Targeted gene deletions in *C. elegans* using transposon excision

Christian Frøkjær-Jensen[1,2], M Wayne Davis[1], Gunther Hollopeter[1], Jon Taylor[3], Todd W Harris[4], Paola Nix[5], Rachel Lofgren[1], Michael Prestgard-Duke[1], Michael Bastiani[5], Donald G Moerman[3] & Erik M Jorgensen[1]

We developed a method, MosDEL, to generate targeted knockouts of genes in *Caenorhabditis elegans* by injection. We generated a double-strand break by mobilizing a *Mos1* transposon adjacent to the region to be deleted; the double-stranded break is repaired using injected DNA as a template. Repair can delete up to 25 kb of DNA and simultaneously insert a positive selection marker.

Gene-specific knockouts are a defining technology of reverse genetics, allowing phenotypes to be assigned to any of the thousands of genes identified in genome sequencing projects. In *Caenorhabditis elegans,* reverse genetics has mainly relied on random chemical mutagenesis to generate loss-of-function mutants[1] or, more recently, random deletions downstream of guanine-quadruplex DNA[2]. In both cases, mutagenized populations are screened by PCR with gene-specific primers for random deletions in target genes. This approach has been used to generate putative loss-of-function alleles in more than 5,000 genes, mainly through the efforts of the *C. elegans* Gene Knockout Consortium in the United States and Canada and the National BioResource Project in Japan[1,3]. Random deletions have a few limitations. First, deletions are typically small and not necessarily molecularly null. Second, chemical mutagenesis invariably leads to background mutations. And third, some deletions involve complex rearrangements[1].

In fruit flies, large deficiencies can be generated by recombination between *FRT* sites in *P* elements[4]. In other genetic model organisms (for example, yeast and mice) transgenic DNA fragments and homologous recombination are used to generate targeted deletions. Bombardment with DNA-coated gold particles can lead to gene replacement in *C. elegans* as well[5]. Unfortunately, the frequency of homologous recombination is low and the technique has not been widely adopted.

Endogenous Tc1 transposons have been used in *C. elegans* to inactivate genes by causing random deletions after excision[6]. More recently the *Drosophila mauritiana* transposon *Mos1* has been adapted for gene targeting[7]. To facilitate the use of *Mos1* elements, approximately 14,000 molecularly mapped transposon inserts have been generated by the NemaGENETAG consortium[8]. Mobilization of transposons generates double-stranded DNA breaks; repair of the breakpoint can lead to targeted modifications when a repair template is present[9]. This repair mechanism has been used to develop *Mos1* excision–induced transgene-instructed gene conversion (MosTIC)[7], a technique that can be used to reliably modify DNA within one kilobase of a *Mos1* insert. Building on these efforts, we had developed *Mos1*-mediated single copy insertion (MosSCI), a technique to insert transgenes into well-defined genomic loci[10]. *Mos1* excision is induced using a simple injection-based method and successful insertions are identified using an inserted selectable marker[10]. Here we demonstrate that *Mos1* excision can be used to generate targeted deletions of up to 25 kb. We call this technique *Mos1*-mediated deletion (MosDEL).

We generated deletions in a strain with a *Mos1* element adjacent to the gene of interest (**Fig. 1a**). The *Mos1* element was excised in the germline by the *Mos1* transposase encoded on an injected helper plasmid. Excision resulted in a double-stranded DNA break, which was repaired using the coinjected repair template. Previously we had demonstrated that single-copy DNA can be incorporated into the genome by flanking the DNA with homology regions from both sides of the double-strand break. Under these conditions, both free ends of the chromosome break have homologous DNA in the repair template to initiate repair. Here we used targeting constructs capable of only one-sided repair, such that one of the broken ends has no homologous sequence to invade (**Fig. 1a**). Repair is initiated from one side by strand invasion of the homology arm on the template (right homology region; typically 2 kb). DNA is then copied from the extrachromosomal array, including sequence from a distal homology arm (left homology region; typically 3 kb). The 3′ end can then invade the other half of the broken chromosome at a distance from the break. A wild-type copy of the *unc-119* gene is simultaneously inserted to provide a positive selection marker for deletions. Red fluorescent markers encoded on a co-injected plasmid are used to visually identify worms rescued by extrachromosomal arrays so that they can be disregarded.

To test the targeting strategy, we selected a *Mos1* element (*cxTi10882*) located 1.1 kb 3′ from *dpy-13* (**Fig. 1b**). *dpy-13* mutants are viable and can be identified based on the dumpy phenotype. We injected 83 *unc-119 cxTi10882* worms with

[1]Howard Hughes Medical Institute, Department of Biology, University of Utah, Salt Lake City, Utah, USA. [2]University of Copenhagen, Department of Biomedical Sciences and Danish National Research Foundation Centre for Cardiac Arrhythmia, Copenhagen, Denmark. [3]University of British Columbia, Department of Zoology, Vancouver, British Columbia, Canada. [4]Hi-Line Informatics, Livingston, Montana, USA. [5]Department of Biology, University of Utah, Salt Lake City, Utah, USA. Correspondence should be addressed to E.M.J. (jorgensen@biology.utah.edu).

RECEIVED 17 DECEMBER 2009; ACCEPTED 29 MARCH 2010; PUBLISHED ONLINE 25 APRIL 2010; CORRECTED ONLINE 6 MAY 2010 (DETAILS ONLINE); DOI:10.1038/NMETH.1454

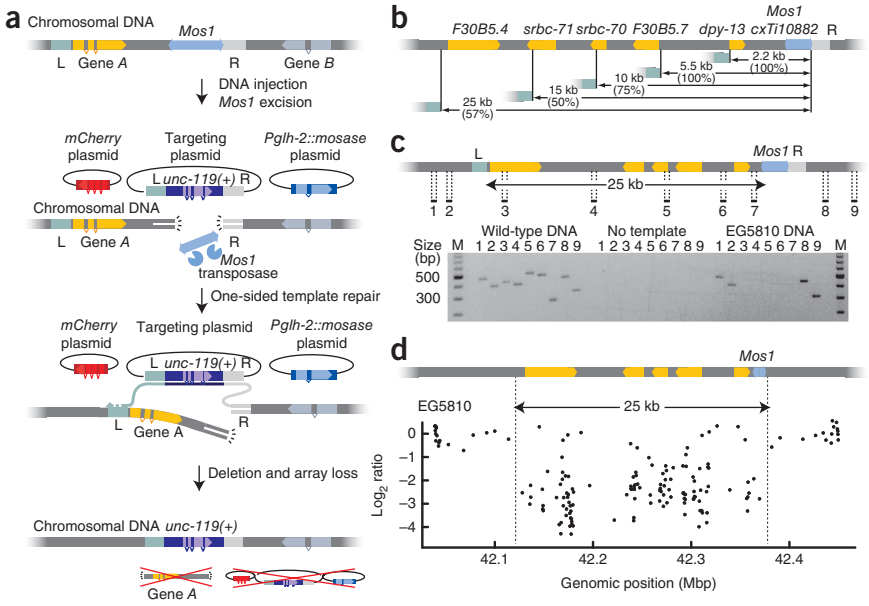


Figure 1 | Using *Mos1* transposons to create targeted deletions. (**a**) Schematic of MosDEL. A *Mos1* transposon near the gene of interest is placed into an *unc-119(ed3)* strain. Injection of a plasmid encoding *Mos1* transposase (*Pglh-2::mosase*) results in excision of the transposon. The resulting double-strand DNA break is repaired by synthesis-dependent strand annealing at the right homology region (labeled "R") on the extrachromosomal targeting plasmid, incorporating the positive selection marker *unc-119(+)* into the chromosomal locus. Genomic DNA between the right homology region and left homology region (labeled "L") is deleted when the nascent repair strand re-invades the genomic DNA and resolves the break by homologous recombination. Red fluorescence (from coinjected plasmids encoding mCherry) marks the extra-chromosomal array; deletion mutants are isolated by screening for *unc-119* rescued worms lacking red fluorescence. *Pglh-2::mosase*, mCherry plasmid and the targeting plasmid are injected as separate plasmids; extrachromosomal arrays are genetically unstable and therefore lost at high frequency. (**b**) Schematic of deletions generated at the *Mos1* insertion *cxTi10882(IV)*. Targeting efficiencies (percentage of deletions per *unc-119* insertion) are indicated in parentheses. (**c**) Schematic of the 25 kb deletion in strain EG5810 (top). PCR oligonucleotides were designed to amplify 200–500 bp fragments outside and inside the targeted region (labeled 1–9). Deletions were validated by PCR (bottom). M, 100 bp marker. (**d**) For CGH verification, EG5810 (25 kb deletion) and wild-type DNAs were hybridized to a *C. elegans* specific CGH chip[12]. A $\log_2$ ratio below −2 indicates deleted DNA. Points within the deletion that show normal hybridization are likely due to nonspecific DNA binding.

a targeting plasmid that deleted the full *dpy-13* coding region (2.2 kb deletion), and inserted a GFP marker expressed in coelomocytes and the *unc-119*(+) selection marker. After two generations (one week), we visually screened plates for the presence of *unc-119*(+) worms that lacked the co-injection markers. We identified three putative deletion strains (3 of 83 injected worms = 3.7%). Each strain was homozygous for *dpy-13* mutations, the *unc-119(+)* marker and expressed GFP in coelomocytes; we confirmed the 2.2 kb deletion by PCR analysis (**Supplementary Fig. 1** and **Supplementary Table 1**).

In most organisms, the efficiency of transgene-instructed gene conversion drops rapidly with distance from the DNA breakpoint; only 25% of conversions extend 1 kb from the breakpoint[7,11]. To determine how much DNA could be feasibly deleted using MosDEL, we injected additional constructs coding for 5, 10, 15, 25, 35 and 50 kb deletions (**Fig. 1b**). The frequency of *unc-119(+)* insertion was largely independent of the targeting construct: we found rescued strains in the progeny of about 5% of the injected worms (**Supplementary Table 1**). For example, we isolated seven stable lines with *unc-119(+)* insertions from 81 injected worms (9%) using the 25 kb construct. Two of the lines did not exhibit a Dpy phenotype and thus were simply insertions of the *unc-119(+)* transgene. Five strains exhibited a Dpy phenotype. One of these five strains did not have a deletion of the entire 25 kb stretch, whereas the other four had deletions of the entire 25 kb region (4/7 lines = 57% deletions correct) based on PCR analysis (**Fig. 1c**). However, the generation of deletions decreased sharply for events larger than 25 kb: we did not recover full-length deletions using the 35 kb and 50 kb templates (**Supplementary Table 1**). For these larger deletions, all *unc-119(+)* insertions (13 strains) were incomplete deletions (data not shown).

We used comparative genome hybridization (CGH) to verify deletion endpoints and confirm PCR results. In CGH, fluorescently labeled mutant DNA is compared to binding of wild-type DNA on a high-density array of gene-specific oligonucleotides[12]. CGH analysis verified the deletion of the targeted genomic regions in the three strains tested: a 25 kb deletion (**Fig. 1d**) and two 15 kb deletions (data not shown).

Deletion of essential genes is lethal when homozygous. Such deletions must be maintained as heterozygotes over a balancer chromosome; the balancer chromosome usually contains a marker that identifies loss of heterozygosity. Balanced strains often degenerate because of recombination between the marker and deletion on the homologous chromosomes. MosDEL leads to perfectly balanced lethal chromosomes by inserting a selectable marker (*unc-119(+)*) at the site of the deletion. Rescue of the *unc-119(−)* uncoordinated phenotype identifies the presence of the deletion in heterozygotes. Because *unc-119* worms are subviable, the homozygous balancer worms are selected against, and the deletion chromosome is maintained as a heterozygote. To test the utility of this feature, we deleted the entire coding region of the *C. elegans* dynamin ortholog *dyn-1* (**Fig. 2a**). From 100 injected worms, we obtained 17 strains that were *unc-119(+)* and did not contain extrachromosomal arrays based on the absence of red fluorescent markers. As expected from a lethal mutation balanced by the insertion of *unc-119(+)*, we could not generate homozygous *unc-119(+)* rescued worms for 11 of 17 strains (65%). These putative deletion strains segregated as 50% wild-type heterozygotes, 25% *unc-119* worms and 25% dead *dyn-1* embryos. We selected five of these 11 strains for verification. PCR analysis confirmed that four of five strains contained the full targeted deletion (**Fig. 2b**), and lethality could be rescued by a *dyn-1(+)* transgene.

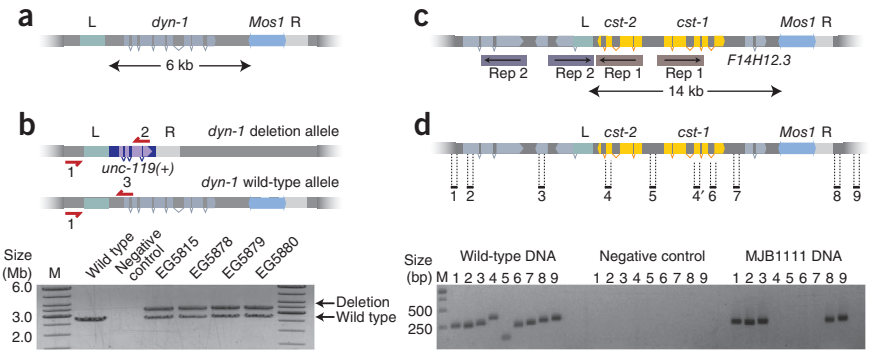


Figure 2 | Deleting lethal genes and multiple similar genes. (**a**) Schematic for the deletion of the essential gene dynamin *dyn-1*. (**b**) Schematic showing the balanced heterozygous deletion mutant. The *dyn-1* deletion allele is balanced with the insertion of *unc-119(+)*. The gel shows PCR verification of deletions in heterozygous *dyn-1* strains (EG5815, EG5878, EG5879 and EG5880) with the three oligos (1, 2 and 3) shown in the schematic. M, 1 kb DNA marker. (**c**) Schematic showing the *cst-1* and *cst-2* genomic region. *cst-1* and *cst-2* are 100% identical in the inverted repeats ("Rep 1"). An adjacent region contains a second inverted repeat ("Rep 2"), which overlaps the left homology region ("L"). The targeting construct removes 14 kb of genomic DNA, including *cst-1*, *cst-2* and *F14H12.3*. (**d**) PCR verification of the strain MJB1111 with deletions of *cst-1*, *cst-2* and *F14H12.3*. Oligos were designed to amplify 200–300 bp fragments inside and outside the deletion. M, 1 kb DNA marker.

Occasionally, the genome contains tandem gene duplications that provide redundant function or operons encoding genes with related functions. The loss-of-function phenotype of such loci requires the deletion of both genes, which can be accomplished using MosDEL. To test our system for this application, we targeted protein kinase genes *cst-1* and *cst-2* (**Fig. 2c**). These genes are adjacent to each other as two identical inverted repeats in a complex genomic region. We designed a deletion template for these two genes and isolated two strains containing *unc-119(+)* insertions; of these, one strain contained the correct deletion (**Fig. 2d**).

These results show that MosDEL can be used to target a gene if there is a *Mos1* insertion within 25 kb of the gene. Our method has several advantages relative to current knockout techniques. First, the technique is relatively fast and efficient. Second, the endpoints of deletions can be specified to completely eliminate the gene so that no partial gene products are generated. Third, lethal mutations are balanced by the positive insertion marker. In cases in which alleles do not have an obvious phenotype, a fluorescent protein marker can be inserted to follow the mutation in crosses. Finally, several adjacent genes can be deleted. This is particularly useful when genes with similar function are grouped.

The technique relies on the prevalence of *Mos1* insertions in the genome. We analyzed the distribution of the 14,305 *Mos1* elements relative to all 20,160 genes in *C. elegans*. Of these, 20,043 genes (99.4%) fell within 25 kb of at least one *Mos1* element (median distance to nearest *Mos1* element was 3.1 kb), so essentially all *C. elegans* genes can be targeted by MosDEL (**Supplementary Fig. 2** and **Supplementary Table 2**). Thus, 15,072 of the 15,157 genes presently lacking a deletion allele can be deleted by MosDEL. Moreover, 8,803 genes have no other genes between them and the *Mos1* element so these genes can be removed as a single-gene deletion.

Plasmid-mediated repair of double-strand breaks is not limited to *C. elegans*. For example, plasmid-driven repair has been described in *Drosophila*[13], and MosDEL could be adapted to make deletions in fruit flies. This technique should be a useful tool for the *C. elegans* research community and possibly for other genetic model organisms.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturemethods/.

Note: Supplementary information is available on the Nature Methods website.

ACKNOWLEDGMENTS

We thank L. Segalat (Centre National de la Recherche Scientifique, Lyon, France) and members of the NemaGENETAG consortium for *Mos1* strains. C.F.-J. is funded by a fellowship from the Lundbeck Foundation and G.H. by a Jane Coffin Childs fellowship funded by Howard Hughes Medical Institute. M.B. is funded by a McKnight Grant. D.G.M. is funded by Genome Canada and Genome British Columbia.

AUTHOR CONTRIBUTIONS

C.F.-J., M.W.D., G.H. and E.M.J. conceived and designed experiments; C.F.-J., G.H., J.T., P.N., R.L. and M.P.-D. performed experiments; T.W.H. performed the bioinformatic analysis of *Mos1* distribution; M.B., D.G.M. and E.M.J. provided supervision and funding; C.F.-J. and E.M.J. wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details accompany the full-text HTML version of the paper at http://www.nature.com/naturemethods/.

Published online at http://www.nature.com/naturemethods/.
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.

1. Moerman, D.G. & Barstead, R.J. *Brief. Funct. Genomics Proteomics* **7**, 195–204 (2008).
2. Pontier, D.B., Kruisselbrink, E., Guryev, V. & Tijsterman, M. *Nat. Methods* **6**, 655–657 (2009).
3. Gengyo-Ando, K. & Mitani, S. *Biochem. Biophys. Res. Commun.* **269**, 64–69 (2000).
4. Ryder, E. *et al. Genetics* **177**, 615–629 (2007).
5. Berezikov, E., Bargmann, C.I. & Plasterk, R.H.A. *Nucleic Acids Res.* **32**, e40 (2004).
6. Zwaal, R.R., Broeks, A., van Meurs, J., Groenen, J.T. & Plasterk, R.H. *Proc. Natl. Acad. Sci. USA* **90**, 7431–7435 (1993).
7. Robert, V. & Bessereau, J. *EMBO J.* **26**, 170–183 (2007).
8. Bazopoulou, D. & Tavernarakis, N. *Genetica* **137**, 39–46 (2009).
9. Plasterk, R.H. & Groenen, J.T. *EMBO J.* **11**, 287–290 (1992).
10. Frøkjaer-Jensen, C. *et al. Nat. Genet.* **40**, 1375–1383 (2008).
11. Gloor, G.B., Nassif, N.A., Johnson-Schlitz, D.M., Preston, C.R. & Engels, W.R. *Science* **253**, 1110–1117 (1991).
12. Maydan, J.S. *et al. Genome Res.* **17**, 337–347 (2007).
13. Keeler, K.J., Dray, T., Penney, J.E. & Gloor, G.B. *Mol. Cell. Biol.* **16**, 522–528 (1996).

ONLINE METHODS

Nematode strains. *Mos1* alleles were selected by visual screening in Wormbase (http://wormbase.org/db/gene/variation?name=cxti10882) for appropriately located transposon insertions and provided by members of the NemaGENETAG consortium. *Mos1* insertions were made homozygous and analyzed in crosses by PCR. Strains were maintained on nematode growth medium (NGM) plates seeded with OP50 or HB101 bacteria. Strains with *Mos1* elements and all deletions generated are listed in **Supplementary Table 3**.

Deletion protocol. *Mos1* alleles were identified in Wormbase and requested from the NemaGENETAG consortium (http://elegans.imbb.forth.gr/nemagenetag/). The presence of the *Mos1* insertion was verified with gene-specific primers, annealing inside and outside of the *Mos1* element. The *Mos1* element was crossed into the *unc-119(ed3)* background to make an injection strain and verified by PCR analysis.

Targeting constructs were designed to contain (i) a right homology region (adjacent to the *Mos1* element), (ii) a left homology region (distant from the *Mos1* element) and (iii) a *Caenorhabditis briggsae unc-119(+)* rescue region. The right homology regions comprised approximately 2 kb of homologous DNA immediately adjacent to the *Mos1* insertion site. The left homology regions comprised 2–3 kb of homologous DNA, which specifies the endpoint of the targeted deletion. These regions were selected to avoid repetitive DNA sequences, in particular short inverted repeats, which are likely to anneal and reduce the frequency of correct deletions. DNA between the right and left homology region is deleted; DNA contained in the left region is retained in the deletion strain. A *C. briggsae unc-119(+)* rescue fragment was chosen because it is smaller than the *C. elegans unc-119* gene.

An injection mix was made that contained the targeting plasmid (50 ng μl^{-1}), *Mos1* transposase helper plasmid pJL43.1 (ref. 14) (50 ng μl^{-1}), plasmids encoding co-injected mCherry markers pGH8 (ref. 10) (P*rab-3::mCherry*, 10 ng μl^{-1}), pCFJ90 (ref. 10) (P*myo-2::mCherry*, 2.5 ng μl^{-1}) and pCFJ104 (ref. 10) (P*myo-3::mCherry*, 5 ng μl^{-1}). The injection strain was maintained at 15 °C on HB101 bacteria. Young adult worms were mounted on an agarose pad under halocarbon oil and injected following standard techniques. Injected worms were left to recover at 15 °C for several hours and then transferred three at a time to HB101- or OP50-seeded NGM plates and placed at 25 °C. In our hands, approximately 70% of all injected worms resulted in transgenic progeny. We found that growing worms on HB101 bacteria at 15 °C considerably improved the health of *unc-119* worms.

After approximately 7 days each plate was screened for deletion mutants. Screening was greatly facilitated by allowing the plate to starve out completely because *unc-119* worms cannot form dauers and are therefore selected against. Strains with an *unc-119(+)* insertion were identified on a fluorescence dissection microscope as first stage larval (L1) worms that move like wild-type worms but have none of the fluorescent co-injection markers. A single rescued, nonfluorescent worm was picked to a new plate and allowed to propagate for one generation. In the case of obvious phenotypes (for example, Dpy-13) a single mutant worm was picked from the progeny; in cases in which the phenotype was wild type, ten worms were picked to individual plates and tested for homozygosity.

It took approximately 2 weeks from injecting the strain to recovering a homozygous deletion worm.

Analysis of *Mos1* distribution. We calculated the distance of every protein-coding gene in the WS205 referential release of WormBase (http://ws205.wormbase.org/)[15] to all current *Mos1* insertion alleles using a state machine algorithm written in Perl. The closest *Mos1* element was defined as the distance from the insertion site to the ATG start codon. The number of *Mos1* elements in the vicinity of each gene was determined by extracting a sequence segment upstream and downstream of the ATG and summing the number of elements contained within that span. The number of intervening genes between a given gene and its nearest *Mos1* element was determined by extracting the segment ranging from the ATG to the insertion site and tallying the number of genes present, including genes that partially reside within the interval. The analysis was repeated against all genes lacking a deletion allele from either the *C. elegans* Gene Knockout Consortium in the US and Canada (*ok* alleles) and the National BioResource Project in Japan (*tm* alleles).

Comparative genome hybridization. Genomic DNA from worms was isolated with the Gentra Puregene Tissue kit (Qiagen) following the manufacturer's supplementary protocol, "Purification of archive-quality DNA from nematodes or nematode suspensions using the Gentra Puregene Tissue Kit". DNA labeling, sample hybridization, image acquisition and determination of fluorescence were all performed as previously described[12]. We used a 3× high-density (HD) chip divided into three 720,000 whole-genome sections for all experiments. The chip design was based on our original 385,000 whole genome chip[12]. All microarrays were manufactured by Roche-NimbleGen with oligonucleotides synthesized at random positions on the arrays. Chip design name is 90420_Cele_RZ_CGH_HX3. For all experiments, normalization of intensity ratios were performed with a local scatterplot smoothing (LOESS) regression as previously described[12]. Three strains, EG5810 (25 kb deletion), EG5620 (15 kb) and EG5621 (15 kb), were tested against wild-type DNA. All strains had the targeted deletions. All samples also had two identical untargeted deletions: an approximately 8 kb deletion of *pgp-6* and *pgp-7* on chromosome X and an approximately 4 kb deletion of the telomeric region *cTelX3.1* at the left end of chromosome V. We verified by CGH that these deletions were present in the parent strain (EG5003), and the deletions therefore do not represent second site mutations caused by the MosDEL technique.

Molecular biology. Targeting vectors typically consist of three distinct fragments: a right homology region, a positive selection marker (*cb-unc-119(+)*) and a left homology region. In some cases the positive selection marker was flanked by the fluorescent marker *Punc-122::GFP*, which is dimly expressed in the coelomocytes. See **Figure 1a** for a schematic overview of the components of the targeting vector. All targeting vectors were made using the MultiSite Gateway Three-Fragment Vector Construction kit (Invitrogen).

To verify deletions by PCR, we designed oligos that would amplify short genomic DNA fragments inside and outside the targeted regions. These reactions were reproducible and robust; PCR amplification was successful on crude genomic lysates

doi:10.1038/nmeth.1454

from five to ten worms or from high quality DNA samples prepared with a Gentra Puregene kit (Qiagen).

***dyn-1* heterozygous verification primers.** Complete *dyn-1* deletions were verified at the 5′ end by PCR amplification with the three primers: oGH154, oGH133 and oCF125. Oligos oGH154 and oGH133 gave a 2.7 kb PCR product when the wild-type copy of *dyn-1* is present. Oligos oGH154 and oCF125 gave a 3.4 kb PCR product when the *dyn-1* locus is deleted. The deletion was verified at the 3′ end by PCR with oligos oGH155, oGH156 and oCF400. Oligos oGH155 and oGH156 produce a 2.5 kb PCR product when the wild-type *dyn-1* locus and the *ttTi14024 Mos1* element are present. Oligos oGH155 + oCF400 produce a 3.0 kb PCR product when the *dyn-1* gene is deleted.

All molecular biology analysis and design was done with the program ApE, a plasmid editor, that is freely available at http://www.biology.utah.edu/jorgensen/wayned/ape/.

All oligos and plasmids are listed in **Supplementary Table 3**.

14. Bessereau, J.L. *et al. Nature* **413**, 70–74 (2001).
15. Harris, T.W. *et al. Nucleic Acids Res.* **38**, 463–467 (2010).

nature | **methods**

Targeted gene deletions in *C. elegans* using transposon excision

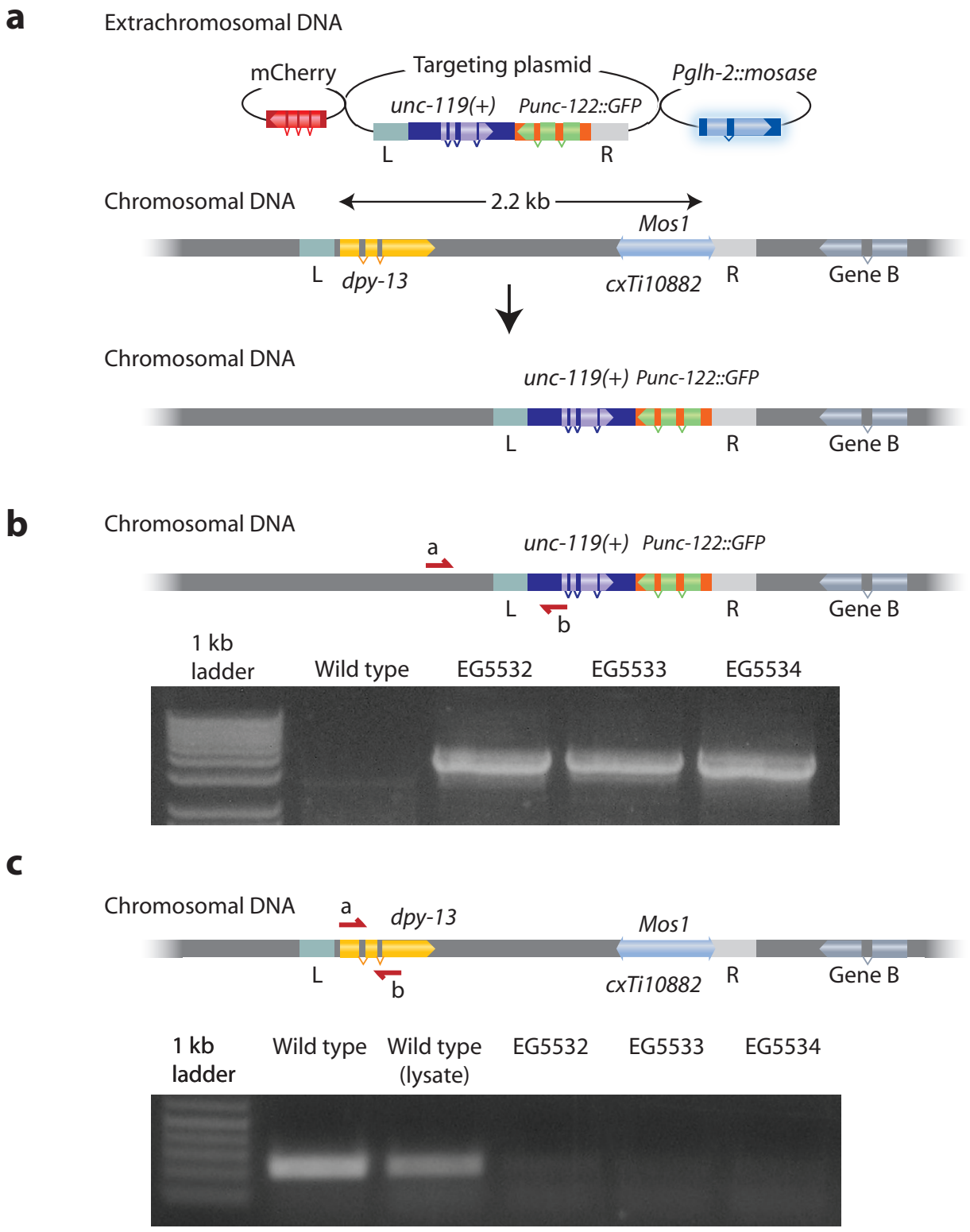
Christian Frøkjær-Jensen, M Wayne Davis, Gunther Hollopeter, Jon Taylor, Todd W Harris, Paola Nix, Rachel Lofgren, Michael Prestgard-Duke, Michael Bastiani, Donald G Moerman & Erik M Jorgensen

Supplementary figures and text:

Supplementary Figure 1	PCR verification of *dpy-13* deletion.
Supplementary Figure 2	*Mos1* distribution in the genome.
Supplementary Table 1	Deletion frequency.
***Supplementary Table 2**	Mos1 distribution.
Supplementary Table 3	Strains, plasmids and oligos.

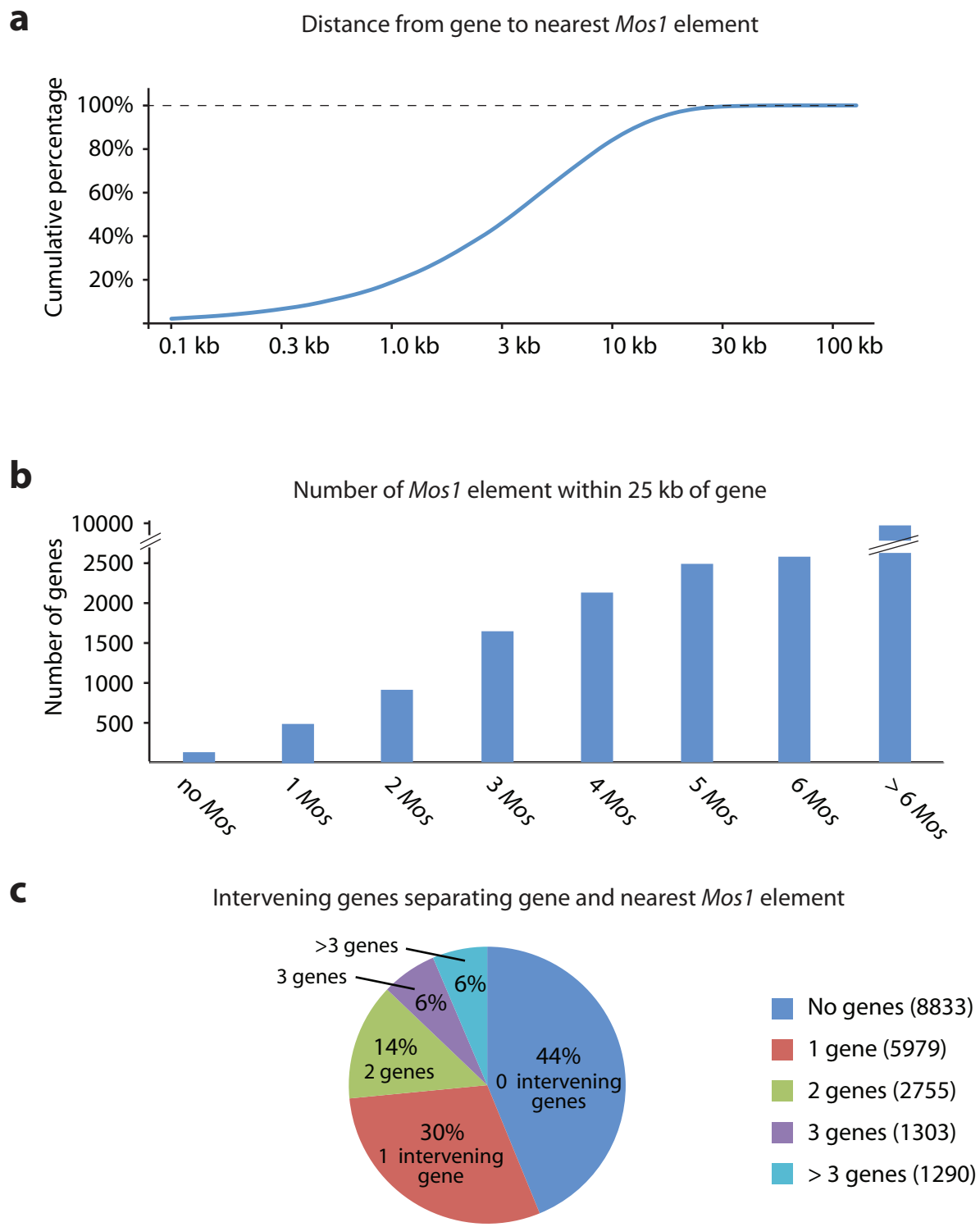
Note: Supplementary Table 2 is available on the Nature Methods website.

Supplementary Figure 1. PCR verification of *dpy-13* deletion



Supplementary Figure 1: PCR verification of *dpy-13* deletion.
(**a**) Schematic of targeting construct to delete a 2.2 kb genomic fragment adjacent to *cxTi10882* including *dpy-13*. The *Mos1* allele was crossed into the *unc-119(ed3)* background (not shown). Three independent deletion strains were generated: EG5532, EG5533 and EG5534. (**b**) PCR verification of deletions. Oligos annealing outside the displaced targeting homology and inside the *cb-unc-119(+)* were used to amplify across the junction. PCR products were subsequently sequenced to verify correct deletion junctions. (**c**) PCR verification of deletions. Oligos were used to amplify a 245 bp fragment from *dpy-13*. The fragment could be detected from high quality wild type genomic DNA ("wild-type") and single worm lysates from wild type ("wild type lysate") but not from any of the deletion mutants.

Supplementary Figure 2. *Mos1* distribution in the genome



Supplementary Figure 2: *Mos1* distribution in the genome.
We analyzed the distribution of *Mos1* elements in the *C.elegans* genome (WS205) with a total of 14,305 *Mos1* elements and 20,160 coding genes. See Supp. Table 2 for a comprehensive list of the data. (**a**) Distance from start codon to closest *Mos1* element. Shown is a cumulative plot of the percentage of coding genes within a given distance from the nearest Mos1 element. 99.4 % of genes are within 25 kb. (**b**) Number of *Mos1* elements within 25 kb. Shown is a histogram of the number of *Mos1* elements within 25 kb of the start codon. The *Mos1* strain collection was frozen as pools of animals and therefore not all inserts can be recovered. The presence of more than one *Mos1* element close to a given gene should minimize the chances of not recovering a *Mos1* allele within 25 kb. (**c**) Intervening genes. Shown is a pie diagram of the number of genes between the closest *Mos1* element and the start codon. Shown in parenthesis to the right are the number of genes for each category. For example, 8833 genes can be deleted without perturbing coding sequence from other genes.

Nature Methods: doi: 10.1038/nmeth.1454

Supplementary Table 1. Deletion frequency

Deletion size	Insert	Displaced ('left') homology size	Worms injected	Stable unc-119(+) inserts	Insertions per injected animal	Complete deletions	Successful deletions per *unc-119(+)* insertion	Deletions per injected animal
***cxTi10882* locus (*dpy-13*)**								
2.2 kb	*Cb-unc-119(+)* and *Punc-122::GFP*	1.5 kb	83	3	3.6 % (3/83)	3	100% (3/3)	**3.6 %** (3/83)
5.5 kb	*Cb-unc-119(+)* and *Punc-122::GFP*	2.8 kb	66	2	3.0 % (2/66)	2	100 % (2/2)	**3.0 %** (2/66)
10 kb	*Cb-unc-119(+)* and *Punc-122::GFP*	2.5 kb	108	4	3.7 % (4/108)	3	75 % (3/4)	**2.8 %** (3/108)
15 kb	*Cb-unc-119(+)*	3.1 kb	45	2	4.4 % (2/45)	1	50 % (1/2)	**2.2 %** (1/45)
25 kb	*Cb-unc-119(+)*	3.2 kb	81	7	8.6 % (7/81)	4	57 % (4/7)	**4.9 %** (4/81)
35 kb	*Cb-unc-119(+)*	3.0 kb	no data	8	no data	0	0 % (0/8)	**0%**
50 kb	*Cb-unc-119(+)*	3.3 kb	162	5	3.1 % (5/162)	0	0 % (0/5)	**0%**
***ttTi14024* locus (*dyn-1*)**								
6.0	*Cb-unc-119(+)*	2.4 kb	100	17	17 % (17/100)	4 of 5 tested	80 % (4/5)	**8-9 %**[$] (8-9/100)

Supplementary Table 1. **Deletion frequency**. Please note that the frequency of deletions from the different injections are not directly comparable. In some cases (2.2 kb, 5.5 kb and 10 kb) a larger fragment was inserted. In other cases, two different experimenters performed injections. In our hands, approximately 70 % of all injected worms give rescued transgenic progeny. All deletion strains were isolated from plates that had a mixed population of fluorescent animals rescued by extra-chromosomal arrays and non-fluorescent animals rescued by stable insertion of *unc-119*(+). [$] 11 of 17 strains (65 %) could not be homozygosed for the *unc-119(+)* marker as expected from a lethal mutation balanced by the insertion of *unc-119(+)*. 4 of 5 tested strains (80 %) had a full *dyn-1* deletion. We therefore estimate that 8 or 9 of the 11 strains (80 % * 11 strains = 8.8 strains) were full deletions.

Supplementary Table 3 - Strains

	Strain	Genotype
Mos1 strains		
	EG5003	*unc-119(ed3)* III; *cxTi10882* IV
	EG5817	*unc-119(ed3)* III ; *ttTi14024* X
	EG5618	*unc-119(ed3)* III ; *ttTi44501* X
Deletion strains		
		cxTi10882 mos1 allele
2.2 kb	EG5532	*unc-119(ed3)* III ; *dpy-13(ox448::cb-unc-119(+) Punc-122::GFP)* IV
	EG5533	*unc-119(ed3)* III ; *dpy-13(ox449::cb-unc-119(+) Punc-122::GFP)* IV
	EG5534	*unc-119(ed3)* III ; *dpy-13(ox450::cb-unc-119(+) Punc-122::GFP)* IV
5.0 kb	EG5637	*unc-119(ed3)* III ; *dpy-13(ox434::cb-unc-119(+) Punc-122::GFP)* IV
	EG5638	*unc-119(ed3)* III ; *dpy-13(ox435[cb-unc-119(+) Punc-122::GFP])* IV
10 kb	EG5620	*unc-119(ed3)* III ; *oxDf19* IV. EG5621: *unc-119(ed3)* III ; *oxDf20* IV
	EG5862	*unc-119(ed3)* III ; *oxIs562[Punc-122::GFP cb-unc-119(+)]* IV
	EG5863	*unc-119(ed3)* III ; *oxDf21* IV
15 kb	EG5864	*unc-119(ed3)* III ; *dpy-13(ox451::cb-unc-119(+))*
	EG5865	*unc-119(ed3)* III ; *oxDf22* IV
25 kb	EG5810	*unc-119(ed3)* III ; *oxDf13* IV.
	EG5866	unc-119(ed3) III ; oxDf23 IV
	EG5867	*unc-119(ed3)* III ; *oxIs563[cb-unc-119(+)]* IV
	EG5858	*unc-119(ed3)* III ; *oxIs564[cb-unc-119(+)]* IV
	EG5811	*unc-119(ed3)* III ; *oxDf14* IV
	EG5812	*unc-119(ed3)* III ; *oxDf15* IV
	EG5869	*unc-119(ed3)* III ; *dpy-13(ox452::cb-unc-119(+))* IV
		ttTi14024 mos1 allele
dyn-1 deletions	EG5815	*unc-119(ed3)* III ; *dyn-1(ox447::cb-unc-119(+))* X
	EG5878	*unc-119(ed3)* III ; *dyn-1(ox456::cb-unc-119(+))* X
	EG5879	*unc-119(ed3)* III ; *dyn-1(ox457::cb-unc-119(+))* X
	EG5880	*unc-119(ed3)* III ; *dyn-1(ox458::cb-unc-119(+))* X
	EG5881	*unc-119(ed3)* III ; *dyn-1(ox459::cb-unc-119(+))* X
dyn-1 rescue strain	EG6028	*unc-119(ed3)* III ; *dyn-1(ox447::cb-unc-119(+)*) X ; *oxEx1473[*pMPD13(*dyn-1(+) unc-119(+)) Punc-122::GFP litmus38i]*
		ttTi44501 mos1 allele
cst-1/cst-2 deletions	MJB1110	*unc-119(ed3)* III ; *basIs10[cb-unc-119(+) Punc-122::GFP]* X
	MJB1111	*unc-119(ed3)* III ; *basDf1* X

Supplementary Table 3 - Oligos

Uppercase sequences are gene specific; lowercase sequences are Gateway linkers.

Oligo	Sequence (5'->3')
oBN1	ggggacagctttcttgtacaaagtggATAGAATCAAGCATGCTCCG
oBN2	ggggacaactttgtataataaagttgGGCCCATTAGGTCAGACAAA
oCF125	GGGTGCCAAATAACCAGCTA
oCF168	CAGGAGAGCAAGGACCAAAG
oCF250	GTTGTCCTGAAAGATGTGGTGA
oCF400	CCGAATTCACGAAACAACTG
oCF570	ggggacaactttgtatagaaaagttgGGCGAGGAGGTACCAAAGAT
oCF571	ggggactgctttttttgtacaaacttgCTTTCGTCCGTTGACAAGGT
oCF607	ggggacaactttgtatagaaaagttgGTGAATTGTAATACGACTCACTATAGGG
oCF608	ggggactgctttttttgtacaaacttgCTGCCAAAGTTGAGCGTTTATTCTGA
oCF673	ggggacaactttgtatagaaaagttgCACGGATTGACGGACAATAA
oCF674	ggggactgctttttttgtacaaacttgGCACATCCCTTTGCTGTAAAA
oCF677	ggggacagctttcttgtacaaagtggTGCACCAGAGAAACACAAAGA
oCF678	ggggacaactttgtataataaagttgCCGTAAACAAATAAAATTGTGTGG
oCF681	ggggacaagtttgtacaaaaaagcaggctTGAGCCAATTTATCCAAGTCC
oCF682	ggggaccactttgtacaagaaagctgggtCAGTTGAAATTGAAAATGAGTTAAAG
oCF692	ATAGACGGGGCCAAATTTTC
oCF693	TGTCAACCATTGGTGCTTGT
oCF711	GGAAGACCTTAAGCCGGTTC
oCF712	AGCGAGAGCCTGAAGAAGTG
oCF713	GAACAGGTAGCCCATTGGAA
oCF714	ATTGGCACGTTTCATTAGCC
oCF715	GTTTTCTTGGTGCTGGCCTA
oCF716	AAAACGGGGGAAAACTGACT
oCF717	TTTGGAGCCCATAAAATTGC
oCF718	AATGCAAAACTGCTGTCACG
oCF719	ATGTTGGCGGCTATTCAAAC
oCF720	GTGCAGAAATTTGGGGCTTA
oCF721	GTAGGCGGAGCATACATCGT
oCF722	TAGAACGCCCTTAACCATGC
oCF723	GGCGATTTGTCGGTACTTTG
oCF724	CTTTCACGGGAATGTCTGGT
oCF725	GCGGATTGGAAGGTCATAAG
oCF726	GGAAACCTTGCTCAAATCCA
oCF727	CAGGCAAGGGAGTGCTTTTA
oCF728	GCGTGGAATCTGTCCAAAAT
oCF729	TCCCCATTTCACCAGAGAAC
oCF730	GCCTACTGGGTCAACGACAC
oCF731	TGGGCAGTAGTATGGCAAAA
oCF732	TCAAATTTTCCCGTAGACGA
oCF733	CGCCATAATTGCTTGGCTAC
oCF734	GCCACTTCAAGGGAATTCAG
oCF735	AAGGATGCTGCTTACCTGCT
oCF736	AAGCGGATCGCCTTTAATTT
oCF737	ACAGATTCCACGCTCTCCTG
oCF738	ACGGAACCCCAGGTAATCAT
oCF739	GATCTCCTCGCAAGGCTTCT
oCF740	CTGGAAACGGACCAAACTTC

oCF741	CAGAGATTCCGTGTGCAAGA
oCF742	GGTTTCGGGGGAAAAGAATA
oCF743	TTCACACATAGGCGCTGAAG
oCF744	GTCCCTCCCTCCCTAAACTG
oCF745	TTATTGCCGCACACACCTAA
oCF746	ATTACCGAAACATCGCCAAG
oCF747	GGTGGGAGGCAATTGTTAAA
oCF748	ATTCCAGATTGCCGAGTTTG
oCF749	AGTGGTCGATGGAGATTTCG
oCF750	AATTTCCGGCTTTTTGAGGT
oCF751	TGGGGAGGTTGGAGTAGTTG
oCF752	TTCTCACAGCAGTCCAGGTG
oCF753	TTGCTTGGTGCTTGGTGTTA
oCF754	GCGGCCTTATTTTTGAAATG
oCF755	TGCCTTCCCTGCTTTTCTAA
oCF756	ATATGAACGTTGGCGAGCTT
oCF786	ggggacaactttgtatagaaaagttgGGCTTCTCTGTCGTTCTTCG
oCF787	ggggactgctttttttgtacaaacttgCCATTAAACCTGAAAAGACTCTGA
oRL1	ggggacaactttgtatagaaaagttgTCCACTACCCAAATGTGCCATC
oRL2	ggggactgctttttttgtacaaacttgGCTGGCGATGCTGGTTATAGTT
oRL4	ggggacaactttgtatagaaaagttgGATAAATGTTCCAGCGGCAAGGG
oRL5	ggggactgctttttttgtacaaacttgAACAGTTCAGTCCCTGCGTG
oRL12	ggggacaactttgtatagaaaagttgGGGCTGAAACCATTTCAAGA
oRL13	ggggactgctttttttgtacaaacttgGGGATTGTCCGTGCTACACT
oRL14	ggggacaactttgtatagaaaagttgGAGGTGGTGGCTTGGAATTA
oRL15	ggggactgctttttttgtacaaacttgAAGTGCCGAAAACCACTTTG
oRL16	ggggacaactttgtatagaaaagttgCTGCAATGTTTTGCGGACTA
oRL17	ggggactgctttttttgtacaaacttgGGAAAAATGCCCAGAGTTCA
oGH128	ggggacagctttcttgtacaaagtggTTTGTTTGTGAGTTTCAGGAAACTTG
oGH129	ggggacaactttgtataataaagttgGGGGATGAAGTTTAAAGGTTGAGTAGC
oGH130	ggggacaactttgtatagaaaagttgCAGTCCAGCTATCCGTCCGTC
oGH131	ggggactgctttttttgtacaaacttgGTCGATTAATGATATGAGGAAGAGCATTG
oGH132	ggggacaactttgtatagaaaagttgCTGCTCCTGGTATCCTTAACCAG
oGH133	ggggactgctttttttgtacaaacttgATTTCGCAAAGTTATGTTTTTTTTCTTC
oGH134	ggggacaagtttgtacaaaaaagcaggctAAAATCACAATTTTTTTCCTTTCCAG
oGH135	ggggaccactttgtacaagaaagctgggtGGTTTTTCTGTCTTACAAGTCATTAAAGAAG
oGH136	ggggacagctttcttgtacaaagtggTTTGTGTTGTCGTCCTTGTGATC
oGH137	ggggacaactttgtataataaagttgCCGTTTTAACTTTATTCCACTTTCACG
oGH154	GTGTTCTAGAACTCAATGGCACAG
oGH155	CTATGGTTTTGGAACACTGGCTAG
oGH156	AAACGACATTTCATACTTGTACACCTG

Supplementary Table 3 - PCR verification

	Annealing	Forward Oligo	Reverse Oligo	PCR product size
cxTi10882 locus (Figure 1c)	39 kb upstream	oCF713	oCF714	451 bp
	29 kb upstream	oCF715	oCF716	382 bp
	19 kb upstream	oCF717	oCF718	409 bp
	14 kb upstream	oCF719	oCF720	385 bp
	9 kb upstream	oCF721	oCF722	500 bp
	4 kb upstream	oCF723	oCF724	483 bp
	1.3 kb upstream	oCF168	oCF250	245 bp
	0.5 kb downstream (control)	oCF725	oCF726	447 bp
	5.7 kb downstream (control)	oCF727	oCF728	312 bp
ttTi44501 (cst-1/cst-2) (Figure 2d)	27.5 kb upstream	oCF741	oCF742	263 bp
	24 kb upstream	oCF743	oCF744	253 bp
	18 kb upstream	oCF745	oCF746	268 bp
	13.0 and 5.5 kb upstream (inverted repeat)	oCF692	oCF693	327 bp
	9.4 kb upstream	oCF747	oCF748	105 bp
	4.2 kb upstream	oCF749	oCF750	232 bp
	2.2 kb upstream	oCF751	oCF752	245 bp
	1.7 kb downstream(control)	oCF753	oCF754	272 bp
	5 kb downstream	oCF755	oCF756	290 bp

Supplementary Table 3 - Plasmids

Entry vectors

Positive selection vectors

Name	Forward Oligo	Reverse Oligo	Template	DONR vector	Description
pCFJ66					[1-2] Entry vector with *C.briggsae-unc-119(+)* and *Punc-122::GFP*. Described in Frøkjær-Jensen *et al.*, (2008
pRL8	oCF681	oCF682	pCFJ66	pDONR221	[1-2] Entry vector with *C.briggsae-unc-119(+)*
cxTi10882 locus					
pBN03	oBN1	oBN2	N2 DNA	pDONRP2-P3	Right homology vectors
pCFJ215	oCF570	oCF571	N2 DNA	pDONRP4-P1R	2.2 kb deletion
pRL3	oRL4	oRL5	N2 DNA	pDONRP4-P1R	5.0 kb deletion
pRL4	oRL1	oRL2	N2 DNA	pDONRP4-P1R	10 kb deletion
pRL11	oRL12	oRL13	N2 DNA	pDONRP4-P1R	15 kb deletion
pRL19	oRL14	oRL15	N2 DNA	pDONRP4-P1R	25 kb deletion
pCFJ281	oCF786	oCF787	N2 DNA	pDONRP4-P1R	35 kb deletion
pRL12	oRL16	oRL17	N2 DNA	pDONRP4-P1R	50 kb deletion
ttTi44501 locus (cst-1/cst-2)					
pCFJ251	oCF677	oCF678	N2 DNA	pDONRP2-P3	Right homology region
pCFJ249	oCF673	oCF674	N2 DNA	pDONRP4-P1R	Left homology region
***ttTi14024* locus (*dyn-1*)**					
pMPD6	oGH128	oGH129	N2 DNA	pDONRP2-P3	Right homology region
pMPD8	oGH130	oGH131	N2 DNA	pDONRP4-P1R	Left homology region
[4-1] dyn-1(+)	oGH132	oGH133	N2 DNA	pDONRP4-P1R	Part rescue fragment dyn-1
[1-2] dyn-1(+)	oGH134	oGH135	N2 DNA	pDONR221	Part rescue fragment dyn-1
[2-3] dyn-1(+)	oGH136	oGH137	N2 DNA	pDONRP2-P3	Part rescue fragment dyn-1

Expression vectors

Name	[4-1]	[1-2]	[2-3]	pDEST	Description
***cxTi10882* targeting plasmids**					
pCFJ216	pCFJ215	pCFJ66	pBN03	pDESTR4-R3	2.2 kb deletion
pRL5	pRL3	pCFJ66	pBN03	pDESTR4-R3	5.0 kb deletion
pRL6	pRL4	pCFJ66	pBN03	pDESTR4-R3	10 kb deletion
pRL13	pRL11	pRL8	pBN03	pDESTR4-R3	15 kb deletion
pCFJ260	pRL19	pRL8	pBN03	pDESTR4-R3	25 kb deletion
pCFJ287	pCFJ281	pRL8	pBN03	pDESTR4-R3	35 kb deletion
pRL14	pRL12	pRL8	pBN03	pDESTR4-R3	50 kb deletion
***ttTi44501* (*cst-1/cst-2*) targeting plasmids**					
pCFJ255	pCFJ249	pCFJ66	pCFJ251	pDESTR4-R3	cst-1/cst-2 deletion
***ttTi14024* (*dyn-1*) targeting plasmids**					
pGH238	pMPD6	pRL8	pMPD8	pDESTR4-R3	dyn-1 deletion
pMPD13	[4-1] dyn-1(+)	[1-2] dyn-1(+)	[2-3] dyn-1(+)	pCFJ150	dyn-1 rescue plasmid