Published in final edited form as:

Nat Genet. 2008 November ; 40(11): 1375–1383. doi:10.1038/ng.248.

Single copy insertion of transgenes in *C. elegans*

Christian Frøkjær-Jensen^{1,2}, M. Wayne Davis¹, Christopher E. Hopkins¹, Blake Newman¹, Jason M. Thummel¹, Søren-Peter Olesen^{2,3}, Morten Grunnet^{2,3}, and Erik M. Jorgensen¹

¹Department of Biology and Howard Hughes Medical Institute, University of Utah, Salt Lake City, UT 84112

²The Danish National Research Foundation Centre for Cardiac Arrhythmia, University of Copenhagen, Denmark

³NeuroSearch A/S, Ballerup, Denmark

Abstract

Currently transgenes in *C. elegans* are generated by injecting DNA into the germline. The DNA assembles into a semi-stable 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 called MosSCI (Mos1-mediated Single Copy Insertion) that inserts a single copy of a transgene into a defined site. Mobilization of a Mos1 transposon generates a double strand break in non-coding 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 two weeks by injecting approximately twenty animals. 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.

Keywords

C. elegans; homologous recombination; gene targeting; transposon; Mos1

Introduction

The first method for introducing DNA into *C. elegans* was developed by Craig Mello *et al.* 1. The methodology has not changed in the last seventeen years and is elegant in its simplicity: DNA is injected into the gonad of a hermaphrodite. Injected DNA 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, and thus 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 bonafide 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

Correspondence: Erik M. Jorgensen, email: jorgensen@biology.utah.edu, phone: (801) 585-3517, Department of Biology, University of Utah, 257 South 1400 East, Salt Lake City, UT 84112.

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.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.

tissues including muscles^{3,4} and the germline⁵. The arrays can be silenced even after they are integrated into a chromosome by irradiation, presumably because of transcriptional silencing of arrays⁶. 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 transformation⁹ or, more effectively, by template-directed repair following 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. The Bessereau group recently demonstrated that specific DNA changes can be targeted to loci with Mos1 insertions¹³. This technique is called Mos1 excision-induced transgene-instructed gene conversion (MosTIC). These authors demonstrated that they could 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 NemaGENETAG consortium has generated a large library of Mos1 inserts with known locations in the genome¹⁴.

In this study, we adapt intergenic Mos1 elements for the routine insertion of transgenes using a variation of the MosTIC technique. We call this technique Mos1-mediated single copy insertion (MosSCI). We show that transgenes are inserted as single copies at a defined chromosomal locus. This locus supports expression in a broad range of tissues at apparently endogenous levels. Importantly, 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 animals.

Results

Insertion method

A perfect 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 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 *tTi5605* Mos1 allele, near the center of chromosome II as a test case (map position +0.78).

In overview, the goal is to generate a double-strand break in the chromosome via Mos1 excision and provide a homologous template for repair of the break. We generated an extrachromosomal array that contains roughly 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 provided a positive selection marker. The *tTi5605* Mos1 insertion was crossed into an *unc-119(ed3)* mutant background. *unc-119(ed3)* mutant animals are small, almost paralyzed, have small brood sizes and are incapable of forming dauer larvae when starved¹⁵. By placing the wild-type *C. 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 inducing gene conversion the only animals capable of

surviving starvation will be animals 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 coding for the *Mos1* transposase under the control of the heat shock promoter *hsp-16-48* 12. Heat shock activates synthesis of the transposase, which in turn excises the transposon. The double strand break is then healed via gene conversion - at least in some cases - using the extrachromosomal array as the template.

The main challenge is to distinguish targeted insertions from worms carrying the extrachromosomal array since both will be rescued for the positive *unc-119* co-insertion marker (Fig. 1). To differentiate integrations from arrays, we co-injected the DNA of interest with negative selection markers: genes expressing red fluorescent proteins and *TK-18(gf)*, an activated K^+ channel that causes muscle paralysis at elevated temperatures (Supplementary Fig. 1). These markers are incorporated into the transgene arrays, but are not copied into the targeted integration. Animals carrying the array are active at 15° and can be propagated. At 25° these animals are paralyzed; only animals 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 contains, as the gene of interest, a construct expressing GFP in coelomocytes (*Punc-122::GFP*) along with the positive selection marker that rescues *unc-119* (Fig. 2a and Supplementary Fig. 2). To direct gene conversion, these two genes were flanked by 1.4 kb genomic sequences on each side that are homologous to the flanking sequences of the *ttTi5605* *Mos1* site. The targeting vector, the transposase construct, the *twk-18* negative selection marker and fluorescent mCherry markers were co-injected into *unc-119(ed3) ttTi5605mos*. Injected animals were then cultured at 15°C. Five independent *unc-119* rescued lines were selected and propagated to expand the population. Rescued animals displayed all of the fluorescent markers and were strongly paralyzed at 25° C. A semi-synchronous population of young adult animals was heat-shocked to induce *Mos1* excision. The F2 generation was screened for transgene instructed repair (see Supplementary Methods for a detailed protocol). In total, we heat-shocked 1000 animals from 5 independent transgenic lines and recovered 10 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 animals that never segregated progeny exhibiting the *Unc-119* mutant phenotype. We confirmed dim expression of GFP in the coelomocytes in 7 of 10 insert lines (Fig. 2b).

The polymerase chain reaction (PCR) was used to demonstrate 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 the original targeting strain (genotype: *unc-119(ed3); ttTi5605*) confirmed that our PCR reaction is specific.

The presence of the *Punc-122::GFP* transgene was determined by Southern analysis using a GFP-specific probe (Fig. 2d). Six of ten insertion strains exhibited 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, has what appears to be a tandem insertion. As expected, the three non-fluorescent strains showed either no band or an aberrant band size. We further characterized these three non-fluorescent 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).

Consistent with the lack of fluorescence and the Southern 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 sequenced the product. EG4441 is a 1800 bp deletion within the *Punc-122::GFP* transgene and EG4448 is an 803 bp deletion accompanied by a 1713 bp insertion from elsewhere on the extrachromosomal array (Supplementary Fig. 3 and Supplementary Fig. 4). Similar deletions were observed with MosTIC¹³.

The *Punc-122::GFP* transgene was specifically expressed in the coelomocytes, suggesting that the *ttTi5605* site is permissive for specific tissue 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). These constructs are appropriately expressed in just the 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 *hsp:Mos transposase* and negative selection markers, to form an extrachromosomal array. We selected a single transgenic line and heat-shocked an adult population. We recovered 8 transgene insertions from 800 heat-shocked worms, a frequency that is identical to the *ttTi5605* site. We confirmed dim coelomocyte fluorescence in five of six lines that we studied in detail.

Together, these results demonstrate that on average 1 in every 100 heat-shocked animals resulted in an insertion event at two distinct genomic loci. 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 analysis. The gene was inserted into a targeting plasmid along with *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 transposase expression was induced by heatshock in a population of adults. From 500 heat-shocked animals, we recovered four targeted integrants (Supplementary Table 1). From this limited data set, insertion frequency does not appear 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 particles¹⁶. 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)* animals. We confirmed integration by genetic mapping. UZ566 exhibited 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 genes¹⁶. We performed a Southern blot on all of the *unc-18::mCherry* integrants to test if the differences in fluorescence

intensity reflected differences in transgene copy number (Fig. 3c). The blot was probed 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.

These single copy insertions are expressed at similar levels as the endogenous gene as analyzed by Western blot (Fig. 3d). We detected a single band of the appropriate size from whole worm lysate of wild-type animals. The UNC-18::mCherry fusion protein was observed in lysates from three of the four single copy insertions, as expected from fluorescence microscopy. 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 over-expressed 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 notoriously difficult to express transgenes in the *C. elegans* germline using standard methodologies. Repetitive arrays are efficiently silenced by RNA interference in the germline^{5,17}. Even when arrays are made less repetitive by co-injecting 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 if 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 3060 heat-shocked P0 worms. This corresponds to an average insertion frequency of 1 in every 80 heatshocked worms with some variability between parent strains (Table 2). Of the 37 lines containing an insert, only 10 exhibited robust germline fluorescence (Fig. 4a). Interestingly, 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 by Vastenhouw *et al.* 18. Silencing is eventually lost during passaging of the strain. Once fully de-silenced animals were picked clonally, we did not observe re-silencing of GFP expression in the germline after observing 2 lines for more than ten generations. The reduced frequency of successful expressing transgenes 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 5 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). 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 Orfeome¹⁹ and Promoterome²⁰ 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 the *myo-2* promoter. *oxIs322* expresses mCherry in the pharyngeal muscle cell nuclei and is relatively bright so that 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.

Direct insertions

We occasionally observed putative direct insertions while screening the starved F2 progeny of injected animals. These animals had never been heatshocked, yet they were rescued for *unc-119* and failed to express the mCherry markers or the *twk-18(ts)* paralyzed phenotype from the negative selection markers. Importantly, 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 P0 onto a plate. In the F1 generation we counted the number of P0s that generated rescued progeny; we considered these 'successfully injected animals'. In the F2 generation we determined how many of these P0s generated direct insertion events. Note that the P0s that gave rise to direct insertions also gave rise to extrachromosomal arrays. We recovered 3 direct insertion events from 17 successfully injected animals (18 %), when we injected the *Pspe-11::GFP::H2B* transgene. We verified that these transgenes were true insertion events by PCR and Southern analysis (Table 3 and Supplementary Fig. 7).

The frequency of insertions was high enough that direct screening of individual plates with injected animals without using the negative selection marker was feasible. Although *twk-18(ts)* provides a powerful negative selection against the extrachromosomal array, it is not completely benign at 15°C or room temperature and therefore makes direct insertions more difficult. From 63 successfully injected animals 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% vs 18 % in the control, Table 3).

These animals had never been heatshocked yet spontaneous expression of the transposase gene was able to stimulate excision of Mos1. A germline promoter might express even better in the gonad than the heatshock 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 a heatshock promoter (*Phsp::transposase*). In experiments that are directly comparable, the germline-expressed Mos1 transposase is significantly more efficient at generating direct inserts (*Pglh-2::transposase*: 26/138 P0s (18.8%) vs *Phsp::transposase*: 13/152 (8.6%), $p = 0.015$, Fisher's exact). Direct integrations,

using either the heatshock or germline promoter to express the transposase, resulted in F2 progeny that were often homozygous. The presence of F2 homozygotes suggests that the integration occurs in the germline of the injected P0 animal rather than in the germline of the F1 progeny.

Direct integration appears to generate a high fraction of perfect insertions. Visual screening of 24 *unc-119(+)* strains injected with the *Punc-122::GFP* targeting vector, demonstrated that only 2 strains failed to express GFP in the coelomocytes, 2 strains expressed GFP more brightly, and 20 strains appeared to be single copy insertions (83% vs. 60% by heatshock induction). Further analyses were performed on 8 of the strains expressing GFP in the coelomocytes, one of which was a 'bright' expressor. PCR confirmed that *Cbr-unc-119(+)* was inserted at the correct genomic locus (Fig. 5b). Southern analysis demonstrated that 7 of these strains inserted a single copy of the transgene and one, which also expressed relatively brighter GFP (EG4893), had inserted two copies of the transgene (Fig. 5c).

Direct integration events are not specific to the *tTi5605* 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 animals we obtained 12 insertions, corresponding to a success rate of 17.9 %. In conclusion, these data demonstrate that single copy DNA can be introduced into the genome with high fidelity by simply injecting DNA into the gonad.

Discussion

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

We describe two protocols to generate MosSCI insertions: by direct insertion from injected DNAs, or by insertion from an extrachromosomal array following transposon mobilization. The heat-shock protocol is slower and requires 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 animals with single-copy inserts.

Not all insertions are perfect. The presence of errors suggest that insertion proceeds via synthesis-dependent strand annealing (SDSA)²¹. The insertion mechanism used in MosSCI has been characterized by Robert and Bessereau¹³. The repair process appears to share many features with gap repair in *Drosophila*^{22,23}. Following P-element excision, insertion frequency is largely independent of transgene size for constructs up to 8 kb, and approximately 25% of insertions are complex conversion events containing duplications or deletions²³. Our results are similar; insertion frequencies are similar for 4 kb and 9 kb insertions (*Punc-122::GFP* and *Punc-18::mCherry*, respectively), and 10-40% of the transgenes contain deletions or insertions. How do these aberrant structures arise? The structures we observe (Supplementary Fig. 4) suggest that at least some of the repair proceeds through SDSA²¹. 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 break²³. Premature termination or inappropriate DNA synthesis followed by non-homologous end-joining of the break will generate deletions or insertions, respectively. On the other hand, these complex structures may reflect the structure of DNA in the extrachromosomal template, since approximately 20% of plasmids re-isolated from extrachromosomal *C. elegans* arrays contain insertions and deletions²⁴. 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 less than 25% of insertions contained errors. However, the *Ppie-1* promoter construct resulted in a large number of non-fluorescing inserts (73 %) which 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 animals that have carried a repetitive transgene array¹⁸.

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 can be compared since copy number and DNA context will be identical. In the cases where a *C. elegans* mutant can be rescued by its human ortholog, this will allow the harnessing the substantial advantages of worm genetics to analysis of human genes.

Materials and 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

Nematode strains: Mos1 alleles were selected by visual screening in Wormbase (www.wormbase.org) 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 NGM plates seeded with OP50 bacteria, except where Ahringer lab bacterial RNA interference clone X-4F11 against *twk-18* was used to increase growth rate. RNAi plates were prepared as described previously²⁵.

Insertion technique

Transgenic worms were made by injection into EG4322 (*ttTi5605; unc-119(ed3)*) or EG4316/EG5003 (*unc-119(ed3) III; cxTi10882 IV*) animals¹. 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* animals are severely paralyzed and egg-laying defective. Therefore, L1-L2 animals were manually distributed across a lawn of OP50 and very young adults were selected for injection. Injected animals were individually transferred to standard NGM plates and placed at 15°C. Plates were scored for the number of phenotypically rescued F1 animals 3 days after injection.

Heat-shock protocol—Clonal populations of stable array-transmitting lines were picked from the F2 progeny. To increase the speed of population expansion, lines were grown on *twk-18* RNAi plates at room temperature. We tested the negative temperature selection caused by *Pmyo-3::twk-18(cn110)* by propagating animals on OP50 for two generations and shifting them to 25°C. Animals with good negative selection were almost fully paralyzed and unable to lay eggs after 1-2 days 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 hour at 34°C in a water bath and allowed to recover at 15°C for several hours. Sets of 20 adult heat-shocked animals were transferred to 10 cm NGM plates seeded with OP50 bacteria and propagated at room temperature. When animals 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 non-paralyzed, wild-type animals. Insertion strains were verified on a fluorescence dissection microscope by the lack of fluorescent mCherry co-injection markers and subsequently homozygosed.

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 co-injection markers. Plates containing insertion events typically had a large proportion of non-fluorescent moving animals, although some plates only had a few.

Length of homology regions: For most experiments we inserted transgenes flanked by approximately 1.5 kb of homology to each side of the *Mos1* insertion¹³. 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 appeared to be reduced by a factor of almost five with shorter homology arms. Since 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 Biorad PDS/HE-1000 as described¹⁶.

Molecular biology

Many of the plasmids were constructed using the Invitrogen (Carlsbad, CA, USA) MultiSite Gateway Three Fragment kit (cat. no. 12537-023). Reactions were done according to manufacturer's instructions and the enzymes were purchased directly from Invitrogen.

All cloning PCR amplifications were done with a high fidelity Phusion polymerase (Finnzymes, Espoo, Finland).

Southern blot

Worms were grown on 2YT agarose plates seeded with NA22 bacteria. Genomic DNA was isolated with a Qiagen (Hilden, Germany) 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 electrophoreses at low voltage (50mV). DNA bands were transferred to a Millipore (Bedford, MA, USA) Immobilon-NY+ membrane. Chemiluminescent probes were synthesized with New England Biolabs (Ipswich, MA, USA) NEblot kit. Hybridization and washes were performed according to manufactory's instruction and detected with New England Biolabs (Ipswich, MA, USA) Phototope-Star detection kit for nucleic acids.

Western blot

Worms were harvested from plates with 50 to 75% food consumption by rinsing with M9. Worms were allowed to settle and supernatant was removed. Worms were washed 3x with M9. Worms were re-suspended in M9 to give 50% worm pellet volume. An equal volume of 2x SDS-page sample buffer was added, and samples boiled for 5 min. 50 µl aliquots of boiled lysates were drop dialyzed (Millipore, VSWP02500) against 50% M9 media for 10 min. Dialyzed lysates (~50 µl) were recovered, an equal volume of 2x SDS-page sample buffer added, samples boiled for 5 minutes, and 20 µl samples loaded on 10 and 15% SDS-page gels (Biorad, mini gel). 20 µl aliquots of sample were loaded and ran 50 mV for 30 min then 150 mV for 1 hr. Transfers were made to PVDF membrane using semi-dry apparatus (20 mV, 2 hr). Membranes were probed with appropriate primary antibodies: anti-UNC-18 (gift James Rand, anti rabbit serum, affinity purified) and anti-Tubulin (DSHB, anti-mouse monoclonal, "12G10" supernatant) were used at 1:2000 in 1x PBS-Tween (Tween 20 at 0.1%). Secondary probing was performed at 1:10,000 with anti-mouse IgG-HRP and anti-rabbit IgG-HRP (GELifescience). Membranes were developed with ECL reagent (GELifescience.com), imaged on film (Piercenet.com), and processed (ImageJ, gel analysis plug-in).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors would like to especially thank Laurent Segalat and the NemaGENETAG consortium for generating and supplying Mos1 insertion alleles, Larry Salkoff for *twk-18* cDNA clones, Geraldine Seydoux for promoter plasmids and Jim Rand for UNC-18 antibodies. We thank Colin Thacker at the Utah Worm core facility for generating biolistic transformants and Gillian Stanfield for help confirming sperm expression. We thank Jean-Louis Bessereau and Valerie Robert for sharing unpublished data. We are in debt for scientific input from Michael Ailion and Gunther Hollopeter, as well as the rest of the Jorgensen lab. EMJ is an Investigator of the Howard Hughes Medical Institute. CFJ was funded by fellowships from Fonden af 17.12.1981 and the Lundbeck Foundation.

Literature Cited

1. Mello CC, Kramer JM, Stinchcomb D, Ambros V. Efficient gene transfer in *C.elegans*: extrachromosomal maintenance and integration of transforming sequences. *EMBO J* 1991;10:3959–3970. [PubMed: 1935914]
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* 2003;130:4057–71. [PubMed: 12874127]
3. Hsieh J, Fire A. Recognition and silencing of repeated DNA. *Annu Rev Genet* 2000;34:187–204. [PubMed: 11092826]
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* 1999;13:2958–70. [PubMed: 10580003]
5. Kelly WG, Xu S, Montgomery MK, Fire A. Distinct requirements for somatic and germline expression of a generally expressed *Caenorhabditis elegans* gene. *Genetics* 1997;146:227–238. [PubMed: 9136012]
6. Mello C, Fire A. DNA transformation. *Methods Cell Biol* 1995;48:451–82. [PubMed: 8531738]
7. Hammarlund M, Palfreyman MT, Watanabe S, Olsen S, Jorgensen EM. Open syntaxin docks synaptic vesicles. *PLoS Biol* 2007;5:e198. [PubMed: 17645391]
8. Sha K, Fire A. Imprinting capacity of gamete lineages in *Caenorhabditis elegans*. *Genetics* 2005;170:1633–1652. [PubMed: 15944356]
9. Berezikov E, Bargmann CI, Plasterk RH. Homologous gene targeting in *Caenorhabditis elegans* by biolistic transformation. *Nucleic Acids Res* 2004;32:e40. [PubMed: 14982959]

10. Barrett PL, Fleming JT, Gobel V. Targeted gene alteration in *Caenorhabditis elegans* by gene conversion. *Nat.Genet* 2004;36:1231–1237. [PubMed: 15502826]
11. Plasterk RH, Groenen JT. Targeted alterations of the *Caenorhabditis elegans* genome by transgene instructed DNA double strand break repair following Tc1 excision. *EMBO J* 1992;11:287–290. [PubMed: 1740110]
12. Bessereau JL, et al. Mobilization of a *Drosophila* transposon in the *Caenorhabditis elegans* germ line. *Nature* 2001;413:70–74. [PubMed: 11544527]
13. Robert V, Bessereau JL. Targeted engineering of the *Caenorhabditis elegans* genome following Mos1-triggered chromosomal breaks. *EMBO J* 2007;26:170–183. [PubMed: 17159906]
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* 2007;35:e11. [PubMed: 17164286]
15. Maduro M, Pilgrim D. Identification and cloning of *unc-119*, a gene expressed in the *Caenorhabditis elegans* nervous system. *Genetics* 1995;141:977–988. [PubMed: 8582641]
16. Praitis V, Casey E, Collar D, Austin J. Creation of low-copy integrated transgenic lines in *Caenorhabditis elegans*. *Genetics* 2001;157:1217–1226. [PubMed: 11238406]
17. Kim JK, et al. Functional genomic analysis of RNA interference in *C. elegans*. *Science* 2005;308:1164–1167. [PubMed: 15790806]
18. Vastenhouw NL, et al. Gene expression: long-term gene silencing by RNAi. *Nature* 2006;442:882. [PubMed: 16929289]
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* 2003;34:35–41. [PubMed: 12679813]
20. Dupuy D, et al. A first version of the *Caenorhabditis elegans* Promoterome. *Genome Res* 2004;14:2169–75. [PubMed: 15489340]
21. Chen JM, Cooper DN, Chuzhanova N, Ferec C, Patrinos GP. Gene conversion: mechanisms, evolution and human disease. *Nat Rev Genet* 2007;8:762–75. [PubMed: 17846636]
22. Gloor GB, Nassif NA, Johnson-Schlitz DM, Preston CR, Engels WR. Targeted gene replacement in *Drosophila* via P element-induced gap repair. *Science* 1991;253:1110–1117. [PubMed: 1653452]
23. Nassif N, Penney J, Pal S, Engels WR, Gloor GB. Efficient copying of nonhomologous sequences from ectopic sites via P-element-induced gap repair. *Mol.Cell Biol* 1994;14:1613–1625. [PubMed: 8114699]
24. Stinchcomb DT, Shaw JE, Carr SH, Hirsh D. Extrachromosomal DNA transformation of *Caenorhabditis elegans*. *Mol.Cell Biol* 1985;5:3484–3496. [PubMed: 3837845]
25. Kamath RS, Martinez-Campos M, Zipperlen P, Fraser AG, Ahringer J. Effectiveness of specific RNA-mediated interference through ingested double-stranded RNA in *Caenorhabditis elegans*. *Genome Biol* 2001;2:RESEARCH0002

Figure 1.

Schematic presentation of *Mos1* mediated *Single Copy Insertion* (MosSCI). A *Mos1* transposon located at a non-coding 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 animals 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 animals. Following loss of the array, single-copy transgenic animals are isolated. L and R homologous regions = 1.4 kb.

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 is restricted to the coelomocytes (arrows). Remaining visible fluorescence is non-specific gut granule fluorescence. Image is of an adult hermaphrodite, anterior is left. DIC above, fluorescence image below. (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 (1.8kb) from all MosSCI insertion strains

(EG4441-EG4450) confirm insertions are at the targeted locus. We interpret 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 bridging¹³. (d) Southern analysis confirms 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 non-fluorescent (indicated by asterisks) and the Southern blot demonstrated that these transgenes contained short deletions and insertions. The molecular changes in EG4441 and EG4448 were further characterized (Supplementary Fig. 3 and Supplementary Fig. 4). Strain EG4449 was visibly more fluorescent and contains two copies of GFP.

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 and by biolistic bombardment. *unc-18::mCherry* was inserted by MosSCI (left) and by biolistic bombardment (right). Confocal images were taken with identical settings to compare fluorescence intensity. Expression is specific to the nervous system but brighter in the biolistic strain UZ566. (c) Transgene copy number is uniform in fluorescent MosSCI lines. Southern blot hybridized with mCherry specific probe shows 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 non-fluorescent and is marked by an asterisk. (d) Western blot with UNC-18 polyclonal antibody shows 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 an UNC-18 antibody. 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 with variable, higher expression from biolistic strains. *unc-18::mCherry* degradation bands are marked by asterisks.

Figure 4.

MosSCI inserts are expressed in the female and male germlines. (a) A *Ppie-1::GFP::histone* transgene is expressed throughout the female germline. Scale bar = 100 μ m. (b) A *Pspe-11::mCherry::histone* transgene is expressed in hermaphrodite sperm. Left, mCherry expression from *Pspe-11::mCherry::histone* is specific to hermaphrodite sperm. Right, overlay with DIC image. Most sperm are found in the spermatheca, although a few can also be seen in the gonad (left) and uterus (right). Scale bar = 10 μ m. Right inset, mCherry expression is localized to nucleus. 4x magnification.

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 the P0 animals by co-injection with germline specific *Pglh-2::transposase* and in the absence of *twk-18(gf)* negative selection marker. (b) Transgene insertion at target site is verified by PCR. A PCR product of the predicted size is obtained in all strains generated by direct injection. (c) A Southern blot confirms single-copy insertion at the target site. Genomic DNA digested with *EcoRI* was probed with a GFP specific probe and confirms that seven of eight strains contain the predicted single-copy insert. EG4893 contains two GFP fragments and GFP fluorescence was visibly brighter.

Table 1
Insertion of the coelomocyte-specific *Punc-122::GFP* transgene
Insertion of coelomocyte-specific *Punc-122::GFP* transgene at *ttf5605*. (Left panel) 5 independent transgenic strains carrying extrachromosomal arrays were generated and the frequency of insertions was determined after heat-shock. (Right panel) A total of ten MosSCI inserts were recovered and listed adjacent to their parent strain in the left panel. Strains were scored for rescue of locomotion and GFP fluorescence. PCR was used to verify insertion at the *ttf5605* target site and to test for the presence of Mos1 transposon elements after insertion (See Supplementary Methods). Southern blot was used to verify the integrity of the *Punc-122::GFP* transgene insert.

Extragenic strains		MosSCI strains			GFP expression	PCR <i>ttf5605</i> site	Transgene insert	Mos1 element in strain?
Parent strain	Insertion frequency	Strain	Phenotype					
EG4380	2/200	EG4441 EG4442	wild-type weak Unc	No No	Yes Yes	1.8 kb deletion N.D.	N.D. No	
EG4381	1/240	EG4443	wild-type	Yes	Yes	Full	No	
EG4382	2/240	EG4444 EG4445	wild-type wild-type	Yes Yes	Yes Yes	Full Full	No N.D.	
EG4383	2/60	EG4446 EG4447	wild-type wild-type	Yes Yes	Yes Yes	Full Full	N.D. No	
EG4385	3/260	EG4448 EG4449 EG4450	wild-type wild-type wild-type	No Yes (brighter) Yes	Yes Yes Yes	0.8 kb deletion, 0.7 kb insertion Tandem? Full	No No No	
Average	1/100		90 % wild-type	70% expression	100% correct	60 % correct	0% Mos1 element	

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

Insertion of germline-specific *Ppie-1::GFP::histone* transgene at *tTi5605*. 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 after 3-4 generations at 25°C on a dissection microscope.

Parent strain	Transmission	Insertion frequency	Germ line fluorescence?
EG4855	69%	12/800 = 1.50% 5/500 = 1.00 %	4/12 0/5
EG4856	38%	4/400 = 1.00 % 1/500 = 0.20%	1/4 1/1
EG4857	44%	5/360 = 1.39% 10/500 = 2.00%	1/5 3/10
Average	50%	37/3060 = 1.2%	10/37 = 27%

Direct inserts generated by injection

Germline-expressed transposase is more effective at generating direct insertions. Injected P0 animals were individually placed on a small NGM plate, F1 progeny was scored and *unc-119* rescued animals counted. The F2 progeny was scored for direct insertion events based on rescue and the absence of co-injection 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*) (*Pglh-2* vs *Phsp-16-48*, $p = 0.015$, two-sided Fischer's exact test). The insertion frequency was calculated as the fraction of plates containing rescued F1 progeny that resulted in an insert.

Construct	Mos T'ase source	Negative selection?	RNAi (<i>wk-18</i>)	Number injected	P0s w/resc'd F1s	Avg. resc'd F1s / injected P0	Direct integrants	Insertion freq (Integr/ P0)	GFP?	Mos1 present?
<i>Ppie-1::GFP</i>	<i>Phsp-16-48</i>	y	y	20	N.D.	N.D.	1	N.D.	100%	N.D.
		y	y	27	17	N.D.	3	18% (3/17)	N.D.	2 of 3
<i>Pspe-11::GFP</i>	<i>Phsp-16-48</i>	n	y	100	63	10.6	11	17% (11/63)	N.D.	6 of 10
		n	y	137	91	15.8	8	8% (8/91)	100% (5/5)	1 of 2
<i>Punc-122::GFP</i>	<i>Phsp-16-48</i>	n	n	110	61	13.2	5	7% (5/61)	100% (2/2)	N.D.
			y	102	63	13.2	13	19% (13/63)	83% (10/12)	N.D.
			n	130	75	13.0	13	17% (13/75)	100% (4/4)	N.D.