Engineering the *Caenorhabditis elegans* genome using Cas9-triggered homologous recombination

Daniel J Dickinson[1,2], Jordan D Ward[3], David J Reiner[2,4,5] & Bob Goldstein[1,2]

Study of the nematode *Caenorhabditis elegans* has provided important insights in a wide range of fields in biology. The ability to precisely modify genomes is critical to fully realize the utility of model organisms. Here we report a method to edit the *C. elegans* genome using the clustered, regularly interspersed, short palindromic repeats (CRISPR) RNA-guided Cas9 nuclease and homologous recombination. We demonstrate that Cas9 is able to induce DNA double-strand breaks with specificity for targeted sites and that these breaks can be repaired efficiently by homologous recombination. By supplying engineered homologous repair templates, we generated *gfp* knock-ins and targeted mutations. Together our results outline a flexible methodology to produce essentially any desired modification in the *C. elegans* genome quickly and at low cost. This technology is an important addition to the array of genetic techniques already available in this experimentally tractable model organism.

The ability to precisely modify the genome of an organism by adding, deleting or mutating genes is a critical tool for experimental biology. The type II CRISPR-associated (Cas) system is a powerful tool for genome editing in a variety of experimental systems. The Cas9 nuclease and two small noncoding RNAs comprise an adaptive immune system in prokaryotes[1]. A chimeric fusion of the two RNAs, termed a single guide RNA (sgRNA), supports site-specific cleavage of target DNA by Cas9, with target specificity determined by base-pairing between the 5′ end of the sgRNA and the target DNA[1] (**Fig. 1a**). The only specific sequence requirement for cleavage is an NGG nucleotide sequence (the protospacer adjacent motif) at the 3′ end of the target DNA sequence (**Fig. 1a**). Changing the targeting sequence at the 5′ end of the sgRNA can allow an exceptionally broad variety of DNA substrates to be targeted without the need to reengineer the Cas9 nuclease[1–13]. Compared to zinc-finger nucleases and transcription activator–like effector nucleases (TALENs) that can be used to produce double-strand breaks[14], the CRISPR-Cas9 system is substantially less expensive and is easier to program for new target sites. Cas9 has been used to produce targeted insertion or deletion (indel) mutations, which are generated via error-prone repair mechanisms, in a wide range of species[2–13]. In addition, homology-directed repair of Cas9-induced double-strand breaks has been demonstrated in bacteria[7], yeast[3], cultured human and mouse cells[5,11], fruit flies[4], zebrafish[10] and mice[13].

The nematode *C. elegans* is a valuable and widely used experimental system owing to its rapid growth, ease of handling and transparency (which facilitates microscopy). Recent reports demonstrated that Cas9 can induce double-strand breaks in the *C. elegans* germ line, which leads to mutations via error-prone repair mechanisms[8,15]. Here we demonstrate that Cas9-induced double-strand breaks can be repaired efficiently by homologous recombination. By supplying engineered homologous repair templates, we generated in-frame *gfp* insertions and targeted mutations. We refer to this method as Cas9-triggered homologous recombination.

RESULTS

Design of a CRISPR-Cas9 system for *C. elegans*

To establish Cas9 as a tool for genome editing in *C. elegans*, we expressed Cas9 and sgRNA in the *C. elegans* germ line. In mammalian systems, sgRNAs were expressed from a *U6* small nuclear RNA promoter, which drives transcription by RNA polymerase III[5,11]. To our knowledge, no germline RNA polymerase III promoters had been described in *C. elegans* when we initiated these experiments. Aligning ten *C. elegans U6* genes revealed a region of 200–300 base pairs (bp) upstream of the transcriptional start site that is partially conserved and may function as the promoter (**Fig. 1b** and **Supplementary Fig. 1**). We therefore expressed sgRNAs using 350 bp upstream of the highly expressed *R07E5.16 U6* gene[16]. A different *U6* promoter was independently identified by Friedland *et al.*[8]. To express Cas9 in the germ line, we used the *eft-3* promoter and *tbb-2* 3′ UTR, which have been used successfully for genome modifications employing the *Mos1* transposon[17]. We built a Cas9-sgRNA plasmid containing both the P*eft-3*::*Cas9*::*tbb-2*–3′ UTR and P*U6*::*sgRNA* constructs (**Fig. 1c**). This plasmid, available through Addgene, can be engineered to target any desired sequence by using site-directed mutagenesis to insert the appropriate targeting sequence.

[1]Department of Biology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA. [2]Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA. [3]Department of Cellular and Molecular Pharmacology, University of California, San Francisco, San Francisco, California, USA. [4]Department of Pharmacology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA. [5]Present address: Center for Translational Cancer Research, Institute of Biosciences and Technology, Texas A&M University, Houston, Texas, USA. Correspondence should be addressed to D.J.D. (ddickins@live.unc.edu).

RECEIVED 29 JUNE; ACCEPTED 12 AUGUST; PUBLISHED ONLINE 1 SEPTEMBER 2013; CORRECTED ONLINE 16 SEPTEMBER 2013; DOI:10.1038/NMETH.2641

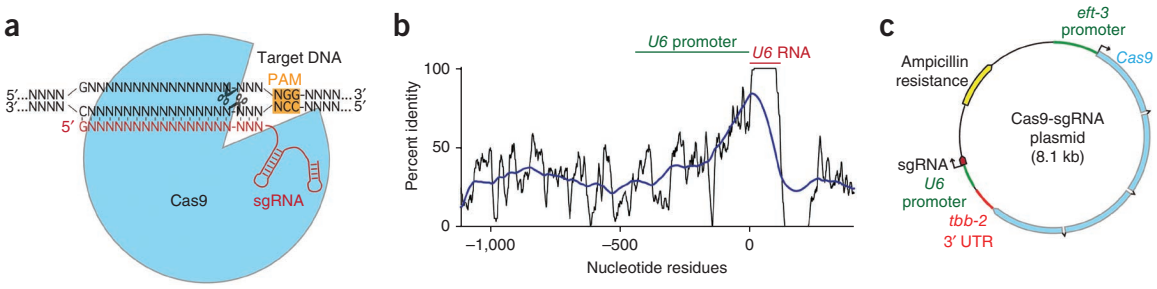

Figure 1 | Adaptation of the CRISPR-Cas9 system for *C. elegans*. (**a**) Schematic of the Cas9 nuclease and sgRNA. Formation of a double-strand break requires (i) base-pairing between the sgRNA and the target DNA sequence and (ii) the presence of the NGG motif (protospacer adjacent motif, PAM) immediately adjacent to the target sequence[1]. Cleavage occurs 3 bp 5′ of the PAM. The guanine (G) residue at the 5′ end of the sgRNA is required for transcription initiation by the *U6* promoter. (**b**) Sequence conservation of the ten *U6* RNA genes that we identified in *C. elegans*. The blue trace is a rolling average produced using locally weighted scatter-plot smoothing. The green bar indicates the region of *R07E5.16* that we used as the promoter in the Cas9-sgRNA construct. (**c**) Schematic of the Cas9-sgRNA plasmid.

Homologous repair of Cas9-induced double-strand breaks

Several existing methods for modifying the *C. elegans* genome rely on homologous repair of double-strand breaks generated by excision of a *Mos1* transposon[18–21], but these techniques are limited by the relative scarcity of *Mos1* insertion sites in the genome[22]. In principle, a much wider range of genome modifications could be made by using Cas9 to generate double-strand breaks. We therefore tested whether Cas9-induced double-strand breaks could be repaired by homologous recombination in *C. elegans*. We designed an sgRNA targeting a sequence adjacent to the *ttTi5605 Mos1* insertion site on chromosome II and compared the efficiency of single-copy transgene insertion into this site using either *Mos1*-mediated single-copy insertion (MosSCI[20]) or Cas9-triggered homologous recombination (**Fig. 2a**). The efficiency of both approaches varied among individual experiments, but the overall efficiency of the two methods was similar (**Fig. 2b** and **Supplementary Table 1**). We examined expression of the resulting single-copy *gfp* transgenes under the control of the *mex-5* promoter and *tbb-2* 3′ UTR. The transgenes were expressed in the germ line, and the pattern of expression was indistinguishable regardless of the method used for transgene insertion (**Fig. 2c**). These data demonstrate that Cas9-induced double-strand breaks can stimulate homologous recombination in the *C. elegans* germ line.

Integration of *gfp* into endogenous loci

In *C. elegans*, fluorescent fusion proteins are often expressed via microinjection of DNA into the gonad, which generates semi-stable extrachromosomal arrays that contain many copies of the injected DNA[23]. Transgenes generated in this way are typically overexpressed in somatic tissues and silenced in the germ line and early embryo[24]. Microparticle bombardment can be used to generate low-copy transgenes, which are expressed at closer to endogenous levels[25,26], but this approach is expensive and time consuming. MosSCI can be used to generate single-copy transgenes[20], but for many genes, the regulatory sequences needed to recapitulate the native expression pattern are unknown. All of these approaches also leave the endogenous copy of the gene of interest intact, which makes it difficult to assess the function of the fusion protein genetically and can introduce complications in quantitative experiments because less than 100% of molecules of the protein of interest are labeled. Inserting genetically encoded tags

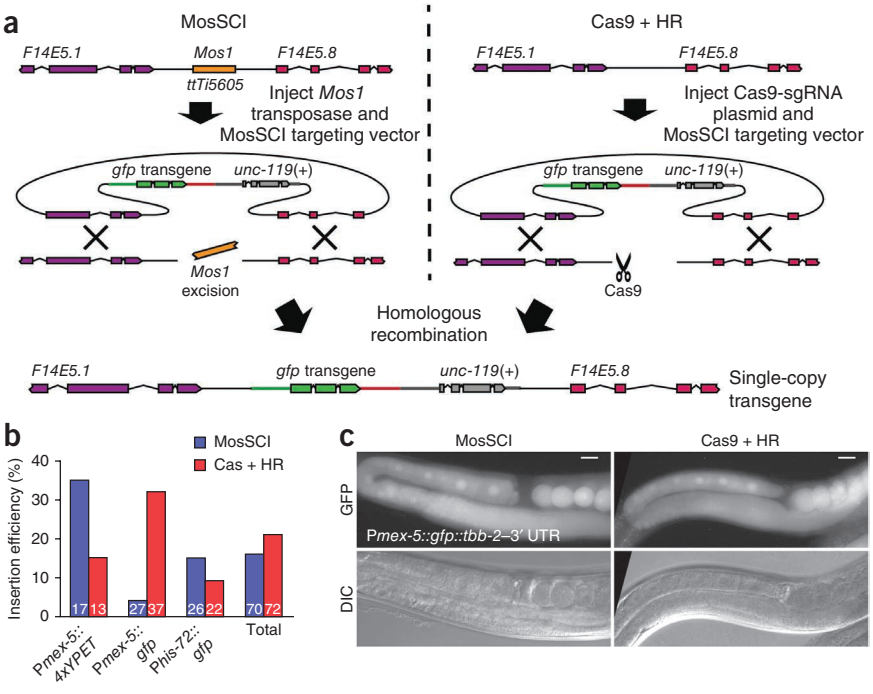

Figure 2 | Efficiency of Cas9-triggered homologous recombination in *C. elegans*. (**a**) Schematic for homologous recombination (HR) mediated by either *Mos1* transposon excision (left) or Cas9 (right). (**b**) Efficiency of single-copy transgene insertion for three different transgenes using either *Mos1*-mediated single-copy insertion (MosSCI) or Cas9. *n* values at the bottom of each bar indicate the number of successfully injected animals (those that yielded non-Unc progeny). Percent efficiency is the fraction of successfully injected animals that yielded integrated transgenes. See **Supplementary Table 1** for raw data. (**c**) Images of germline GFP expression from *Pmex-5::gfp::tbb-2*–3′ UTR transgenes generated using MosSCI or Cas9. Images were acquired, processed and displayed with identical settings. Results are representative of five animals of each strain. Scale bars, 20 μm.

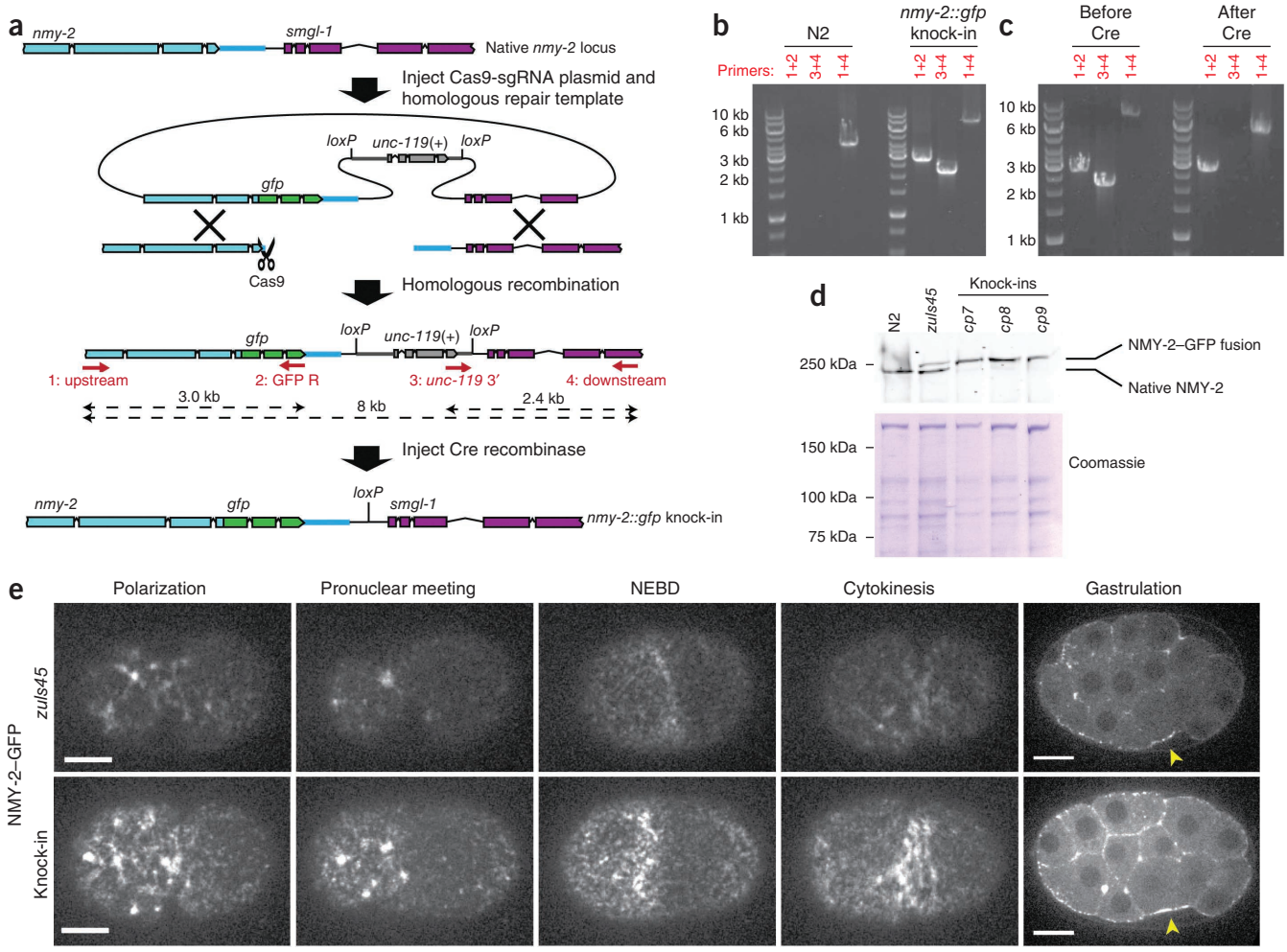

Figure 3 | Tagging of endogenous *nmy-2* with *gfp*. (**a**) Strategy for producing *nmy-2::gfp* knock-ins. Cas9 cleavage of the 3′ end of *nmy-2* stimulates homologous recombination, resulting in insertion of *gfp* and *unc-119*(+) into the genome. After isolating recombinants, we excised the *unc-119*(+) selectable marker by expressing Cre recombinase. (**b**) PCR genotyping of the *nmy-2* locus in the indicated strains, using primer pairs as indicated and as schematized in **a**. Results are representative of three independently isolated knock-in strains. (**c**) PCR genotyping of the *nmy-2* locus before and after excision of the *unc-119*(+) marker with Cre. Results are representative of five independent Cre-mediated *unc-119*(+) excision experiments. (**d**) Western blot showing NMY-2 levels in embryonic lysates in N2 (wild type), a strain carrying *zuIs45* and strains carrying three independent knock-in alleles. Coomassie staining of total protein is shown as a loading control. Results are representative of three independent experiments. (**e**) Stage-matched images of NMY-2–GFP localization in an *nmy-2::gfp* knock-in strain compared to in *zuIs45*. The embryos shown were placed side-by-side on the same coverslip and imaged simultaneously. The images in the four left columns are maximum-intensity projections of two 0.5-µm sections at a cortical focal plane and are taken from **Supplementary Video 1**. The far-right panels are single confocal sections from a different pair of embryos at gastrulation stage. Arrowheads indicate apical accumulation of NMY-2–GFP in gastrulating endodermal precursors. Results are representative of 14 independent experiments. Scale bars, 10 µm.

such as *gfp* into endogenous genes ensures 100% labeling and expression under the control of native regulatory elements and in the normal chromatin context.

To test whether Cas9-triggered homologous recombination could be used to insert protein tags into endogenous genes, we targeted the *nmy-2* gene, which encodes nonmuscle myosin II (**Fig. 3**). We built a homologous repair template comprising the C-terminal 1.5 kilobases (kb) of *nmy-2* fused in-frame to *gfp*, which was followed by the *nmy-2* 3′ UTR, an *unc-119*(+) selectable marker and 1.5 kb of downstream genomic sequence (**Fig. 3a**). The *unc-119*(+) gene was flanked by *loxP* sites, which allowed it to be removed by subsequent expression of Cre recombinase. We also generated a Cas9-sgRNA plasmid targeted to cleave at the 3′ end of *nmy-2*. Co-injection of the Cas9-sgRNA plasmid and homologous repair template into *unc-119* worms resulted in integration of *gfp* and *unc-119*(+) into the 3′ end of the *nmy-2* locus (three independent knock-in alleles from 60 total injected animals). We confirmed the correct integration of *gfp* at the 3′ end of the *nmy-2* gene in all three lines by PCR (**Fig. 3b**) and sequencing.

We examined the expression and localization of the NMY-2–GFP fusion protein in these three homozygous knock-in lines. For comparison, we analyzed a strain carrying *zuIs45*, a well-established transgene generated by microparticle bombardment[27]. The knock-in strains expressed NMY-2–GFP at levels similar to those of endogenous NMY-2, whereas the *zuIs45* strain expressed NMY-2–GFP at lower levels (**Fig. 3d**). NMY-2–GFP localized to the cell cortex in *nmy-2::gfp* knock-in embryos, with a pattern that was indistinguishable from that in embryos carrying *zuIs45*

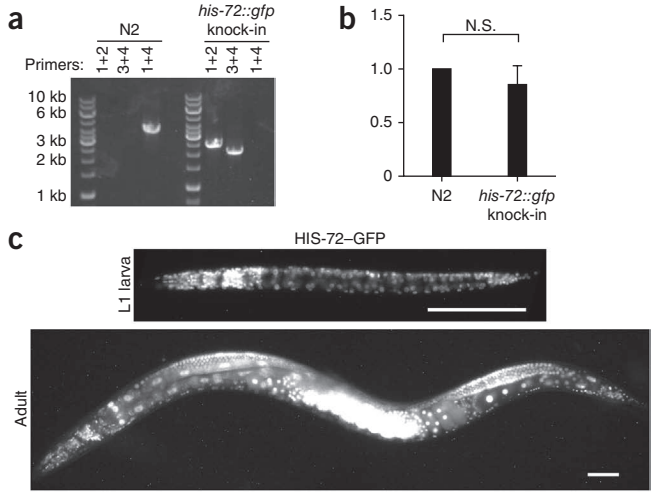

Figure 4 | Tagging of endogenous *his-72* with *gfp*. (**a**) PCR genotyping of the *his-72* locus in the indicated strains using a PCR strategy similar to that outlined in **Figure 3a,b**. (**b**) *his-72* mRNA expression levels in the indicated strains, as measured by quantitative reverse transcription PCR (qRT-PCR). Results are the average of three independent experiments, and error bar shows 95% confidence interval. N.S., not significant ($P = 0.07$, two-tailed *t*-test). (**c**) HIS-72–GFP fluorescence in whole worms at the indicated stages. Results are representative of seven animals imaged. Scale bars, 50 µm.

(**Fig. 3e** and **Supplementary Video 1**). However, the knock-in embryos showed consistently brighter fluorescence, which is consistent with the higher expression of NMY-2–GFP in the knock-ins.

Because *nmy-2* is an essential gene[28], we tested whether *gfp* insertion disrupted protein function by assaying for lethality. Two of three *nmy-2::gfp* knock-in strains were 100% homozygous viable, and a third strain was 99% viable ($n > 100$ embryos for each strain). Animals of all three knock-in strains displayed wild-type movement and had no discernable phenotypes. We conclude that insertion of *gfp* at the endogenous locus does not affect *nmy-2* gene function.

For some applications, the insertion of the *unc-119*(+) selectable marker into the genome may be problematic. We therefore developed a simple procedure to remove the *unc-119*(+) marker by injecting a plasmid encoding Cre recombinase under the control of the *eft-3* promoter and *tbb-2* 3′ UTR (Online Methods) and picking uncoordinated (Unc) worms from the F_2 progeny (**Fig. 3c**). We isolated Unc animals in 5 of 5 independent experiments (12–20 animals injected in each experiment). The high success rate achieved with relatively few injected animals indicates that Cre-mediated excision of *unc-119*(+) is highly efficient. We then outcrossed to wild-type worms to remove the unlinked *unc-119* mutation required for *unc-119*(+) selection. This resulted in a strain containing no known genomic modifications except for insertion of *gfp* into the *nmy-2* gene and a 34-bp *loxP* site in the intergenic region downstream of *nmy-2* (**Fig. 3a**). NMY-2–GFP expression in these animals was not altered by removal of the *unc-119*(+) selection marker (data not shown), suggesting that this marker does not affect fluorescent fusion protein expression when it is inserted into an intergenic region.

To test whether this strategy is likely to be broadly applicable, we used a similar approach to endogenously tag the *his-72* gene, which encodes histone H3, with GFP. We obtained one knock-in strain from ten successfully injected animals. We were able to amplify the left and right insertion junctions, confirming insertion of *gfp* into the *his-72* locus (**Fig. 4a**). However, we were unable to amplify across the insertion, and a more detailed PCR characterization showed that a rearrangement involving a duplication of the *unc-119*(+) cassette had occurred (data not shown). Such rearrangements have been reported to occur in other studies involving homologous recombination[20,29] and occurred at a low frequency in our study (1 of 16 strains generated by Cas9-triggered homologous recombination, 1 of 7 MosSCI strains). Of note, *his-72* mRNA levels in the *his-72::gfp* knock in-strain were indistinguishable from those of wild type, indicating that the rearrangement did not affect the *his-72* gene itself (**Fig. 4b**). These animals were healthy and showed bright nuclear GFP fluorescence in a wide range of tissues including the germ line, consistent with labeling of endogenous histone (**Fig. 4c**).

Generation of multiple point mutations in a single step

Another application of Cas9-triggered homologous recombination is the generation of targeted mutations at endogenous loci. To demonstrate this, we made mutations in *lin-31*, which encodes a FOXB transcription factor required for vulval development[30]. LIN-31 forms a complex with a LIN-1 (a homolog of mammalian Ets transcription factor), and this complex is thought to repress the primary vulval fate[31]. MPK-1, a mitogen-activated protein kinase (MAPK), can phosphorylate LIN-31 on four C-terminal threonines, and addition of active MAPK disrupts the interaction between LIN-31 and LIN-1 *in vitro*[31]. Overexpression of nonphosphorylatable LIN-31 causes a vulval phenotype[31], but whether MAPK phosphorylation affects the function of endogenous LIN-31 has not been tested directly.

We used Cas9-triggered homologous recombination to mutate the four MAPK phosphorylation sites at the C terminus of LIN-31 to either alanine (nonphosphorylatable) or glutamic acid (phosphomimetic) (**Fig. 5a**,**b**). We used a Cas9-sgRNA plasmid targeting a site 5′ of these four residues to ensure that the Unc-119(+) animals we isolated would contain mutations at all four sites (**Fig. 5b**). We obtained two independent *lin-31*(*4T→A*) alleles from 60 total injected animals and three independent *lin-31*(*4T→E*) alleles from 62 total injected animals. We confirmed the desired mutations by PCR (not shown) and sequencing (**Fig. 5c**).

Mutation of the four MAPK phosphorylation sites to alanine caused a distorted vulval morphology at the L4 stage and a partially penetrant protruding vulva phenotype in the adult. A similar phenotype was observed when these residues were mutated to glutamic acid, mimicking constitutive phosphorylation (**Fig. 5d**,**e** and **Supplementary Table 2**). Identical phenotypes were observed for two independently isolated *lin-31*(*4T→A*) lines and for three independently isolated *lin-31*(*4T→E*) lines. These results suggest that dynamic regulation of LIN-31 phosphorylation is important for normal vulval development.

Assessment of Cas9 specificity in *C. elegans*

Cas9 has been reported to produce off-target mutations in mammalian cells, which raises concerns about the specificity of this enzyme in genome-editing applications[32]. To assess the specificity of Cas9 in *C. elegans*, we identified the genomic sequences most similar to the targeting sequences that we used to modify *nmy-2*

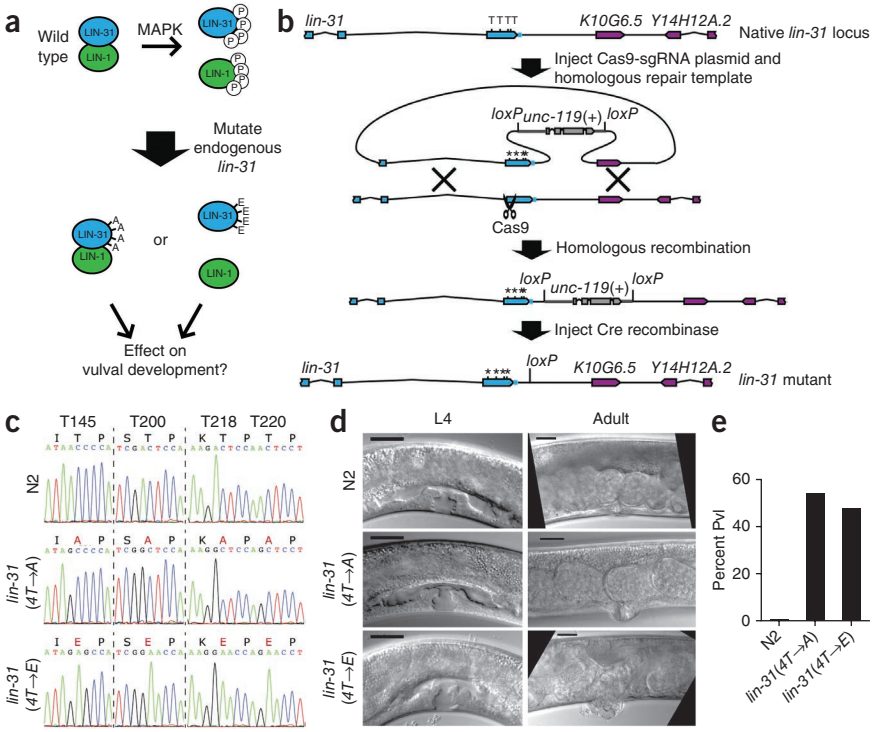

Figure 5 | Targeted mutations in an endogenous gene. (**a**) Model for how MAPK phosphorylation affects LIN-31 function[31] and the predicted effects of mutating these residues to either alanine or glutamic acid. (**b**) Strategy for simultaneous mutagenesis of four threonine residues (T145, T200, T218 and T220) in *lin-31* to either alanine or glutamic acid (asterisks). Cas9 is targeted to the 5′ end of the last exon of *lin-31*, ensuring mutation of all four threonine residues. (**c**) Sequence confirmation of the induced mutations in *lin-31* mutant strains. Results are representative of two independently isolated *lin-31*(*4T→A*) strains and three independently isolated *lin-31*(*4T→E*) strains. (**d**) Vulval morphology in L4 and adult animals of the indicated genotypes. Images are representative of 15 N2 animals, 35 *lin-31*(*4T→A*) animals and 36 *lin-31*(*4T→E*) animals. Scale bars, 20 µm. (**e**) Quantification of the Pvl phenotype of *lin-31* adults. Scoring of phenotypes was done in a blinded fashion. Results from multiple isolates of each *lin-31* mutation are combined; $n = 406$ for N2, $n = 267$ for *lin-31*(*4T→A*) and $n = 217$ for *lin-31*(*4T→E*). See **Supplementary Table 2** for the raw data.

and *lin-31*. We focused on candidate off-target sequences that closely matched the 3′ end of our targeting sequences, as Cas9 activity has been shown to be most sensitive to mismatches in the 3′ half of sgRNA targeting sequences[1,5,7,32]. We PCR-amplified and sequenced ten candidate off-target sites for the *nmy-2* sgRNA and four for the *lin-31* sgRNA (**Supplementary Table 3**). We found no mutations at any of these sites in any of the strains we isolated. We note that although Fu *et al.* detected off-target activity of Cas9 toward sequences with up to five mismatches to the sgRNA sequence[32], the closest matches that existed to our *nmy-2* and *lin-31* sgRNAs contained six or more mismatches each (**Supplementary Table 3**). The small size of the *C. elegans* genome compared to that of mammalian genomes may reduce the odds of closer off-target matches to particular target sequences.

DISCUSSION

We have demonstrated a method to efficiently induce essentially any desired modification in the *C. elegans* genome. This approach was robust and cost-effective: for example, we obtained three independent *nmy-2*::*gfp* knock-in lines from a single set of injections, requiring less than 4 weeks total time (of which less than 2 d was hands-on time) and less than $200 worth of materials (**Supplementary Fig. 2** and **Supplementary Protocol**). In 6 of 8 experiments, we obtained multiple independent lines on the first attempt. In the other two cases, the first set of injections failed for trivial technical reasons (Online Methods), and we readily obtained homologous recombinants upon reinjection. Thus, we achieved an overall success rate of 100% without extensive reengineering or optimization, which implies that our strategy is likely to be broadly applicable. In addition, the ability to remove the selectable marker in a single step using Cre recombinase should facilitate the use of *unc-119* or other selectable markers for a wide variety of genome-editing strategies.

Our method relies on double-strand break repair using an engineered homologous template, similar to earlier methods that used *Mos1* transposon excision to generate double-strand breaks[18–21]. The use of Cas9 in our system overcomes several important limitations of *Mos1*-based methods. First, Cas9 target sites occur once every 32 bp in random DNA sequence; we estimate that there are over 1 million potential Cas9 target sites in the *C. elegans* genome, compared to approximately 14,000 *Mos1* insertion sites[22,33]. This greatly expands the range of modifications that can be made because homologous recombination is most efficient within 0.5 kb of a double-strand break[18]. Indeed, none of the alleles presented in this study could have been made using *Mos1*-based methods because appropriate transposon insertion alleles do not exist. Second, by carefully choosing the positions of the Cas9 cleavage site and selectable marker relative to the desired modification, the investigator can ensure that the desired genome modification is present in every isolated recombinant (**Fig. 5b** and **Supplementary Protocol**), in contrast to *Mos1*-based methods, in which the length of a gene conversion track is stochastic and decays rapidly with increasing distance from the transposon insertion site[18]. Third, Cas9 could in principle be used in any genetic background, whereas *Mos1*-based genome modifications must be generated in a strain carrying a *Mos1* insertion. Finally, mobilization of *Mos1* generates an average of 2–3 new *Mos1* insertions at random sites elsewhere in the genome[33,34], which could cause undesired phenotypes.

Cas9-triggered homologous recombination also has advantages over conventional transgenic approaches for generating fluorescent-protein fusions. Generating and maintaining a knock-in line with our approach is cheaper and less labor intensive than either microparticle bombardment or extrachromosomal arrays. Also, it is expected to maintain the endogenous pattern of expression in most cases, and fusion-protein function can be easily assessed genetically owing to the absence of untagged protein. Knock-in strategies are the standard approach for

generating fluorescent-protein fusions in yeast for these reasons, and our methodology will allow widespread application of this approach in *C. elegans*.

Despite these advantages, there are some potential limitations associated with the use of Cas9-triggered homologous recombination to endogenously tag genes. First, fusion of *gfp* to the 3′ end of a gene may compromise protein function, though this can be detected if it causes a loss-of-function phenotype. Second, GFP-coding sequences are sometimes recognized as 'nonself' and silenced by the Piwi Argonaute RNA pathway[35]. Such a silencing event should again be detectable if it produces a loss-of-function phenotype and/or loss of GFP fluorescence. Third, the DNA repair mechanism involved in homologous recombination can sometimes generate complex rearrangements, as we observed for *his-72::gfp*[20,29]. Because these events are uncommon, a straightforward solution is to isolate multiple knock-in alleles for each gene of interest. We obtained multiple independent alleles from a single set of injections for 7 of the 8 Cas9-triggered homologous recombination experiments performed for this study (*his-72::gfp* was the exception). Finally, for some genes with low endogenous expression levels, fluorescence of *gfp* knock-ins may be too dim to visualize. Development of brighter fluorescent proteins[36–38] may facilitate knock-in approaches for genes with lower endogenous expression levels. If fluorescence in a knock-in strain is too dim to be useful, then the investigator may choose to use an overexpression strategy, accepting the caveat that overexpression artifacts are possible.

In using Cas9-triggered homologous recombination to generate targeted mutations in endogenous genes, we chose to make point mutants; but we expect that insertions, deletions or other modifications could be made with similar ease, as has been done previously using *Mos1* excision[20,21]. Of note, single point mutants have also been made in *C. elegans* using TALEN cleavage and ssDNA oligo-mediated repair[15], but this approach is limited by the length of a synthetic DNA oligonucleotide and would have required at least two sequential injection steps to produce the four *lin-31* point mutations that we were able to generate in a single step.

Interestingly, the phenotype we observed in *lin-31* mutants was different from that reported by Tan *et al.*[31], who found that expression of *lin-31*(*4T→A*) from an extrachromosomal array inhibited vulval fate specification. We suggest that repression of vulval fate in earlier experiments may have been due to overexpression of LIN-31 protein rather than solely its inability to be phosphorylated by MPK-1. The ability to quickly and efficiently induce mutations in endogenous genomic loci renders the use of multicopy extrachromosomal arrays unnecessary and should greatly simplify the interpretation of reverse genetic experiments.

In summary, we have developed a flexible, inexpensive and robust strategy for genome editing in *C. elegans* using Cas9 targeted cleavage and homologous recombination. Given the ease with which our approach can be adapted to new targets, we suggest that the ability to modify the *C. elegans* genome is now limited only by the imagination of the investigator.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. Addgene: Cas9-sgRNA plasmid targeting a site near *ttTi5605*, 47550; Cas9-sgRNA plasmid with no targeting sequence, 47549; P*eft-3::Cre::tbb-2*–3′ UTR construct, 47551.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank C. Frøkjær-Jensen (University of Utah) for sharing strains, plasmids and protocols; K. Kemphues (Cornell University) for antibodies; and K. Bloom, A. Maddox, G. Monsalve, N. Pujol, K. Slep, S. Taubert, K. Yamamoto and members of the Goldstein lab for helpful suggestions and comments on the manuscript. Some strains were provided by the *Caenorhabditis* Genetics Center, which is funded by the US National Institutes of Health (NIH) Office of Research Infrastructure Programs (P40 OD010440). This work was supported by NIH T32 CA009156 and a Howard Hughes postdoctoral fellowship from the Helen Hay Whitney Foundation (D.J.D.); a postdoctoral fellowship from the Canadian Institutes of Health Research (award #234765) (J.D.W.); NIH R01 GM085309 (D.J.R.); NIH CA20535 and US National Science Foundation (NSF) MCB 1157767 (K. Yamamoto); and NIH R01 GM083071 and NSF IOS 0917726 (B.G.).

AUTHOR CONTRIBUTIONS

D.J.D. and J.D.W. jointly conceived of the project, and all authors discussed and contributed to the experimental design. D.J.D. performed the experiments. D.J.D. and D.J.R. analyzed the data. D.J.D. prepared the manuscript, and all authors discussed and contributed to the final version.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
2. Hwang, W.Y. *et al.* Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* **31**, 227–229 (2013).
3. DiCarlo, J.E. *et al.* Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic Acids Res.* **41**, 4336–4343 (2013).
4. Gratz, S.J. *et al.* Genome engineering of *Drosophila* with the CRISPR RNA-guided Cas9 nuclease. *Genetics* **194**, 1029–1035 (2013).
5. Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013).
6. Wang, H. *et al.* One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* **153**, 910–918 (2013).
7. Jiang, W., Bikard, D., Cox, D., Zhang, F. & Marraffini, L.A. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotechnol.* **31**, 233–239 (2013).
8. Friedland, A.E. *et al.* Heritable genome editing in *C. elegans* via a CRISPR-Cas9 system. *Nat. Methods* **10**, 741–743 (2013).
9. Bassett, A.R., Tibbit, C., Ponting, C.P. & Liu, J.-L. Highly efficient targeted mutagenesis of *Drosophila* with the CRISPR/Cas9 system. *Cell Rep.* **4**, 220–228 (2013).
10. Chang, N. *et al.* Genome editing with RNA-guided Cas9 nuclease in zebrafish embryos. *Cell Res.* **23**, 465–472 (2013).
11. Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826 (2013).
12. Cho, S.W., Kim, S., Kim, J.M. & Kim, J.-S. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat. Biotechnol.* **31**, 230–232 (2013).
13. Shen, B. *et al.* Generation of gene-modified mice via Cas9/RNA-mediated gene targeting. *Cell Res.* **23**, 720–723 (2013).
14. Wood, A.J. *et al.* Targeted genome editing across species using ZFNs and TALENs. *Science* **333**, 307 (2011).
15. Lo, T.-W. *et al.* Heritable genome editing using TALENs and CRISPR/Cas9 to engineer precise insertions and deletions in evolutionarily diverse nematode species. *Genetics* doi:10.1534/genetics.113.155382 (9 August 2013).
16. Gerstein, M.B. *et al.* Integrative analysis of the *Caenorhabditis elegans* genome by the modENCODE project. *Science* **330**, 1775–1787 (2010).
17. Frøkjær-Jensen, C., Davis, M.W., Ailion, M. & Jorgensen, E.M. Improved Mos1-mediated transgenesis in *C. elegans*. *Nat. Methods* **9**, 117–118 (2012).

18. Robert, V. & Bessereau, J.-L. Targeted engineering of the *Caenorhabditis elegans* genome following Mos1-triggered chromosomal breaks. *EMBO J.* **26**, 170–183 (2007).
19. Robert, V.J., Davis, M.W., Jorgensen, E.M. & Bessereau, J.-L. Gene conversion and end-joining-repair double-strand breaks in the *Caenorhabditis elegans* germline. *Genetics* **180**, 673–679 (2008).
20. Frøkjær-Jensen, C. *et al.* Single-copy insertion of transgenes in *Caenorhabditis elegans*. *Nat. Genet.* **40**, 1375–1383 (2008).
21. Frøkjær-Jensen, C. *et al.* Targeted gene deletions in *C. elegans* using transposon excision. *Nat. Methods* **7**, 451–453 (2010).
22. Vallin, E. *et al.* A genome-wide collection of Mos1 transposon insertion mutants for the *C. elegans* research community. *PLoS ONE* **7**, e30482 (2012).
23. Mello, C.C., Kramer, J.M., Stinchcomb, D. & Ambros, V. Efficient gene transfer in *C.elegans*: extrachromosomal maintenance and integration of transforming sequences. *EMBO J.* **10**, 3959–3970 (1991).
24. Kelly, W.G., Xu, S., Montgomery, M.K. & Fire, A. Distinct requirements for somatic and germline expression of a generally expressed *Caenorhabditis elegans* gene. *Genetics* **146**, 227–238 (1997).
25. Praitis, V., Casey, E., Collar, D. & Austin, J. Creation of low-copy integrated transgenic lines in *Caenorhabditis elegans*. *Genetics* **157**, 1217–1226 (2001).
26. Sarov, M. *et al.* A genome-scale resource for *in vivo* tag-based protein function exploration in *C. elegans*. *Cell* **150**, 855–866 (2012).
27. Nance, J., Munro, E.M. & Priess, J.R. *C. elegans* PAR-3 and PAR-6 are required for apicobasal asymmetries associated with cell adhesion and gastrulation. *Development* **130**, 5339–5350 (2003).
28. Guo, S. & Kemphues, K.J. A non-muscle myosin required for embryonic polarity in *Caenorhabditis elegans*. *Nature* **382**, 455–458 (1996).
29. Berezikov, E., Bargmann, C.I. & Plasterk, R.H.A. Homologous gene targeting in *Caenorhabditis elegans* by biolistic transformation. *Nucleic Acids Res.* **32**, e40 (2004).
30. Ferguson, E.L. & Horvitz, H.R. Identification and characterization of 22 genes that affect the vulval cell lineages of the nematode *Caenorhabditis elegans*. *Genetics* **110**, 17–72 (1985).
31. Tan, P.B., Lackner, M.R. & Kim, S.K. MAP kinase signaling specificity mediated by the LIN-1 Ets/LIN-31 WH transcription factor complex during *C. elegans* vulval induction. *Cell* **93**, 569–580 (1998).
32. Fu, Y. *et al.* High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* doi:10.1038/nbt.2623 (23 June 2013).
33. Granger, L., Martin, E. & Ségalat, L. Mos as a tool for genome-wide insertional mutagenesis in *Caenorhabditis elegans*: results of a pilot study. *Nucleic Acids Res.* **32**, e117 (2004).
34. Williams, D.C., Boulin, T., Ruaud, A.-F., Jorgensen, E.M. & Bessereau, J.-L. Characterization of Mos1-mediated mutagenesis in *Caenorhabditis elegans*: a method for the rapid identification of mutated genes. *Genetics* **169**, 1779–1785 (2005).
35. Shirayama, M. *et al.* piRNAs initiate an epigenetic memory of nonself RNA in the *C. elegans* germline. *Cell* **150**, 65–77 (2012).
36. Lam, A.J. *et al.* Improving FRET dynamic range with bright green and red fluorescent proteins. *Nat. Methods* **9**, 1005–1012 (2012).
37. Nguyen, A.W. & Daugherty, P.S. Evolutionary optimization of fluorescent proteins for intracellular FRET. *Nat. Biotechnol.* **23**, 355–360 (2005).
38. Shcherbo, D. *et al.* Far-red fluorescent tags for protein imaging in living tissues. *Biochem. J.* **418**, 567–574 (2009).

ONLINE METHODS

Strains and nomenclature. New genetic nomenclature for genome editing applications has been developed by the WormBase Gene Name Curators (J. Hodgkin, University of Oxford, and T. Schedl, Washington University School of Medicine, personal communication). Briefly, edited loci are assigned conventional allele designations, with the nature of the modification described in brackets after the allele name. For example, one of our *nmy-2::gfp* knock-in alleles is *nmy-2(cp7[nmy-2::gfp + LoxP unc-119(+) LoxP])* I, and a *lin-31* mutant *lin-31(cp2[T145A T200A T218A T220A + LoxP unc-119(+) LoxP])* II. Note that each independently isolated mutant line is given its own allele designation, even if the molecular lesion is identical. When the *unc-119*(+) selectable marker has been removed using Cre recombinase, a new allele designation is assigned. For example, an *nmy-2::gfp* knock-in allele with the *unc-119*(+) cassette removed is designated *nmy-2(cp13[nmy-2::gfp + LoxP])* I. The *cp13* allele was derived from *cp7* by Cre-mediated recombination.

Supplementary Table 4 lists all strains generated and used in this study. All strains were kept at 25 °C and fed *Escherichia coli* strain OP50 except where noted below and were handled using standard techniques[39].

Plasmid construction. Plasmids have been deposited in Addgene with the following accession numbers: Cas9-sgRNA plasmid targeting a site near *ttTi5605*, 47550; Cas9-sgRNA plasmid with no targeting sequence, 47549; and the P*eft-3::Cre::tbb-2*–3′ UTR construct, 47551. All other plasmids used in this study are available from the authors upon request.

To construct the Cas9-sgRNA expression plasmid shown in **Figure 1c**, we first designed a synthetic gene encoding Cas9, with *C. elegans* coding bias and synthetic *C. elegans* introns, using the *C. elegans* Codon Adaptor[40]. Our Cas9 sequence includes a nuclear localization signal and an HA tag at the C terminus. The synthetic gene was produced as a series of overlapping 500-bp gBlocks (Integrated DNA Technologies), assembled using Gibson Assembly (New England BioLabs) and inserted into the vector pCFJ601 (P*eft-3::Mos1 Transposase::tbb-2*–3′ UTR)[17] in place of the *Mos1* transposase. Next, a gBlock containing the *U6* promoter and sgRNA sequence was inserted 3′ of the *tbb-2* 3′ UTR. Genomic targets of Cas9 conform to the target sequence $GN_{19}NGG$, where N is any base. The initial G is a requirement for transcription initiation by the *U6* promoter, and the NGG (PAM) motif is required for Cas9 activity (note that the NGG motif must be present in the genomic target but is not included in the sgRNA sequence). To target Cas9 to different genomic sequences, we inserted the desired targeting sequence into the Cas9-sgRNA construct using the Q5 site-directed mutagenesis kit (New England BioLabs) with forward primer 5′-N_{19}GTTTTAGAGCTAGAAATAGCAAGT-3′, where N_{19} is replaced by the desired 19-bp targeting sequence, and reverse primer 5′-CAAGACATCTCGCAATAGG-3′.

Supplementary Table 5 lists the targeting sequences used in this study.

Targeting vectors for single-copy transgene insertion on chromosome II were constructed in the pCFJ150 vector backbone[20] using Gateway cloning. We used site-directed mutagenesis with the Q5 site-directed mutagenesis kit to delete a short region of the 3′ recombination arm comprising the Cas9 target sequence to prevent the homologous repair templates from being cleaved by Cas9.

Homologous repair templates for *gfp* insertion and *lin-31* mutagenesis were constructed in two steps. First, we PCR-amplified a 3- to 4-kb region centered on the desired modification from N2 genomic DNA and cloned the resulting fragment into the pCR-Blunt vector using the ZeroBlunt TOPO Cloning Kit (Life Technologies). Second, we modified this genomic clone by inserting *gfp* (for *gfp* knock-ins) or a 3′ exon containing point mutations (for *lin-31* mutagenesis), along with the *unc-119*(+) rescue gene flanked by *loxP* sites. *gfp* and *unc-119*(+) fragments were generated by PCR, and *LoxP* sites were included in the *unc-119*(+) primers. The mutated *lin-31* 3′ exons were synthesized as gBlocks. These fragments were integrated into the genomic clones using Gibson assembly, which allows for seamless fusion of DNA fragments without the need to include any extra sequence (for example, restriction sites). To avoid cleavage of the repair templates by Cas9, we deleted or mutated the Cas9 target site in all repair templates. Complete plasmid sequences of all targeting vectors are available from the authors upon request.

To construct the P*eft-3::Cre::tbb-2*–3′ UTR plasmid used for removal of selectable markers with Cre recombinase, we first amplified the Cre ORF from the plasmid pEM3 (ref. 41) and cloned it into the Gateway donor vector pDONR221. We then performed a three-fragment gateway reaction using our Cre donor vector, pCFJ386 (P*eft-3*; a gift from C. Frøkjær-Jensen), pCM1.36 (*tbb-2* 3′ UTR)[42] and the destination vector pCFJ212 (ref. 17), which contains an *unc-119*(+) rescue gene.

Supplementary Table 6 lists all primers used in this study.

Single-copy transgene insertion with MosSCI. We inserted transgenes into the *ttTi5605 Mos1* site by following a published MosSCI protocol[17]. We prepared an injection mix containing 10 ng/μL targeting vector, 50 ng/μL pCFJ601 (P*eft-3::Mos1 Transposase*), 10 ng/μL pMA122 (heat-shock driven PEEL-1 negative selection), 10 ng/μL pGH8 (P*rab-3::mCherry* neuronal co-injection marker), 5 ng/μL pCFJ104 (P*myo-3::mCherry* body-wall muscle co-injection marker) and 2.5 ng/μL pCFJ90 (P*myo-2::mCherry* pharyngeal co-injection marker). The mixture was microinjected into the gonads of Unc young adults of strain EG6699 (*ttTi5605* II; *unc-119*(*ed3*) III), which were raised on HB101 bacteria at 15 °C. After injection, single worms were picked to new plates and maintained at 25 °C until starvation (10–12 d). Plates containing non-Unc worms were counted as successfully injected. Occasionally, a batch of 50–60 injected animals yielded fewer than five successful injections, and we concluded that the injections had failed (usually for technical reasons, such as a bad needle) and repeated the injections. Following successful injections, we applied heat shock to plates with non-Unc worms at 34 °C for 4 h in an air incubator to activate the PEEL-1 negative-selection marker, which kills animals carrying extrachromosomal arrays. After overnight recovery at 25 °C, plates were visually screened to identify non-Unc animals that survived heat shock and did not express the red fluorescent co-injection markers. Single worms from these plates were picked to establish lines, and the presence of single-copy inserts was confirmed by PCR using primers listed in **Supplementary Table 6**.

Cas9-triggered homologous recombination. To modify the genome using Cas9-triggered homologous recombination, we followed a protocol similar to that for MosSCI (above). A mix

containing 10 ng/µL homologous repair template, 50 ng/µL Cas9-sgRNA plasmid, and the negative selection and co-injection markers listed above was injected into young adults of strain DP38 (*unc-119*(*ed3*) III)[43]. Note that although we used DP38 for the experiments reported in this study, we are recommending the use of the outcrossed derivative HT1593 for future experiments (see **Supplementary Protocol**). The procedure for selecting insertions and eliminating extrachromosomal arrays was identical to that described above. We were able to isolate strains that were either homozygous or heterozygous for all of our modifications (D.J.D., unpublished observations); for the experiments presented in this study, only the homozygous lines were kept.

Removal of *unc-119*(+) using Cre recombinase. An injection mix containing 50 ng/µL of pDD104 (P*eft-3*::*Cre*::*tbb-2*–3′ UTR) and 2.5 ng/µL pCFJ90 (P*myo-2*::*mCherry* pharyngeal marker) was injected into the gonads of young adult animals carrying an *unc-119*(+) cassette flanked by *loxP* sites. In each experiment, 15–20 animals were injected and placed at 25 °C. We picked single F_1 progeny (10–20 per experiment) expressing the red pharyngeal marker, which represent progeny of successful injections. We then selected Unc animals from among the F_2 progeny of these mCherry-positive animals. Because our Cre expression construct also carries *unc-119*(+), only animals that have (i) excised both genomic copies of the *unc-119*(+) cassette and (ii) lost the extrachromosomal array generated by injecting the Cre expression construct will be Unc. We verified that animals picked during this step were mCherry-negative and segregated only Unc progeny. Excision of *unc-119*(+) was then confirmed by PCR (see **Fig. 3c**).

Antibodies and western blotting. Embryos were isolated from gravid adult worms by bleaching and were lysed by sonication for 20 min in a bath sonicator filled with boiling water. Lysates were separated on 3–8% NuPAGE Tris-acetate gels (Life Technologies) and transferred to nitrocellulose. The following antibodies were used: rabbit anti-NMY-2 (ref. 28) at 1:1,000 dilution of crude serum and Alexa Fluor 790 goat anti-rabbit (Life Technologies, catalog number A11369) at 1:1,000 dilution. Blots were scanned on an Odyssey imaging system (LI-COR Biosciences).

Microscopy. DIC and fluorescence imaging of whole worms was performed using a Nikon Eclipse E800 microscope equipped with epifluorescence and Nomarski DIC optics. Worms were mounted on 2.5% agar pads containing 10 mM sodium azide as a paralytic.

For NMY-2–GFP imaging, early one-cell embryos were mounted on poly(L-lysine)-coated coverslips and gently flattened using 2.5% agar pads. Images were captured using a Nikon Eclipse Ti microscope equipped with a 60×, 1.4-NA objective and a Yokogawa CSU-X1 spinning disk head.

Maximum-intensity projection and adjustment of brightness and contrast were done using Fiji. No other image manipulations were performed.

qRT-PCR. RNA was isolated from gravid adult worms as follows. Worms were picked into Trizol reagent (Life Technologies) and lysed by repeated freeze-thaw cycles. Chloroform was added to a final concentration of 20% to separate phases, and the upper aqueous phase was mixed with an equal volume of ethanol and loaded on a RNeasy spin column (Qiagen). On-column DNase digestion was performed with the Qiagen on-column DNase digestion kit, and then RNA was washed and eluted according to the manufacturer's instructions. 25–75 ng total RNA was used for cDNA synthesis with the Superscript III reverse transcriptase kit (Life Technologies). qPCR was performed using a Viia 7 real-time PCR instrument and SYBR Green PCR Master Mix (Life Technologies). *his-72* transcripts were detected with forward primer 5′-TCGTTCGTGAGATTGCCCAG-3′ and reverse primer 5′-GAGTCCGACGAGGTATGCTT-3′. *Y45F10D.4* was used for normalization[44,45]. Data were analyzed to determine *his-72* expression levels in Viia 7 software; we used the default settings for a relative standard-curve experiment.

Screening for off-target mutations induced by Cas9. Candidate off-target cleavage sites for each sgRNA were identified by BLAST searches[46] against the *C. elegans* genomic sequence. For each candidate off-target site, we PCR-amplified an ~1-kb fragment centered on the candidate site from genomic DNA isolated from N2 control animals and from each modified strain we generated. The PCR products were sequenced by Eton Bioscience using primers binding to each end of the fragment. The sequence reads were aligned to the genomic sequence to look for insertion or deletion mutations at the putative cleavage site.

Reproducibility. Our sample sizes were chosen to allow confidence in the results while maintaining feasibility, and they are consistent with established norms for *C. elegans* research and the developmental biology community more broadly. No randomization was necessary for these studies. Phenotypes presented in **Figure 5d**,**e** were scored blindly.

39. Stiernagle, T. in *WormBook* (ed. The *C. elegans* Research Community) doi:10.1895/wormbook.1.101.1 (2006).
40. Redemann, S. *et al.* Codon adaptation-based control of protein expression in *C. elegans*. *Nat. Methods* **8**, 250–252 (2011).
41. Macosko, E.Z. *et al.* A hub-and-spoke circuit drives pheromone attraction and social behaviour in *C. elegans*. *Nature* **458**, 1171–1175 (2009).
42. Merritt, C., Rasoloson, D., Ko, D. & Seydoux, G. 3′ UTRs are the primary regulators of gene expression in the *C. elegans* germline. *Curr. Biol.* **18**, 1476–1482 (2008).
43. Maduro, M. & Pilgrim, D. Identification and cloning of *unc-119*, a gene expressed in the *Caenorhabditis elegans* nervous system. *Genetics* **141**, 977–988 (1995).
44. Zhang, Y., Chen, D., Smith, M.A., Zhang, B. & Pan, X. Selection of reliable reference genes in *Caenorhabditis elegans* for analysis of nanotoxicity. *PLoS ONE* **7**, e31849 (2012).
45. Hoogewijs, D., Houthoofd, K., Matthijssens, F., Vandesompele, J. & Vanfleteren, J.R. Selection and validation of a set of reliable reference genes for quantitative sod gene expression analysis in *C. elegans*. *BMC Mol. Biol.* **9**, 9 (2008).
46. Altschul, S.F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).

doi:10.1038/nmeth.2641