nature genetics

Single-copy insertion of transgenes in *Caenorhabditis elegans*

Christian Frøkjær-Jensen[1,2], M Wayne Davis[1], Christopher E Hopkins[1], Blake J Newman[1], Jason M Thummel[1], Søren-Peter Olesen[2], Morten Grunnet[2,3] & Erik M Jorgensen[1]

At present, transgenes in *Caenorhabditis elegans* are generated by injecting DNA into the germline. The DNA assembles into a semistable extrachromosomal array composed of many copies of injected DNA. These transgenes are typically overexpressed in somatic cells and silenced in the germline. We have developed a method that inserts a single copy of a transgene into a defined site. Mobilization of a Mos1 transposon generates a double-strand break in noncoding DNA. The break is repaired by copying DNA from an extrachromosomal template into the chromosomal site. Homozygous single-copy insertions can be obtained in less than 2 weeks by injecting approximately 20 worms. We have successfully inserted transgenes as long as 9 kb and verified that single copies are inserted at the targeted site. Single-copy transgenes are expressed at endogenous levels and can be expressed in the female and male germlines.

The method for introducing DNA into *C. elegans*[1] has not changed in the last 17 years and is elegant in its simplicity. DNA injected into the gonad of a hermaphrodite concatenates to form an extrachromosomal array and is eventually incorporated into the nucleus. Because chromosomes in *C. elegans* are holocentric in mitosis, any piece of DNA can serve as a centromere, so these extrachromosomal arrays are duplicated and segregated to daughter cells in mitosis. However, this method for generating transgenic lines suffers from several limitations. First, these minichromosomes do not behave like *bona fide* chromosomes; they are not perfectly stable in mitosis or meiosis. Thus, transgenic animals are mosaic: some cells carry the transgene, whereas others have lost the array. Second, such arrays contain hundreds of copies of the injected DNA, and the genes are overexpressed. This high copy number can cause dominant-negative or toxic effects[2]. Third, these repetitive arrays are silenced in some tissues, including muscles[3,4] and the germline[5]. The arrays can be silenced even after they are integrated into a chromosome by irradiation, presumably because of transcriptional silencing of arrays[6]. Finally, arrays change and show 'drift' of expression over many generations[7,8]; drift may arise from changes in the structure of the arrays or by heritable silencing. These limitations complicate studies relying on stable, tissue-specific expression of transgenes.

Stable changes can be generated at chromosomal sites in rare instances by homologous recombination after biolistic transformation[9] or, more effectively, by template-directed repair after excision of a transposon. For example, mobilizing a Tc1 transposon induces a double-strand break at a defined location in a chromosome; the break can be repaired by copying DNA from a transgenic template[10,11]. The disadvantage of these mutator strains is that there are hundreds of copies of the transposon in the genome; breaks will be induced at many sites, and the frequency of events at any particular site can be quite low. To generate single-copy transposon insertions, the *Drosophila* Mos1 element was introduced into *C. elegans*[12]. It was recently shown that specific DNA changes can be targeted to loci with Mos1 insertions[13]. This technique, called Mos1 excision-induced transgene-instructed gene conversion (MosTIC), has been used to insert tags or engineer deletions in particular genes. MosTIC relies on the presence of a Mos1 insertion at the genetic locus to be modified. The nematode gene-tagging tools and resources (NemaGENETAG) consortium has generated a large library of Mos1 inserts with known locations in the genome[14].

In this study, we adapted intergenic Mos1 elements for the routine insertion of transgenes using a variation of the MosTIC technique called Mos1-mediated single-copy insertion (MosSCI). This technique inserts transgenes as single copies at a defined chromosomal locus, which supports expression in a broad range of tissues at apparently endogenous levels. Stable expression is observed in tissues that frequently silence transgenes, including the male and female germlines. Insertions can be induced efficiently in transgenic strains or can be obtained directly from injected worms.

RESULTS

Insertion method

An ideal integration site would be genetically neutral, so we picked a Mos1 insertion that matched the following criteria. First, the insertion should not disrupt the function of neighboring genes. Second, nearby promoters and enhancers should not affect expression of the inserted

[1]Department of Biology and Howard Hughes Medical Institute, University of Utah, 257 South 1400 East, Salt Lake City, Utah 84112, USA. [2]The Danish National Research Foundation Centre for Cardiac Arrhythmia, University of Copenhagen, Blegdamsvej 3, 2200 Copenhagen N, Denmark. [3]NeuroSearch A/S, Pederstrupvej 93, 2750 Ballerup, Denmark. Correspondence should be addressed to E.M.J. (jorgensen@biology.utah.edu).

Received 14 February; accepted 14 August; published online 26 October 2008; doi:10.1038/ng.248

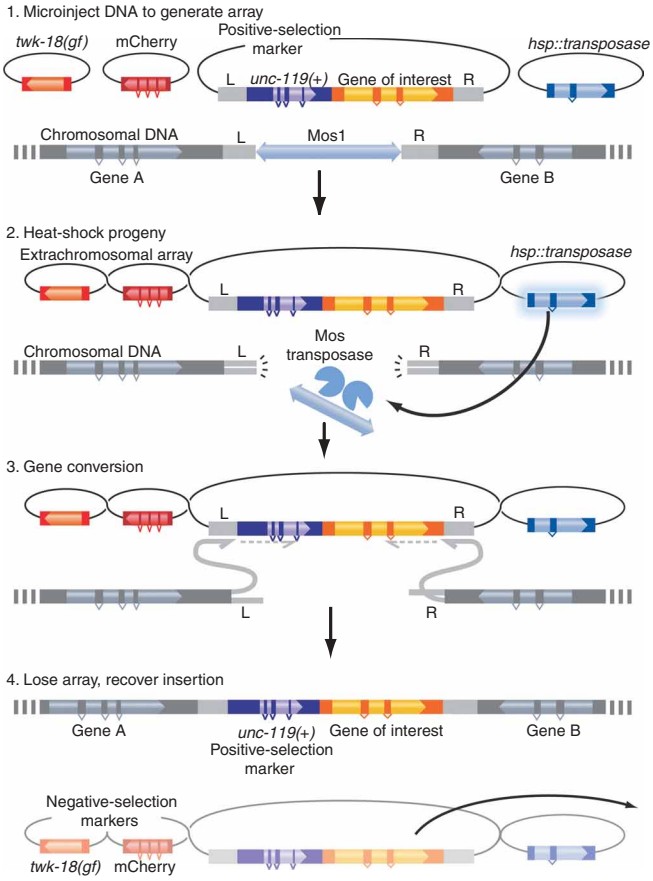

Figure 1 Schematic overview of MosSCI. A Mos1 transposon located at a noncoding locus was isolated by the NemaGENETAG consortium. The Mos1 element can be excised by transposase expression, resulting in a double-strand break in the chromosome. Presumably, the 3′ ends from the left (L) and right (R) flanks invade and anneal to homologous regions in the extrachromosomal array. The break can then be repaired by synthesis-dependent strand annealing. The positive-selection marker *unc-119*(+) and the gene of interest are inserted into the genome by gene conversion. The extrachromosomal array contains a source of transposase (*hsp::transposase*) and two negative-selection markers, *twk-18(gf)* and fluorescent mCherry markers. *twk-18(gf)* is a temperature-sensitive dominant mutation in a potassium channel, which paralyzes the worms at 25 °C but not at 15 °C. mCherry markers are expressed in the pharynx, body muscle and nervous system for visual identification of array-carrying worms. After loss of the array, single-copy transgenic worms are isolated. L and R homologous regions are 1.4 kb each.

transgene. For these reasons, genomic regions 3′ to coding regions were selected. We identified several Mos1 elements that were inserted in tail-to-tail gene regions and settled on the *ttTi5605* Mos1 allele, near the center of chromosome II, as a test case (map position +0.78).

Our goal was to generate a double-strand break in the chromosome through Mos1 excision and provide a homologous template for repair of the break. We generated an extrachromosomal array that contained ~1.4 kb of homologous chromosomal DNA from each flanking side of the Mos1 element. Between the left and right flanks, we inserted the gene to be integrated (**Fig. 1**).

To identify insertions in the chromosome, we also included a positive-selection marker. The *ttTi5605* Mos1 insertion was crossed into an *unc-119(ed3)* mutant background. *unc-119(ed3)* mutant worms are small and almost paralyzed, have small brood sizes and are incapable of forming dauer larvae when starved[15]. By placing the wild-type *Caenorhabditis briggsae unc-119* gene between the flanking DNA, the rescuing construct will also be incorporated into the site along with the gene of interest. Thus, after gene conversion is induced, the only worms capable of surviving starvation will be those carrying the extrachromosomal array or those that integrated the *unc-119(+)* gene.

The strategy requires the excision of the Mos1 element in the germline to generate a double-strand break. When making the extrachromosomal array, we coinjected the sequence encoding the Mos1 transposase under the control of the heat-shock promoter *hsp-16-48* (ref. 12). Heat shock activates synthesis of the transposase, which in turn excises the transposon. The double-strand break is then healed by gene conversion, at least in some cases, using the extrachromosomal array as the template.

The main challenge is to distinguish worms with targeted insertions from worms carrying the extrachromosomal array, as both will be rescued for the *unc-119* marker (**Fig. 1**). To differentiate integrations from arrays, we coinjected the DNA of interest with negative-selection markers: genes expressing red fluorescent proteins and TWK-18(gf), an activated K^+ channel that causes muscle paralysis at elevated temperatures (**Supplementary Fig. 1** online). These markers are incorporated into the transgene arrays but are not copied into the targeted integration. At 15 °C, worms carrying the array are active and can be propagated. At 25 °C, these worms are paralyzed; only worms that have lost the array are not paralyzed.

Insertion frequency and transgene copy number

To test the insertion strategy, we built a targeting construct that contained, as the gene of interest, a construct expressing green fluorescent protein (GFP) in coelomocytes (*Punc-122::GFP*) along with the positive-selection marker that rescues *unc-119* (**Fig. 2a** and **Supplementary Fig. 2** online). To direct gene conversion, these two genes were flanked by 1.4-kb genomic sequences on each side that were homologous to the flanking sequences of the *ttTi5605* Mos1 site. We coinjected the targeting vector, the transposase construct, the *twk-18* negative-selection marker and fluorescent mCherry marker into *unc-119(ed3) ttTi5605mos*. Injected worms were then cultured at 15 °C. We then selected five independent *unc-119*-rescued lines and propagated them to expand the population. Rescued worms showed all of the fluorescent markers and were strongly paralyzed at 25 °C. We heat-shocked a semisynchronous population of young-adult worms to induce Mos*1* excision and then screened the F_2 generation for transgene-instructed repair (for a detailed protocol, see **Supplementary Methods** online). In total, we heat-shocked 1,000 worms from five independent transgenic lines and recovered ten putative targeted insertions (**Table 1**). Consistent with a loss of the transgenic array, insertion lines were not paralyzed at 25 °C, and none expressed mCherry protein at detectable levels. As expected from a genomic insert, we could isolate putatively homozygous worms that never segregated progeny with the Unc-119 mutant phenotype. We confirmed dim expression of GFP in the coelomocytes in seven of ten insertion lines (**Fig. 2b**).

We used PCR to show that DNA from the transgenic array had inserted into the *ttTi5605* locus (**Fig. 2c**). We isolated genomic DNA from the ten lines and amplified sequences spanning the left junction. The *Cbr unc-119(+)* DNA was inserted adjacent to the left flank of the *ttTi5605* Mos1 site in all ten lines. The absence of product in the wild-type and original targeting strain (genotype: *unc-119(ed3); ttTi5605*) confirmed that our PCR reaction was specific.

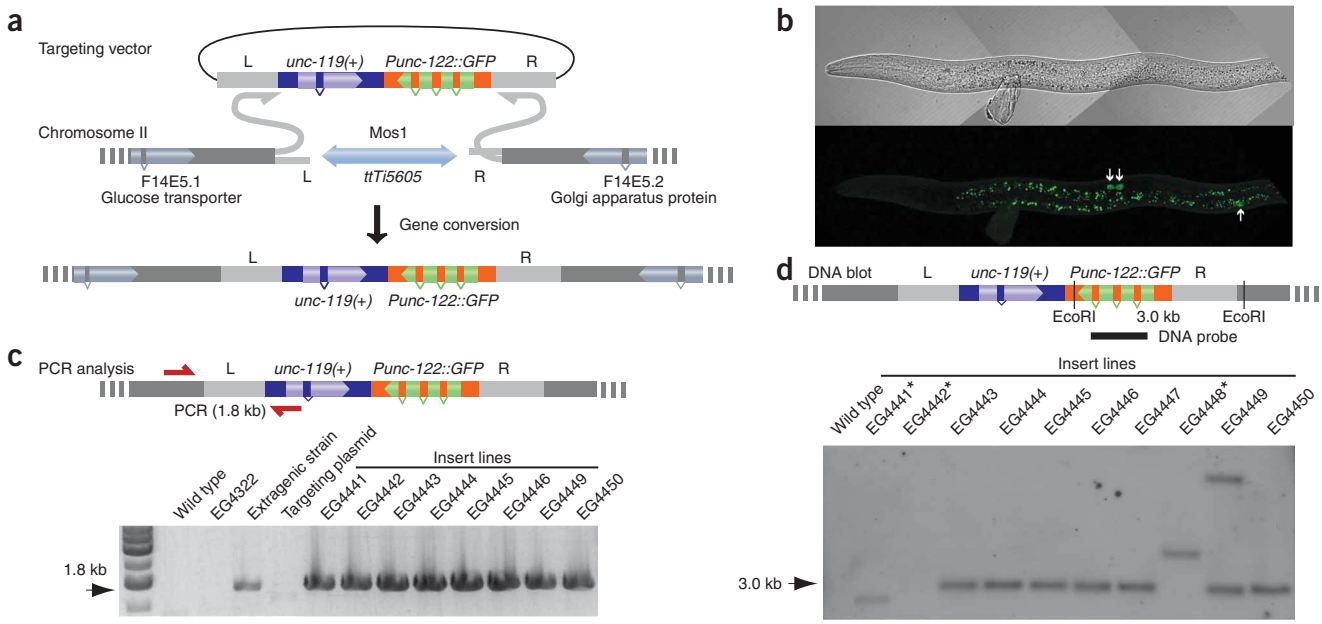

Figure 2 Single-copy insertions of transgenes. (**a**) Schematic of the targeting construct containing the *Punc-122::GFP* transgene and *unc-119*(+) rescue marker (4.3 kb total) flanked by DNA homologous to the *ttTi5605mos* insertion site. (**b**) Expression of the *Punc-122::GFP* transgene in MosSCI insertion strains was restricted to the coelomocytes (arrows). Remaining visible fluorescence is nonspecific gut granule fluorescence. Image is of an adult hermaphrodite; anterior is left. Top, differential interference contrast image; bottom, fluorescence image. (**c**) PCR verification of inserted *Punc-122::GFP* transgenes. The forward primer anneals to the genomic region outside of that contained in the targeting construct, and the reverse primer is in the *C. briggsae unc-119*(+) selectable marker. A PCR band of expected size (1.8 kb) from all MosSCI insertion strains (EG4441-EG4450) confirmed insertions at the targeted locus. We interpreted the presence of a dim band from the parent strain carrying the extrachromosomal array as evidence for either low levels of somatic gene conversion or PCR bridging[13]. (**d**) DNA analysis confirmed single-copy insertion in 60% of the strains. Genomic DNA was digested with EcoRI and hybridized with a GFP-specific probe. Six of ten strains showed a single, specific band of appropriate size, verifying single-copy transgene insertions. Three strains were nonfluorescent (indicated by asterisks), and the DNA blot showed that these transgenes contained short deletions and insertions. The molecular changes in EG4441 and EG4448 were further characterized (**Supplementary Figs. 3** and **4**). Strain EG4449 was visibly more fluorescent and contained two copies of GFP.

We next analyzed the presence of the *Punc-122::GFP* transgene by Southern blotting using a GFP-specific probe (**Fig. 2d**). Six of ten insertion strains showed the predicted band corresponding to a single, targeted insertion of the *Punc-122::GFP* transgene. One strain, EG4449, which we had noticed to be moderately more fluorescent by visual inspection, had what seemed to be a tandem insertion. As expected, the three nonfluorescent strains showed either no band or an aberrant band size. We further characterized these three nonfluorescent insertion events by 'PCR walking': one primer was fixed in the *Cbr-unc-119* rescue fragment, and other primers were staggered every 500 bp along the transgene (**Supplementary Fig. 3** online). Consistent with the lack of fluorescence and the Southern blot data, these strains deleted primers in the *Punc-122::GFP* transgene. For two of these strains, we were able to amplify across the deletion and sequence the product. EG4441 was a 1,800-bp deletion within the *Punc-122::GFP* transgene, and EG4448 was an 803-bp deletion accompanied by a 1,713-bp insertion from elsewhere on the extrachromosomal array (**Supplementary Figs. 3** and **4** online). Similar deletions were observed with MosTIC[13].

The *Punc-122::GFP* transgene was specifically expressed in the coelomocytes, suggesting that the *ttTi5605* site is permissive for tissue-specific expression. To further test tissue-specific expression, we generated three insertions using *Punc-47::mCherry* and two *Pdpy-30::H2B::mCherry* insertions (**Supplementary Fig. 5** online). These constructs were appropriately expressed in only GABA neurons or ubiquitously, respectively.

In some cases, it would be advantageous to have an insertion site located on a different chromosome. An *unc-119(+)* targeting vector was developed for a Mos1 insertion on chromosome IV (*cxTi10882*; map position IV, –0.05). The *Punc-122::GFP* construct was used as the transgene. As in the previous experiments, we injected this plasmid together with DNAs encoding the transposase construct and negative-selection markers to form an extrachromosomal array. We selected a single transgenic line and heat-shocked an adult population. We recovered eight transgene insertions from 800 heat-shocked worms, a frequency that was identical to that for the *ttTi5605* site. We confirmed dim coelomocyte fluorescence in five of six lines that we studied in detail.

Together, these results indicate that, on average, heat-shocking resulted in an insertion event at two distinct genomic loci in 1 in every 100 worms. Of these, 60% were functional single-copy insertions of the targeted DNA.

Larger transgenes

In the initial experiments, we chose to insert *Punc-122::GFP* because of its small size (3 kb) and restricted expression pattern. In most cases, a larger transgene is desirable. To test the frequency of integration for larger transgenes, we constructed a 6.8-kb *unc-18* gene fragment containing upstream and downstream regulatory elements as well as a C-terminal mCherry tag (**Fig. 3a**). We chose *unc-18* because the mutant phenotype is easy to score, the expression pattern is restricted to neurons and expression levels can be determined by western blot analysis. The gene was inserted into a targeting plasmid along with

Table 1 Insertion of the coelomocyte-specific *Punc-122::GFP* transgene at *ttTi5605*

Extragenic strains		MosSCI strains					
Parent strain	Insertion frequency	Strain	Phenotype	GFP expression	PCR *ttTI5605* site	Transgene insert	Mos1 element in strain
EG4380	2/200	EG4441	Wild-type	No	Yes	1.8-kb deletion	ND
		EG4442	Weak Unc	No	Yes	ND	No
EG4381	1/240	EG4443	Wild-type	Yes	Yes	Full	No
EG4382	2/240	EG4444	Wild-type	Yes	Yes	Full	No
		EG4445	Wild-type	Yes	Yes	Full	ND
EG4383	2/60	EG4446	Wild-type	Yes	Yes	Full	ND
		EG4447	Wild-type	Yes	Yes	Full	No
EG4385	3/260	EG4448	Wild-type	No	Yes	0.8-kb deletion, 0.7-kb insertion	No
		EG4449	Wild-type	Yes (brighter)	Yes	Tandem?	No
		EG4450	Wild-type	Yes	Yes	Full	No
Average	1/100		90% wild-type	70% expression	100% correct	60% correct	0% Mos1 element

Five independent transgenic strains carrying extrachromosomal arrays were generated, and the frequency of insertions was determined after heat-shock. Ten MosSCI inserts were recovered and listed adjacent to their parent strains in the left two columns. Strains were scored for rescue of locomotion and GFP fluorescence. PCR was used to verify insertion at the *ttTi5605* target site and to test for presence of Mos1 transposon elements after insertion (see **Supplementary Methods**). DNA blotting was used to verify the integrity of the *Punc-122::GFP* transgene insert. ND, not determined.

unc-119(+) to generate a final insertion length of 9 kb (**Fig. 3a**). An array was generated with the negative-selection markers. We selected a single transgenic line and induced transposase expression by heat-shock in a population of adults. From 500 heat-shocked worms, we recovered four targeted integrants (**Supplementary Table 1** online). From this limited dataset, insertion frequency does not seem to be adversely affected by increased transgene size. Three of the four strains had uniform, dim red fluorescence specific to the nervous system (**Fig. 3b** and **Supplementary Table 1**). To test whether the *unc-18::mCherry* transgenes were functional, we crossed the fluorescent lines to a loss-of-function allele, *unc-18(e81)*; all three insertions rescued the uncoordinated mutant phenotype.

An alternative method for generating stable transgenes is biolistic transformation using DNA-coated gold particles[16]. To compare single-copy insertions to those generated by biolistic transformation, we generated two *unc-18::mCherry* integrants (UZ566 and UZ567) by biolistic transformation of *unc-119(ed3)* worms. We confirmed integration by genetic mapping. UZ566 showed brighter mCherry fluorescence than the targeted insertion strains and was slow growing (**Fig. 3b**). UZ567 was very brightly fluorescent and dauer constitutive. These results agree with previous observations in which biolistic transformation generated variable transgene expression and occasionally disrupted endogenous genes[16]. We used Southern blot analysis on all of the *unc-18::mCherry* integrants to determine whether the differences in fluorescence intensity reflected differences in transgene copy number (**Fig. 3c**). We probed the blot with labeled DNA specific to mCherry and detected a single band of the appropriate size in the three successfully targeted insertion strains, indicating single-copy transgene insertions. In the biolistic strains, we detected multiple copies of the transgene.

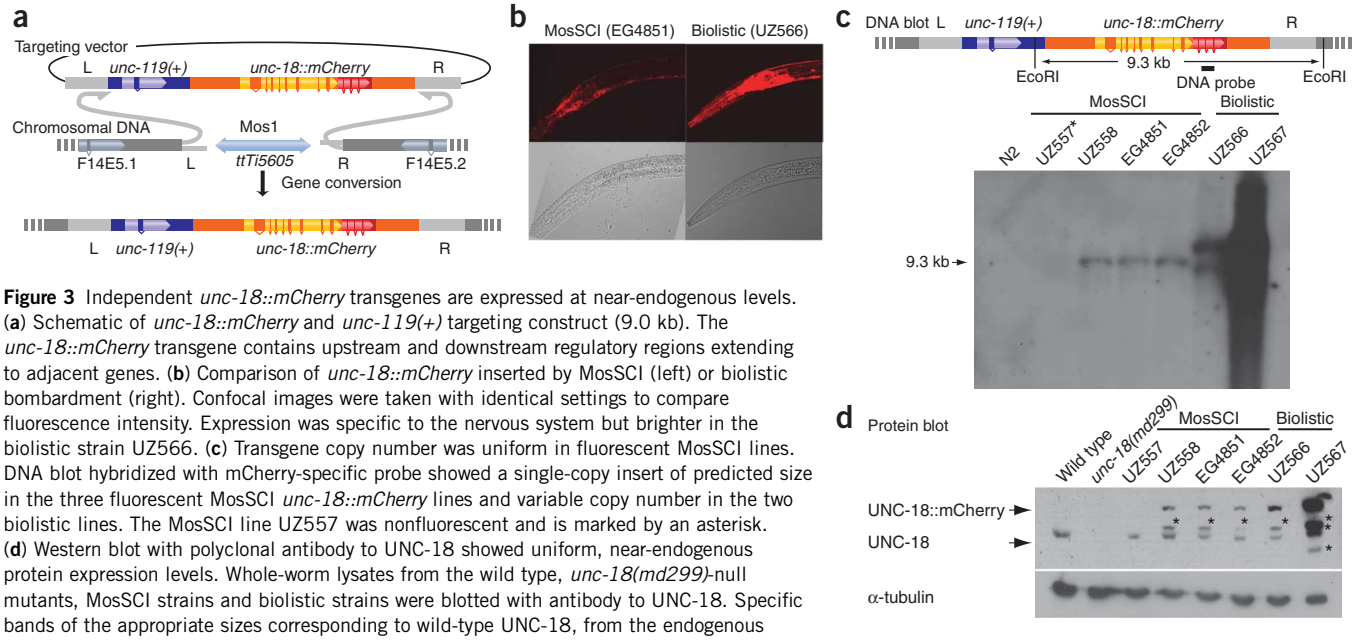

Figure 3 Independent *unc-18::mCherry* transgenes are expressed at near-endogenous levels. (**a**) Schematic of *unc-18::mCherry* and *unc-119(+)* targeting construct (9.0 kb). The *unc-18::mCherry* transgene contains upstream and downstream regulatory regions extending to adjacent genes. (**b**) Comparison of *unc-18::mCherry* inserted by MosSCI (left) or biolistic bombardment (right). Confocal images were taken with identical settings to compare fluorescence intensity. Expression was specific to the nervous system but brighter in the biolistic strain UZ566. (**c**) Transgene copy number was uniform in fluorescent MosSCI lines. DNA blot hybridized with mCherry-specific probe showed a single-copy insert of predicted size in the three fluorescent MosSCI *unc-18::mCherry* lines and variable copy number in the two biolistic lines. The MosSCI line UZ557 was nonfluorescent and is marked by an asterisk. (**d**) Western blot with polyclonal antibody to UNC-18 showed uniform, near-endogenous protein expression levels. Whole-worm lysates from the wild type, *unc-18(md299)*-null mutants, MosSCI strains and biolistic strains were blotted with antibody to UNC-18. Specific bands of the appropriate sizes corresponding to wild-type UNC-18, from the endogenous locus, and mCherry-tagged UNC-18, from the transgene, were observed in all fluorescent strains. Transgenic strains generated by MosSCI showed uniform, near-endogenous protein levels compared to variable, higher expression from biolistic strains. *unc-18::mCherry* degradation bands are marked by asterisks.

Table 2 Insertion of the germline-specific *Ppie-1::GFP::histone* transgene at *ttTi5605*

Parent strain	Transmission	Insertion frequency	Germline fluorescence
EG4855	69%	12/800 (1.50%)	4/12
		5/500 (1.00%)	0/5
EG4856	38%	4/400 (1.00%)	1/4
		1/500 (0.20%)	1/1
EG4857	44%	5/360 (1.39%)	1/5
		10/500 (2.00%)	3/10
Average	50%	37/3,060 (1.2%)	10/37 (27%)

We generated three independent strains containing the *Ppie-1::GFP::histone* on an extrachromosomal array. The transmission rate of each extrachromosomal array was quantified, and each strain was tested twice for transgene insertion by heat-shock. We could not detect any obvious correlation between transmission rate and insertion frequency. Insertion strains were scored for germline GFP fluorescence on a dissection microscope after 3 or 4 generations at 25 °C.

These single-copy insertions were expressed at similar levels as the endogenous gene, as detected by western blotting (**Fig. 3d**). We detected a single band of the appropriate size from whole-worm lysates of wild-type worms. As expected from fluorescence microscopy, DNA blotting revealed expression of UNC-18::mCherry fusion protein in lysates from three of the four single-copy insertions. The protein levels were uniform across the three fluorescent strains and comparable in intensity to the wild-type band. By contrast, the levels of protein expression in the biolistic strains were different from each other and overexpressed compared to the wild type.

We conclude that transgenes up to at least 7 kb can be inserted as single copies without any obvious decrease in insertion frequency or fidelity. Moreover, gene expression more closely mimics endogenous levels compared to multicopy biolistic insertions.

Germline expression

It is difficult to express transgenes in the *C. elegans* germline using standard methodologies. Repetitive arrays are efficiently silenced by RNA interference (RNAi) in the germline[5,17]. Even when arrays are made less repetitive by coinjecting complex carrier DNA or by integrating the DNA by biolistic transformation, it is difficult to obtain germline expression that persists for many generations. A single-copy targeted insertion would be expected to circumvent this limitation, although it is not clear a priori that the *ttTi5605* locus is permissive for germline expression.

To determine whether the region surrounding *ttTi5605* is permissive for germline expression, we inserted transgenes with germline-specific promoter elements and examined expression patterns. Hermaphrodite germline expression was tested by inserting a transgene containing a GFP-histone fusion protein under the control of the *pie-1* promoter (*Ppie-1::GFP::H2B*). From three independent extrachromosomal array lines, we generated 37 independent inserts from 3,060 heat-shocked P_0 worms. This corresponds to an average insertion frequency of 1 in every 80 heat-shocked worms, with some variability between parent strains (**Table 2**). Of the 37 lines containing an insert, only 10 showed robust germline fluorescence (**Fig. 4a**). Notably, germline fluorescence gradually increased over three to four generations after isolation of the insertions in some strains. Low initial expression is consistent with residual silencing from the repetitive array, as previously described[18]. Silencing was eventually lost during passaging of the strain. Once fully desilenced animals were picked clonally, we did not observe resilencing of GFP expression in the germline after observing two lines for more than ten generations. The reduced frequency of successful transgene expression using the *Ppie-1::GFP* construct might be caused by errors in gene conversion. Unlike previous constructs, the *pie-1* promoter contains several inverted repeats, which could interfere with synthesis-dependent strand annealing.

To test for expression in sperm, we inserted an mCherry-histone fusion under the control of the *spe-11* promoter (*Pspe-11::mCherry::H2B*). We obtained five independent insertions. Three of five strains expressed mCherry-tagged histone 2B in the hermaphrodite sperm (**Fig. 4b**). mCherry::H2B expression in male sperm was restricted to the distal tip of the gonad (data not shown). In conclusion, the *ttTi5605* locus is permissive for robust expression in the female and male germlines.

Additional reagents

To facilitate gene insertion and cloning, we developed standard cloning and Gateway-compatible vectors targeting the *ttTi5605* locus (**Supplementary Fig. 6** online). The standard cloning vector pCFJ151 contains a multiple cloning site together with the elements that are necessary for targeting and selection (flanking regions for recombination and *Cbr-unc-119(+)*). The Gateway Multisite vector pCFJ150 contains an attR4-attR3 destination cassette between the targeting and selection sequences. This targeting vector is compatible with the Orfeome[19] and Promoterome[20] vector kits.

To facilitate genetic manipulations, we made a strain (EG4887) that can be used for balancing inserted transgenes at the *ttTi5605* locus. We inserted a transgene expressing mCherry-tagged histone under control of the *myo-2* promoter. *oxIs322* expresses mCherry in the pharyngeal muscle cell nuclei and is relatively bright, so it can easily be identified on a fluorescence dissection microscope. This marked chromosome can be used to follow the presence of nonfluorescent transgenes inserted into the *ttTi5605* locus because it perfectly balances the locus in crosses.

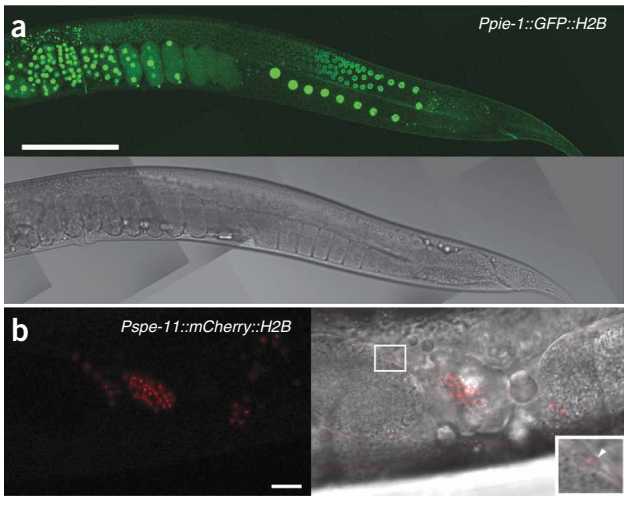

Figure 4 MosSCI inserts are expressed in the female and male germlines. (**a**) The *Ppie-1::GFP::H2B* transgene was expressed throughout the female germline. Scale bar = 100 μm. (**b**) The *Pspe-11::mCherry::H2B* transgene was expressed in hermaphrodite sperm. Left, mCherry expression from *Pspe-11::mCherry::H2B* was specific to hermaphrodite sperm. Scale bar = 10 μm. Right, overlay with differential interference contrast image. Most sperm were found in the spermatheca, although a few were also seen in the gonad and uterus. Right inset shows larger view of boxed area, in which mCherry expression is localized to nucleus (arrowhead). Magnification, ×4.

Table 3 Direct insertions generated by injection

Construct	Mos transposase source	Negative selection	RNAi (*twk-18*)	Number injected	P_0 worms with rescued F_1 progeny	Mean rescued F_1/injected P_0 worms	Direct integrants	Insertion frequency (Integrations/ P_0 worms)	GFP	Mos1 present
Ppie-1::GFP	*Phsp-16-48*	Yes	Yes	20	ND	ND	1	ND	100%	ND
Pspe-11::GFP	*Phsp-16-48*	Yes	Yes	27	17	ND	3	3/17 (18%)	ND	2/3
		No		100	63	10.6	11	11/63 (17%)	ND	6/10
Punc-122::GFP	*Phsp-16-48*	No	Yes	137	91	15.8	8	8/91 (8%)	5/5 (100%)	1/2
			No	110	61	13.2	5	5/61 (7%)	2/2 (100%)	ND
	Pglh-2	No	Yes	102	63	13.2	13	13/63 (19%)	10/12 (83%)	ND
			No	130	75	13.0	13	13/75 (17%)	4/4 (100%)	ND

Germline-expressed transposase is more effective than heat-shock–induced transposase at generating direct insertions. Injected P_0 worms were individually placed on a small NGM plate. F_1 progeny were scored and *unc-119*-rescued worms counted. F_2 progeny were scored for direct insertion events based on rescue and absence of coinjection mCherry markers. A subset of insertion strains were homozygosed and scored for GFP fluorescence in the appropriate tissue. All selected strains were readily homozygosed. In matched experiments with the *Punc-122::GFP* transgene, there was no effect of RNAi bacteria on direct insertion. It was significantly more effective to use germline-expressed transposase (*Pglh-2*) compared to heat-shock–induced transposase (*Phsp-16-48*; $P = 0.015$ by two-sided Fisher's exact test for *Pglh-2* versus *Phsp-16-48*). The insertion frequency was calculated as the fraction of plates containing rescued F_1 progeny that resulted in an insert.

Direct insertions

We occasionally observed putative direct insertions while screening the starved F_2 progeny of injected worms. These worms had never been heat-shocked, yet they were rescued for *unc-119* and did not express the mCherry markers or the *twk-18(ts)* paralyzed phenotype from the negative-selection markers. Notably, these strains showed specific expression of the relevant inserted transgene, including germline expression of a *Ppie-1::GFP::H2B* transgene. We confirmed that these events had occurred by targeted insertion using PCR (**Table 3**). Direct insertions have the considerable advantage that they can be isolated in only a week and with significantly fewer steps than the heat-shock protocol (**Fig. 5a**). We therefore characterized direct events in detail and tested conditions to optimize the frequency of insertions.

To determine the frequency of direct insertions, we singled each injected P_0 worm onto a plate. In the F_1 generation, we counted the number of P_0 worms that generated rescued progeny; we considered these 'successfully injected worms'. In the F_2 generation, we determined how many of these P_0 worms generated direct insertion events.

The P_0 worms that gave rise to direct insertions also gave rise to extrachromosomal arrays. We recovered three direct insertion events from 17 successfully injected worms (18%) when we injected the *Pspe-11::GFP::H2B* transgene. We verified that these transgenes were true insertion events by PCR and Southern blot analysis (**Table 3** and **Supplementary Fig. 7** online).

The frequency of insertions was high enough that it was feasible for us to directly screen individual plates with injected worms without using the negative-selection marker. Although *twk-18(ts)* provides powerful negative selection against the extrachromosomal array, it is not completely benign at 15 °C or 20 °C and therefore makes direct insertions more difficult. From 63 successfully injected worms without use of the negative-selection marker, we recovered 11 verified inserts (17%). The negative selection-marker is therefore not necessary to recover direct inserts (17% versus 18% in the control; **Table 3**).

These worms had never been heat-shocked, yet spontaneous expression of the transposase gene was able to stimulate excision of Mos1. We hypothesized that a germline promoter might provide even

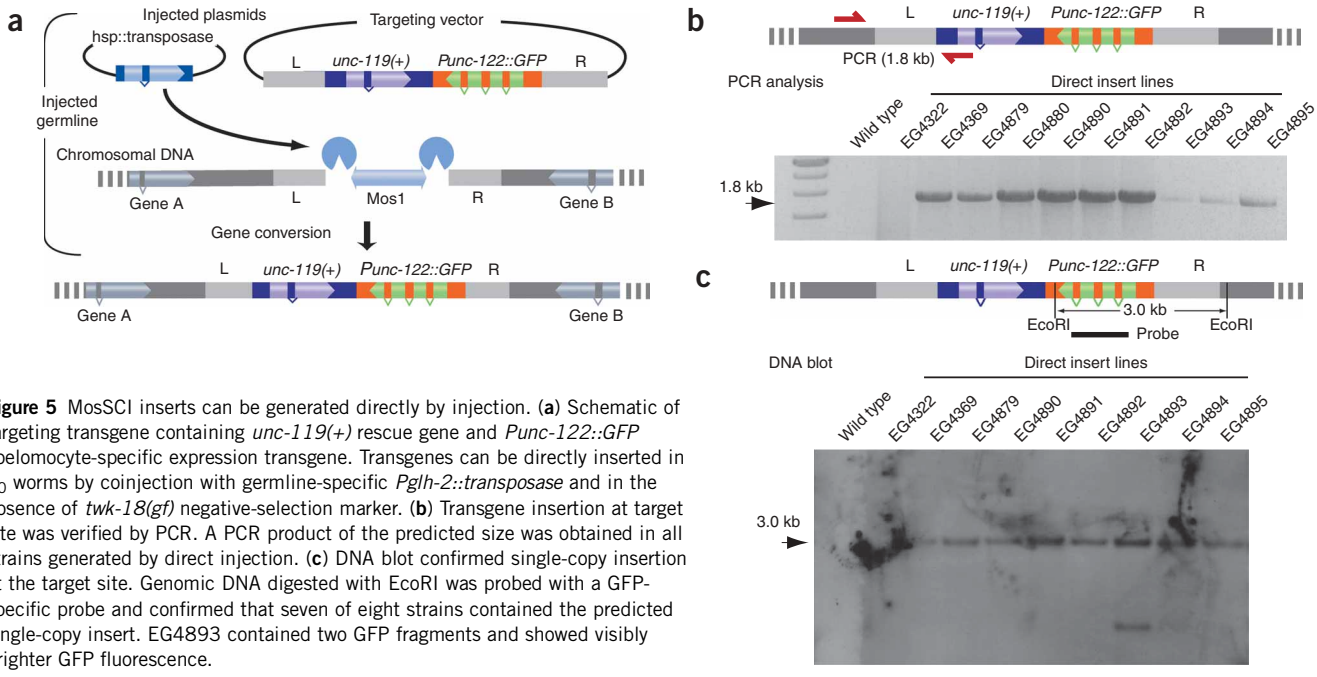

Figure 5 MosSCI inserts can be generated directly by injection. (**a**) Schematic of targeting transgene containing *unc-119(+)* rescue gene and *Punc-122::GFP* coelomocyte-specific expression transgene. Transgenes can be directly inserted in P_0 worms by coinjection with germline-specific *Pglh-2::transposase* and in the absence of *twk-18(gf)* negative-selection marker. (**b**) Transgene insertion at target site was verified by PCR. A PCR product of the predicted size was obtained in all strains generated by direct injection. (**c**) DNA blot confirmed single-copy insertion at the target site. Genomic DNA digested with EcoRI was probed with a GFP-specific probe and confirmed that seven of eight strains contained the predicted single-copy insert. EG4893 contained two GFP fragments and showed visibly brighter GFP fluorescence.

greater expression in the gonad than the heat-shock promoter. We coinjected the targeting construct with either a plasmid containing the Mos1 transposase under control of the germline-specific *glh-2* promoter (*Pglh-2::transposase*) or, as previously, under control of a heat-shock promoter (*Phsp::transposase*). In experiments that were directly comparable, the germline-expressed Mos1 transposase was significantly more efficient at generating direct inserts (*Pglh-2::transposase*, 26 (18.8%) of 138 P_0 worms; *Phsp::transposase*, 13 (8.6%) of 152 worms; $P = 0.015$ by Fisher's exact test). Direct integrations, using either the heat-shock or germline promoter to express the transposase, resulted in F_2 progeny that were often homozygous. The presence of F_2 homozygotes suggested that the integration occurred in the germline of the injected P_0 worms rather than in the germline of the F_1 progeny.

Direct integration seemed to generate a high fraction of perfect insertions. Visual screening of 24 *unc-119(+)* strains injected with the *Punc-122::GFP* targeting vector revealed that only two strains did not express GFP in the coelomocytes, two strains expressed GFP more brightly and 20 strains seemed to be single-copy insertions (83% versus 60% by heat-shock induction). Further analyses were conducted on eight of the strains expressing GFP in the coelomocytes, one of which was a 'bright' expresser. PCR confirmed that *Cbr-unc-119(+)* was inserted at the correct genomic locus (**Fig. 5b**). Southern blot analysis showed that seven of these strains had a single-copy insert of the transgene and one, which also expressed relatively brighter GFP (EG4893), had a two-copy insert of the transgene (**Fig. 5c**).

Direct integration events are not specific to the *ttTi5605* Mos1 allele. We tested insertion of *Punc-122::GFP* at the *cxTi10882* Mos1 allele on chromosome IV with *Pglh-2::transposase*. From 67 successfully injected worms, we obtained 12 insertions, corresponding to a success rate of 17.9%. In conclusion, these data show that single-copy DNA can be introduced into the genome with high fidelity simply by injecting DNA into the gonad.

DISCUSSION

We have developed a technique, MosSCI, that inserts single-copy transgenes into a well-defined genomic environment in *C. elegans*. At this locus, inserted transgenes did not cause obvious mutant phenotypes, nor did they seem to be influenced by endogenous promoters, as specific expression of fluorescent markers was achieved in coelomocytes, the hermaphrodite germline, sperm and the nervous system. Insertion frequency was efficient: ~1 in every 20 injected worms or 1 in every 100 heat-shocked worms. For most constructs tested, a high proportion (>60%) were complete single-copy insertions at the targeted location. Insertions were generated at similar frequencies for at least two different genomic loci.

We described two protocols to generate MosSCI insertions: by direct insertion from injected DNAs or by insertion from an extrachromosomal array after transposon mobilization. The heat-shock protocol was slower and required more experimental steps but might be advantageous in some cases. First, injections are inefficient when one is learning the technique; for the beginner, a single transgenic line is a victory. Second, some transgenes might be difficult to insert correctly, for example, if the transgene is large or if repetitive elements affect gene conversion. Third, for transgenes that confer a dominant phenotype, the negative selection in the array will provide a selective advantage to worms with single-copy inserts.

Not all insertions are perfect. The presence of errors suggests that insertion proceeds by synthesis-dependent strand annealing[21]. The insertion mechanism used in MosSCI was previously characterized[22]. The repair process seems to share many features with gap repair in *Drosophila*[23,24]. After P-element excision, insertion frequency is largely independent of transgene size for constructs up to 8 kb, and ~25% of insertions are complex conversion events containing duplications or deletions[24]. Our results were similar; insertion frequencies were similar for 4-kb and 9-kb insertions (*Punc-122::GFP* and *Punc-18::mCherry*, respectively), and 10–40% of the transgenes contained deletions or insertions. How do these aberrant structures arise? The structures we observed (**Supplementary Fig. 4**) suggested that at least some of the repair proceeds through synthesis-dependent strand annealing[21]. In this mechanism, the two broken DNA ends independently synthesize DNA from homologous repair templates until these single strands overlap and can anneal to bridge the double-strand break[24]. Premature termination or inappropriate DNA synthesis followed by nonhomologous end-joining of the break will generate deletions or insertions, respectively. In contrast, these complex structures may reflect the structure of DNA in the extrachromosomal template, as ~20% of plasmids reisolated from extrachromosomal *C. elegans* arrays contain insertions and deletions[25]. In this case, the gene conversion simply duplicates errors that pre-exist in the extrachromosomal array.

The structure of the template DNA might also affect the frequency of errors. In most cases, we saw that fewer than 25% of insertions contained errors. However, the *Ppie-1* promoter construct resulted in a large number of nonfluorescing inserts (73%) that are likely to be defective. The *Ppie-1* promoter contains many simple inverted repeats from the large intron of the *pie-1* gene. Inverted repeats can form hairpins in single-stranded DNA and might disrupt annealing of the two repaired strands. Alternatively, the *pie-1* promoter might have simply maintained the inherited silencing effects from the transgene array. Hereditary silencing occurs in worms that have carried a repetitive transgene array[18].

The MosSCI technique opens up a number of experimental possibilities. First, transgene expression and rescue that depend on germline expression can be achieved faster and in a more controlled fashion than by biolistic bombardment. Second, structure-function studies will benefit from single-copy insertions at identical genomic contexts. Strains carrying different structural variants could be compared, as copy number and DNA context would be identical. In the cases where a *C. elegans* mutant can be rescued by its human ortholog, this technique will allow the substantial advantages of worm genetics to be harnessed for the analysis of human genes.

METHODS

Reagents. Strains EG4322, EG5003 and EG4887 have been deposited with the Caenorhabditis Genetics Center. Plasmids necessary for transgene insertion have been deposited with Addgene.

Genetics. Mos1 alleles were selected by visual screening in WormBase for appropriately located transposon insertions and provided by the NemaGENETAG consortium. Mos1 insertions were homozygosed and followed in crosses by PCR. Strains were maintained on nematode growth medium (NGM) plates seeded with OP50 bacteria, except where Ahringer laboratory bacterial RNAi clone X-4F11 against *twk-18* was used to increase growth rate. RNAi plates were prepared as described previously[26].

Insertion technique. Transgenic worms were made by injection into EG4322 (*ttTi5605; unc-119(ed3)*) or EG4316/EG5003 (*unc-119(ed3) III; cxTi10882 IV*) worms[1]. The standard injection mix consisted of 50 ng/µl repair template, 50 ng/µl Mos1 transposase (either pJL44 (*Phsp-16-48::transposase*) or pJL43.1 (*Pglh-2::transposase*)), 10 ng/µl pCFJ70 (*Pmyo-3::twk-18(cn110)*), 5 ng/µl pGH8 (*Prab-3::mCherry*), 5ng/µl pCFJ104 (*Pmyo-3::mCherry*) and 2.5 ng/µl pCFJ90 (*Pmyo-2::mCherry*). In later direct insertion experiments, pCFJ70 (*Pmyo-3::twk-18(cn110)*) was omitted from the injection mix. *unc-119* worms are severely paralyzed and egg-laying defective, so L_1–L_2 worms were manually

distributed across a lawn of OP50, and very young adults were selected for injection. Injected worms were individually transferred to standard NGM plates and placed at 15 °C. Plates were scored for the number of phenotypically rescued F_1 worms 3 d after injection.

For the heat-shock protocol, clonal populations of stable array-transmitting lines were picked from the F_2 progeny. To increase the speed of population expansion, lines were grown on *twk-18* RNAi plates at 20 °C. We tested the negative temperature selection caused by *Pmyo-3::twk-18(cn110)* by propagating worms on OP50 for two generations and then shifting them to 25 °C. Worms with good negative selection were almost fully paralyzed and unable to lay eggs after 1–2 d at 25 °C. Once a transgenic line with sufficient negative selection and visible fluorescent markers had been established, a population of young adults was heat-shocked for 1 h at 34 °C in a water bath and allowed to recover at 15 °C for several hours. Sets of 20 adult heat-shocked worms were transferred to 10-cm NGM plates seeded with OP50 bacteria and propagated at 20 °C. When worms on these plates became starved, roughly a quarter of the plate was chunked to a fresh, seeded 10-cm NGM plate and placed at 25 °C. Two to five days later (but before starvation), these plates were visually screened for insertion events based on the presence of nonparalyzed, wild-type worms. Insertion strains were verified on a fluorescence dissection microscope by the lack of fluorescent mCherry coinjection markers and subsequently homozygosed.

For the direct insertion protocol, individual injected worms were allowed to exhaust the food source. Once starved, plates containing transgenic lines were screened for insertion events on a fluorescence dissection microscope based on wild-type movement but complete lack of fluorescent coinjection markers. Plates containing insertion events typically had a large proportion of non-fluorescent moving worms, although some plates only had a few.

For most experiments, we inserted transgenes flanked by ~1.5 kb of homology to each side of the Mos1 insertion[13]. We also tested constructs with a shorter 500-bp homology region to minimize the cloning vector. Transgene insertions were possible, but pilot experiments showed that the frequency seemed to be reduced by a factor of almost five with shorter homology arms. Because only a marginal decrease in vector size was achieved, we did not continue these experiments past pilot experiments.

Biolistic transformation. Integrated strains were made by biolistic bombardment with a Bio-Rad PDS/HE-1000 as described[16].

Molecular biology. Many of the plasmids were constructed using the Invitrogen MultiSite Gateway Three-Fragment kit (cat. no. 12537-023). Reactions were done according to the manufacturer's instructions, and the enzymes were purchased directly from Invitrogen. All cloning PCR amplifications were done with a high-fidelity Phusion polymerase (Finnzymes).

DNA blotting. Worms were grown on 2YT agarose plates seeded with NA22 bacteria. Genomic DNA was isolated with a Qiagen genomic tip 100/G kit or DNeasy kit. Standard techniques were used for overnight genomic restriction digest with EcoRI and subsequent agarose (0.7%) gel electrophoresis at low voltage (50 mV). DNA bands were transferred to a Millipore Immobilon-NY+ membrane. Chemiluminescent probes were synthesized with a New England Biolabs NEBlot kit. Hybridization and washes were done according to the manufacturer's instructions and detected with a New England Biolabs Phototope-Star detection kit for nucleic acids.

Protein blotting. Worms were collected from plates with 50–75% food consumption by rinsing with M9 medium. Worms were allowed to settle, and the supernatant was removed. Worms were then washed 3× with M9 medium and resuspended in M9 medium to give a 50% worm pellet volume. An equal volume of 2× SDS-PAGE sample buffer was added, and samples were boiled for 5 min. Aliquots (50 µl) of boiled lysates were drop-dialyzed (VSWP02500, Millipore) against 50% M9 medium for 10 min. Dialyzed lysates (~50 µl) were recovered, an equal volume of 2× SDS-page sample buffer was added, samples were boiled for 5 min and 20 µl of samples were loaded on 10% and 15% SDS-PAGE gels (Mini-Gels, Bio-Rad). Aliquots (20 µl) of sample were loaded and run at 50 mV for 30 min then at 150 mV for 1 h. Transfers were made to polyvinylidene fluoride membrane using semidry apparatus (20 mV, 2 h). Membranes were probed with affinity-purified rabbit antibody to UNC-18 (gift from J. Rand, Oklahoma Medical Research Foundation) and mouse monoclonal antibody to tubulin (12G10 supernatant; Developmental Studies Hybridoma Bank) at 1:2,000 in 1× PBS-Tween (Tween 20 at 0.1%). Secondary probing was done at 1:10,000 with horseradish peroxidase–conjugated antibodies to mouse IgG and rabbit IgG (GE Healthcare). Membranes were developed with ECL reagent (GE Healthcare), imaged on film (Pierce Biotechnology) and processed (ImageJ, gel analysis plug-in).

Statistical analysis. We used a two-sided Fisher's exact test to determine significance of *Pglh-2* versus *Phsp16-48* for direct insertions; 0.05 was used as the cutoff for statistical significance.

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

ACKNOWLEDGMENTS

We thank L. Segalat and the NemaGENETAG consortium for generating and supplying *Mos1* insertion alleles; L. Salkoff for *twk-18* cDNA clones; G. Seydoux for promoter plasmids; J. Rand for UNC-18 antibodies; C. Thacker at the Utah worm core facility for generating biolistic transformants; G. Stanfield for help in confirming sperm expression; J.-L. Bessereau and V. Robert for sharing unpublished data; and M. Ailion, G. Hollopeter and the rest of the Jorgensen laboratory for scientific input. E.M.J. is an Investigator of the Howard Hughes Medical Institute. S.-P.O. is an Investigator of the Danish National Research Foundation. C.F.-J. was funded by fellowships from Fonden af 17.12.1981 and the Lundbeck Foundation.

AUTHOR CONTRIBUTIONS

C.F.-J., M.W.D., C.E.H. and E.M.J. designed the experiments. C.F.-J., M.W.D., C.E.H., B.J.N. and J.M.T. carried out the experiments. E.M.J., M.G. and S.-P.O. supervised and funded the experiments. C.F.-J., M.W.D. and E.M.J. wrote the manuscript.

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

1. 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).
2. Thellmann, M., Hatzold, J. & Conradt, B. The Snail-like CES-1 protein of *C. elegans* can block the expression of the BH3-only cell-death activator gene *egl-1* by antagonizing the function of bHLH proteins. *Development* **130**, 4057–4071 (2003).
3. Hsieh, J. & Fire, A. Recognition and silencing of repeated DNA. *Annu. Rev. Genet.* **34**, 187–204 (2000).
4. Hsieh, J. *et al.* The RING finger/B-box factor TAM-1 and a retinoblastoma-like protein LIN-35 modulate context-dependent gene silencing in *Caenorhabditis elegans*. *Genes Dev.* **13**, 2958–2970 (1999).
5. Kelly, W.G., Xu, S., Montgomery, M.K. & Fire, A. Distinct requirements for somatic and germline expression of a generally expressed *Caernorhabditis elegans* gene. *Genetics* **146**, 227–238 (1997).
6. Mello, C. & Fire, A. DNA transformation. *Methods Cell Biol.* **48**, 451–482 (1995).
7. Hammarlund, M., Palfreyman, M.T., Watanabe, S., Olsen, S. & Jorgensen, E.M. Open syntaxin docks synaptic vesicles. *PLoS Biol.* **5**, e198 (2007).
8. Sha, K. & Fire, A. Imprinting capacity of gamete lineages in *Caenorhabditis elegans*. *Genetics* **170**, 1633–1652 (2005).
9. Berezikov, E., Bargmann, C.I. & Plasterk, R.H. Homologous gene targeting in *Caenorhabditis elegans* by biolistic transformation. *Nucleic Acids Res.* **32**, e40 (2004).
10. Barrett, P.L., Fleming, J.T. & Gobel, V. Targeted gene alteration in *Caenorhabditis elegans* by gene conversion. *Nat. Genet.* **36**, 1231–1237 (2004).
11. Plasterk, R.H. & Groenen, J.T. Targeted alterations of the *Caenorhabditis elegans* genome by transgene instructed DNA double strand break repair following Tc1 excision. *EMBO J.* **11**, 287–290 (1992).
12. Bessereau, J.L. *et al.* Mobilization of a *Drosophila* transposon in the *Caenorhabditis elegans* germ line. *Nature* **413**, 70–74 (2001).
13. Robert, V. & Bessereau, J.L. Targeted engineering of the *Caenorhabditis elegans* genome following Mos1-triggered chromosomal breaks. *EMBO J.* **26**, 170–183 (2007).
14. Duverger, Y. *et al.* A semi-automated high-throughput approach to the generation of transposon insertion mutants in the nematode *Caenorhabditis elegans*. *Nucleic Acids Res.* **35**, e11 (2007).
15. Maduro, M. & Pilgrim, D. Identification and cloning of *unc-119*, a gene expressed in the *Caenorhabditis elegans* nervous system. *Genetics* **141**, 977–988 (1995).
16. Praitis, V., Casey, E., Collar, D. & Austin, J. Creation of low-copy integrated transgenic lines in *Caenorhabditis elegans*. *Genetics* **157**, 1217–1226 (2001).
17. Kim, J.K. *et al.* Functional genomic analysis of RNA interference in *C. elegans*. *Science* **308**, 1164–1167 (2005).

18. Vastenhouw, N.L. *et al.* Gene expression: long-term gene silencing by RNAi. *Nature* **442**, 882 (2006).
19. Reboul, J. *et al. C. elegans* ORFeome version 1.1: experimental verification of the genome annotation and resource for proteome-scale protein expression. *Nat. Genet.* **34**, 35–41 (2003).
20. Dupuy, D. *et al.* A first version of the *Caenorhabditis elegans* Promoterome. *Genome Res.* **14**, 2169–2175 (2004).
21. Chen, J.M., Cooper, D.N., Chuzhanova, N., Ferec, C. & Patrinos, G.P. Gene conversion: mechanisms, evolution and human disease. *Nat. Rev. Genet.* **8**, 762–775 (2007).
22. Robert, V., Davis, M.W., Jorgensen, E.M. & Bessereau, J.L. Gene conversion and end-joining repair double-strand breaks in the *C. elegans* germline. *Genetics* **180**, 673–679 (2008).
23. Gloor, G.B., Nassif, N.A., Johnson-Schlitz, D.M., Preston, C.R. & Engels, W.R. Targeted gene replacement in *Drosophila* via P element-induced gap repair. *Science* **253**, 1110–1117 (1991).
24. Nassif, N., Penney, J., Pal, S., Engels, W.R. & Gloor, G.B. Efficient copying of nonhomologous sequences from ectopic sites via P-element-induced gap repair. *Mol. Cell. Biol.* **14**, 1613–1625 (1994).
25. Stinchcomb, D.T., Shaw, J.E., Carr, S.H. & Hirsh, D. Extrachromosomal DNA transformation of *Caenorhabditis elegans*. *Mol. Cell. Biol.* **5**, 3484–3496 (1985).
26. Kamath, R.S., Martinez-Campos, M., Zipperlen, P., Fraser, A.G. & Ahringer, J. Effectiveness of specific RNA-mediated interference through ingested double-stranded RNA in. *Caenorhabditis elegans*. *Genome Biol.* **2**, RESEARCH0002 (2001).