, but this variability is unlikely due to germline silencing. The three genes that were edited most efficiently on both primary and secondary culture plates are essential genes, whereas the three genes that were edited less efficiently are non-essential. One potential explanation for this observation is that the target sites are being cut repeatedly until either a templated or untemplated mutagenic repair occurs. For essential genes, the untemplated repairs result in lethal alleles that are lost from the population. For non-essential genes, untemplated repairs create viable mutations that lack the targeted cut site. Indeed, limited sequencing of 'dead-end' strains revealed mutated guide-RNA binding sites (S3A and S3B Fig).

In addition to damaged target sites, there are two potential outcomes that are false positives. We call these (1) 'dark arrays' and (2) 'dark inserts'. Both these modes mimic *bona fide* inserts, because they exhibit *Cbr-unc-119(+)* rescue and lack fluorescent array markers. Dark arrays lack detectable fluorescent markers, either because the markers were never incorporated or because the markers were excised or silenced during propagation. Lines harboring dark arrays become apparent in the F1 because of low transmission of *unc-119* rescue; whereas genomic inserts are transmitted to 75% or 100% of the progeny by chromosomal segregation. Arrays lacking one or more injected plasmids have been observed in other forms of array-based genome editing ('false positives', wormbuilder.org).

Dark inserts, on the other hand, are insertions in the genome in which GFP is not expressed. Dark inserts lack array markers and are rescued for *unc-119*, but do not express GFP and the genotype cannot be confirmed by PCR. These are due to mutated or partial edits, or off-target insertions of the *Cbr-unc-119(+)* cassette into the genome. Of 206 lines fully analyzed in this study, we identified 3 plates containing animals lacking array markers but also lacking a GFP insert. Thus, the observed rate of these failure modes is below 5% of injected P0's. Through whole genome sequencing we characterized two off-target inserts. Both are inserted non-specifically in the genome, likely into off-target Cas9 cut sites or other double-strand breaks (S3C Fig).

Neither dark arrays nor dark inserts are specific to integrated Cas9. Both are observed at similar frequencies when Cas9 is expressed from an injected plasmid. Importantly, both of these failure modes can be easily identified and excluded from further processing.

To build a versatile genetic toolkit for CRISPR modifications, we inserted Cas9 transgenes at various sites in the genome. We tested 15 integration sites, 5 of these insertions expressed sufficient levels of Cas9 for efficient gene tagging with GFP and were on diverse locations in the genome (S4 Fig). Four of these loci were modified to include a Cre transgene to remove the selection marker from the edited gene, and 3 modified to include a fluorescent marker so that the Cas9 transgene could be easily outcrossed to generate the final edited strain (S5 Fig). We include a protocol to illustrate the steps for successful gene editing and outcrossing (Fig 4 and S1 File).
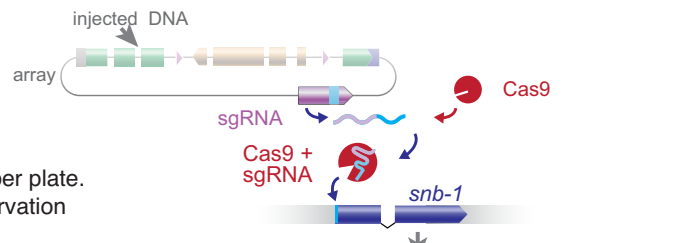
Methods

Reagents

All chemicals were purchased from Sigma-Aldrich (St. Louis). All enzymes were purchased from New England Biolabs (Beverly, MA). All synthetic DNAs were purchased from Integrated DNA Technologies.
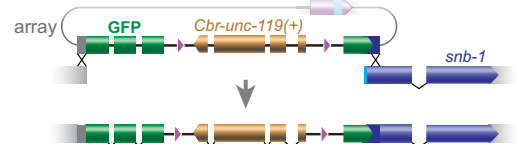
1. starting strain
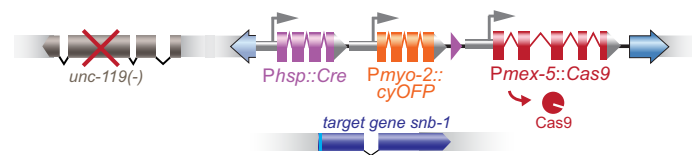
Day 0: grow at 15°C on HB101 for optimal injection quality.

2. inject DNA

Day 1: Inject 10-20 P0s. Pool 2-3 injected animals per plate.
Day 2: Incubate at 18-20°C until plates are near starvation (~2-3 generations).

3. select non-Unc, array(-) progeny

Day 8: Screen plates for *unc-119*(+) and array(-) animals. Clone up to 2 animals from each plate to new HB101 plates.
Day 8-10: Incubate at 15°C to homozygose *unc-119*(+) tag.
Day 10: Genotype homozygous strains by GFP PCR and sequencing to confirm correct repair.
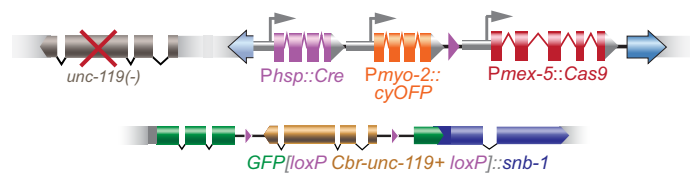
4. heat shock to excise selection

Day 11: Pool 20 L1-L2 animals and heatshock at 32°C for 2 hours.
Day 11-14: Incubate heatshocked plates at 15°C
Day 14: Isolate *unc-119*(-) animals. *NB: these animals grow slower than wildtype.*

5. outcross unc-119(-) and Cas9

Day 21: Cross N2 or balancer males to GFP(+) *unc-119*(-) L4s.
Day 25: Clone GFP(+) unc-119(+) CyOFP(-) animals. *Enjoy!*

Fig 4. Gene tagging with GFP using integrated Cas9. Protocol for tagging *snb-1* using integrated Cas9. (1) Select a strain with integrated Cas9 that is not tightly linked to the gene of interest, and contains *unc-119(ox819)* if a selectable marker is desirable. (2) Inject P0s with plasmid DNAs that encode the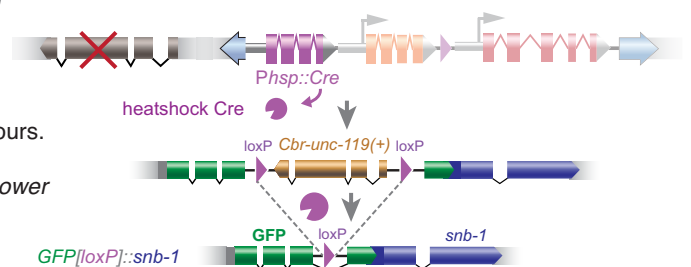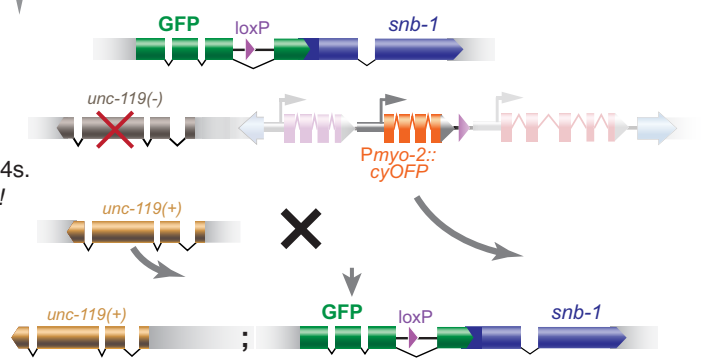 guide RNA, your DNA template for gene tagging, and 'red' fluorescent markers to mark the array. (3) When P0 plates are starved, single non-Unc, non-'Red' animals, and confirm the presence of the edit by fluorescence or PCR. (4) Heatshock the line express Cre and excise the *Cbr-unc-119(+)* selectable marker. In the next generation, single Unc progeny. (5) Cross the strain to wild type, single heterozygotes, and then single non-orange individuals from the progeny to establish the tagged strain.

C. elegans strains

C. elegans strains were cultured using standard methods [25] on nematode growth media (NGM) feeding on OP50 or HB101 bacteria. Animals were maintained at 15˚C, 18˚C, and 20˚C. A list of worm strains created in this study can be found in S3 Table.

Molecular biology and cloning

A list of plasmids used in this study can be found in S4 Table. All novel plasmids were generated using standard molecular biology techniques. Annotated sequences for all plasmids used in this work are included in S1 File.

General injection procedures. All injections were conducted into young adult (<24 hours) hermaphrodites reared at 15˚C on either OP50 (wild type worm strains) or HB101 (*unc-119* mutant worm strains). For plasmid injections, plasmids were purified using the *Purelink* kit (Thermofisher). Unless noted otherwise, the final total concentration of all injected plasmids was 100 ng/μl. Arrays were marked by inclusion of either 1xRed co-injection markers [2 ng/μl pCFJ90 (P*myo-2*::*mCherry*), 4 ng/μl pGH8 (P*rab-3*::*mCherry*) + 4 ng/μl pCFJ104 (P*myo-3*::*mCherry*)] or 1xGreen co-injection markers [2 ng/μl pCFJ91(P*myo-2*::*gfp*) + 8 ng/μl pCFJ421 (P*eft-3*::*gfp*::*h2b*)]. For CRISPR injections with a single plasmid containing both the repair template and guide RNA expression cassette, the plasmid was included at 65 ng/μl. For CRISPR injections with the repair template and guide RNA expression cassette on different plasmids, the repair template was included at 60 ng/μl and the guide RNA expression plasmid was included at 30 ng/μl. When included, Cas9 expression plasmids were included at 30 ng/μl (cDNA) or 2 ng/μl (with smu-2 introns); higher levels of the germline-licensed Cas9 transgene were toxic. If needed to bring injection mix concentrations to 100 ng/μl, pBluescript(sk+) was used as "stuffer" DNA. After injection, P_0s were either singled to fresh OP50 plates (when quantifying event frequency per injected P_0) or pooled in groups of 3–5 (for construction operations) and incubated at 25˚C for 3–10 days prior to screening.

Transposon-mediated Cas9 insertion. All MosSci injections were conducted with young adult EG6249 worms. MosSci injection mixes contained 25–40 ng/μl ttTi5605-targeting vector, 50 ng/μl pCFJ601 (P*eft-3*::*Mosase*), 10 ng/μl 1xRed markers, and 10 ng/μl pMA122 (P*hsp-16.41*::*peel-1*) (95–110 ng/μl total). MiniMos injections were conducted with young adult EG9814 *[unc-119(ox819) III]* worms. MiniMos injection mixes contained 50 ng/μl pCFJ601 (P*eft-3*::*Mosase*), 40 ng/μl pMLS714 (miniMos targeting vector), and 10 ng/μl 1xRed markers. Initial MosSci inserts (oxSi1196 and oxSi1091) were built using the standard pCFJ150 targeting vector. To allow removal of *Cbr-unc-119* after transgene insertion, we built general targeting vectors for MosSci at the ttTi5605 locus (pMLS640) and for MiniMos (pMLS713) containing a *lox2272*-flanked *Cbr-unc-119* cassette. In all cases, genomic inserts were identified by screening for Unc-119+ animals that lacked extrachromosomal array markers. The floxed *Cbr-unc-119* was removed from strains by injecting an array containing 50 ng/μl pDD104 (P*eft-3*::*Cre*) and selecting for *unc* F_2 progeny.

Cas9 locus modification. Cas9+, *unc-119*- animals were injected with 90 ng/μl pMLS719 (miniMos-targeting sgRNA), 10 ng/μl either pMLS716 (P*hsp-16.41*::*Cre* + *lox2272* flanked

Cbr-unc-119) or 2 ng/μl pMLS791 (P*hsp-16.41::Cre* + P*myo-2::2x-nls-cyOFP* + *lox2272* flanked *Cbr-unc-119*) + 10 ng/μl 1xRed marker. Cas9 activity was provided by the integrated Cas9 transgene. Injected P_0 were incubated at 25°C for 10 days. Inserts were selected by *unc-119* rescue and the absence of fluorescence array markers. To remove the *Cbr-unc-119* marker, larval animals were heat shocked at 32°C for 2 hours. *unc-119-* animals were selected by phenotype from the F_1 progeny of heat-shocked animals.

Measuring Cas9 activity with dpy-10 sgRNA. For initially screening MosSci Cas9 strains, worms were injected with a low concentration of *dpy-10* sgRNA plasmid to improve our dynamic range: 10 ng/μl pMLS597 (P*U6::sgRNA(dpy-10)*), 10 ng/μl 1xRed markers, and 80 ng/μl stuffer DNA. For screening miniMos-based Cas9 strains, worms were injected with 90 ng/μl pMLS597 (P*U6::sgRNA(dpy-10)*) and 10 ng/μl 1xRed markers. The higher concentration of pMLS597 resulted in a higher rate of germline mutagenesis of *dpy-10* among array(+) animals (18 Dpy or Rol animals /69 array+ animals for 10 ng/μl pMLS597; 71 Dpy or Rol animals/94 total array+ animals for 90 ng/μl pMLS597). Injected P0 animals were pooled on fresh OP50 plates and incubated at 25°C for 2–3 days. Array(+) F_1 animals were selected by fluorescence from the co-injection markers. Selected worms were later scored for Dpy and Rol phenotypes under white light.

Measuring tag (GFP) insertion frequencies. For quantifying insertion frequency at multiple target loci in EG9747, worms were injected with 65 ng/μl targeting vector, 10 ng/μl 1xRed markers, and 25 ng/μl suffer DNA. For quantifying insertion frequency in miniMos insertion strains, worms were injected with 90 ng/μl targeting vector (pMLS338) and 10 ng/μl 1xRed markers. We used only 65 ng/μl targeting vector when quantifying insertion frequency in EG9747 so that the targeting vector concentration in these injections matched the targeting vector concentration used for quantifying insertion frequency at these sites using plasmid-borne Cas9 (Schwartz and Jorgensen, 2016; data reproduced with permission). After injection, P0 animals were singled to OP50 plates and incubated at either 25°C or room temperature (20°C) until starvation. Starved plates were inspected for array+ and GFP+ animals. Insertion frequency was calculated as (number of plates with GFP + worms) / (number of successful injections), where a 'successfully injection' is an injection resulting in array+ worms in the F_2 generation.

Cas9 activity from the miniMos insertion at *W03F9.11* on chromosome V was silenced by insertion of the fluorescent marker P*myo-2::2xNLS-cyOFP::let-858 3'UTR*. To genetically mark the Cas9 transgene we inserted the cyOFP marker at two sites linked to *W03F9.11*. Both insertions resulted in silencing of Cas9 activity. The first landing site was 13.7 kb away between *W03F9.4* and *ttb-1* (bp154310). The second landing site was 141 kb away between *nra-4* and *flp-34* (bp1037379).

Quantifying floxed marker excision rates. For quantifying Cre activity, we used *unc-32::GFP(Cbr-unc-119+)* alleles created in either the EG9886, EG9875, or EG9880 background. Cohorts of ~20 homozygous young larvae (L1 –L2) of each strain were either incubated at room temperature or at 32°C for 2 hours ("heat shock"). After adulthood, eggs were collected from each cohort, and each hatched animal was scored for the presence of Unc-119(-) phenotype.

Supporting information

S1 Table. Post-silencing edits. Number of injected P0 worms (EG9747) that yielded at least one GFP insertion event on primary or secondary plates. 'Primary + Secondary' tallies P0s for which GFP+ worms were found on both the primary plate founded by the single injected P0 and on the secondary plate founded by 20 array+ GFP- progeny of the originally injected P0

(~F3). 'Primary Only' tallies P0s for which GFP+ worms were found on the primary plate but not on the secondary plate. 'Secondary Only' tallies P0s for which GFP+ worms were not found on the primary plate but were found on the secondary plate. 'Neither' tallies P0s for which GFP+ worms were not found among the progeny on either the primary of secondary plate.
(PDF)

S2 Table. Cas9 insertions by miniMos. Summary of analyses of 14 Cas9 miniMos integration strains. 'Strain' is the isolation name. 'chrom' and 'gene' identify the location of the insert determined by inverse PCR, 'unmapped' means the inverse PCR failed or was ambiguous. PCR confirmation: '+' indicates position was confirmed by locus-specific PCR, '-' indicates confirmatory PCR failed. 'n.t.', not tested. 'cuts *dpy-10*': P0s injected with a *dpy-10* sgRNA expressing plasmid, '+' indicates the presence, and '-' indicates the absence of Dpy and Rol worms among the progeny of those P0s. 'GFP inserts': number of lines yielding *unc-32*::GFP insertions over total injected lines.
(PDF)

S3 Table. Strains used in this study. Bolded strains are the most useful for CRISPR modifications using *unc-119* selections: EG9747 chromosome II, EG9881 chromosome III, EG9887 chromosome I, EG9891 chromosome V.
(PDF)

S4 Table. Plasmids used in this study.
(PDF)

S1 Fig. Generating *unc-119(ox819)*. The classic *unc-119(ed3)* allele was generated using EMS, and is likely to have other mutations in its background. Because CRISPR edits should be generated in a wild-type background that has not been exposed to a chemical mutagen, we generated a new *unc-119* mutation in a wild-type background. N2 worms were injected with a plasmid encoding Cas9 and a guide RNA targeting the same region of *unc-119* as *ed3*. Sequencing of *ox819* revealed an 11 base-pair deletion starting at V110 (UNC-119a) straddling the targeted cut site which leads to a frameshift and stop after 27aa.
(PDF)

S2 Fig. Expression pattern of Cas9 locus fluorescent marker. (Top) Fluorescence pattern of nuclear localized 'cyan-excitable Orange Fluorescent Protein' (cyOFP) in pharynx of an L4 larva from the P*myo-2*::*2xNLS-cyOFP*::*let-858 3'UTR* transgene (EG9882). Fluorescence is brightest in L3 and L4 larvae and is dim in adults. (Bottom) cyOFP has a broad excitation range (maximum from ~480–520) and orange fluorescence emission (590 nm peak) that can be observed using a mercury lamp and standard GFP longpass filter sets on a dissecting scope or using 488 nm or 514 nm lasers with an orange-red bandpass or long-pass emission filter (tagRFP, mScarlet, mRuby) on a confocal microscope. Emission and excitation spectra from fpbase.org.
(PDF)

S3 Fig. Unintended CRISPR outcomes: dead-end targets and dark inserts. (A and B) Dead-end target sites: 20 EG9881 animals were injected with a guide RNA and repair template for tagging *unc-32* with *gfp*. At starvation of the primary plates, array+, *gfp*- animals were singled to generate secondary plates, and the plates were incubated until starvation. Some lines failed to generate GFP+ animals (8/20 lines) by the time the secondary plate starved. These lines were analyzed by PCR amplification of the targeted insertion site (A) and then followed by sequencing (B). **(A)** PCR amplification of the *unc-32* target locus. Expected amplicon sizes for

the native locus and for a complete *gfp + Cbr-unc-119* insert are indicated. **(B)** Sequencing results for the two *unc-32* loci in each strain. (Top) the native *unc-32* target site is shown with the guide RNA binding site and Cas9 cut site highlighted. For each strain, the sequences for both chromosomes are shown. Unique alleles are individually color coded to highlight sibling relationships. Strains 1, 2, 4, 6 and 7 likely originated from a single parent, and strains 3 and 8 likely originated from a single parent. These sequences document six different healing events: three chromosomes healed by NHEJ (red, blue, purple), one chromosome (green) incorporated a truncated GFP, one chromosome (brown) inserted 4 nucleotides at the cut site, and one chromosome (pink) inserted a novel segment with sequence variations (underlined). It is possible this sequence represents a *de novo* event segregating in this strain. **(C)** Location of two 'dark inserts'. Two strains with off target insertions of *unc-119(+)* were located by whole genome sequencing (Illumina). After aligning all reads to the reference *C. elegans* genome, reads that failed to align to the genome were aligned to the sequences of all injected plasmid species. From this mapping, we determined which array sequences were present in the genome. To locate the junctions between the plasmid and the genome, we identified regions by inspection where the plasmid reads contained unmapped tails. BLAST analysis of the unmapped tails was used to determine the insertion site. The insertion site of one of the off-target events is in a central intron of the *rict-1* gene. The sequences flanking the second off-target insertion are present at multiple loci in the genome, preventing exact mapping. **(i).** Schematic representation of the CRISPR targeting injection mix containing a targeting plasmid and multiple array marker plasmids. **(ii).** Schematic representation of the off-target insertion of the entire *gfp*::*snb-1* targeting plasmid into the gene *rict-1* on chromosome II. **(iii).** Schematic representation of the off-target insertion of a portion of the *unc-32*::*gfp* targeting plasmid into a non-unique genomic location. The insertions were not at sites obviously similar to the predicted guide RNA binding sites.
(PDF)

S4 Fig. Genetic map of Cas9 and MosSCI target sites. Insertion sites of the integrated Cas9 alleles (red) in the genome. The locus into which the miniMos-Cas9 transposon inserted is written above (for example 'W01A8.6'), and the original, unmodified, transposon insertions are named ('*oxTi*'). *oxSi1106* is a single-copy insertion at the Mos1 insertion site *ttTi5605*. Also depicted are prominent MosSCI sites (light blue) and 'universal MosSCI 'sites (dark blue) [26,27]. 'Universal MosSCI' sites use the *ttTi5605* arms for homologous recombination insertion and are usually flanked by NeoR and *unc-18(+)* transgenes (see www.wormbuilder.org website).
(PDF)

S5 Fig. Gene models for Cas9 loci. Insertion sites are ordered by chromosome number ('Chr') and site into which the miniMos inserted (red). Strain and name of modified miniMos element are provided below (black). The chromosome II site is a MosSci insertion into the ttTi5605 site and does not have a Cre or orange fluorescent protein ('cyOFP') transgene.
(PDF)

S1 File. Annotated protocol for CRISPR gene editing in *C. elegans*.
(PDF)

S2 File. Text file containing plasmid sequences used in this study.
(TXT)

S3 File. Text file containing sequences of off-target CRISPR insertions.
(TXT)

Acknowledgments

Author Contributions

Conceptualization: Matthew L. Schwartz, M. Wayne Davis, Erik M. Jorgensen.

Formal analysis: Matthew L. Schwartz, Matthew S. Rich.

Funding acquisition: Erik M. Jorgensen.

Investigation: Matthew L. Schwartz, M. Wayne Davis, Matthew S. Rich.

Methodology: Matthew L. Schwartz, M. Wayne Davis, Matthew S. Rich.

Project administration: Erik M. Jorgensen.

Software: M. Wayne Davis.

Supervision: Erik M. Jorgensen.

Validation: Matthew L. Schwartz, Matthew S. Rich.

Visualization: Erik M. Jorgensen.

Writing – original draft: Matthew L. Schwartz.

Writing – review & editing: M. Wayne Davis, Matthew S. Rich, Erik M. Jorgensen.

References

1. Wang H, La Russa M, Qi LS. CRISPR/Cas9 in Genome Editing and Beyond. Annu Rev Biochem. 2016 Jun 2; 85(1):227–64. https://doi.org/10.1146/annurev-biochem-060815-014607 PMID: 27145843

2. Cho SW, Lee J, Carroll D, Kim J-S, Lee J. Heritable gene knockout in Caenorhabditis elegans by direct injection of Cas9-sgRNA ribonucleoproteins. Genetics. 2013 Nov; 195(3):1177–80. https://doi.org/10.1534/genetics.113.155853 PMID: 23979576

3. Paix A, Folkmann A, Rasoloson D, Seydoux G. High Efficiency, Homology-Directed Genome Editing in Caenorhabditis elegans Using CRISPR-Cas9 Ribonucleoprotein Complexes. Genetics. 2015 Sep; 201 (1):47–54. https://doi.org/10.1534/genetics.115.179382 PMID: 26187122

4. Chen C, Fenk LA, de Bono M. Efficient genome editing in Caenorhabditis elegans by CRISPR-targeted homologous recombination. Nucleic Acids Res. 2013 Nov; 41(20):e193. https://doi.org/10.1093/nar/gkt805 PMID: 24013562

5. Dickinson DJ, Ward JD, Reiner DJ, Goldstein B. Engineering the Caenorhabditis elegans Genome Using Cas9-Triggered Homologous Recombination. Nat Methods. 2013 Oct; 10(10):1028–34. https://doi.org/10.1038/nmeth.2641 PMID: 23995389

6. Friedland AE, Tzur YB, Esvelt KM, Colaiácovo MP, Church GM, Calarco JA. Heritable genome editing in C. elegans via a CRISPR-Cas9 system. Nat Methods. 2013 Aug; 10(8):741–3. https://doi.org/10.1038/nmeth.2532 PMID: 23817069

7. Dokshin GA, Ghanta KS, Piscopo KM, Mello CC. Robust Genome Editing with Short Single-Stranded and Long, Partially Single-Stranded DNA Donors in Caenorhabditis elegans. Genetics. 2018 Nov; 210 (3):781–7. https://doi.org/10.1534/genetics.118.301532 PMID: 30213854

8. Ghanta KS, Mello CC. Melting dsDNA Donor Molecules Greatly Improves Precision Genome Editing in Caenorhabditis elegans. Genetics. 2020 Nov; 216(3):643–50. https://doi.org/10.1534/genetics.120.303564 PMID: 32963112

9. Paix A, Schmidt H, Seydoux G. Cas9-assisted recombineering in C. elegans: genome editing using in vivo assembly of linear DNAs. Nucleic Acids Research. 2016 Sep 6; 44(15):e128. https://doi.org/10.1093/nar/gkw502 PMID: 27257074

10. Dickinson DJ, Pani AM, Heppert JK, Higgins CD, Goldstein B. Streamlined Genome Engineering with a Self-Excising Drug Selection Cassette. Genetics. 2015 Aug; 200(4):1035–49. https://doi.org/10.1534/genetics.115.178335 PMID: 26044593

11. Norris AD, Kim H-M, Colaiácovo MP, Calarco JA. Efficient Genome Editing in Caenorhabditis elegans with a Toolkit of Dual-Marker Selection Cassettes. Genetics. 2015 Oct; 201(2):449–58. https://doi.org/10.1534/genetics.115.180679 PMID: 26232410

12. Schwartz ML, Jorgensen EM. SapTrap, a Toolkit for High-Throughput CRISPR/Cas9 Gene Modification in Caenorhabditis elegans. Genetics. 2016 Apr; 202(4):1277–88. https://doi.org/10.1534/genetics.115.184275 PMID: 26837755

13. Aljohani MD, El Mouridi S, Priyadarshini M, Vargas-Velazquez AM, Frøkjær-Jensen C. Engineering rules that minimize germline silencing of transgenes in simple extrachromosomal arrays in C. elegans. Nat Commun. 2020 Dec 9; 11(1):6300. https://doi.org/10.1038/s41467-020-19898-0 PMID: 33298957

14. Schaner CE, Kelly WG. Germline chromatin. WormBook [Internet]. 2006 [cited 2021 Jul 20]; Available from: http://www.wormbook.org/chapters/www_germlinechromatin/germlinechromatin.html https://doi.org/10.1895/wormbook.1.73.1 PMID: 18050477

15. Kelly WG, Xu S, Montgomery MK, Fire A. Distinct requirements for somatic and germline expression of a generally expressed Caernorhabditis elegans gene. Genetics. 1997 May; 146(1):227–38. https://doi.org/10.1093/genetics/146.1.227 PMID: 9136012

16. Frøkjaer-Jensen C, Davis MW, Hopkins CE, Newman BJ, Thummel JM, Olesen S-P, et al. Single-copy insertion of transgenes in Caenorhabditis elegans. Nat Genet. 2008 Nov; 40(11):1375–83. https://doi.org/10.1038/ng.248 PMID: 18953339

17. Zhang D, Tu S, Stubna M, Wu W-S, Huang W-C, Weng Z, et al. The piRNA targeting rules and the resistance to piRNA silencing in endogenous genes. Science. 2018 Feb 2; 359(6375):587–92. https://doi.org/10.1126/science.aao2840 PMID: 29420292

18. Frøkjær-Jensen C, Jain N, Hansen L, Davis MW, Li Y, Zhao D, et al. An Abundant Class of Non-coding DNA Can Prevent Stochastic Gene Silencing in the C. elegans Germline. Cell. 2016 Jul 14; 166(2):343–57. https://doi.org/10.1016/j.cell.2016.05.072 PMID: 27374334

19. Waaijers S, Portegijs V, Kerver J, Lemmens BBLG, Tijsterman M, van den Heuvel S, et al. CRISPR/Cas9-Targeted Mutagenesis in *Caenorhabditis elegans*. Genetics. 2013 Nov 1; 195(3):1187–91. https://doi.org/10.1534/genetics.113.156299 PMID: 23979586

20. Doench JG, Hartenian E, Graham DB, Tothova Z, Hegde M, Smith I, et al. Rational design of highly active sgRNAs for CRISPR-Cas9–mediated gene inactivation. Nat Biotechnol. 2014 Dec; 32(12):1262–7. https://doi.org/10.1038/nbt.3026 PMID: 25184501

21. Farboud B, Meyer BJ. Dramatic Enhancement of Genome Editing by CRISPR/Cas9 Through Improved Guide RNA Design. Genetics. 2015 Apr 1; 199(4):959–71. https://doi.org/10.1534/genetics.115.175166 PMID: 25695951

22. Mello CC, Kramer JM, Stinchcomb D, Ambros V. Efficient gene transfer in C.elegans: extrachromosomal maintenance and integration of transforming sequences. EMBO J. 1991 Dec; 10(12):3959–70. PMID: 1935914

23. Stinchcomb DT, Shaw JE, Carr SH, Hirsh D. Extrachromosomal DNA transformation of Caenorhabditis elegans. Molecular and Cellular Biology. 1985 Dec 1; 5(12):3484–96. https://doi.org/10.1128/mcb.5.12.3484-3496.1985 PMID: 3837845

24. Fire A, Alcazar R, Tan F. Unusual DNA Structures Associated With Germline Genetic Activity in *Caenorhabditis elegans*. Genetics. 2006 Jul 1; 173(3):1259–73. https://doi.org/10.1534/genetics.106.057364 PMID: 16648589

25. Brenner S. The genetics of Caenorhabditis elegans. Genetics. 1974 May; 77(1):71–94. https://doi.org/10.1093/genetics/77.1.71 PMID: 4366476

26. Frøkjær-Jensen C, Davis MW, Sarov M, Taylor J, Flibotte S, LaBella M, et al. Random and targeted transgene insertion in *Caenorhabditis elegans* using a modified *Mos1* transposon. Nature Methods. 2014 May; 11(5):529–34. https://doi.org/10.1038/nmeth.2889 PMID: 24820376

27. Frøkjær-Jensen C, Davis MW, Ailion M, Jorgensen EM. Improved Mos1-mediated transgenesis in C. elegans. Nat Methods. 2012 Jan 30; 9(2):117–8. https://doi.org/10.1038/nmeth.1865 PMID: 22290181

Gene	Primary + Secondary	Primary Only	Secondary Only	Neither	Total
sng-1	1	4	1	2	8
snb-1	5	0	0	0	5
rab-3	6	3	1	5	15
unc-32	4	3	0	1	8
unc-17	19	0	0	0	19
snt-1 (short)	0	0	0	6	6
snt-1 (long)	4	2	2	13	21

S1 Table Post-silencing edits. Number of injected P0 worms (EG9747) that yielded at least one GFP insertion event on primary or secondary plates. 'Primary + Secondary' tallies P0s for which GFP+ worms were found on both the primary plate founded by the single injected P0 and on the secondary plate founded by 20 array+ GFP- progeny of the originally injected P0 (~F3). 'Primary Only' tallies P0s for which GFP+ worms were found on the primary plate but not on the secondary plate. 'Secondary Only' tallies P0s for which GFP+ worms were not found on the primary plate but were found on the secondary plate. 'Neither' tallies P0s for which GFP+ worms were not found among the progeny on either the primary of secondary plate.

| Strain	chrom	gene	PCR confirmation	cuts dpy-10	GFP inserts
Strain 1	**III**	***unc-119***	**+**	**+**	**1/2**
Strain 2	IV	*pes-1*	n.t.	+	3/7
Strain 3	V	Y69H2.3	+	+	0/6
Strain 4	**III**	**F53A2.9**	**+**	**+**	**5/5**
Strain 5	X	K05G3.1	n.t.	-	n.t.
Strain 6	II	*clec-119*	n.t.	n.t.	1/8
Strain 7	I	*eipr-1*	+	+	3/5
Strain 8	?	unmapped	-	+	**high**
Strain 9	V	F21D9.6	+	+	**low**
Strain 10	**I**	**W01A8.6**	**+**	**+**	**7/11**
Strain 11	?	unmapped	n.t.	n.t.	0/6
Strain 12	X	C23H4.6	n.t.	n.t.	n.t.
Strain 13	V	*srab-21*	n.t.	n.t.	0/1
Strain 14	**V**	**W03F9.11**	**+**	**n.t.**	**4/5**

S2 Table Cas9 insertions by miniMos. Summary of analyses of 14 Cas9 miniMos integration strains. 'Strain' is the isolation name. 'chrom' and 'gene' identify the location of the insert determined by inverse PCR, 'unmapped' means the inverse PCR failed or was ambiguous. PCR confirmation: '+' indicates position was confirmed by locus-specific PCR, '-' indicates confirmatory PCR failed. 'n.t.', not tested. 'cuts *dpy-10*': P0s injected with a *dpy-10* sgRNA expressing plasmid, '+' indicates the presence, and '-' indicates the absence of Dpy and Rol worms among the progeny of those P0s. 'GFP inserts': number of lines yielding *unc-32*::GFP insertions over total injected lines.

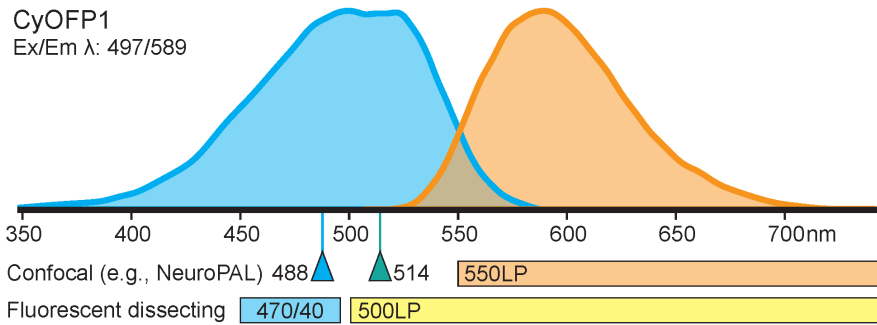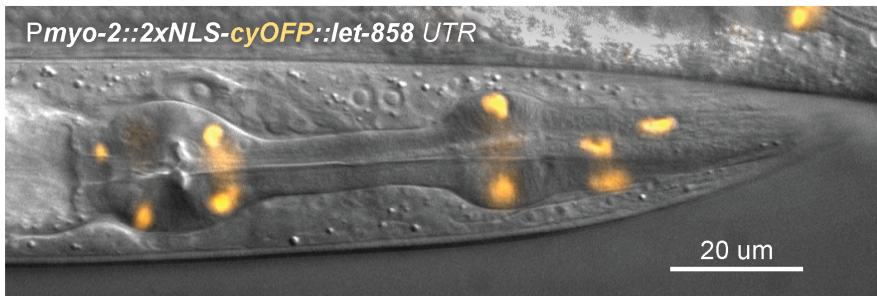| Strain	Genotype
N2	-
EG6249	*ttTi5605* II ; *unc-119*(ed3) III
none	*oxSi1196 [Pmex-5::Cas9(cDNA) :tbb-2 3'UTR Cbr-unc-119+; *ttTi5605] II ; unc-119(ed3)* III
EG9615	*oxSi1091[Pmex-5::cas9(+smu-2 introns)::tbb-2 3'UTR unc-119+; *ttTi5605] II; unc-119(ed3)* III
EG9747	*oxSi1106[Pmex-5:Cas9:tbb-2 3'UTR +lox2272; *ttTi5605] II ; unc-119(ed3)* III
EG9814	*unc-119(ox819)* III
EG9809	*unc-119(ox819 oxTi1083[Pmex-5::cas9(+smu-2 introns) + lox2272])* III
EG9873	*unc-119(ox819 oxTi1113[Pmex-5::cas9(+ smu-2 introns), Phsp-16.41::Cre + lox2272 Cbr-unc-119 lox2272])* III
EG9874	*unc-119(ox819 oxTi1111[Pmex-5::cas9(+ smu-2 introns), Phsp-16.41::Cre + lox2272])* III
EG9875	*unc-119(ox819 oxTi1122[Pmex-5::cas9(+ smu-2 introns), Phsp-16.41::Cre, Pmyo-2::2xNLS-CyOFP + lox2272 Cbr-unc-119 lox2272])* III
EG9876	*unc-119(ox819 oxTi1126[Pmex-5::cas9(+ smu-2 introns), Phsp-16.41::Cre, Pmyo-2::2xNLS-CyOFP + lox2272])* III
EG9877	*F53A2.9(oxTi1112[Pmex-5::cas9(+smu-2 introns) + lox2272]), unc-119(ox819)* III
EG9878	*F53A2.9(oxTi1118[Pmex-5::cas9(+ smu-2 introns), Phsp-16.41::Cre + lox2272 Cbr-unc-119 lox2272], unc-119(ox819)* III
EG9879	*F53A2.9(oxTi1118[Pmex-5::cas9(+ smu-2 introns), Phsp-16.41::Cre + lox2272]), unc-119(ox819)* III
EG9880	*F53A2.9(oxTi1123[Pmex-5::cas9(+ smu-2 introns), Phsp-16.41::Cre, Pmyo-2::2xNLS-CyOFP + lox2272 Cbr-unc-119 lox2272]), unc-119(ox819)* III
EG9881	*F53A2.9(oxTi1127[Pmex-5::cas9(+ smu-2 introns), Phsp-16.41::Cre, Pmyo-2::2xNLS-CyOFP + lox2272]), unc-119(ox819)* III
EG9882	*F53A2.9(oxTi1127[Pmex-5::cas9(+ smu-2 introns), Phsp-16.41::Cre, Pmyo-2::2xNLS-CyOFP + lox2272])* III
EG9883	*W01A8.6(oxTi1114[Pmex-5::cas9(+smu-2 introns) + lox2272]) I ; unc-119(ox819)* III
EG9884	*W01A8.6(oxTi1120[Pmex-5::cas9(+ smu-2 introns), Phsp-16.41::Cre + lox2272 Cbr-unc-119 lox2272] I ; unc-119(ox819)* III
EG9885	*W01A8.6(oxTi1120[Pmex-5::cas9(+ smu-2 introns), Phsp-16.41::Cre + lox2272]) I ; unc-119(ox819)* III
EG9886	*W01A8.6(oxTi1124[Pmex-5::cas9(+ smu-2 introns), Phsp-16.41::Cre, Pmyo-2::2xNLS-CyOFP + lox2272 Cbr-unc-119 lox2272]) I ; unc-119(ox819)* III
EG9887	*W01A8.6(oxTi1128[Pmex-5::cas9(+ smu-2 introns), Phsp-16.41::Cre, Pmyo-2::2xNLS-CyOFP + lox2272]) I ; unc-119(ox819)* III
EG9888	*W01A8.6(oxTi1128[Pmex-5::cas9(+ smu-2 introns), Phsp-16.41::Cre, Pmyo-2::2xNLS-CyOFP + lox2272]) I*
EG9889	*unc-119(ox819) III ; W03F9.11(oxTi1116[Pmex-5::cas9(+smu-2 introns) + lox2272])* V
EG9890	*unc-119(ox819) III ; W03F9.11(oxTi1121[Pmex-5::cas9(+ smu-2 introns), Phsp-16.41::Cre + lox2272 Cbr-unc-119 lox2272]* V
EG9891	*unc-119(ox819) III ; W03F9.11(oxTi1121[Pmex-5::cas9(+ smu-2 introns), Phsp-16.41::Cre + lox2272])* V
EG9892	*unc-119(ox819) III ; W03F9.11(oxTi1125[Pmex-5::cas9(+ smu-2 introns), Phsp-16.41::Cre, Pmyo-2::2xNLS-CyOFP + lox2272 Cbr-unc-119 lox2272])* V
EG9893	*unc-119(ox819) III ; W03F9.11(oxTi1129[Pmex-5::cas9(+ smu-2 introns), Phsp-16.41::Cre, Pmyo-2::2xNLS-CyOFP + lox2272])* V
EG9894	*W03F9.11(oxTi1129[Pmex-5::cas9(+ smu-2 introns), Phsp-16.41::Cre, Pmyo-2::2xNLS-CyOFP + lox2272])* V

S3 Table Strains used in this study. Bolded strains are the most useful for CRISPR modifications using *unc-119* selections: EG9747 chromosome II, EG9881 chromosome III, EG9887 chromosome I, EG9891 chromosome V.

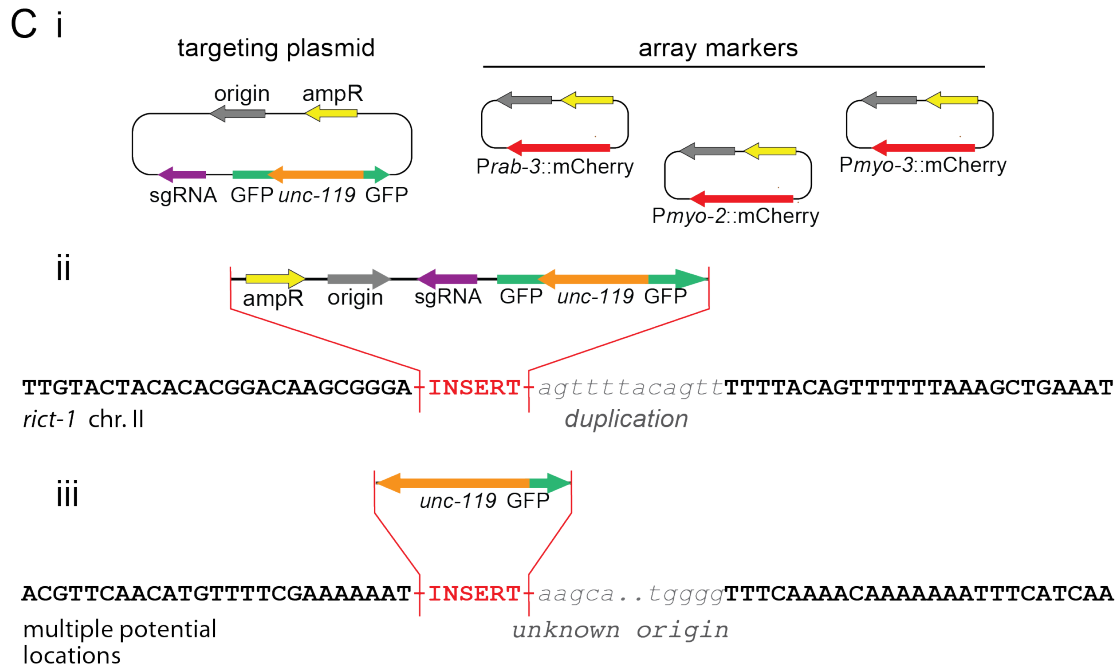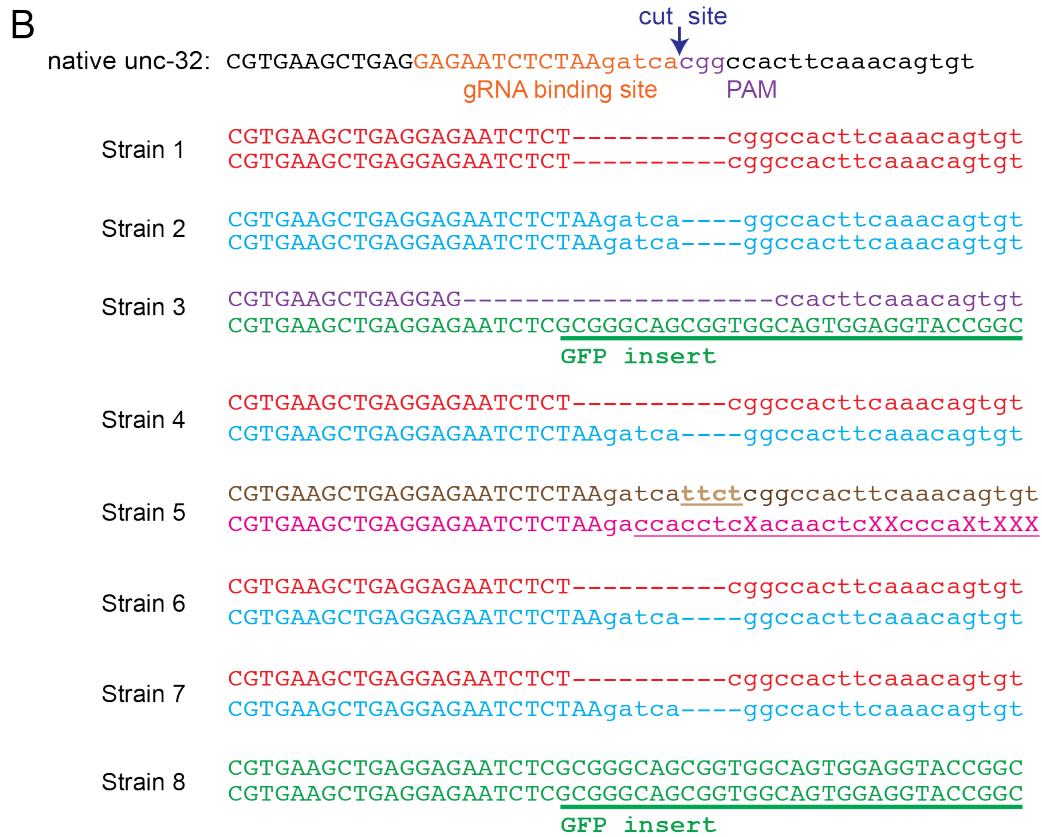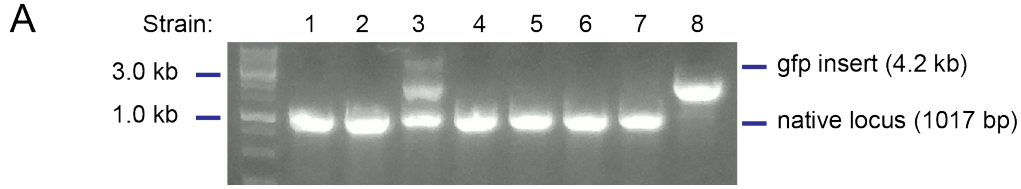| Plasmid	Description	Availability	Reference
pMLS538	pDONR-221:2xNLS-Cas9(+*smu-2* introns)	Addgene	this study
pMLS539	pDONR-221:2xNLS-Cas9(cDNA + syntron)	-	this study
pMLS544	pEXPR(5605):P*mex-5::Cas9*(+*smu-2* introns)::*tbb-2* 3'UTR; *Cbr-unc-119*	Addgene	this study
pMLS546	pEXPR(5605):P*mex-5::Cas9*(cDNA + syntron)::*tbb-2* 3'UTR; *Cbr-unc-119*	Addgene	this study
pMLS389	pEXPR:P*U6::sgRNA(unc-119)U6*-terminator	-	this study
pMLS597	pEXPR:P*U6::sgRNA*(*dpy-10*)*U6*-terminator	Addgene	this study
pMLS640	pEXPR(5605):P*mex-5::Cas9*(+*smu-2* introns)::*tbb-2* 3'UTR; lox2272 *Cbr-unc-119* lox2272	Addgene	this study
pMLS621	pDEST(5605):lox2272 *Cbr-unc-119* lox2272	Addgene	this study
pMLS713	pDEST(miniMos):lox2272 *Cbr-unc-119* lox2272	Addgene	this study
pMLS714	pEXPR(5605):P*mex-5::Cas9*(+*smu-2* introns)::*tbb-2* 3'UTR ; lox2272 *Cbr-unc-119* lox2272	Addgene	this study
pMLS715	pDEST(miniMos-5'Arm):lox2272 *Cbr-unc-119* lox2272	Addgene	this study
pMLS716	pEXPR(miniMos-5'Arm):P*hsp16.41::Cre::tbb-2* 3'UTR ; lox2272 *Cbr-unc-119* lox2272	Addgene	this study
pMLS719	pEXPR:P*U6::sgRNA*(miniMos-5'Arm)*U6*-terminator	Addgene	this study
pMLS791	pEXPR:P*hsp16.41::Cre::tbb-2* 3'UTR ; P*myo-2::2xNLS-cyOFP::let-858* 3'UTR ; lox2272 *Cbr-unc-119* lox2272	Addgene	this study
pMLS338	pCRISPR:*unc-32::gfp + loxp Cbr-unc-119 loxp*	Addgene	Schwartz and Jorgensen, 2016
pMLS255	pCRISPR:*gfp::unc-17 + loxp Cbr-unc-119 loxp*	-	Schwartz and Jorgensen, 2016
pSYC33	pCRISPR:*snt-1::gfp + loxp Cbr-unc-119 loxp*	-	this study
pMLS357	pCRISPR:*sng-1::gfp + loxp Cbr-unc-119 loxp*	-	Schwartz and Jorgensen, 2016
pMLS276	pCRISPR:*gfp::snb-1 + loxp Cbr-unc-119 loxp*	-	Schwartz and Jorgensen, 2016
pMLS817	pCRISPR:*gfp::rab-3 + loxp Cbr-unc-119 loxp*	-	this study
pMLS485	pEXPR:P*eft-3::Cas9(+smu-2* introns)::*tbb-2* 3'UTR	Addgene	this study
pMLS1272	pEXPR:*ric-4 sgRNA*	Addgene	This study

Supplemental Table 4. Plasmids used in this study.
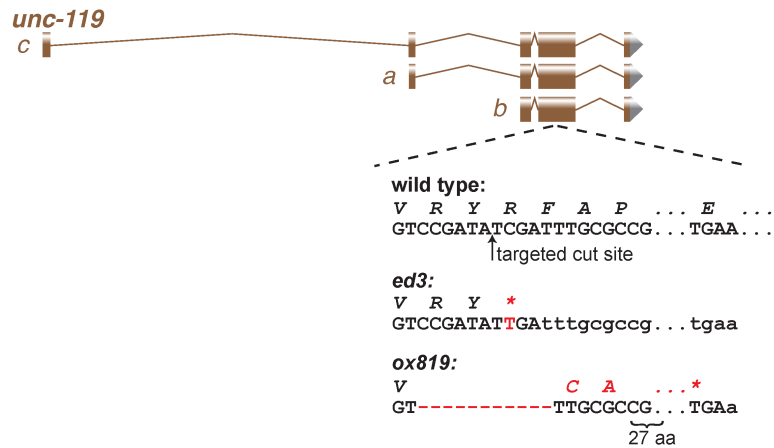
S1 Fig Generating *unc-119(ox819)*. The classic *unc-119(ed3)* allele was generated using EMS, and is likely to have other mutations in its background. Because CRISPR edits should be generated in a wild-type background that has not been exposed to a chemical mutagen, we generated a new *unc-119* mutation in a wild-type background. N2 worms were injected with a plasmid encoding Cas9 and a guide RNA targeting the same region of *unc-119* as *ed3*. Sequencing of *ox819* revealed an 11 base-pair deletion starting at V110 (UNC-119a) straddling the targeted cut site which leads to a frameshift and stop after 27aa.
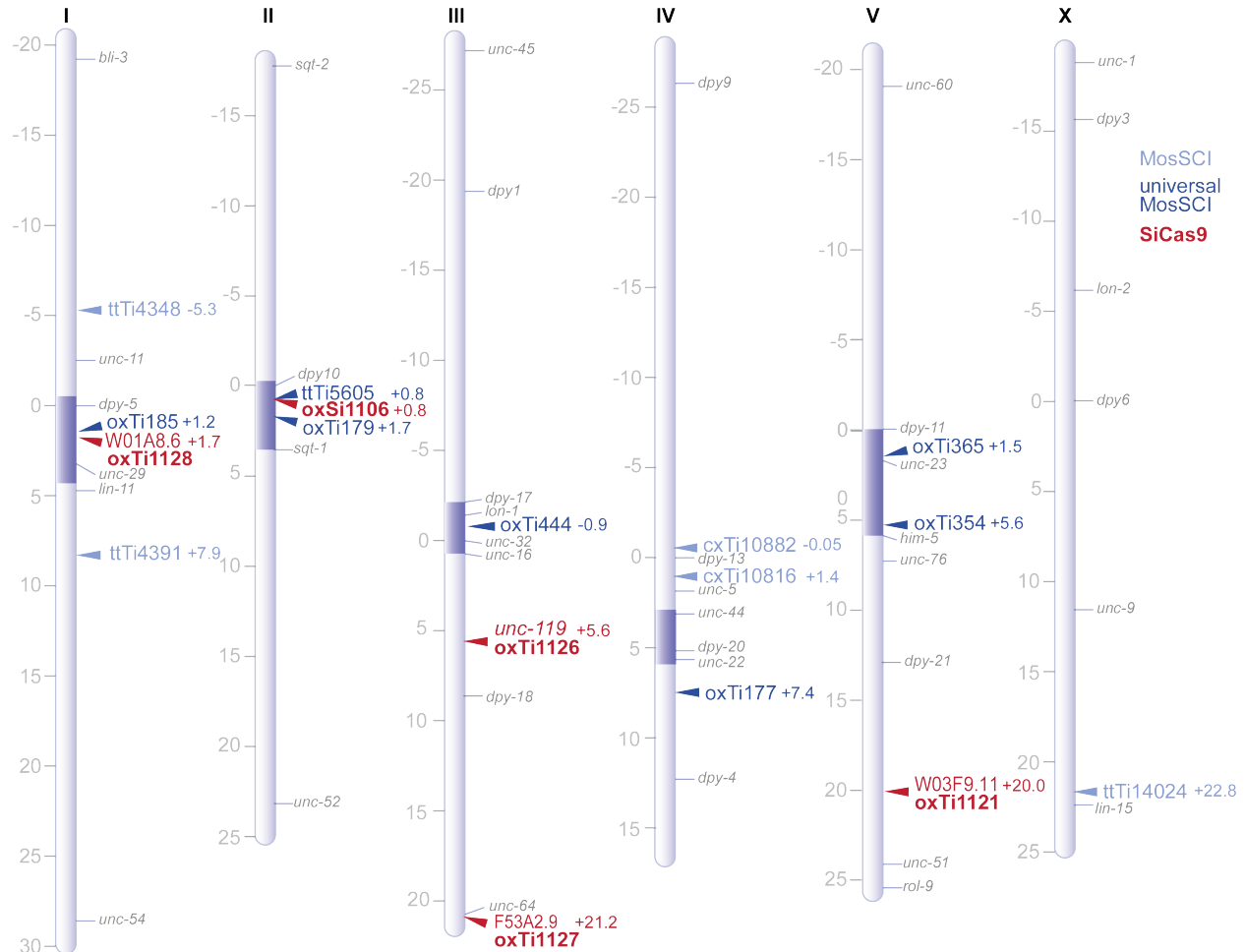
S2 Fig Expression pattern of Cas9 locus fluorescent marker. (Top) Fluorescence pattern of nuclear localized 'cyan-excitable Orange Fluorescent Protein' (cyOFP) in pharynx of an L4 larva from the P*myo-2::2xNLS-cyOFP::let-858 3'UTR* transgene (EG9882). Fluorescence is brightest in L3 and L4 larvae and is dim in adults. (Bottom) cyOFP has a broad excitation range (maximum from ~480-520) and orange fluorescence emission (590 nm peak) that can be observed using a mercury lamp and standard GFP longpass filter sets on a dissecting scope or using 488 nm or 514 nm lasers with an orange-red bandpass or long-pass emission filter (tagRFP, mScarlet, mRuby) on a confocal microscope. Emission and excitation spectra from fpbase.org.

A

Strain: 1 2 3 4 5 6 7 8

3.0 kb —

gfp insert (4.2 kb)

1.0 kb —

native locus (1017 bp)

B

cut site

native unc-32: CGTGAAGCTGAGGAGAATCTCTAAgatcacggccacttcaaacagtgt

gRNA binding site   PAM

Strain 1
CGTGAAGCTGAGGAGAATCTCT----------cggccacttcaaacagtgt
CGTGAAGCTGAGGAGAATCTCT----------cggccacttcaaacagtgt

Strain 2
CGTGAAGCTGAGGAGAATCTCTAAgatca----ggccacttcaaacagtgt
CGTGAAGCTGAGGAGAATCTCTAAgatca----ggccacttcaaacagtgt

Strain 3
CGTGAAGCTGAGGAG------------------ccacttcaaacagtgt
CGTGAAGCTGAGGAGAATCTCGCGGGCAGCGGTGGCAGTGGAGGTACCGGC
GFP insert

Strain 4
CGTGAAGCTGAGGAGAATCTCT----------cggccacttcaaacagtgt
CGTGAAGCTGAGGAGAATCTCTAAgatca----ggccacttcaaacagtgt

Strain 5
CGTGAAGCTGAGGAGAATCTCTAAgatcattctcggccacttcaaacagtgt
CGTGAAGCTGAGGAGAATCTCTAAgaccacctcXacaactcXXcccaXtXXX

Strain 6
CGTGAAGCTGAGGAGAATCTCT----------cggccacttcaaacagtgt
CGTGAAGCTGAGGAGAATCTCTAAgatca----ggccacttcaaacagtgt

Strain 7
CGTGAAGCTGAGGAGAATCTCT----------cggccacttcaaacagtgt
CGTGAAGCTGAGGAGAATCTCTAAgatca----ggccacttcaaacagtgt

Strain 8
CGTGAAGCTGAGGAGAATCTCGCGGGCAGCGGTGGCAGTGGAGGTACCGGC
CGTGAAGCTGAGGAGAATCTCGCGGGCAGCGGTGGCAGTGGAGGTACCGGC
GFP insert

C i

targeting plasmid   array markers

origin   ampR

sgRNA GFP unc-119 GFP

Prab-3::mCherry   Pmyo-3::mCherry

Pmyo-2::mCherry

ii

ampR origin sgRNA GFP unc-119 GFP

TTGTACTACACACGGACAAGCGGGA—INSERT—agttttacagtt TTTTACAGTTTTTTTAAAGCTGAAAT
rict-1 chr. II   duplication

iii

unc-119 GFP

ACGTTCAACATGTTTTCGAAAAAAT—INSERT—aagca..tgggg TTTCAAAACAAAAAAATTTCATCAA
multiple potential
locations   unknown origin

**S3 Fig Unintended CRISPR outcomes: dead-end targets and dark inserts**.
(A and B) Dead-end target sites: 20 EG9881 animals were injected with a guide RNA and repair template for tagging *unc-32* with *gfp.* At starvation of the primary plates, array+, *gfp-* animals were singled to generate secondary plates, and the plates were incubated until starvation. Some lines failed to generate GFP+ animals (8/20 lines) by the time the secondary plate starved. These lines were analyzed by PCR amplification of the targeted insertion site (A) and then followed by sequencing (B). **(A)** PCR amplification of the *unc-32* target locus. Expected amplicon sizes for the native locus and for a complete *gfp+ Cbr-unc-119* insert are indicated. **(B)** Sequencing results for the two *unc-32* loci in each strain. The native *unc-32* target site is shown above with the guide-RNA binding site and Cas9 cut site highlighted. For each strain, the sequences for both chromosomes are shown. Unique alleles are individually color coded to highlight sibling relationships. Strains 1, 2, 4, 6 and 7 likely originated from a single parental, and strains 3 and 8 likely originated from a single parent. These sequences document six different healing events: three chromosomes healed by NHEJ (red, blue, purple), one chromosome (green) incorporated a truncated GFP, one chromosome (brown) inserted 4 nucleotides at the cut site, and one chromosome (pink) inserted a novel segment with sequence variations (underlined). 'X' indicates that two nucleotides were present at this position. It is possible this sequence represents a *de novo* event segregating in this strain.
(C) Location of two 'dark inserts'. Two strains with off target insertions of *unc-119(+)* were located by whole genome sequencing (Illumina). After aligning all reads to the reference *C. elegans* genome, reads that failed to align to the genome were aligned to the sequences of all injected plasmid species. From this mapping, we determined which array sequences were present in the genome. To locate the junctions between the plasmid and the genome, we identified regions by inspection where the plasmid reads contained unmapped tails. BLAST analysis of the unmapped tails was used to determine the insertion site. The insertion site of one of the off-target events is in a central intron of the *rict-1* gene. The sequences flanking the second off-target insertion are present at multiple loci in the genome, preventing exact mapping. **(i)** Schematic representation of the CRISPR targeting injection mix containing a targeting plasmid and multiple array marker plasmids. **(ii)** Schematic representation of *rict-1(ox773),* an off-target insertion of the entire *gfp::snb-1* targeting plasmid into intron #9 of *rict-1a* on chromosome II. **(iii)** Schematic representation of the off-target insertion of a portion of the *unc-32::gfp* targeting plasmid into a non-unique genomic location. The insertions were not at sites obviously similar to the predicted guide RNA binding sites.

S4 Fig Genetic map of Cas9 and MosSCI target sites. Insertion sites of the integrated Cas9 alleles (red) in the genome. The locus into which the miniMos-Cas9 transposon inserted is written above (for example 'W01A8.6'), and the original, unmodified, transposon insertions are named ('*oxTi*'). *oxSi1106* is a single-copy insertion at the Mos1 insertion site *ttTi5605*. Also depicted are prominent MosSCI sites (light blue) and 'universal MosSCI 'sites (dark blue) (Frøkjær-Jensen et al., 2014, 2012). 'Universal MosSCI' sites use the *ttTi5605* arms for homologous recombination insertion and are usually flanked by NeoR and *unc-18(+)* transgenes (see www.wormbuilder.org website).
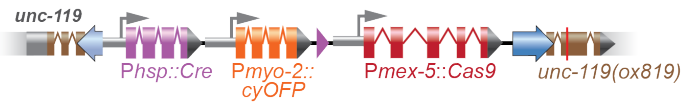
Chr I W01A8.6
EG9887
oxTi1128
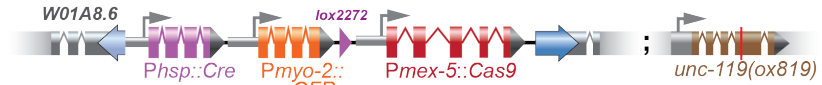
EG9888
oxTi1128
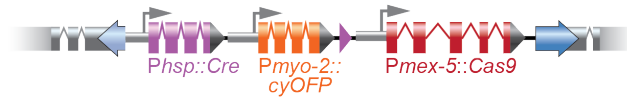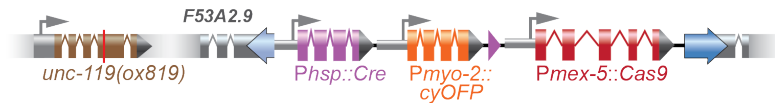
EG9885
oxTi1120

Chr II ttTi5605
EG9747
oxSi1106
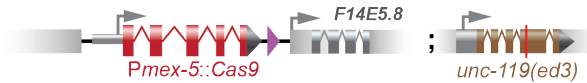
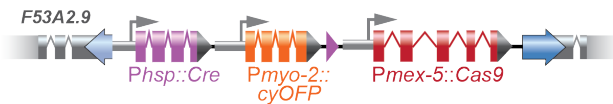**Chr III** unc-119
EG9876
oxTi1126
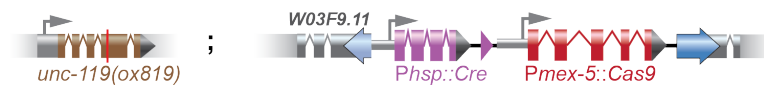
EG9874
oxTi1111

Chr III F53A2.9
EG9881
oxTi1127

EG9882
oxTi1127

Chr V W03F9.11
EG9891
oxTi1121

S5 Fig Gene models for Cas9 loci. Insertion sites are ordered by chromosome number ('Chr') and site into which the miniMos inserted (red). Strain and name of modified miniMos element are provided below (black). The chromosome II site is a MosSci insertion into the ttTi5605 site and does not have a Cre or orange fluorescent protein ('cyOFP') transgene.