Caenorhabditis elegans reporter fusion genes generated by seamless modification of large genomic DNA clones

Colin T. Dolphin* and Ian A. Hope[1]

Pharmaceutical Science Research Division, King's College London, Franklin-Wilkins Building, 150 Stamford Street, London SE1 9NH, UK and [1]Institute of Integrative and Comparative Biology, Faculty of Biological Sciences, University of Leeds, Leeds LS2 9JT, UK

ABSTRACT

By determining spatial-temporal expression patterns, reporter constructs provide significant insights into gene function. Although additionally providing information on subcellular distribution, translational reporters, where the reporter is fused to the gene coding sequence, are used less frequently than simpler constructs containing only putative promoter sequences. Because these latter constructs may not contain all necessary regulatory elements, resulting expression patterns must be interpreted cautiously. To ensure inclusion of all such elements and provide details of subcellular localization, construction of translational reporters would, preferably, utilize genomic clones, containing the complete locus plus flanking regions and permit seamless insertion of the reporter anywhere within the gene. We have developed such a method based upon λ Red-mediated recombineering coupled to a robust two-step counter-selection protocol. We have inserted either *gfp* or *cfp* precisely at the C-termini of three *Caenorhabditis elegans* target genes, each located within different fosmid clones, and examined previously with conventional reporter approaches. Resulting transgenic lines revealed reporter expression consistent with previously published data for the tagged genes and also provided additional information including subcellular distributions. This simple and straightforward method generates reporters highly likely to recapitulate endogenous gene expression and thus represents an important addition to the functional genomics toolbox.

INTRODUCTION

Functional analysis of the *Caenorhabditis elegans* genome, the first of any metazoan animal to be sequenced (1), has played a major role in helping shape and develop the experimental and informatics tools now available for large-scale functional genomics projects in this and other species. Most notably, the application of RNAi on large, even genome-wide scales (2,3) has revealed gene associations with physiological or developmental processes some with human disease relevance (4,5). Another approach that provides key functional insight is the determination of spatial-temporal patterns of gene expression. Expression patterns can be determined on a large-scale at the transcriptional level in *C.elegans* either with microarrays (6) or via serial analysis of gene expression (SAGE) (7,8). Although microarrays can determine mRNA abundance for many genes simultaneously the assay is restricted to known, characterized genes. In contrast, SAGE has the potential to both quantify mRNA abundance for known genes and identify novel transcripts. Although both have been applied to certain specific cell types (9,10) or developmental stages (7), the techniques lack spatial resolution due, essentially, to the difficulty of isolating cell- or tissue-specific RNA. *In situ* hybridization has also been adapted for relatively large-scale analyses (11) and is better at providing spatial aspects of mRNA distribution, although, for *C.elegans*, sensitivity and resolution remain poor. An alternative to such mRNA-based methods would be to reveal the location of *C.elegans* proteins by immunohistochemistry although such an approach is not amenable to high-throughput, large-scale assay and would also be prohibitively expensive.

A commonly used alternative to either RNA- or protein-focussed approaches for investigating gene expression in *C.elegans* is reporter technology. This approach cannot only provide direct information on the promoter activity of a gene but potentially other aspects, such as subcellular distribution,

---

*To whom correspondence should be addressed. Tel: +44 207 848 4806; Fax: +44 207 848 4800; Email: colin.dolphin@kcl.ac.uk

can be investigated if a translational reporter, in which the reporter is fused, in-frame to the gene of interest, is used. However, more commonly a transcriptional reporter is constructed in which the promoter sequence alone is cloned upstream of the reporter, typically *LacZ* or *gfp*, in an appropriate base plasmid [e.g. Ref. (12)]. If a precise promoter region has not been defined then it is usual to clone all, or part, of the 5′ flanking sequence on the basis that, for many *C.elegans* genes, most of the relevant *cis*-acting regulatory elements are contained within the upstream intergenic sequence (IGS). Recently, a project designed to catalogue *C.elegans* gene expression patterns on a genome-wide scale, the so-called worm promoterome project, based on high-throughput assembly of IGS-containing transcriptional reporters cloned with the Gateway system (13), has been described (14). However, irrespective of the scale or manner of construction, such transcriptional reporters containing a IGS alone lack, by definition, potential regulatory elements positioned more distally beyond the 5′ end of the cloned fragment or located outside this region within, for example, introns or the 3′ flanking or 3′-untranslated region (3′-UTR). In particular, the absence of a 3′-UTR may be of particular significance due to potential post-transcriptional regulation by microRNAs that repress translation of target mRNAs by binding specifically within this region. Consequently, expression patterns generated with such reporters have to be interpreted with caution and include the caveat that, because the cloned sequence may not contain all necessary regulatory elements, the pattern may not represent fully that of the gene under study.

An improved translational-style reporter would contain not only the complete genomic locus under study but also significant stretches of the 5′ and 3′ flanking sequence to preserve higher-order chromatin structural context, that may also influence transcriptional activity, and in case there are more distantly located regulatory elements. Because such a construct would almost certainly contain all associated regulatory elements the resulting expression pattern would be highly likely to recapitulate that of the endogenous gene. Furthermore, it would be desirable to be able to insert the reporter sequence, seamlessly and in-frame, at any specific position within the target gene. Lastly, the construction method should not be overly complicated and be potentially adaptable for high-throughput construct generation. We describe such a method here. Based upon homologous recombination mediated by the bacteriophage λ Red system, commonly referred to as recombineering (15), the method, when coupled with a simple and robust two-step counter-selection protocol, allows straightforward, flexible construction of translational reporters directly from *C.elegans* fosmid clones.

MATERIALS AND METHODS

General molecular methods

Classical restriction enzyme-based genetic engineering techniques, preparation of media, etc. were undertaken as described (16) unless otherwise stated. Solid and liquid cultures of *Escherichia coli* strain EL350 (17) were incubated at 32°C apart from induction of Red functions when the temperature was briefly shifted to 42°C. All PCRs were performed in volumes of 30 µl containing 15 pmol of each oligonucleotide (ODN) primer and 200 µM of each dNTP and were catalysed with the high-fidelity DNA polymerase ExpandHF (Roche Diagnostics) with conditions designed to minimize mis-incorporation. PAGE-purified ODN primers for generating recombineering cassettes by PCR ranged from 72 to 76 nt in length and were from Integrated DNA Technologies (Coralville, USA). ODNs for sequencing and priming non-recombineering PCRs were desalted and from TAGN (Newcastle, UK). Primer design, recombineering strategies and sequence alignments were performed with MacVector or GCG (both from Accelrys).

Fosmid clones, copy-number induction and DNA isolation

C.elegans genomic fosmid clones WRM0640dE04, WRM0614cG08 and WRM0624dB10, constructed in pCC1Fos (Epicentre) and maintained in *E.coli* strain EPI300-T1^R (Epicentre), and containing the respective target genes *F02A9.3* (*far*-2), *F40E10.6* and *F54C9.11* were identified via the web-based searching facility (http://elegans.bcgsc.bc.ca/perl/fosmid/CloneSearch) at the Michael Smith Genome Sciences Centre. Fosmid DNAs were isolated from cultures, including from those that had been treated with CopyControl Induction Solution (Epicentre) to induce clone copy-number, with a FosmidMAX DNA purification kit (Epicentre) according to the manufacturer's instructions. Following induction ∼5 µg of fosmid DNA was obtained per 1.5 ml culture volume.

PCR-amplification of the *rps*L-*tet*A(C) and *gfp*/*cfp* recombineering cassettes

For each gene the corresponding *rps*L-*tet*A(C) recombineering cassette (RT-cassette) was generated in a PCR with, as template, ∼5 ng of a gel-purified 2.7 kb NcoI fragment of pBAC-RT (18), containing the *omp*F promoter and *rps*L-*tet*A(C) coding sequences. PCRs were primed with primers 4282/4283, 4286/4287 or 4290/4291 for target genes *F02A9.3*, *F40E10.6* and *F54C9.11*, respectively (Table 1). Similarly, ∼5 ng of a gel-purified 1.8 kb XbaI–SpeI restriction fragment of either pPD95.77 or pPD136.61 (12) was used as the template to amplify, respectively, *gfp* or *cfp* cassettes minus the termination codons. PCRs were primed with primer pairs 4284/4285, 4288/4289 or 4292/4293 for target genes *F02A9.3*, *F40E10.6* and *F54C9.11*, respectively (Table 1). The first 50 nt at the 5′ end of each recombineering primer comprised the left, or right, homology arm while the remaining 22–26 nt annealed to the PCR template (Table 1). PCRs were incubated for 10 cycles at 95°C for 5 s, 50°C for 30 s and 68°C for 3 min followed by 20 cycles at 95°C for 5 s, 50°C for 30 s and 68°C for 3 min with an additional 5 s extension per cycle. The concentration of each purified PCR product (High Pure PCR Purification Kit, Roche Diagnostics) was determined electrophoretically by visual comparison to a 1 kb mass ladder (NEB).

Preparation of electrocompetent cells and two-step counter-selection recombineering

Recombineering, performed essentially as described (17,19), was carried out as follows. First, aliquots of *E.coli* strain

Table 1. Sequences of recombineering and general PCR/sequencing primers

Fosmid	Target gene	Cassette/flanking[a]	Primer number	Sequence (5′–3′)[b]
WRM0640dE04	F02A9.3	RT	4282	taattcaaacaataatttattatttcagacgaaaagattcaagccatcTCGCTGTCGAGATATGACGGTG
		RT	4283	ttgaaaaccatatcataattttattatttcagacgaaaagattccgacGATGATAAGCTGTCAAACATGAG
		gfp/cfp	4284	taattcaaacaataatttattatttcagacgaaaagattcaagccatcATGAGTAAAGGAGAAGAACTTTTC
		gfp/cfp	4285	ttgaaaaccatatcataattttattatttcttagttgacgtattccgacTTTGTATAGTTCATCCATGCCATGTG
		Flanking	4264	CCAAAACCACAAGGAATACAAG
		Flanking	4265	CCAGTTCTGAATTACGTTGTTGG
WRM0614cG08	F40E10.6	RT	4286	tgggagcggcaccaatgatagctcatcatcagctccggctcaccTCGCTGTCGAGATATGACGGTG
		RT	4287	gagagagagaggtgtacgaataaataactatgcgtcgagaagttgccGATGATAAGCTGTCAAACATGAG
		gfp/cfp	4288	tgggagcggcaccaatgatagctcatcatcagctccggctcaagccaccATGAGTAAAGGAGAAGAACTTTTC
		gfp/cfp	4289	gagagagagaggtgtacgaataaataactatgcgtcgagaggttgccTTTGTATAGTTCATCCATGCCATGTG
		Flanking	4268	TTACTTTCAGCACTCCAACGCTC
		Flanking	4269	ATAAGTATATTGGACAGGTGAATGC
WRM0624dB10	F54C9.11	RT	4290	cgttaattgccaggaataggtgagttattccactttctcttggtctcggaTCGCTGTCGAGATATGACGGTG
		RT	4291	ctttaaagcgcgcacctttggcagtctcaattaaaattcatcgtGATGATAAGCTGTCAAACATGAG
		gfp/cfp	4292	cgttaattgccaggaataggtgagttattccactttctcttggtctcggaATGAGTAAAGGAGAAGAACTTTTC
		gfp/cfp	4293	ctttaaagcgcgcacctttggcagtctcaattatcaatcgtTTTGTATAGTTCATCCATGCCATGTG
		Flanking	4272	CAAATGCAGTTGAAGAGATCGATCC
		Flanking	4273	ATGTTGGACGCGGGACAAATCC
		Flanking	4237	ATGAGTAAAGGAGAAGAACTTTTC
		Flanking	4238	TTTGTATAGTTCATCCATGCCATGTG

[a] RT = *rpsL-tetA*(C) counter-selection cassette; *gfp/cfp* = GFP or CFP reporter cassette; flanking = primer flanks cassette insertion site.
[b] Homology arm and PCR template annealing sequences given in lower and upper case, respectively.

EL350 (17) were transformed, by electroporation, with DNA of each of the three fosmid clones and glycerol stocks prepared. For each clone, 1 ml of an overnight mini-culture (SOB[-Mg] plus chloramphenicol [Cm, 12.5 μg/ml], 32°C, 220 r.p.m.), seeded with a single colony from a plate streaked with the appropriate EL350 transformant, was used to inoculate 100 ml of the same media plus Cm (12.5 μg/ml) in a 500 ml flask and incubated (32°C, 220 r.p.m.) until the OD_{600} was ∼0.6. At this point 50 ml was transferred to a sterile, pre-warmed (42°C) 500 ml flask and, to induce Red functions, incubated at 42°C for a further 20 min in a gently shaking water bath (New Brunswick Scientific Innova 3100; 100 r.p.m.) after which both cultures were chilled on ice (15 min). Electrocompetent bacteria were prepared from each culture by twice pelleting and washing the cells in ice-cold 10% glycerol (50 ml). After the final wash all but ∼500 μl of each supernatant was aspirated, the bacterial pellets gently resuspended and aliquots (100 μl) either stored at −80°C or used immediately for the first recombineering step as follows.

Aliquots (100 μl) of the non-induced control or induced electrocompetent bacterial cells were electroporated (Eppendorf electroporator 2510, 2.3 kv, 0.2 mm cuvette) with the appropriate purified PCR-generated RT-cassette (100–150 ng) and cells allowed to recover by addition of 1 ml of SOB[-Mg] and incubation (32°C, 220 r.p.m.) in a 25 ml tube for 2–3 h. Recovered cells were serially diluted in M9 salts and aliquots (50 μl) of either undiluted and 10^{-1} or undiluted, 10^{-1} and 10^{-2} dilutions, for non-induced and induced cells, respectively, spread on Luria–Bertani (LB)/agar plates containing tetracycline (Tc; 5 μg/ml). Aliquots (50 μl) of 10^{-4} and 10^{-5} dilutions of non-induced and induced cells were each spread on LB/agar plates without selection for determination of total viable cell numbers and recombineering efficiency. After 36–48 h incubation (32°C) colony numbers were recorded for each plate. Six single discrete colonies from one of the diluted induced plates were each re-streaked onto a fresh LB/agar plate containing Tc (5 μg/ml) and incubated (32°C) for a further 24–36 h. A single colony from each plate was then expanded by overnight mini-culture (LB plus Tc [5 μg/ml], 32°C, 220 r.p.m.) and glycerol stocks made. Fosmid DNAs, isolated from each of these mini-cultures, were used as templates in PCRs with flanking primer pairs 4264/4265, 4268/4269 or 4272/4273 for target genes *F02A9.3*, *F40E10.6* and *F54C9.11*, respectively, (Table 1), to confirm insertion of the RT-cassette. PCR conditions were as above.

In the second recombineering step, electrocompetent bacteria prepared, as above, from both non-induced and Red function-induced cultures (SOB[-Mg] plus Tc [5 μg/ml], 32°C, 220 r.p.m.) of the appropriate RT-containing recombinant line, were electroporated with the corresponding purified PCR-generated cassette (100–150 ng) encoding either the *gfp* or *cfp* reporter sequence and recovered, as above, except SOC replaced SOB[-Mg]. Recovered cells were serially diluted in M9 and aliquots (50 μl) of either undiluted and 10^{-1} or 10^{-1}, 10^{-2} and 10^{-3} dilutions, for non-induced and induced cells, respectively, spread on NSLB/agar (18) plates containing Cm (12.5 μg/ml) and streptomycin (Str; 500 μg/ml). As above, aliquots (50 μl) of 10^{-4} and 10^{-5} dilutions of both non-induced and induced cells were also spread

on LB/agar plates without selection. After 36–48 h incubation (32°C) colony numbers were recorded for all plates. As above, six discrete colonies from one of the induced plates were each re-streaked onto fresh Cm- (12.5 µg/ml) and Str- (500 µg/ml) containing NSLB/agar plates and incubated further (24–36 h, 32°C). Glycerol stocks were made from overnight mini-cultures (LB plus Cm [12.5 µg/ml], 32°C), seeded with a single colony from each of these re-streaked plates. PCRs, using the appropriate target-flanking primer pair (above and Table 1) and fosmid DNA isolated from the corresponding mini-culture as template, confirmed the replacement, or otherwise, of the RT-cassette with the *gfp* or *cfp* reporter sequence. Two of these PCR products were sequenced, with the PCR and *gfp*- and *cfp*-specific primers 4237 and 4238 (Table 1), to confirm successful recombineering.

Finally, fosmid DNA was re-transformed, by electro-poration, into the CopyControl strain EPI300-T1^R and high-quality fosmid DNA prepared from an overnight mini-culture (LB plus Cm [12.5 µg/ml], 37°C) following induction of clone copy number. An aliquot of each recombineered fosmid DNA was digested with an appropriate restriction enzyme and the resulting fragment patterns examined visually for insert integrity. Where determined the percent recombination efficiency was calculated as ('number of colonies with selection'/'total viable cell count') × 100.

C.elegans culture, strains, transformation and observation

Worm culture, handling and manipulations were performed according to Materials and Methods described (20). The wild-type Bristol N2 strain (21) was transformed by microin-jection (22) with fosmid DNA at 50 ng/µl and pRF4 plasmid DNA at 100 ng/µl. pRF4 contains *rol-6(su1006)* which confers a rolling phenotype by which transformants can be recognized and transgenic lines can be maintained. The extra-chromosomal arrays created in the transformation events contain both plasmid and fosmid DNA. Each transgenic strain was established from a different injected animal and therefore is independently generated. Green fluorescent protein (GFP) and CFP expression patterns were observed in herma-phrodites only, by fluorescence microscopy with Chroma Technology Corp. filter set 41 012 and Zeiss filter set 47, respectively, on a Zeiss axioplan microscope equipped with DIC optics. Images were collected with a Photometrics CoolSNAP camera and Improvision Openlab software.

## RESULTS

Choice of target genes

Three *C.elegans* genes, *far-2/F02A9.3*, *F40E10.6* and *F54C9.11*, were selected on which to test the recombineering procedures. Reporter expression patterns had been deter-mined for these genes previously (23) during earlier projects investigating expression of the *C.elegans* genome. For each the reporter fusion was generated simply by conventional ligation of a relatively small DNA fragment, containing the region upstream of the protein-coding region of the *C.elegans* gene and predicted to contain the promoter, onto the start of

the reporter gene. In addition, there were distinct reasons for questioning the complete validity of the expression patterns observed. Furthermore, a β-*galactosidase* reporter had been used, rather than a fluorescent reporter protein. The recombi-neering approach would allow both GFP and CFP fusions to be made for all of these genes and the expression patterns generated, which were expected to be more reliable, could then be compared with those described previously.

Two-step counter-selection recombineering protocol

We first attempted to modify a number of *C.elegans* genomic cosmid clones (24) directly in their original *E.coli* host strain(s) by supplying λ Red functions either via plasmid pKD46 (25) or by integrating mini-λ (26) into the bacterial chromosome. Neither approach proved successful (data not shown) due to likely incompatibilities between the cosmid clone and/or host strain with the incoming Red-supplying DNA and also inherent stability problems with the cosmid clones. We subsequently explored an alternative approach in which fosmid clone DNA (see below) was moved, by effi-cient electroporation, into the recombineering-competent *E.coli* strain EL350 (17). EL350 contains a chromosomally-integrated defective λ prophage containing the Red *gam*, *bet* and *exo* genes arranged in their natural context and expressed from a native *PL* promoter tightly regulated by the temperature-sensitive cI857 repressor (19). Inactivating the repressor, by brief temperature shift to 42°C, permits high-level, coordinated expression of *gam*, *bet* and *exo* that together mediate homologous recombination between the recipient target gene and a double-stranded linear donor DNA. Although the plasmid- and prophage-based systems for delivery of phage recombination functions each have their advantages and disadvantages, for example plasmid delivery is more flexible as it can be transferred readily between strains, the tighter control and coordinated expres-sion of the integrated prophage approach is considered more efficient and controllable.

As we also wished to seamlessly modify our target genes this necessitated the adoption of a two-step recombineering strategy. In this approach a counter-selection cassette, con-taining both positive and negative genetic markers and inserted at the target site by an initial recombineering step, using positive selection for recombinants (Figure 1A), is replaced with the desired sequence in a second round of recombineering using negative selection (Figure 1B). We used a dicistronic counter-selection cassette (the RT-cassette), containing the respective positive and negative markers *tet*A(C) and *rps*L driven by the hypo-osmotically up-regulated *omp*F promoter, as it provides stringent positive and negative selection of single-copy vectors (18). Positive selection is provided by *tet*A(C) conferring Tc resistance (Tc^R) whereas, when expressed in a *rps*L$^-$ host, the wild-type *rps*L$^+$ gene provides the negative marker. Mutations in the chromosomal *rps*L copy, that encodes the ribosomal protein S12, confers Str resistance (Str^R). When both mutant and wild-type *rps*L alleles are co-expressed a dominant Str sensi-tive (Str^S) phenotype results (27). Because EL350 is *rps*L$^-$ and thus Str^R, the introduction, via the RT-cassette, of a *rps*L$^+$ allele confers Str^S. Consequently, non-recombinants following the second recombineering step will remain Str^S
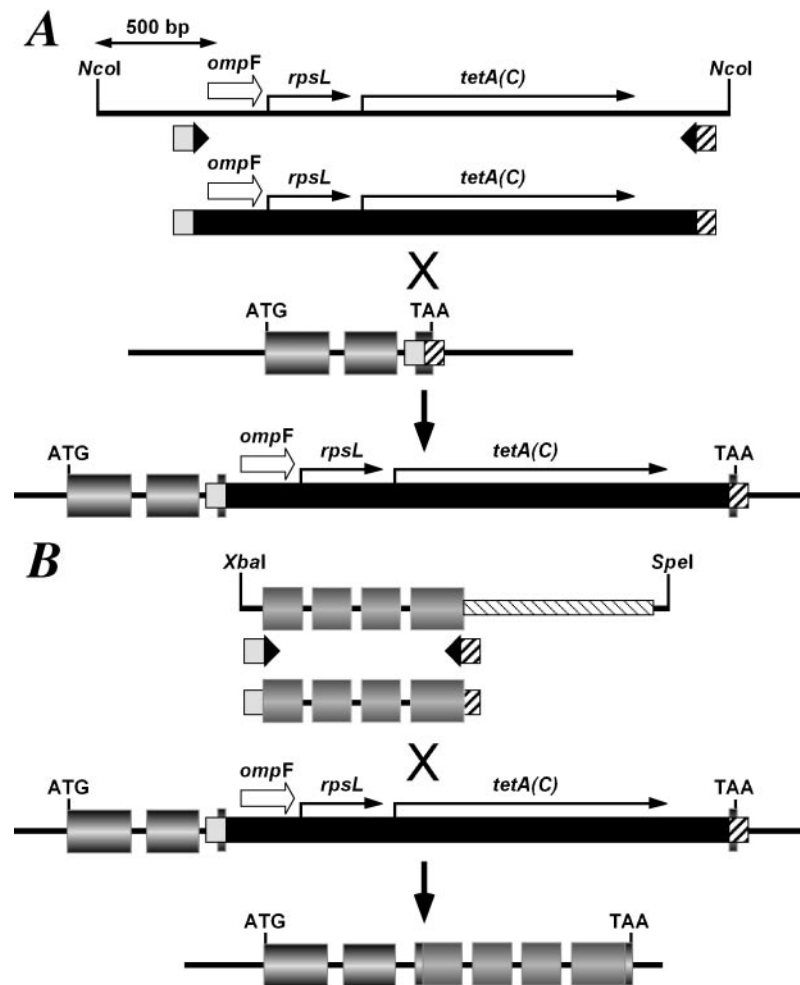
Figure 1. General scheme for seamless introduction of *gfp* (or *cfp*) into *C.elegans* fosmid clones by Red-mediated recombineering. (**A**) Insertion of the RT-cassette into the target site. Two 75 nt recombineering primers, containing 50 nt left (stippled) and right (hatched) homology arms, are used to amplify, using a gel-purified NcoI restriction fragment of pBAC-RT as template, a 2.7 kb dicistronic cassette encoding wild-type *rps*L and the *tet*A(C) efflux pump driven by the *omp*F promoter. Red recombineering-competent EL350 *E.coli*, containing the target fosmid clone are electroporated with the purified PCR product, recovered and recombinant clones selected with tetracylcine (5 µg/ml) on LB agar. (**B**) Replacement of the RT-cassette with the *gfp* coding sequence. A PCR product encoding the *gfp* coding sequence, minus the stop codon, is amplified using a gel-purified XbaI–SpeI restriction fragment of pPD95.77 as template and primed with 75 nt primers containing the same left and right homology arms. Recombineering competent RT-cassette-containing *E.coli* are electroporated with the purified *gfp* PCR product, recovered and recombinant clones selected with chloramphenicol (12.5 µg/ml) and streptomycin (500 µg/ml) on NSLB agar.

and are selected against whereas desired recombinants are Str^R revertants.

It is difficult to negatively select for successful recombinants when the target clone is present in a multi-copy vector because of the background of non-recombinant clones that remain due to the relative inefficiency of recombineering. Thus, our choice of a counter-selection approach limited us to modifying *C.elegans* genes cloned into single- or very low-copy-number vectors, such as BACs, PACs or fosmids. During construction of the *C.elegans* physical map a genomic library was constructed (1) in the fosmid vector pFos1 (28) from which ∼700 clones have been restriction-mapped onto the genome and 100 sequenced in full or part (John Spieth, personal communication). In an initial proof-of-principle experiment we optimized our two-step RT-cassette-based counter-selection approach by fusing a *gfp* reporter sequence in-frame at the C-terminus of *fmo*-5 (29) present in a fosmid

clone (*H24K24*) originating from this library. Although successfully proving the method's validity (data not shown) the lack of genome coverage means that fosmid clones from this library will be of limited use as a recombineering resource. However, shortly after the success with the pFos1-based clone, the availability of an additional *C.elegans* genomic library, comprising ∼16 000 clones and constructed in the copy-number-inducible fosmid vector pCC1FOS, was reported (D. Moerman, personal communication). We therefore undertook recombineering of three target genes each located within a different one of these so-called CopyControl fosmids.

Generation of *C.elegans* fosmid-based *gfp/cfp* reporter constructs

Our recombineering strategy was designed to seamlessly introduce a *gfp* or *cfp* reporter sequence into each of three

target genes each located centrally in the insert of a fosmid clone (illustrated in Figure 2A for *far-2/F02A9.3*). For these genes the flanking genomic sequences ranged in length from a minimum of 8.2 kb (flanking the 3′ side of *F40E10.6*) to a maximum of 26 kb (flanking the 5′ side of *F54C9.11*). Genes are tightly packed in the *C.elegans* genome, on average one every 5 kb, and each of these flanking regions contains the entire transcribed region for at least one other *C.elegans* gene. We also wanted the resulting reporter fusions to retain any native signal peptide sequence(s) at the N- and/or C-terminus thus maximizing the likelihood that intracellular localization would reflect that of the endogenous protein. Consequently, the sequences of the left and right homology arms were designed such that the *gfp* or *cfp* reporter coding sequence would insert, in-frame, immediately 5′ to the last six codons of the native gene (Figure 2B and C).

For the first step of our recombineering strategy, insertion of the RT counter-selection cassette, no background of Tc^R colonies were observed on the undiluted non-induced control plate for any of the three target genes. In contrast, hundreds of Tc^R colonies were observed for each gene on the corresponding Red-induced plate. The absence of background colonies was likely due to the use of a diluted restriction fragment of pBAC-RT as the template, rather than the intact plasmid, for amplification of the RT-cassette; in trial experiments the linear template had not produced any Tc^R colonies when electroporated directly into EL350. PCR analysis indicated insertion of the RT-cassette for each of the six clones examined per target gene (Figure 3 and data not shown). In the subsequent replacement step we did observe a background of Str^R colonies on the non-induced control plates, although this was expected because any sequence change resulting in a functional loss of the negative marker will be selected for. However, there were approximately 1-2-orders of magnitude more Str^R colonies on corresponding Red-induced plates and recombineering efficiencies for this step were reproducibly ∼0.1%. For each of the six clones examined per gene, PCR analyses (Figure 3 and data not shown) demonstrated successful replacement of the RT-cassette with the *gfp* or *cfp* sequence. Further, for each recombineered fosmid visual comparison of the EcoRI restriction fragment pattern with the corresponding predicted restriction map (data not shown) provided no evidence for any unwanted insert rearrangement. Finally, DNA sequencing across the recombineering joints confirmed recombination fidelity.

Creation of transgenic worms and gene expression analysis

The fosmid DNA preparations were used in transformation of wild-type *C.elegans* by microinjection using procedures established for plasmid DNAs (22). Although rates of successful transformation by microinjection of plasmid DNA can be variable, the nature of the fosmid DNA preparation did not appear to interfere with the process and transformant lines were generated at a reasonable frequency (approximately one independent line was established per 5 worms injected). As with conventional-plasmid based reporter gene fusions, the recombineered fosmids were mixed with the plasmid pRF4 that contains an allele of the *rol-6* gene to
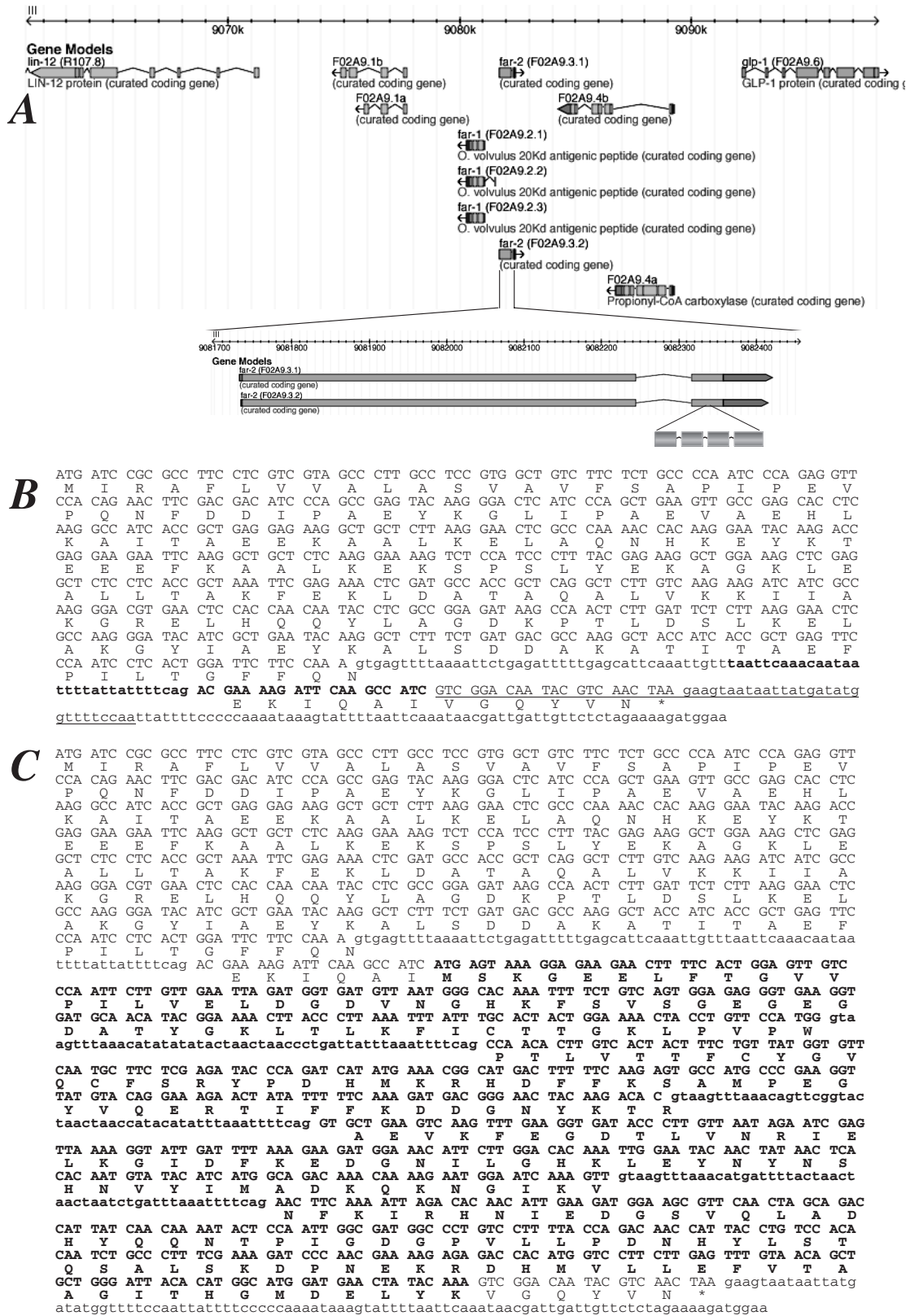
confer a rolling phenotype for recognition of transgenic progeny. In such co-injections the different DNAs recombine to form large extrachromosomal arrays which may then be transmitted through subsequent generations, transgenic lines being maintained by screening for rolling individuals at successive generations. As reporter gene expression was seen (see below) in 10 of the 11 lines propagated the fosmid DNA appeared able to participate in formation of these arrays in the same way as plasmid DNAs.

far-2/F02A9.3

The expression pattern of *far-2/F02A9.3* was previously examined with a conventional *lacZ* reporter fused directly at the translational start codon (23). The 7 kb of upstream DNA that was assayed included two further genes, *far-1*, a close homologue of *far-2*, and *F02A9.1*, both transcribed from the opposite strand. Although the respective initiation codons for *far-1* and *far-2* are only 600 bp apart, the former gene has an alternative transcript, with an untranslated exon, resulting in an effective IGS of less than 200 bp. The β-galactosidase expression [Figure 4A, Ref. (23)] was observed in adults only, in body wall and vulval muscle cells. The nuclear-localization resulted from a nuclear-localization signal on the reporter and revealed nothing of the subcellular distribution of FAR-2 protein. The expression pattern appeared strong and very specific, suggesting that it could be relied upon as an accurate reflection of the endogenous gene. However, there are no particular distinctions between larval and adult *C.elegans* body wall muscle cells and so the adult specificity of the expression pattern was curious. Furthermore, *far-2* encodes one of several homologues of an antigenic peptide from *Onchocerca volvulus* (30) and it was unclear why a muscle protein in *C.elegans* should be an antigenic peptide in *O.volvulus*. The *far-2* gene was named on the basis of sequence similarity to secreted fatty acid and retinol binding proteins.

Both *gfp* and *cfp* fusions generated by fosmid recombineering drove body wall muscle expression, that was also specific to adults (Figure 4B and C) as observed previously with the *lacZ* fusion. The fluorescent fusion proteins were not nuclear-localized, as they did not include a nuclear-localization signal like the β-galactosidase fusion protein, but were observed in striations, suggesting the endogenous FAR-2 protein may be organized within muscle cells in register with myofilaments.

In addition to the muscle cells, GFP was seen throughout the pseudocoelom (Figure 4E and G) and in the six coelomocytes (Figure 4J and K), but again only in adults. These components were even stronger for CFP. The abundant pseudocoelomic fluorescence made it difficult to discern the muscle cell component and, hence, it is unclear if expression also occurs in vulval muscle cells as observed with the *lacZ* fusion. An interpretation is that the gene is expressed in adult body wall muscle cells, and perhaps vulval muscle cells, but the gene product is secreted into the pseudocoelom from where it is scavenged by the coelomocytes. The FAR-2 protein does have a signal sequence at the N-terminus that would be expected to direct secretion and would be present in the *gfp* and *cfp* fusions, but not the *lacZ* fusion. Coelomocytes have been reported to take up GFP from the

A

B

```
ATG ATC CGC GCC TTC CTC GTC GTA GCC CTT GCC TCC GTG GCT GTC TTC TCT GCC CCA ATC CCA GAG GTT
 M   I   R   A   F   L   V   V   A   L   A   S   V   A   V   F   S   A   P   I   P   E   V
CCA CAG AAC TTC GAC GAC ATC CCA GCC GAG TAC AAG GGA CTC ATC CCA GCT GAA GTT GCC GAG CAC CTC
 P   Q   N   F   D   D   I   P   A   E   Y   K   G   L   I   P   A   E   V   A   E   H   L
AAG GCC ATC ACC GCT GAG GAG AAG GCT GCT CTT AAG GAA CTC GCC CAA AAC CAC AAG GAA TAC AAG ACC
 K   A   I   T   A   E   E   K   A   A   L   K   E   L   A   Q   N   H   K   E   Y   K   T
GAG GAA GAA TTC AAG GCT GCT CTC AAG GAA AAG TCT CCA TCC CTT TAC GAG AAG GCT GGA AAG CTC GAG
 E   E   E   F   K   A   A   L   K   E   K   S   P   S   L   Y   E   K   A   G   K   L   E
GCT CTC CTC ACC GCT AAA TTC GGA AAA CTC GAT GCC ACC GCT CAG GCT CTT GTC AAG AAG ATC ATC GCC
 A   L   L   T   A   K   F   E   K   L   D   A   T   A   Q   A   L   V   K   K   I   I   A
AAG GGA CGT GAA CTC CAC CAA CAA TAC CTC GCC GGA GAT AAG CCA ACT CTT GAT TCT CTT AAG GAA CTC
 K   G   R   E   L   H   Q   Q   Y   L   A   G   D   K   P   T   L   D   S   L   K   E   L
GCC AAG GGA TAC ATC GCT GAA TAC AAG GCT CTT TCT GAT GAC GCC AAG GCT ACC ATC ACC GCT GAG TTC
 A   K   G   Y   I   A   E   Y   K   A   L   S   D   D   A   K   A   T   I   T   A   E   F
CCA ATC CTC ACT GGA TTC TTC CAA A gtgagttttaaaattctgagatttttgagcattcaaattgtt taattcaaacaataa
 P   I   L   T   G   F   F   Q   N
ttttattattttcag AC GAA AAG ATT CAA GCC ATC GTC GGA CAA TAC GTC AAC TAA gaagtaataattatgatatg
                    E   K   I   Q   A   I   V   G   Q   Y   V   N   *
gtttttccaattatttttcccccaaaataaagtattttaattcaaataacgattgattgttctctagaaaagatggaa
```

C

```
ATG ATC CGC GCC TTC CTC GTC GTA GCC CTT GCC TCC GTG GCT GTC TTC TCT GCC CCA ATC CCA GAG GTT
 M   I   R   A   F   L   V   V   A   L   A   S   V   A   V   F   S   A   P   I   P   E   V
CCA CAG AAC TTC GAC GAC ATC CCA GCC GAG TAC AAG GGA CTC ATC CCA GCT GAA GTT GCC GAG CAC CTC
 P   Q   N   F   D   D   I   P   A   E   Y   K   G   L   I   P   A   E   V   A   E   H   L
AAG GCC ATC ACC GCT GAG GAG AAG GCT GCT CTT AAG GAA CTC GCC CAA AAC CAC AAG GAA TAC AAG ACC
 K   A   I   T   A   E   E   K   A   A   L   K   E   L   A   Q   N   H   K   E   Y   K   T
GAG GAA GAA TTC AAG GCT GCT CTC AAG GAA AAG TCT CCA TCC CTT TAC GAG AAG GCT GGA AAG CTC GAG
 E   E   E   F   K   A   A   L   K   E   K   S   P   S   L   Y   E   K   A   G   K   L   E
GCT CTC CTC ACC GCT AAA TTC GGA AAA CTC GAT GCC ACC GCT CAG GCT CTT GTC AAG AAG ATC ATC GCC
 A   L   L   T   A   K   F   E   K   L   D   A   T   A   Q   A   L   V   K   K   I   I   A
AAG GGA CGT GAA CTC CAC CAA CAA TAC CTC GCC GGA GAT AAG CCA ACT CTT GAT TCT CTT AAG GAA CTC
 K   G   R   E   L   H   Q   Q   Y   L   A   G   D   K   P   T   L   D   S   L   K   E   L
GCC AAG GGA TAC ATC GCT GAA TAC AAG GCT CTT TCT GAT GAC GCC AAG GCT ACC ATC ACC GCT GAG TTC
 A   K   G   Y   I   A   E   Y   K   A   L   S   D   D   A   K   A   T   I   T   A   E   F
CCA ATC CTC ACT GGA TTC TTC CAA A gtgagttttaaaattctgagatttttgagcattcaaattgtttaattcaaacaataa
 P   I   L   T   G   F   F   Q   N
ttttattattttcag AC GAA AAG ATT CAA GCC ATC ATG AGT AAA GGA GAA GAA CTT TTC ACT GGA GTT GTC
                    E   K   I   Q   A   I   M   S   K   G   E   E   L   F   T   G   V   V
CCA ATT CTT GTT GAA TTA GAT GGT GAT GTT AAT GGG CAC AAA TTT TCT GTC AGT GGA GAG GGT GAA GGT
 P   I   L   V   E   L   D   G   D   V   N   G   H   K   F   S   V   S   G   E   G   E   G
GAT GCA ACA TAC GGA AAA CTT ACC CTT AAA TTT ATT TGC ACT ACT GGA AAA CTA CCT GTT CCA TGG gta
 D   A   T   Y   G   K   L   T   L   K   F   I   C   T   T   G   K   L   P   V   P   W
agtttaaacatatatatactaactaaccctgattatttaaattttcag CCA ACA CTT GTC ACT ACT TTC TGT TAT GGT GTT
                                                   P   T   L   V   T   T   F   C   Y   G   V
CAA TGC TTC TCG AGA TAC CCA GAT CAT ATG AAG CGG CAT GAC TTT TTC AAG AGT GCC ATG CCC GAA GGT
 Q   C   F   S   R   Y   P   D   H   M   K   R   H   D   F   F   K   S   A   M   P   E   G
TAT GTA CAG GAA AGA ACT ATA TTT TTC AAA GAT GAC GGG AAC TAC AAG ACA C gtaagtttaaacagttcggtac
 Y   V   Q   E   R   T   I   F   F   K   D   D   G   N   Y   K   T   R
taactaaccatacatatttaaattttcag GT GCT GAA GTC AAG TTT GAA GGT GAT ACC CTT GTT AAT AGA ATC GAG
                                  A   E   V   K   F   E   G   D   T   L   V   N   R   I   E
TTA AAA GGT ATT GAT TTT AAA GAA GAT GGA AAC ATT CTT GGA CAC AAA TTG GAA TAC AAC TAT AAC TCA
 L   K   G   I   D   F   K   E   D   G   N   I   L   G   H   K   L   E   Y   N   Y   N   S
CAC AAT GTA TAC ATC ATG GCA GAC AAA CAA AAG AAT GGA ATC AAA GTT gtaagtttaaacatgattttactaact
 H   N   V   Y   I   M   A   D   K   Q   K   N   G   I   K   V
aactaatctgatttaaattttcag AAC TTC AAA ATT AGA CAC AAC ATT GAA GAT GGA AGC GTT CAA CTA GCA GAC
                            N   F   K   I   R   H   N   I   E   D   G   S   V   Q   L   A   D
CAT TAT CAA CAA AAT ACT CCA ATT GGC GAT GGC CCT GTC CTT TTA CCA GAC AAC CAT TAC CTG TCC ACA
 H   Y   Q   Q   N   T   P   I   G   D   G   P   V   L   L   P   D   N   H   Y   L   S   T
CAA TCT GCC CTT TCG AAA GAT CCC AAC GAA AAG AGA GAC CAC ATG GTC CTT CTT GAG TTT GTA ACA GCT
 Q   S   A   L   S   K   D   P   N   E   K   R   D   H   M   V   L   L   E   F   V   T   A
GCT GGG ATT ACA CAT GGC ATG GAT GAA CTA TAC AAA GTC GGA CAA TAC GTC AAC TAA gaagtaataattatg
 A   G   I   T   H   G   M   D   E   L   Y   K   V   G   Q   Y   V   N   *
atatggttttccaattattttcccccaaaataaagtattttaattcaaataacgattgattgttctctagaaaagatggaa
```

Figure 2. Genomic environment of the fosmid clone WRM0640dE04 and recombineering of *gfp* into *F02A9.3* (*far*-2). (**A**) Clone WRM0640dE04, encompassing the target gene *far*-2, spans 36.8 kb of chromosome III from position 9 061 282 to 9 098 108 (WormBase freeze WS140). The insertion site of the *gfp* coding sequence into the second exon of *far*-2 is illustrated below the genomic region. (**B**) The 50 bp regions of *far*-2, corresponding to the left and right homology arms mediating recombination, are in bold and underlined, respectively. (**C**) The *far*-2::*gfp* fusion sequence generated by recombineering of fosmid WRM0640dE04. The inserted *gfp* coding sequence, derived from pPD95.77, is in bold.
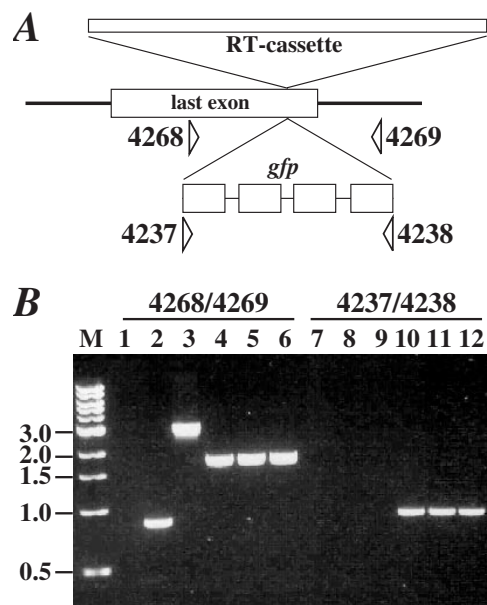
Figure 3. PCR analysis of recombineering products of *C.elegans* gene *F40E10.6*. (**A**) Annealing locations (not to scale) of PCR primers 4268/4269 and 4237/4238 in relation to the insertion site of the RT- and *gfp/cfp* cassettes within the last exon of *F40E10.6* are drawn below their respective templates. (**B**) DNA isolated from clones of either unmodified fosmid DNA WRM0614cG08 (lanes 2 and 8) or following insertion of the RT-cassette (lanes 3 and 9) and subsequent replacement with *gfp* (lanes 4–6, 10–12) was amplified with primers 4268/4269 (lanes 1–6) or 4237/4238 (lanes 7–12). No template control (lanes 1 and 7), M = 1 kb DNA ladder (NEB).

pseudocoelom (31). The reason for the striated appearance of the muscle cell-localized fluorescent protein is not yet apparent but may be an artefact. Nevertheless, the reporter expression patterns generated by the different approaches appear consistent, revealing different aspects of the endogenous *far-2* gene's expression pattern; the transcriptional fusion reveals where the promoter is active and driving production of the gene product while the translational fusion reveals the distribution of the FAR-2 protein.

F54C9.11

The *F54C9.11*::*lacZ* fusion had driven β-galactosidase expression in hypodermal cells and vulval muscle cells when assayed previously [Figure 4N; Ref. (23)]. On that occasion the fragment contained the entire 6 kb IGS, which is relatively large for *C.elegans*, and the fusion point was in the second of the three exons of *F54C9.11*. Nevertheless, expression was very weak and the hypodermal expression, starting in mid-larval stages through to the adult was suspiciously reminiscent of the expression of the co-transformation marker, *rol-6*. It is generally accepted that *rol-6* enhancers, present on the mixed extrachromosomal arrays formed upon microinjection transformation of *C.elegans*, can act on reporter gene fusions in the same array.

Very weak vulval muscle cell expression was observed for the *F54C9.11*::*gfp* fusion generated by fosmid recombineering (Figure 4O and P). The GFP could just be seen despite the strong, background, endogenous, intestinal autofluorescence that appears yellow with the filter set used. CFP

could not be detected for the *F54C9.11*::*cfp* fusion, but the filter set used to observe this fluor meant that the autofluorescence appeared blue, the same colour as the reporter signal, and thus a weak vulval expression may not have been discernable.

Curiously, when examining the worms for CFP fluorescence, a strong fluorescent signal was observed from lateral hypodermal nuclei. However, this was also seen in wild-type *C.elegans* and has nothing to do with the CFP reporter. With very careful observation this autofluorescence could again be seen, exceedingly weakly, with the GFP filter set but not with a variety of other filter sets that were available. This autofluorescence precludes claiming that no hypodermal expression was generated for the *F54C9.11* fluorescent protein reporter fusions. Nevertheless, it would appear that *F54C9.11* is expressed weakly in the vulval muscle cells and the hypodermal expression previously observed was indeed due to activation by regulatory elements of the co-transformation marker. Furthermore, the additional flanking DNA present for reporter gene fusions generated by fosmid recombineering may well insulate the fusion gene from such artifacts.

F40E10.6

One generally accepted potential artifact of reporter gene fusions in *C.elegans* is inappropriate expression in certain tissues. In particular, the high frequency of such expression in large-scale gene expression pattern screens using reporters and in studies seeking to identify transcriptional regulatory elements by serial deletion of promoter regions, has led to the perception that pharyngeal expression can be a consequence of fusing an incomplete promoter region to a reporter. The expression pattern previously observed (23) for *F40E10.6*::*lacZ* included ill-defined expression in the head, around the vulva and around the tail, and, in addition, expression in the pharynx (Figure 4Q). However, it would have been thought that an intact promoter would have been present in this fusion. The DNA fragment assayed contained the entire 3 kb IGS and reached to more than half way through the next upstream gene, *tag-53/F33C8.1*, which is transcribed in the opposite direction. Furthermore, most of the coding region of *F40E10.6* was also included in the fragment.

GFP and CFP expression patterns observed for *F40E10.6* reporter fusions generated by fosmid recombineering were strong and clear, restricted to the nervous system (Figure 4T), from the circumpharyngeal nerve ring in the head (Figure 4U, V and W), down the nerve cord, around the vulva (Figure 4R and S), to the tail ganglia. No expression could be detected in the pharynx. The large amounts of flanking DNA in the recombineered fusions may avoid artifactual expression potentially arising from fusions with incomplete promoters. The more tightly localized distribution of the fluorescent proteins may be because the product of the X-gal staining reaction, used to visualize β-galactosidase, spreads from the source. However, the recombineered fusions result in the fluorescent tag being added to the terminus of a virtually intact *C.elegans* protein, and the fusion protein may then show a subcellular distribution more tightly reflecting the distribution of the endogenous protein.
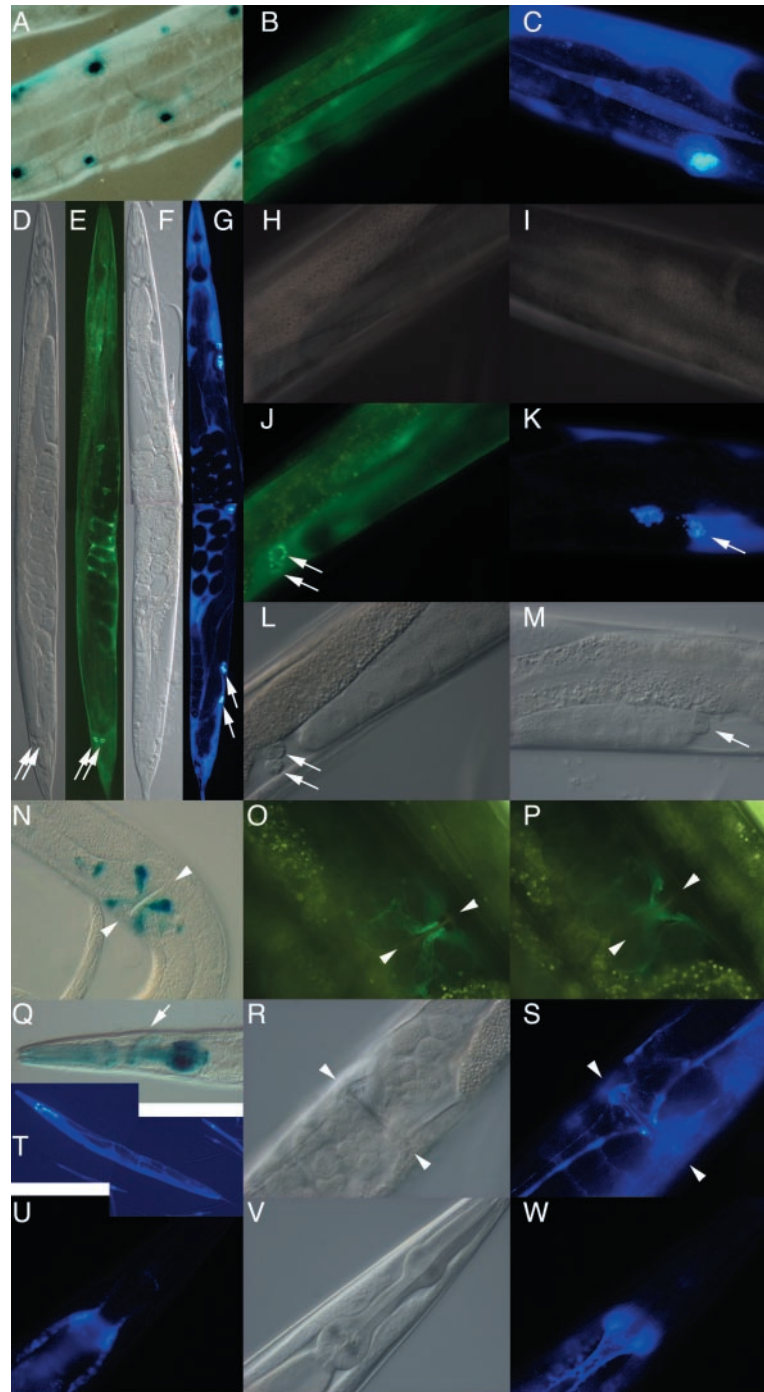
Figure 4. Expression patterns. Reporter expression patterns in adult hermaphrodites for *far-2*/*F02A9.3* (**A–M**), *F54C9.11* (**N–P**) and *F40E10.6* (**Q–W**). β-galactosidase (A, N, Q), GFP (B, E, J, O, P) or CFP (C, G, K, S, T, U, W) distributions are presented. DIC was used to observe anatomical details (D, F, L, M, N, Q, R, V) and polarized light was used to examine muscle cells (A, H, I). The fluorescent images were typically captured with corresponding bright-field images (B & H, C & I, D & E, F & G, J & L, K & M, S & R). B/H & J/L and O & P are different focal planes of the same specimens. U and W are different focal planes around the focal plane of V. Trapezoid shaped body wall muscle cells, with striations due to the regular arrangement of the myofilament lattice, are clearly apparent in panels A, B, C, H and I. Coelomocytes do not occupy fixed positions in the pseudocoelomic cavity and are indicated by arrows in panels D, E, G, J, K, L and M. The vulva is located between the arrowheads in panels N, O, P, R and S, with reporter expression in vulval muscle cells either side of the vulva in panels N, O and P, and in the ventral nerve cord running from lower left to upper right in panel S detouring around the vulva, which is clearly apparent in panel R. In panel Q, β-galactosidase expression can be seen weakly in the nerve ring (arrow) running around the pharynx, which also shows expression, particularly in the terminal bulb. In panel T, the thin line of CFP running the length of the animal is the ventral nerve cord, but reporter expression can also be seen in the nerve ring in the head (left) and tail (right), despite the background autofluorescence in the intestine, the broad band of blue running along the lower half of the worm in the figure. The strongest intensity of fluorescence in panels U and W correspond to sections through the nerve ring at different focal planes around the pharynx which can be clearly seen in panel V. The *C.elegans* strains photographed were UL3 (A), UL1980 (B, D, E, H, J, L), UL1982 (C, F, G, I, K, M), UL501 (N), UL1986 (O, P), UL818 (Q), and UL1985 (R-W). All images were captured at 400× magnification apart from D–G, which were captured at 100× magnification.

DISCUSSION

The use of reporter constructs to investigate gene expression patterns in *C.elegans*, as well as other model organisms, has been a mainstay in gene function studies in both the pre- and post-genomics eras. Such constructs are invariably generated using approaches based either on the traditional use of restriction enzymes and DNA ligase or, more recently, via site-specific recombination-based Gateway Technology (13,32). In recent years the growing availability of long, contiguous fragments of genomic DNA, cloned into vectors, such as BACs or fosmids, has provided the opportunity for modification of potentially complete gene sequences within their local structural context. However, their large size precludes their precise modification by restriction enzymes thus limiting their usefulness as base vectors for, e.g. more biologically relevant reporter constructs. Homologous recombination represents a restriction enzyme-independent alternative to manipulating large, cloned DNA fragments and a number of methodological approaches based upon *E.coli RecA* have been described (33,34). Recently, a more efficient homologous recombination method, commonly referred to as recombineering and based on viral recombinases, such as the phage λ Red proteins (15,34) or equivalent RecE/RecT system of the prophage Rac (33,35), has been developed. No restriction enzyme-based manipulations are needed as only short (∼50 bp) homology arms are required which can be built into the linear donor DNA via PCR. Although, to date, most often used for direct modification of BAC clones recombineering also holds promise as a more general genetic engineering tool.

By combining recombineering of fosmid clones with a two-step counter-selection procedure we have been able to precisely and seamlessly retrofit *C.elegans* fosmid clones with either *gfp* or *cfp* coding sequences to generate C-terminal translational reporters for three genes previously demonstrated to drive reporter expression in muscle cells. The method, which is straightforward and robust, enables precise insertion of a reporter sequence at any desired location within the cloned insert thus providing opportunities for exploring the effects of different fusion protein arrangements upon reporter expression. It could also be used to insert sequences other than a reporter into a gene of interest, for example, tandem affinity purification [TAP-; Ref (36)] or localization and affinity purification [LAP-; Ref (37)] tags. Furthermore, using iterative rounds of recombineering, multiple sequences could be inserted into a single clone, for example to modify a series of individual genes within an operon. Alternatively, instead of inserting sequences, the approach could be used to create point mutations or precise deletions enabling, for example, generation of a series of reporter constructs for *in vivo* promoter analysis.

Other non-restriction enzyme-based methods to generate *C.elegans* reporter constructs have been described including *in vivo* recombination within *C.elegans* itself (22), fusion-PCR to generate linear reporter constructs (38) and, more recently, an approach that exploits the inherent capacity of yeast for homologous recombination (39). The success rate is very low for the first method and the second method utilizes relatively small DNA fragments which will not include any downstream regulatory elements. Although most resembling recombineering, particularly with respect to the final construct generated, the third method is, in comparison to the protocol described here, relatively complex and also requires the shuttling of DNA constructs between yeast and bacterial hosts. Furthermore, in terms of the numbers of clones analysed to identify desired recombinants, it appears significantly less efficient. For example, examination of all our recombineering results (this report and data not shown) indicate that, for the second step, ∼90% of the Str^R clones are positive recombinants in which the desired sequence has replaced the counter-selection RT-cassette. A background of non-recombinants is typical of all negative selection strategies because any sequence change that results in a functional loss of the negative marker will be selected for. With the RT-cassette used here it is likely this background could be further reduced if we were to also include low concentration kanamycin in our negative selection procedure as suggested (18). However, the 90% efficiency we observe with Str alone is more than acceptable. For example, all of the six clones tested for each of the target genes described here contained the desired *gfp* or *cfp* reporter.

Apart from the well-known but relatively inefficient *SacB* system used for negative selection (40), a number of alternative counter-selection protocols, specifically designed for use with recombineering, have been described recently. These include combining the homing endonuclease I-SceI recognition site in the counter-selection cassette with expression of I-SceI in *trans* (41) and two approaches (42,43) that both utilize single genetic markers for both positive and negative selection. Although these two latter approaches have the potential for improved negative selection efficiencies both require growth on minimal media which, when combined with incubation at 32°C, may lead to prolonged incubation periods. To the best of our knowledge, the dicistronic *rps*L-*tet*A(C) cassette used here has only been used once before in combination with recombineering and this was in a limited form to rescue deletion mutants generated in a BAC-based human cytomegalovirus (CMV) replicon (44). Because of the high levels of *rps*L$^+$ expression achieved when driven by the *omp*F promoter on hypo-osmotic media penetrance of the Str^S phenotype is high and provides robust negative selection. As well as modifying the *C.elegans* fosmids described here we have also used the RT-cassette with recombineering to seamlessly modify a number of BAC clones (C.T. Dolphin, unpublished data; P. Haussecker and A. Harris, unpublished data).

Fosmid DNA preparations appear suitable for transformation of *C.elegans* by microinjection into the syncytial germ line. We did not explore the potential of microparticle bombardment for transformation with fosmid DNA. However, an *unc*-119 selectable marker, commonly used with this procedure (45), could be easily introduced into, for example, the vector portion of the fosmid clone, by an additional round of recombineering. It should be noted, however, that the larger size of these fosmid clones, in comparison to the more commonly used plasmid-based reporters, may make them more susceptible to breakage during the bombardment procedure. The CopyControl system permits isolation of high-quality DNA with minimal contamination with *E.coli* chromosomal DNA. The larger size of the fosmids, as compared to plasmids, could mean that arrays would have

fewer copies of the transgenes but, although no direct comparison was made with equivalent *gfp* fusions in plasmids, the levels of GFP expression were good. Expression patterns did correspond to those observed previously for conventional fusions, with the apparent loss of what are generally accepted as typical artifacts of reporter gene fusions generated conventionally. It seems likely that reporter fusions generated by recombineering and containing much larger segments of flanking DNA are more likely to reflect the distribution of the endogenous gene product than more conventional reporter fusions because they are more likely to include all the pertinent regulatory elements. However, as with any fusion reporter system resulting expression patterns may be subject to potential artifactual effects arising, for example, from increased fusion gene dosage and/or germline silencing.

We have described a simple and straightforward approach to constructing translational reporter constructs directly from fosmid genomic clones. Because *C.elegans* genes are relatively compact and the average fosmid clone insert is 35–40 kb, constructs in which the target is positioned centrally will have extensive regions of flanking sequence ensuring that the majority, if not all, the associated regulatory elements, even those acting at some distance, will be included. As such, resulting expression patterns are highly likely to recapitulate that of the endogenous gene. Additionally, the fusion proteins are likely to retain functionality and also be correctly localized within the cell. The method has potential for further simplification and should be readily adaptable to high-throughput assays of large numbers of genes in 96-well format. Other genome-wide projects, such as the promoterome (14), designed to interrogate expression of large numbers of *C.elegans* promoters in parallel are currently underway. The method described here would be complementary to such projects and provide confirmatory and/or additional information on gene expression. Finally, although we have applied the method to *C.elegans* it would be equally applicable to other model animals for which fosmid or BAC clones are available. As such, we envisage the method will appeal not only to researchers in the worm community but also others as well.

ACKNOWLEDGEMENTS

REFERENCES

1. *C.elegans* Sequencing Consortium (1998) Genome sequence of the nematode *C.elegans*: a platform for investigating biology. *Science*, **282**, 2012–2018.
2. Kamath,R.S., Fraser,A.G., Dong,Y., Poulin,G., Durbin,R., Gotta,M., Kanapin,A., Le Bot,N., Moreno,S., Sohrmann,M. *et al.* (2003) Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. *Nature*, **421**, 231–237.
3. Sugimoto,A. (2004) High-throughput RNAi in *Caenorhabditis elegans*: genome-wide screens and functional genomics. *Differentiation*, **72**, 81–91.
4. Ashrafi,K., Chang,F.Y., Watts,J.L., Fraser,A.G., Kamath,R.S., Ahringer,J. and Ruvkun,G. (2003) Genome-wide RNAi analysis of *Caenorhabditis elegans* fat regulatory genes. *Nature*, **421**, 268–272.
5. Poulin,G., Nandakumar,R. and Ahringer,J. (2004) Genome-wide RNAi screens in *Caenorhabditis elegans*: impact on cancer research. *Oncogene*, **23**, 8340–8345.
6. Kim,S.K., Lund,J., Kiraly,M., Duke,K., Jiang,M., Stuart,J.M., Eizinger,A., Wylie,B.N. and Davidson,G.S. (2001) A gene expression map for *Caenorhabditis elegans*. *Science*, **293**, 2087–2092.
7. Jones,S.J., Riddle,D.L., Pouzyrev,A.T., Velculescu,V.E., Hillier,L., Eddy,S.R., Stricklin,S.L., Baillie,D.L., Waterston,R. and Marra,M.A. (2001) Changes in gene expression associated with developmental arrest and longevity in *Caenorhabditis elegans*. *Genome Res.*, **11**, 1346–1352.
8. Holt,S.J. and Riddle,D.L. (2003) SAGE surveys *C.elegans* carbohydrate metabolism: evidence for an anaerobic shift in the long-lived dauer larva. *Mech. Ageing. Dev.*, **124**, 779–800.
9. Zhang,Y., Ma,C., Delohery,T., Nasipak,B., Foat,B.C., Bounoutas,A., Bussemaker,H.J., Kim,S.K. and Chalfie,M. (2002) Identification of genes expressed in *C.elegans* touch receptor neurons. *Nature*, **418**, 331–335.
10. Roy,P.J., Stuart,J.M., Lund,J. and Kim,S.K. (2002) Chromosomal clustering of muscle-expressed genes in *Caenorhabditis elegans*. *Nature*, **418**, 975–979.
11. Tabara,H., Motohashi,T. and Kohara,Y. (1996) A multi-well version of *in situ* hybridization on whole mount embryos of *Caenorhabditis elegans*. *Nucleic Acids Res.*, **24**, 2119–2124.
12. Fire,A., Harrison,S.W. and Dixon,D. (1990) A modular set of *lacZ* fusion vectors for studying gene expression in *Caenorhabditis elegans*. *Gene*, **93**, 189–198.
13. Hope,I.A., Stevens,J., Garner,A., Hayes,J., Cheo,D.L., Brasch,M.A. and Vidal,M. (2004) Feasibility of genome-scale construction of promoter::reporter gene fusions for expression in *Caenorhabditis elegans* using a multisite gateway recombination system. *Genome Res.*, **14**, 2070–2075.
14. Dupuy,D., Li,Q.R., Deplancke,B., Boxem,M., Hao,T., Lamesch,P., Sequerra,R., Bosak,S., Doucette-Stamm,L., Hope,I.A. *et al.* (2004) A first version of the *Caenorhabditis elegans* Promoterome. *Genome Res.*, **10B**, 2169–2175.
15. Copeland,N.G., Jenkins,N.A. and Court,D.L. (2001) Recombineering: a powerful new tool for mouse functional genomics. *Nature Rev. Genet.*, **2**, 769–779.
16. Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*. 2nd Edn, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, Plainview, NY.
17. Lee,E.C., Yu,D., Martinez de Velasco,J., Tessarollo,L., Swing,D.A., Court,D.L., Jenkins,N.A. and Copeland,N.G. (2001) A highly efficient *Escherichia coli*-based chromosome engineering system adapted for recombinogenic targeting and subcloning of BAC DNA. *Genomics*, **73**, 56–65.
18. Stavropoulos,T.A. and Strathdee,C.A. (2001) Synergy between *tetA* and *rpsL* provides high stringency positive and negative selection in bacterial artificial chromosome vectors. *Genomics*, **72**, 99–104.
19. Yu,D., Ellis,H.M., Lee,E.C., Jenkins,N.A., Copeland,N.G. and Court,D.L. (2000) An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **97**, 5978–5983.
20. Sulston,J. and Hodgkin,J. (1988) Methods. In Wood,W.B. (ed.), *The Nematode Caenorhabditis elegans*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, pp. 587–606.
21. Brenner,S. (1974) The genetics of *Caenorhabditis elegans*. *Genetics*, **77**, 71–94.
22. Mello,C.C., Kramer,J.M., Stinchcomb,D. and Ambros,V. (1991) Efficient gene transfer in *C.elegans*: extrachromosomal maintenance and integration of transforming sequences. *EMBO J.*, **10**, 3959–3970.
23. Hope,I.A. (1991) 'Promoter trapping' in *Caenorhabditis elegans*. *Development*, **113**, 399–408.

24. Coulson,A., Sulston,A., Brenner,S. and Karn,J. (1986) Towards a physical map of the genome of the nematode *Caenorhabditis elegans*. *Proc. Natl Acad. Sci. USA*, **83**, 7821–7825.

25. Datsenko,K.A. and Wanner,B.L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl Acad. Sci. USA*, **97**, 6640–6645.

26. Court,D.L., Swaminathan,S., Yu,D., Wilson,H., Baker,T., Bubunenko,M., Sawitzke,J. and Sharan,S.K. (2003) Mini-lambda: a tractable system for chromosome and BAC engineering. *Gene*, **315**, 63–69.

27. Lederberg,J. (1951) Streptomycin resistance; a genetically recessive mutation. *J. Bacteriol.*, **61**, 549–550.

28. Kim,U.-J., Shizuya,H., de Jong,P.J., Birren,B. and Simon,M. (1992) Stable propagation of cosmid sized human DNA inserts in an F factor based vector. *Nucleic Acids Res.*, **20**, 1083–1085.

29. Petalcorin,M.I., Joshua,G.W., Agapow,P.M. and Dolphin,C.T. (2005) The *fmo* genes of *Caenorhabditis elegans* and *C. briggsae*: characterisation, gene expression and comparative genomic analysis. *Gene*, **346**, 83–96.

30. Garofalo,A., Rowlinson,M.C., Amambua,N.A., Hughes,J.M., Kelly,S.M., Price,N.C., Cooper,A., Watson,D.G., Kennedy,M.W. and Bradley,J.E. (2003) The FAR protein family of the nematode *Caenorhabditis elegans*. Differential lipid binding properties, structural characteristics, and developmental regulation. *J. Biol. Chem.*, **278**, 8065–8074.

31. Paupard,M.C., Miller,A., Grant,B., Hirsh,D. and Hall,D.H. (2001) Immuno-EM localization of GFP-tagged yolk proteins in *C. elegans* using microwave fixation. *J. Histochem. Cytochem.*, **49**, 949–956.

32. Hartley,J.L., Temple,G.F. and Brasch,M.A. (2000) DNA cloning using *in vitro* site-specific recombination. *Genome Res.*, **10**, 1788–1795.

33. Muyers,J.P., Zhang,Y. and Stewart,A.F. (2001) Techniques: recombinogenic engineering - new options for cloning and manipulating DNA. *Trends Biochem. Sci.*, **26**, 325–331.

34. Court,D.L., Sawitzke,J.A. and Thomason,L.C. (2002) Genetic engineering using homologous recombination. *Annu. Rev. Genet.*, **36**, 361–388.

35. Zhang,Y., Buchholz,F., Muyrers,J.P. and Stewart,A.F. (1998) A new logic for DNA engineering using recombination in *Escherichia coli*. *Nature Genet.*, **20**, 123–128.

36. Polanowska,J., Martin,J.S., Fisher,R., Scopa,T., Rae,I. and Boulton,S.J. (2004) Tandem immunoaffinity purification of protein complexes from *Caenorhabditis elegans*. *Biotechniques*, **36**, 778–782.

37. Cheeseman,I.M. and Desai,A. (2005) A combined approach for the localization and tandem affinity purification of protein complexes from metazoans. *Sci. STKE.*, **266**, 1.

38. Hobert,O. (2002) PCR fusion-based approach to create reporter gene constructs for expression analysis in transgenic *C. elegans*. *Biotechniques*, **32**, 728–730.

39. Sassi,H.E., Renihan,S., Spence,A.M. and Cooperstock,R.L. (2005) Gene CATCHR-gene cloning and tagging for *Caenorhabditis elegans* using yeast homologous recombination: a novel approach for the analysis of gene expression. *Nucleic Acids Res.*, **33**, e163.

40. Reyrat,J.M., Pelicic,V., Gicquel,B. and Rappuoli,R. (1998) Counterselectable markers: untapped tools for bacterial genetics and pathogenesis. *Infect. Immun.*, **66**, 4011–4017.

41. Jamsai,D., Orford,M., Nefedov,M., Fucharoen,S., Williamson,R. and Ioannou,P.A. (2003) Targeted modification of a human beta-globin locus BAC clone using GET recombination and an I-SceI counterselection cassette. *Genomics*, **82**, 68–77.

42. Warming,S., Costantino,N., Court,D.L., Jenkins,N.A. and Copeland,N.G. (2005) Simple and highly efficient BAC recombineering using *galK* selection. *Nucleic Acids Res.*, **33**, e36.

43. Wong,Q.N., Ng,V.C., Lin,M.C., Kung,H.F., Chan,D. and Huang,J.D. (2005) Efficient and seamless DNA recombineering using a thymidylate synthase A selection system in *Escherichia coli*. *Nucleic Acids Res.*, **33**, e59.

44. Dunn,W., Chou,C., Li,H., Hai,R., Patterson,D., Stolc,V., Zhu,H. and Liu,F. (2003) Functional profiling of a human cytomegalovirus genome. *Proc. Natl Acad. Sci. USA*, **100**, 14223–14228.

45. Praitis,V., Casey,E., Collar,D. and Austin,J. (2001) Creation of low-copy integrated transgenic lines in *Caenorhabditis elegans*. *Genetics*, **157**, 1217–1226.