

# Targeted gene alteration in *Caenorhabditis elegans* by gene conversion

Peter L Barrett, John T Fleming & Verena Göbel

Now that some genomes have been completely sequenced, the ability to direct specific mutations into genomes is particularly desirable. Here we present a method to create mutations in the *Caenorhabditis elegans* genome efficiently through transgene-directed, transposon-mediated gene conversion. Engineered deletions targeted into two genes show that the frequency of obtaining the desired mutation was higher using this approach than using standard transposon insertion-deletion approaches. We also targeted an engineered green fluorescent protein insertion-replacement cassette to one of these genes, thereby confirming that custom alleles of different types can be created *in vitro* to make the corresponding mutations *in vivo*. This approach should also be applicable to heterologous transposons in *C. elegans* and other organisms, including vertebrates.

Sequence analysis has predicted that the *C. elegans* genome contains ~20,800 genes, but mutated alleles have been reported for fewer than 1,000 cloned genes (WormBase; see URLs in Methods). Whereas forward genetics has proved fruitful in identifying genes of functional relevance, reverse genetics in *C. elegans* has been hampered by a practical inability to target specific mutations to a selected gene. Homologous recombination, which is widely used in other systems, seems to occur only rarely in *C. elegans*; it also requires the screening of many worms and special transformation procedures<sup>1–3</sup> and is not yet used routinely.

At present, two methods are commonly used for isolating mutations in specific genes; one uses transposon insertions<sup>4</sup>, and the other uses chemical- or radiation-based mutagenesis<sup>5,6</sup>. Both methods use PCR for the gene-specific detection of deletions. With either technique, the low frequency of random events occurring for any given gene requires that many worms be tested. Both methods are also limited to isolating only one type of mutation: namely, random deletions of a size detectable by PCR that occur between two predetermined primers. The location and size of a deletion can be controlled only imprecisely by the selection of these primers.

Transposon-based approaches take advantage of mobile genetic elements that are present in the genomes of most animals. In *C. elegans*, at least eight distinct transposons (Tc1–Tc8) have been

identified<sup>7–9</sup>. Mutator strains have been isolated that have a transposition efficiency ~400 times higher than that of wild-type strains<sup>10</sup>. From these strains, several groups have generated and then identified randomly inserted transposon alleles<sup>11–13</sup>. An increase in these efforts could establish a *C. elegans* genome-wide transposon library. Such a project has been reported in *Drosophila melanogaster*, for which >50% of the genome has already been covered, and improved transposon-based techniques have been developed to address problems such as the nonrandom insertion bias of the transposons<sup>14,15</sup>.

The gene-specific insertion of a transposon, however, neither guarantees nor even generally confers a null phenotype<sup>16</sup>. It is therefore necessary to generate a second library from the line containing a transposon insertion in the target gene to select further for spontaneous excision events that delete a portion of flanking DNA and that may thus confer a null phenotype<sup>17</sup>. Such events may occur only rarely or not at all, thereby requiring the testing of many animals.

Here we show that a transgene-directed gene conversion approach not only occurs at a frequency high enough for routine use but also can be used to create custom alleles such as defined deletions, insertions and replacements of different sizes.

## RESULTS

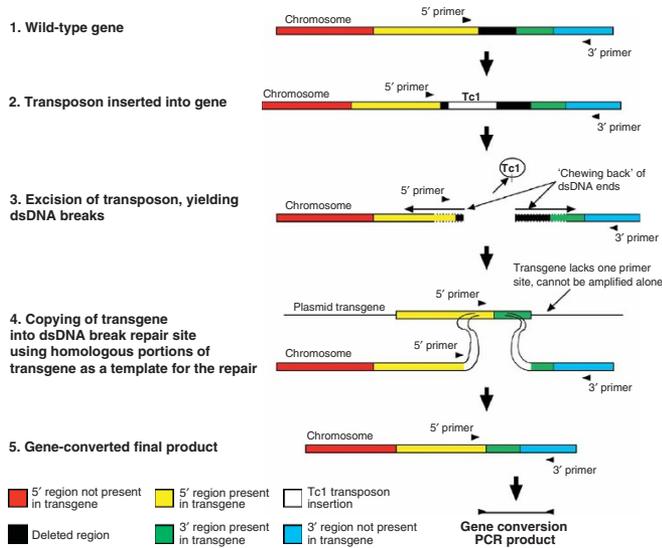
### Strategy for targeted gene alteration by gene conversion

After failed attempts to isolate a deletion of a Tc1-tagged gene by the standard transposon insertion-deletion technique, we investigated gene conversion as a way to create an engineered deletion in the gene. Transposon excision generates double-stranded DNA (dsDNA) breaks, which are thought to be repaired in a template-directed manner by means of the sister strand; spontaneous deletions seem to result from interrupted dsDNA break repair<sup>18</sup>. A transgene can also act as a template for repair after the excision of a Tc1 transposon in *C. elegans*. This type of gene conversion was demonstrated by the introduction of *in vitro*-generated point mutations from a transgene into the genome<sup>19</sup>. The frequency of such an event was found to be very low, however, and use of this technique for gene targeting has not been subsequently reported.

Our general strategy for targeted gene alteration by gene conversion is shown in **Figure 1**. First, we constructed a plasmid containing an engineered deletion of a specific size in the genomic DNA

Laboratory of Developmental Immunology, Massachusetts General Hospital, and Department of Pediatrics, Harvard Medical School, Boston, Massachusetts 02114, USA. Correspondence should be addressed to V.G.

Published online 24 October 2004; doi:10.1038/ng1459



**Figure 1** General strategy for targeted alteration by gene conversion. Homologous regions are shown in the same color. The Tc1 transposon is shown in white, and the sequences deleted in the transgene relative to the chromosomal regions are shown in black. Only one chromosome is shown. Figure is not drawn to scale.

corresponding to the area of the Tc1 insertion site. One endpoint of the deletion was placed near the Tc1 insertion site to facilitate the homology search at the DNA ends during dsDNA break repair<sup>20</sup>. We introduced the plasmid as an extrachromosomal transgene<sup>2</sup> into a mutator strain containing the gene-specific Tc1 insertion. This strain was then expanded into independent populations. We tested a fraction of each population by using a PCR primer pair spanning the deletion, in which one of the two primers was designed to amplify only from the genomic sequence and not from the transgene. Using this

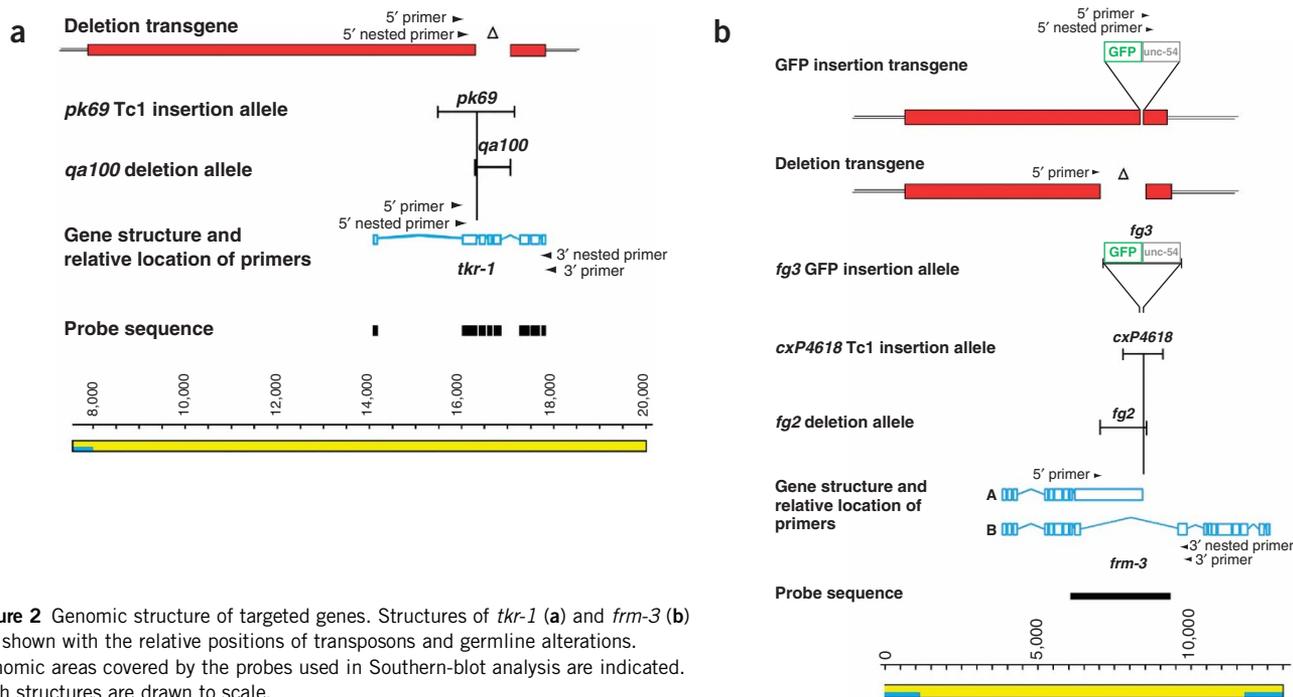
approach, a band of a specific size would be amplified only if the engineered deletion were present at the genomic site. Populations positive for this band were subdivided and tested (sibling selection) until we isolated a single positive worm.

**Isolation of targeted alleles**

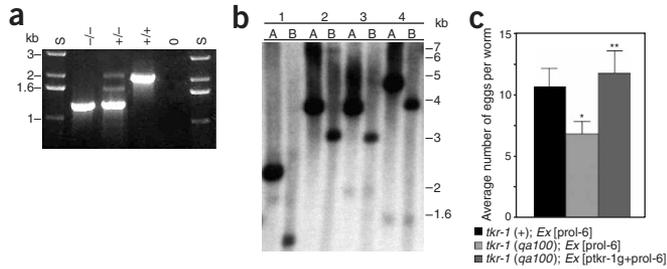
We targeted two different genes containing Tc1 transposons to test the feasibility of this method. One of these genes, *tkr-1*, encodes a tachykinin-like neuropeptide receptor<sup>21</sup> and the other, *frm-3* (P.L.B., J.T.F. and V.G., unpublished data), encodes a member of the FERM domain family of proteins. The conversion plasmid for *tkr-1* contained a 0.85-kb deletion (Fig. 2a), whereas that for *frm-3* contained a 1.5-kb deletion (Fig. 2b). We introduced these plasmids into mutator strains containing their respective Tc1 alleles by microinjection, together with the transformation markers *rol-6d* (ref. 2) or *sur-5-GFP* (ref. 22), which confer, respectively, rolling or green fluorescent protein (GFP) phenotypes. We tested *tkr-1* in the *mut-2* mutator background and *frm-3* in both the *mut-2* and the *mut-7* backgrounds.

We allowed the transgenic strains to expand from 5–10 parent worms to ~500–1,000 worms per population. We sampled a representative portion of each population (approximately one-third) and isolated the genomic DNA, allowing siblings of each population to survive for later isolation. We used gene-specific primer pairs designed to amplify only the gene-converted product and not the transgene (Figs. 2, 3a and 4a) to identify populations containing gene-converted worms by PCR. It was essential in this approach to use primers far enough apart to avoid PCR ‘jumping’ (product generation from annealing of corresponding sequences derived from the individual transgenic and wild-type templates)<sup>23</sup>.

The product amplified from a deletion event will be smaller than the wild-type product and will be a unique size predicted from the engineered deletion; thus, the correct band to pursue is known beforehand. The wild-type product is also usually detected as a result of somatic transposon excision. We designed *tkr-1* primers to generate a predicted product of 1.25 kb from gene conversion of the 0.85-kb



**Figure 2** Genomic structure of targeted genes. Structures of *tkr-1* (a) and *frm-3* (b) are shown with the relative positions of transposons and germline alterations. Genomic areas covered by the probes used in Southern-blot analysis are indicated. Both structures are drawn to scale.



**Figure 3** Gene conversion of a deletion in *tkr-1*. (a) PCR gel showing typical genotyping of the *tkr-1* (*qa100*) deletion allele. Wild-type band, 2.1 kb; deletion band, 1.25 kb; S, DNA standards. Results are shown for homozygous ( $-/-$ ) and heterozygous ( $-/+$ ) *qa100* deletion populations, wild-type ( $+/+$ ) populations and a no-worm negative control (O). (b) Southern hybridization analysis of N2 (A) and *tkr-1* (*qa100*) (B). The probe was the full-length *tkr-1* cDNA (see Fig. 2a and Supplementary Fig. 2 online for expected fragment sizes). (c) Egg-laying assay. Strains wild-type for *tkr-1*, mutant for *tkr-1* without a rescuing plasmid, and mutant for *tkr-1* with a rescuing plasmid were tested for egg-laying behavior as described<sup>37</sup>; all strains were tested in a *rol-6d* background. Error bars indicate the s.e.m. \* $P = 0.0003$  versus *tkr-1 (+)*,  $n = 39/24$ , and \*\* $P < 0.0001$  versus *tkr-1 (qa100)*,  $n = 24/24$ , by a Bonferroni-Dunn multiple comparison test.

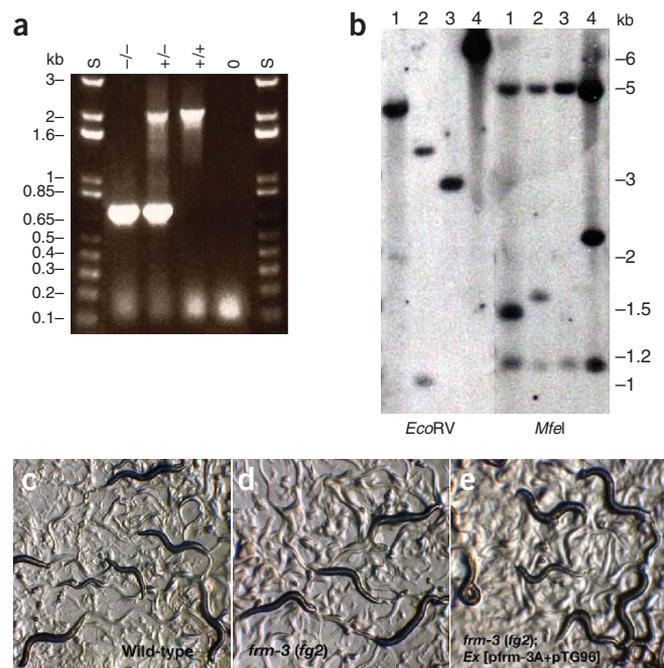
deletion, and a product of 2.1 kb from the wild-type allele, corresponding to a 3.7-kb band in the initial strain containing the Tc1 (1.6 kb). Similarly, we designed *frm-3* primers to amplify a predicted product of 1.5 kb from gene conversion of the 1.5-kb deletion and a product of  $\sim 3$  kb from the wild-type allele, corresponding to a 4.6-kb product when the Tc1 is inserted.

In a pilot study for *tkr-1*, 2 of only 16 populations tested (starting from a total of 80 parental worms) showed a PCR product of the predicted size (Supplementary Fig. 1 online). We assayed approximately one-third of the  $\sim 1,000$  worms per plate, indicating that gene conversion had occurred at a frequency of at least 2 in 5,333 or 1 in 2,666 worms. This number is considerably higher than the rate of  $\sim 1$  in 100,000 previously reported for point mutations<sup>19</sup>. Because the populations were sampled and tested only by PCR and were not phenotypically evaluated, the actual conversion rate may be even higher than this. A PCR product indicative of gene conversion was also amplified from 2 of 45 *frm-3* populations (Table 1 and data not shown). Positive populations were sibling-selected and tested by PCR for the presence of the deletion until single worms containing the deletions were obtained. These deletion alleles were designated *tkr-1* (*qa100*) and *frm-3* (*fg2*).

Sequence analysis of both alleles confirmed that the precise deletion engineered into the plasmids was present at the genomic loci (data not shown). This was also confirmed by Southern hybridization analyses, which produced the expected bands for both alleles and showed no unexpected rearrangements in the vicinity of the target genes (Figs. 3b and 4b and Supplementary Figs. 2 and 3 online). Because the probes included the deleted regions in both cases, this also showed that the deleted sequences were not present elsewhere in the genome. The absence of the deleted regions in both alleles was further verified by the negative PCR results obtained with primers matching sequences inside the deleted region (data not shown). Finally, the Southern-blot analysis also ruled out the possibility of random array integration and

excluded the presence of transgenic DNA, which would have appeared as strongly hybridizing, anomalously migrating bands<sup>24</sup>. Absence of the transgenes was also confirmed by loss of the roller or GFP phenotypes and an inability to amplify transgene vector sequences from the strains (data not shown).

To verify that we had generated gene-specific phenotypes, we rescued both of these gene-converted deletion alleles. The *tkr-1* (*qa100*) deletion removed several transmembrane domains of *tkr-1* and was therefore expected to be a null allele. It conferred a constitutive egg-laying (Egl-c) phenotype that could be rescued by reintroducing the wild-type gene as a transgene<sup>21</sup> (Fig. 3c). The *frm-3* (*fg2*) allele, which was a deletion targeting exon 9 of *frm-3a* (Fig. 2b), conferred a distinctive uncoordinated (Unc) phenotype: compared with wild-type worms (Fig. 4c), *frm-3* (*fg3*) worms were limp (Fig. 4d) and, though capable of movement, moved very little spontaneously (compare Supplementary Videos 1 and 2 online); they also foraged abnormally, forming ‘patches’ of cleared bacteria on a bacterial lawn (Fig. 4d); and their movement in liquid (thrashing) was slightly decreased (Table 2). Reintroduction of the wild-type *frm-3a* genomic sequence on a plasmid into a *frm-3* (*fg2*) strain rescued both the Unc (Fig. 4e and Supplementary Video 3 online) and thrashing phenotypes (Table 2).



**Figure 4** Gene conversion of a deletion in *frm-3*. (a) PCR gel showing typical genotyping of the *frm-3* (*fg2*) deletion allele. Wild-type band, 2.2 kb; deletion band, 0.7 kb; S, DNA standards; primers differ from those used in the original screen. Results are shown for homozygous ( $-/-$ ) and heterozygous ( $-/+$ ) *fg2* deletion populations, wild-type ( $+/+$ ) populations and a no-worm negative control (O). (b) Southern hybridization analysis of N2, *frm-3* (*cxP4618*), *frm-3* (*fg2*) and *frm-3* (*fg3*) (lanes 1–4, respectively; see Supplementary Fig. 3 online for expected fragment sizes). The probe used was a *frm-3* genomic fragment (see Fig. 2b). (c) Wild-type (N2) worms. Note typical sinusoidal S-shaped body posture. (d) *frm-3* (*fg2*) worms. Note limp posture with flattening of body sinusoid. (e) *frm-3* (*fg2*) worms carrying plasmids expressing wild-type *frm-3a* and *sur-5-GFP* as a nonintegrated transgenic array. Note restoration of the sinusoidal body posture.

**Table 1** Frequency of gene conversion events in different mutator backgrounds

Trial number	Gene	Insertion or deletion size (kb)	Strain derivation	Strain viability	Mutator gene and allele	Initial strain genotype	Number of conversion events <sup>a</sup>	Minimum number of positive worms <sup>b</sup>
1	<i>tkr-1</i>	0.85 (deletion)	Crossed	Medium	<i>mut-2 (r459)</i>	<i>mut-2 (r459)</i> I; <i>tkr-1 (pk69)</i> III; <i>qaEx [pTA#6ΔSB+pRF4]</i>	2/16	1/2,666
2	<i>frm-3</i>	1.5 (deletion)	Injected	Medium	<i>mut-7 (pk204)</i>	<i>mut-7 (pk204)</i> <i>dpy-18 (e364)</i> III; <i>frm-3 (cxP4618)</i> X; <i>vjEx [pfrm-3ΔpMfe + pTG96]</i>	2/45	1/7,500
3	<i>frm-3</i>	1.5 (deletion)	Crossed	Good	<i>mut-7 (pk204)</i>	<i>mut-7 (pk204)</i> <i>dpy-18 (e364)</i> III; <i>frm-3 (cxP4618)</i> X; <i>vjEx [pfrm-3ΔpMfe + pTG96]</i>	1/45	1/15,000
4	<i>frm-3</i>	1.5 (deletion)	Injected	Poor	<i>mut-7 (pk204)</i>	<i>mut-7 (pk204)</i> <i>dpy-18 (e364)</i> III; <i>frm-3 (cxP4618)</i> X; <i>vjEx [pfrm-3ΔpMfe + pTG96]</i>	3/45	1/5,000
5	<i>frm-3</i>	1.5 (deletion)	Crossed	Good	<i>mut-2 (r459)</i>	<i>dpy-5 (e61)</i> <i>mut-2 (r459)</i> I; <i>frm-3 (cxP4618)</i> X; <i>vjEx [pfrm-3ΔpMfe + pTG96]</i>	6/45	1/2,500
6	<i>frm-3</i>	1.8 (insertion), 0.025 (deletion)	Crossed	Good	<i>mut-7 (pk204)</i>	<i>mut-7 (pk204)</i> <i>dpy-18 (e364)</i> III; <i>frm-3 (cxP4618)</i> X; <i>vjEx [pfrm-3GFP+3'FL + pTG96]</i>	3/94	1/10,444

<sup>a</sup>Indicates the number of conversion events per population. <sup>b</sup>Indicates the calculated approximate total minimum number of positive worms, determined by assuming that each plate population carried ~1,000 worms and that roughly one-third of these were sampled by PCR. The total number of worms examined was determined by multiplying the number of plate populations by the number of worms that were sampled per plate.

The detailed phenotypes of these two gene conversion alleles will be described elsewhere (P.L.B., J.T.F. and V.G., unpublished data).

Taken together, these data show that engineered deletions can be targeted to transposon-tagged *C. elegans* genes by using transgenes as the templates for dsDNA break repair. They also show that transgene-directed gene conversion occurs at a sufficiently high frequency to be of practical use for targeting any gene.

**Frequencies in different genetic backgrounds**

We compared three independently derived transgenic *mut-7* strains carrying the *frm-3* Tc1 and the conversion plasmid as an extrachromosomal array (Table 1). These strains differed in viability, transgene copy number and transmission rate. To test whether any of these variables affected gene conversion rates, we assayed 45 populations of each strain for conversion frequency and obtained similarly high results: 1–3 of 45 populations contained the expected PCR product, indicative of gene conversion. This finding suggests that the health of the strains, the resulting number of generations needed to populate the plate and the properties of the array are not essential for obtaining gene conversion.

To compare the efficiency of *mut-2* and *mut-7* in the gene conversion process, we tested *frm-3* in both backgrounds. *frm-3* gene conversion occurred at roughly three times the frequency in *mut-2*

as in *mut-7* (Table 1). Comparison of the *tkr-1* and *frm-3* deletions in the same mutant background (*mut-2*) showed that the larger 1.5-kb deletion of *frm-3* was converted almost as well as the smaller 0.85-kb deletion of *tkr-1* (Table 1).

**Custom alleles can be efficiently gene converted**

To extend the utility of this approach further, we generated an insertion-replacement allele (hereafter referred to as an insertion allele) in *frm-3*. We created a GFP translational fusion construct by ligating the GFP sequence to the coding sequence directly preceding the *frm-3a* stop (Fig. 2b). The GFP insertion also contained some artificial (engineered) and 3' untranslated region (UTR) sequences originally derived from the gene *unc-54* (ref. 25), comprising ~1.8 kb of exogenous sequence in total, and was designed to remove ~25 bp of endogenous *frm-3a* sequence in the 3' UTR relative to the wild-type sequence (Fig. 2b).

We introduced this plasmid, along with the transformation markers *rol-6d* and *sur-5-GFP*, into a *mut-7* strain containing the *frm-3* Tc1 allele by microinjection, as described above. This strain was expanded and analyzed by PCR as above, by using 5' *unc-54* primers for the insertion and the same 3' *frm-3* primers, which matched sequences outside the transgene, that were used to detect the *frm-3* deletion allele (Figs. 2b and 5a). As above, a PCR product would be obtained only if the insertion allele were present at the corresponding genomic site. From the DNA analysis of 94 populations, we detected 3 positive populations (Table 1). One of these was followed by sibling selection until we isolated a single positive worm containing the engineered GFP insertion allele; this allele was designated *frm-3 (fg3)*.

In this strategy, the conversion event was identified by PCR; we did not use *frm-3-GFP* itself for screening. The transformation markers *rol-6d* and *sur-5-GFP* were used to distinguish a site-specific conversion from integration events. The desired conversion worm would lose *sur-5-GFP* and *rol-6d* from loss of the transgene but would retain the *frm-3* expression pattern as a result of gene conversion at the locus. As expected, *frm-3 (fg3)* showed the typical expression pattern of *frm-3* but had lost the bright pan-nuclear *sur-5-GFP* expression (and the rolling phenotype) of the parental strain (Fig. 5b,c). The absence of the transgene was also confirmed by an inability to amplify the

**Table 2** *frm-3* thrashing assay

Strain	Genotype	N	Mean number of thrashes per min	s.e.m.
N2	Wild-type	16	209.69	3.66
VJ272	<i>frm-3 (fg2)</i> X	16	189.00*	5.67
VJ283	<i>frm-3 (fg2)</i> X; <i>vjEx200</i> [pfrm-3A + pTG96]	8	226.88**	6.01

Strains N2, VJ272 (*frm-3 (fg2)* X) and VJ283 (*frm-3 (fg2)* X; *vjEx200* [pfrm-3A + pTG96]) were tested for thrashing behavior as described<sup>36</sup>. Thrashing was measured on a Macintosh computer running the Etho program written by J. Thomas. Statistical analysis was done by a Tukey-Kramer multiple comparison test (\**P* < 0.01 versus N2; \*\**P* < 0.001 versus VJ272).



transgene vector sequence from the strain (data not shown) and by Southern-blot analysis with gene-specific and GFP probes, which confirmed the structure of the *frm-3* (*fg3*) allele (Fig. 4b and Supplementary Figs. 3 and 4 online). Use of a GFP probe detected a unique band at the expected size and again established that a singular gene insertion event had occurred (Supplementary Fig. 4 online). This finding showed that the GFP sequence was inserted as a single copy only in the *frm-3* genomic site and not elsewhere in the genome.

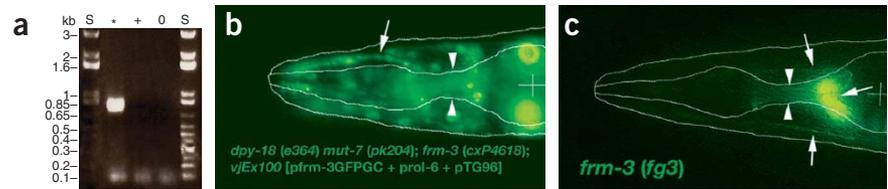
The *fg3* allele expressed GFP primarily in a triplet of circumpharyngeal cells, which send forward processes into the head (Fig. 5b,c), and in the isthmus of the pharynx. This expression pattern is consistent with the expression of *frm-3*, as assessed in a previous transgenic analysis (data not shown), and is consistent with *in situ* RNA hybridization data for *frm-3* (see URLs in Methods). Detailed analysis of the *frm-3* expression pattern and the *fg3* allele will be described elsewhere (P.L.B., J.T.F. and V.G., unpublished data).

## DISCUSSION

Transgene-dependent gene conversion has not been routinely used as an approach to site-specific *in vivo* mutagenesis in *C. elegans*, presumably because of the low frequencies previously observed for the conversion of point mutations<sup>19</sup>. We found, however, that deletions and insertions can be gene converted at relatively high frequencies. There are several differences between our study and the previous work<sup>19</sup>.

First, rather than being assayed by PCR, worms were previously inspected visually for reversion of the *unc-22* twitching phenotype by a wild-type transgene and then analyzed for conversion at the *unc-22* locus. Second, the previous study used a *mut-6* background, which has been reported to have a transposition rate that is 10–100 times lower than that of *mut-2* (refs. 19,26). We found higher frequencies of gene conversion events in *mut-2* than in *mut-7* (Table 1), suggesting that the choice of mutator is important. Third, the total amount of homologous flanking sequence in the conversion template was larger in our study (9.0 kb for *tkr-1* and 7.2 kb for *frm-3* versus 3 kb for *unc-22*), and this difference was accentuated on one side of the deletion (8.3 kb and 6.4 kb versus 1.5 kb). Last, deletions and insertions might be converted more easily than point mutations. This last possibility seems less likely, considering that, first, evidence from P element-induced gene conversion in *D. melanogaster* suggests that deletions are converted at a lower rate than are point mutations and insertions<sup>27,28</sup>; and second, point mutations in a conversion template should contain more contiguous homologous sequence near the transposon insertion site than should deletions, and the nearer a polymorphism in the homologous sequence is to the strand break site, the more likely it is to be converted<sup>20</sup>.

The high frequencies that we obtained for gene-specific deletions compare favorably with the larger-scale Tc1 insertion-deletion and chemical mutagenesis approaches that are currently in use, as well as with attempts at inducing homologous recombination<sup>3</sup>. All these knockout technologies require the screening of large numbers of worms, one to three orders of magnitude higher than required here. The additional effort required to isolate a transposon in the gene (if not already available) and to engineer a mutation-containing



**Figure 5** Gene conversion of an insertion in *frm-3*. (a) PCR gel showing typical genotyping of the *frm-3* (*fg3*) insertion allele. Insertion band, 0.8 kb; S, DNA standards; primers differ from those used in the original screen. Results are shown for an *fg3* insertion worm (\*), a wild-type worm (+) and a no-worm negative control (0). Wild-type worms are negative because the *frm-3* and *unc-54* primers generate a product only on proper insertion of the *unc-54* 3' UTR sequence adjacent to the endogenous *frm-3* gene. (b) Head of worm coexpressing *frm-3-GFP* and *sur-5-GFP*, used as the parental strain for gene conversion of *frm-3-GFP*. Expression of *frm-3-GFP* can be discerned in the isthmus and processes in the head (arrowheads, arrow) but is largely obscured by the pan-nuclear expression of *sur-5-GFP*. (c) Head of a *frm-3* (*fg3*) adult. Three circumpharyngeal cells express *frm-3-GFP* and send forward processes to the head (arrows); isthmus of the pharynx also faintly expresses *frm-3-GFP* (arrowheads). The outlines of the worm and of the pharynx are shown in b and c.

plasmid (also required for homologous recombination approaches) is offset by the high probability of isolating the engineered mutant allele.

Obtaining transposon-derived deletion alleles by the standard insertion-deletion approach seems to have been difficult, and some researchers have resorted to irradiation to encourage the formation of deletions after transposon excision<sup>29</sup> (M. Labouesse, personal communication). When we tested ~200 populations of the otherwise-identical nontransgenic *mut-2*; *tkr-1* insertion strain by the standard Tc1 excision protocol, we isolated no germline deletions (data not shown). By contrast, we isolated the *tkr-1* deletion allele from 16 populations and, on average, fewer than 25 populations per gene had to be tested to isolate alleles (Table 1). Therefore, the presence of the homologous transgene seems to facilitate the generation of deletion alleles. The transgene may be preferentially used as the conversion template in this setting or it may directly affect the excision frequency. Alternatively, the frequency of gene conversion from the sister chromatid or a transgene might be normally higher than that of interrupted dsDNA break repair after transposition, but only conversion from the transgene is ordinarily detected.

A chief advantage of targeted gene conversion is that it can be used to generate custom alleles, as well as knockouts, for any transposon-tagged gene. We used this approach to generate an engineered GFP insertion allele in *C. elegans*, allowing for the examination of gene expression in single copy number, in its native genomic milieu and under physiological conditions, while probably retaining normal gene activity. Gene expression in *C. elegans* has been typically examined with transgene reporter constructs, which may produce nonphysiological expression data<sup>2</sup>. In a general sense, however, any desired mutation could be created *in vitro* and then converted to the corresponding genetic locus *in vivo*, such as activating, dominant-negative, misexpression and other alleles with point mutations, specific substitutions or similar alterations in their genes. With regard to deletions, our finding that the 1.5-kb and 0.85-kb deletions are converted with roughly equal frequencies suggests that the homology scanning window may require that only one end of the dsDNA break be located and secured, as has been suggested for *D. melanogaster*<sup>28</sup>. To our knowledge, the 1.5-kb deletion is larger than any deletion so far reported in *D. melanogaster*. Therefore, considerably larger deletions might be generated by this approach.

In this study, we inserted the GFP sequence at the 3' end of a gene, which was expected to preserve function. The same approach could be used to generate either a different type of knock-in allele, in which the

gene function is abrogated by insertion of the GFP sequence containing a stop codon at the 5' end of the gene, or a full replacement allele, in which the whole gene is replaced by the reporter. The technique could also be further improved by the use of selectable markers. Here we used PCR to detect a GFP insertion allele, making no prior assumptions of GFP expression. In situations where low copy number expression is sufficiently robust, GFP expression might be used directly for detection. In this approach, expression elements (such as the promoter) would be excluded from the conversion template, such that expression of GFP becomes contingent on site-specific gene conversion. But we have observed that, owing to intragenic gene expression elements, even after removal of the endogenous promoter of a gene, it may be unexpectedly difficult to find any sequence that does not confer substantial expression. Finally, the presence of a transposon at the target site may not be strictly required for gene conversion: an apparent instance of transposon-independent gene conversion has been obtained for the gene *rol-6*, albeit at an extremely low frequency<sup>30</sup>.

An expansion of the available transposon insertion libraries<sup>11–13</sup> could conceivably tag every gene in *C. elegans* with a transposon and function as a basis for the detailed *in vivo* analysis of all genes by mutagenesis, as is currently underway in *D. melanogaster*<sup>14,15</sup>. In addition, intergenic regions that are not accessible by classical genetic analysis could be readily targeted. We compiled and categorized, from all available sources, more than 2,000 endogenous transposon insertion alleles for which the exact insertion site in the *C. elegans* genome is known. In addition, a heterologous transposon insertion system has been introduced in *C. elegans*<sup>31</sup>. The advantage of this system is the absence of a background of endogenous transposons. A total of 849 additional heterologous transposon insertions have been identified (WormBase; see URLs in Methods), and efforts are underway to tag the genome with ~45,000 additional insertions by using the *Mos* transposon (see URLs in Methods). Heterologous transposon insertions will probably have similar or greater difficulties with regard to transposition and the generation of imprecise excision alleles. The gene conversion technique reported here should also be effective in this heterologous system, because the process of gene conversion, as shown by ourselves and others<sup>20</sup>, seems merely to require the presence of dsDNA breaks and the homologous template DNA. Similarly, the technique should work for generating targeted gene alterations in other systems, such as zebrafish and mouse, where heterologous insertion systems are also being developed<sup>32–34</sup>.

## METHODS

**Isolation of Tc1 alleles.** R. Plasterk isolated a strain containing a Tc1 in the gene *tkr-1* (C38C10.1) in a *mut-2* (*r459*) background with primers PTR1 and PTR2, and Tc1 right 1 and right 2 (see **Supplementary Methods** online for all PCR conditions and primer sequences) as described<sup>4</sup>. This allele was backcrossed to N2, generating strain XA101 (*tkr-1* (*pk69*) III). A Tc1 insertion allele for *frm-3* (H05G16.1A/B) was provided by L. Ségalat as a part of the *C. elegans* Transposon Insertion Project<sup>13</sup>. This allele was isolated in a *mut-7* (*pk204*) *dpy-18* (*e364*) background with primers S46/18 and Tc1 left 2. We backcrossed this insertion strain with strain CB3855 (*plg-1* (*e2001*) III; *him-5* (*e1490*) V) to generate strain VJ136 (*frm-3* (*cxP4618*) X).

**Gene conversion plasmids and transgenic mutator worms.** We constructed the *tkr-1* gene conversion plasmid pTA#6ΔSB by amplifying *tkr-1* from positions 7,930 to 17,819 (numbering according to the C38C10 cosmid) by LA-PCR (Takara). We cloned this fragment into vector pCR II (Invitrogen), and the resultant plasmid was digested with enzymes *SexAI* and *BaeI* (provided by H. Kong, New England Biolabs) to remove 0.85 kb of sequence, blunt-ended with Klenow polymerase and religated. This plasmid, along with the marker plasmid pRF4 (*rol-6d*), was introduced as a transgene by microinjection into XA101 as described<sup>2</sup>. Each strain generated in this way was crossed

into the *mut-2* mutator background by using *dpy-5* (*e61*) *unc-29* (*e1072*) *in trans* to *mut-2*. We prepared genomic DNA from populations and single worms as described<sup>17</sup>.

We constructed the *frm-3* gene conversion plasmid pfrm-3ΔpMfe by amplifying *frm-3* from positions 635 to 9,312 (numbering according to the H05G16 cosmid) by LA-PCR. We cloned this fragment into vector pCR4-TOPO (Invitrogen), and the resultant plasmid was partially digested with *MfeI* to remove ~1.5 kb of sequence and then religated. This plasmid was introduced as a transgene into strains VJ136 and VJ151 (*dpy-18* (*e364*) *mut-7* (*pk204*) III; *frm-3* (*cxP4618*) X), along with the marker plasmids pRF4 and pTG96 (a *sur-5*–GFP fusion<sup>22</sup>), by injection<sup>2</sup>. The transgenic VJ136 strain generated in this way was crossed into the mutator backgrounds TW404 (*dpy-5* (*e61*) *mut-2* (*r459*); a gift from Q. Boese and J. Collins) and VJ151. The conversion plasmid was therefore introduced into VJ151 by both injection and crossing to compare the effects of relative strain health and transgene structure on the gene conversion process (over time, viability of mutator strains may decline owing to accumulated mutations).

We constructed the *frm-3*–GFP translational fusion gene conversion plasmid pfrm-3GFP+3'FL by amplifying *frm-3* from positions 635 to 8,398 (numbering according to the H05G16 cosmid) by LA-PCR. We cloned this fragment into the β-gal vector pPD95.07 (all pPD vectors; A. Fire, S. Xu, J. Ahnn and G. Seydoux, personal communication) using *XbaI* and *XmaI* as linkers. The GFP and *unc-54* 3' UTR sequences from vector pPD95.79 were excised with the enzymes *BshI* and *SpeI* and used to replace the equivalent sequences of the pPD95.07/*frm-3* plasmid. To obtain *frm-3* homologous sequence 3' to the GFP and *unc-54* genes, we amplified the *frm-3* sequence from positions 8,422 to 9,162 by using *XbaI* linkers and cloned this fragment in the correct orientation into the pPD95.07/GFP/*frm-3* plasmid via the *SpeI* site. This translational fusion places the GFP sequence in-frame directly preceding the *frm-3a* stop codon. This plasmid was injected into strain VJ179 (*dpy-18* (*e364*) *mut-7* (*pk204*) III; *frm-3* (*cxP4618*) X), along with the marker plasmids pRF4 and pTG96.

**Isolation of gene conversion alleles.** We placed 5–10 transgenic mutator worms onto small agarose plates seeded with OP50 and allowed them to grow to saturation (~500–1,000 worms per plate). We prepared genomic DNA from populations and single worms essentially as described<sup>17</sup>. We washed each plate with ~0.75 ml of distilled deionized water and placed one-third of this in a PCR tube. After snap-freezing on liquid nitrogen, the contents of the tubes were desiccated to dryness (*tkr-1*); alternatively, the worms were allowed to settle at 4 °C for 30 min, the supernatant was removed and the worms were snap-frozen (*frm-3*). We added 50–100 μl of worm lysis buffer containing proteinase K<sup>35</sup>. The tubes were heated at 65 °C for 4 h, and proteinase K was inactivated by heating at 95 °C for 15 min. Typically 2 μl of this crude DNA was added to a 25-μl PCR reaction. Phenol-chloroform extraction and ethanol precipitation were used initially but were found not to be absolutely required before PCR.

After PCR, the reactions were separated by electrophoresis on a 0.8–1% agarose gel and analyzed. Positive results were confirmed by making relaysates of the original plate and duplicating the result, which indicated that the event had occurred in at least two independent worms on the same plate and suggested that a germline event had occurred<sup>17</sup>. Sibling selection was done as described<sup>4</sup> to the single-worm level, and homozygous strains were backcrossed to the wild type.

**Southern hybridization analysis.** See **Supplementary Methods** online.

**Construction of rescue strains.** For *tkr-1*, we amplified the full-length wild-type gene by LA-PCR and cloned it into pTA (Promega) to produce plasmid ptkr-1g. This plasmid was coinjected with pRF4 into an N2 background to yield strain XA118. Strain JF2 contains the pRF4 plasmid alone injected into an N2 background. Strain XA156 (*tkr-1* (*qa100*) III; *Ex* [pRF4]) was generated by crossing the *qa100* allele into JF2, and strain XA157 (*tkr-1* (*qa100*) III; *Ex* [ptkr-1g + pRF4]) was generated by crossing in parallel with XA118; this ensured that the same array was used in all experiments. For *frm-3a*, we amplified the full-length wild-type genomic sequence by LA-PCR and cloned it into pCR4-TOPO (Invitrogen) to produce plasmid pfrm-3A. This plasmid was coinjected with pTG96 into strain VJ272 (*frm-3* (*fg2*) X) to generate strain VJ283 (*frm-3* (*fg2*) X; *vjEx200* [pfrm-3A + pTG96]).

URLs. WormBase release WS120, 3 July 2004, was used (<http://www.wormbase.org/>). Y. Kohara's *in situ* hybridization data are available at <http://nematode.lab.nig.ac.jp/>. Information on efforts to 'transposon-tag' the worm genome with *Mos* is available at NemaGeneTag (<http://elegans.imbb.forth.gr/nemagenetag/>).

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

#### ACKNOWLEDGMENTS

We thank J. Collins, C. Almeida, Q. Boese and the other members of the laboratory of J. Collins for sharing unpublished results and providing the TW404 strain; R. Plasterk for his advice and effort in isolating the *tkr-1* (*pk69*) allele; L. Ségalat for providing the *frm-3* (*cxP4618*) allele; C. Gerard for support during initial stages of the project; E. Schwarz for help with Wormbase; A. Fire and M. Han for plasmids; T. Stiernagle and the *Caenorhabditis* Genetics Center for strains; the Sanger Centre for cosmids; and H. Weinstein for continued support. This work was partially supported by a grant from the US National Institutes of Health to V.G.

#### COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Received 5 April; accepted 27 September 2004

Published online at <http://www.nature.com/naturegenetics/>

1. Broverman, S., MacMorris, M. & Blumenthal, T. Alteration of *Caenorhabditis elegans* gene expression by targeted transformation. *Proc. Natl. Acad. Sci. USA* **90**, 4359–4363 (1993).
2. Mello, C. & Fire, A. DNA transformation. *Methods Cell Biol.* **48**, 451–482 (1995).
3. Berezikov, E., Bargmann, C.I. & Plasterk, R.H. Homologous gene targeting in *Caenorhabditis elegans* by biolistic transformation. *Nucleic Acids Res.* **32**, e40 (2004).
4. Zwaal, R.R., Broeks, A., van Meurs, J., Groenen, J.T. & Plasterk, R.H. Target-selected gene inactivation in *Caenorhabditis elegans* by using a frozen transposon insertion mutant bank. *Proc. Natl. Acad. Sci. USA* **90**, 7431–7435 (1993).
5. Jansen, G., Hazendonk, E., Thijssen, K.L. & Plasterk, R.H. Reverse genetics by chemical mutagenesis in *Caenorhabditis elegans*. *Nat. Genet.* **17**, 119–121 (1997).
6. Liu, L.X. *et al.* High-throughput isolation of *Caenorhabditis elegans* deletion mutants. *Genome Res.* **9**, 859–867 (1999).
7. Plasterk, R.H.A. & van Leunen, H.G.A.M. Transposons. in *C. elegans II* vol. 33 (eds. Riddle, D.L., Blumenthal, T., Meyer, B.J. & Priess, J.R.) 97–116 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1997).
8. Rezsóhazy, R., van Luenen, H.G., Durbin, R.M. & Plasterk, R.H. Tc7, a Tc1-hitch hiking transposon in *Caenorhabditis elegans*. *Nucleic Acids Res.* **25**, 4048–4054 (1997).
9. Le, Q.H., Turcotte, K. & Bureau, T. Tc8 a tourist-like transposon in *Caenorhabditis elegans*. *Genetics* **158**, 1081–1088 (2001).
10. Anderson, P. Mutagenesis. *Methods Cell Biol.* **48**, 31–58 (1995).
11. Korschwagen, H.C., Durbin, R.M., Smits, M.T. & Plasterk, R.H. Transposon Tc1-derived, sequence-tagged sites in *Caenorhabditis elegans* as markers for gene mapping. *Proc. Natl. Acad. Sci. USA* **93**, 14680–14685 (1996).
12. Williams, B.D. Genetic mapping with polymorphic sequence-tagged sites. *Methods Cell Biol.* **48**, 81–96 (1995).
13. Martin, E. *et al.* Identification of 1088 new transposon insertions of *Caenorhabditis elegans*: a pilot study toward large-scale screens. *Genetics* **162**, 521–524 (2002).
14. Parks, A.L. *et al.* Systematic generation of high-resolution deletion coverage of the *Drosophila melanogaster* genome. *Nat. Genet.* **36**, 288–292 (2004).
15. Thibault, S.T. *et al.* A complementary transposon tool kit for *Drosophila melanogaster* using P and piggyBac. *Nat. Genet.* **36**, 283–287 (2004).
16. Rushforth, A.M. & Anderson, P. Splicing removes the *Caenorhabditis elegans* transposon Tc1 from most mutant pre-mRNAs. *Mol. Cell. Biol.* **16**, 422–429 (1996).
17. Plasterk, R.H. Reverse genetics: from gene sequence to mutant worm. *Methods Cell Biol.* **48**, 59–80 (1995).
18. Engels, W.R. P elements in *Drosophila*. *Curr. Top. Microbiol. Immunol.* **204**, 103–123 (1996).
19. 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).
20. 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).
21. Barrett, P.L. Ph.D. Thesis. (Harvard University, 1999).
22. Gu, T., Orita, S. & Han, M. *Caenorhabditis elegans* SUR-5, a novel but conserved protein, negatively regulates LET-60 Ras activity during vulval induction. *Mol. Cell. Biol.* **18**, 4556–4564 (1998).
23. Liu, S., Thaler, D.S. & Libchaber, A. Signal and noise in bridging PCR. *BMC Biotechnol.* **2**, 13 (2002).
24. Stinchcomb, D.T., Shaw, J.E., Carr, S.H. & Hirsh, D. Extrachromosomal DNA transformation of *Caenorhabditis elegans*. *Mol. Cell. Biol.* **5**, 3484–3496 (1985).
25. Fire, A., Harrison, S.W. & Dixon, D. A modular set of lacZ fusion vectors for studying gene expression in *Caenorhabditis elegans*. *Gene* **93**, 189–198 (1990).
26. Collins, J., Saari, B. & Anderson, P. Activation of a transposable element in the germ line but not the soma of *Caenorhabditis elegans*. *Nature* **328**, 726–728 (1987).
27. Johnson-Schlitz, D.M. & Engels, W.R. P-element-induced interallelic gene conversion of insertions and deletions in *Drosophila melanogaster*. *Mol. Cell. Biol.* **13**, 7006–7018 (1993).
28. 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).
29. Sengupta, P., Chou, J.H. & Bargmann, C.I. *odr-10* encodes a seven transmembrane domain olfactory receptor required for responses to the odorant diacetyl. *Cell* **84**, 899–909 (1996).
30. 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).
31. Bessereau, J.L. *et al.* Mobilization of a *Drosophila* transposon in the *Caenorhabditis elegans* germ line. *Nature* **413**, 70–74 (2001).
32. Raz, E., van Luenen, H.G., Schaeffer, B., Plasterk, R.H.A. & Driever, W. Transposition of the nematode *Caenorhabditis elegans* Tc3 element in the zebrafish *Danio rerio*. *Curr. Biol.* **8**, 82–88 (1998).
33. Dupuy, A.J. *et al.* Mammalian germ-line transgenesis by transposition. *Proc. Natl. Acad. Sci. USA* **99**, 4495–4499 (2002).
34. Horie, K. *et al.* Characterization of *Sleeping Beauty* transposition and its application to genetic screening in mice. *Mol. Cell. Biol.* **23**, 9189–9207 (2003).
35. Williams, B.D., Schrank, B., Huynh, C., Shownkeen, R. & Waterston, R.H. A genetic mapping system in *Caenorhabditis elegans* based on polymorphic sequence-tagged sites. *Genetics* **131**, 609–624 (1992).
36. Miller, K.G. *et al.* A genetic selection for *Caenorhabditis elegans* synaptic transmission mutants. *Proc. Natl. Acad. Sci. USA* **93**, 12593–12598 (1996).
37. Koelle, M.R. & Horvitz, H.R. EGL-10 regulates G protein signaling in the *C. elegans* nervous system and shares a conserved domain with many mammalian proteins. *Cell* **84**, 115–125 (1996).