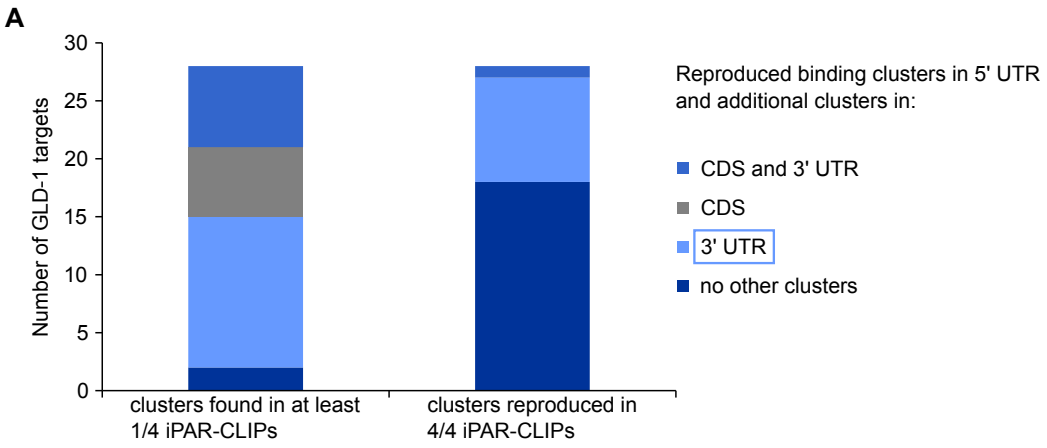
Cell Reports, Volume 22

Supplemental Information

Post-transcriptional Regulation by 3′ UTRs Can Be Masked by Regulatory Elements in 5′ UTRs

Kathrin Theil, Margareta Herzog, and Nikolaus Rajewsky



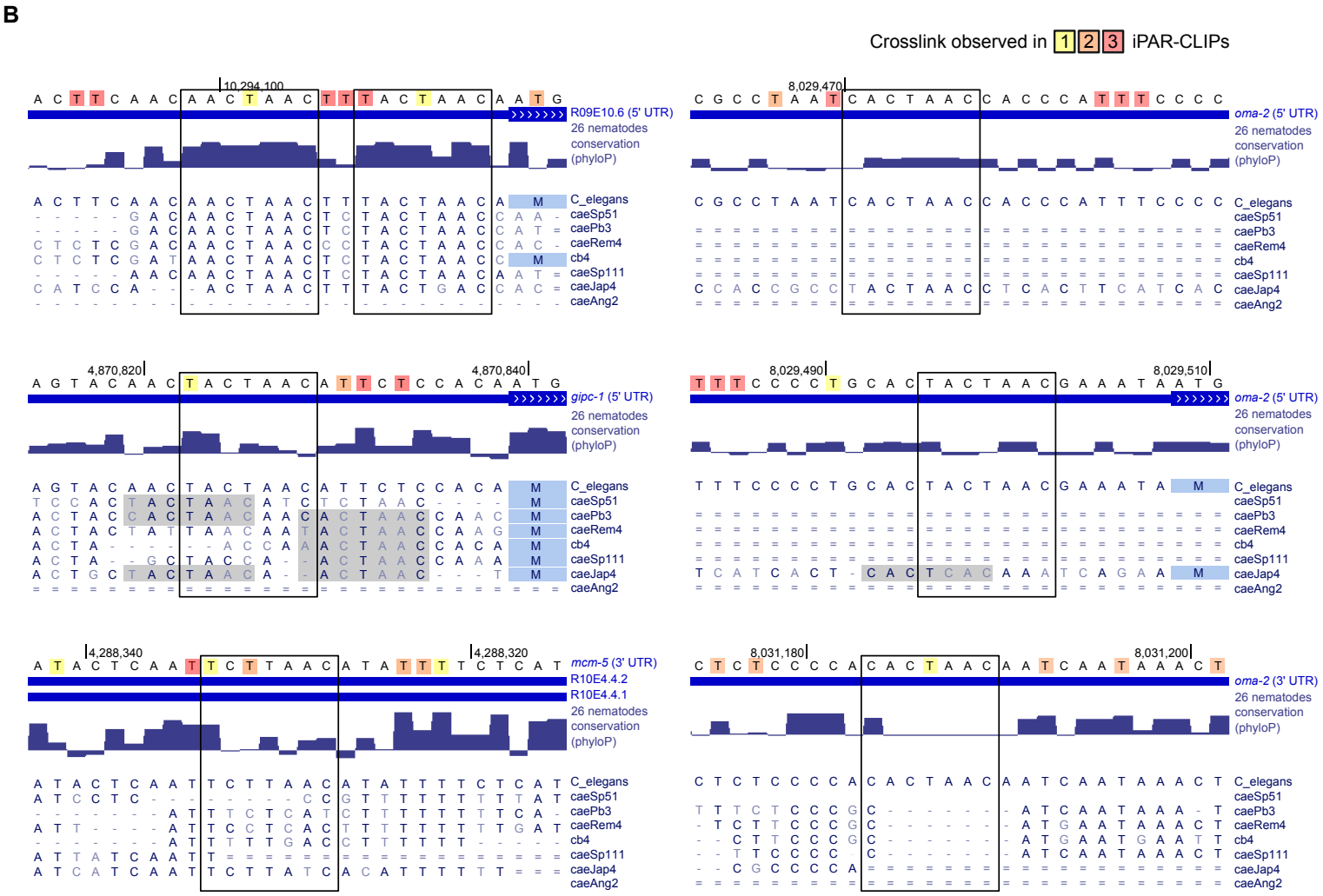


Figure S1. GLD-1 binds to conserved and non-conserved sites within *C. elegans* transcript UTRs. Related to Figure 1.

(A) Number of GLD-1 targets that have reproduced clusters in 5' UTRs (4/4 iPAR-CLIP experiments) and additional reproduced or non-reproduced clusters in other transcript regions (Jungkamp *et al*., 2011).

(B) GLD-1 binding sites investigated in this study (see also **Table 1** and **Figure 1B**). Crosslinking sites from iPAR-CLIP experiments, indicating direct protein-RNA contacts, were found in immediate vicinity of GLD-1 binding motifs (Jungkamp *et al*., 2011). Genome browser screen shots adapted from http://genome.ucsc.edu (Kent *et al*., 2002). Boxes indicate GLD-1 binding motifs; shifted motifs are highlighted in grey.

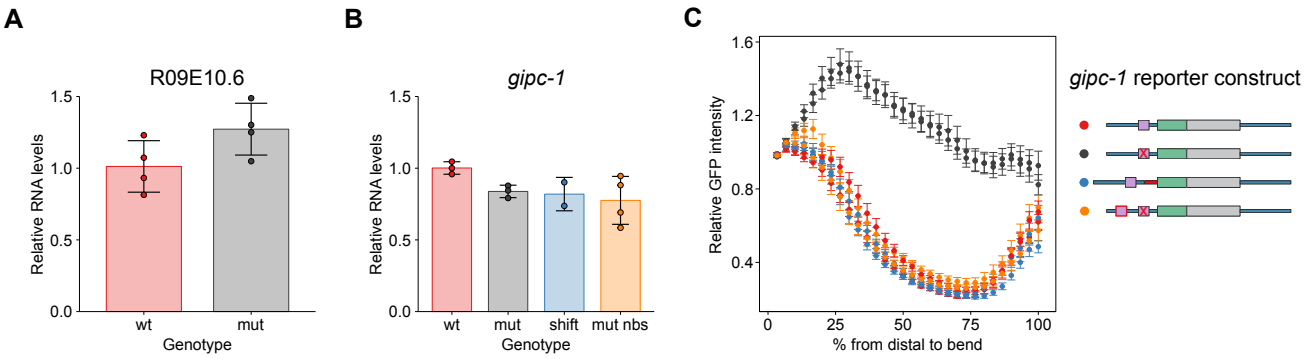


Figure S2. Reporters with single 5' UTR binding sites show comparable RNA levels and reproducible GFP expression along the germline. Related to Figure 2.

(A and B) Reporter RNA levels, as determined by RT-qPCR, for the R09E10.6 (**A**) and *gipc-1* (**B**) constructs. C_t values were normalized to three reference genes (*pmp-3*, *tba-1*, Y45F10D.4; Zhang *et al*., 2012), and are plotted as ratio to wild-type reporter levels. Individual data points correspond to biological replicates of the same or independent transgenic lines. Bars depict mean ± SD.

(C) Quantification of *gipc-1* reporter expression (analogous to **Figure 2E**) along the germline resulted in highly reproducible patterns for independent lines of the same constructs (plotted in same color). Measured GFP pixel intensities were normalized to the mean of the first 5 % of each gonad (where GLD-1 is not expressed), and grouped into 30 bins. Displayed are mean ± SEM for each bin. Per reporter construct, two independent transgenic lines were analyzed (number of gonads per line: wt, n = 23; mut, n = 28; mut nbs, n = 22; shift, n = 25).

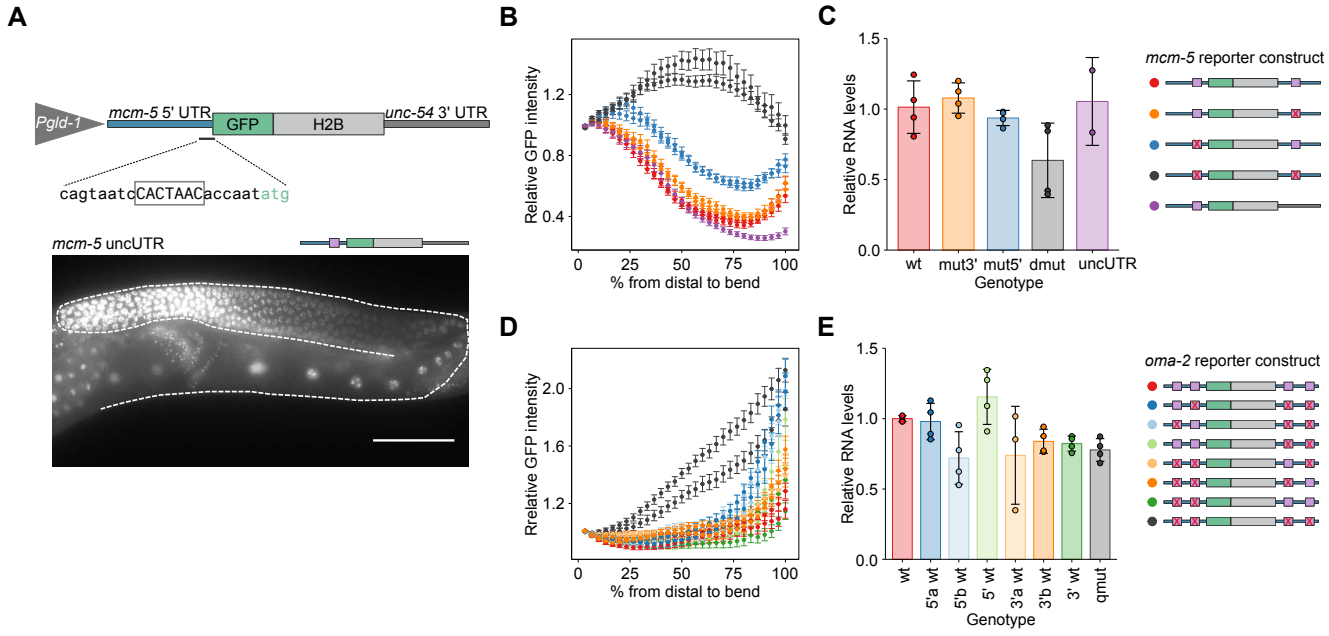


Figure S3. The *mcm-5* 5' UTR is sufficient for regulation, and independent lines of 5' and 3' UTR reporter constructs show reproducible expression. Related to Figure 3 and Figure 4.

(A) Replacement of the endogenous *mcm-5* 3' UTR by a non-regulated 3' UTR retains regulation. Transgene design with blow-up of GLD-1 binding site. *Pgld-1, gld-1* promoter; capital letters indicate GLD-1 binding motif; box indicates intact site. Dashed line encircles gonad. Scale bar: 50 µm.

(B and D) Quantifications of reporter expression (analogous to **Figure 2E**) along the germline for independent lines (plotted in same color) of the *mcm-5* (**B**) and *oma-2* (**D**) constructs show reproducible patterns. Displayed are mean ± SEM for each bin. Per reporter construct, two independent transgenic lines were analyzed, except for *mcm-5* uncUTR where only one line was analyzed (number of gonads per line of *mcm-5* reporters: wt, n = 28; dmut, n = 32; mut3', n = 27; mut5', n = 24; uncUTR, n = 31; number of gonads analyzed per line of *oma-2* reporters: wt, n = 27; qmut, n = 32; 5'a wt, n = 28; 5'b wt, n = 29; 3'a wt, n = 24; 3'b wt, n = 28; 3' wt, n = 28; 5' wt, n = 20). Colors as in **C** and **E**.

(C and E) Reporter RNA levels, as determined by RT-qPCR, for the *mcm-5* (**C**) and *oma-2* (**E**) constructs. C_t values were normalized to three reference genes (*pmp-3*, *tba-1*, Y45F10D.4; Zhang *et al*., 2012), and are plotted as ratio to wild-type reporter levels. Individual data points correspond to biological replicates of the same or independent transgenic lines. Bars depict mean ± SD.

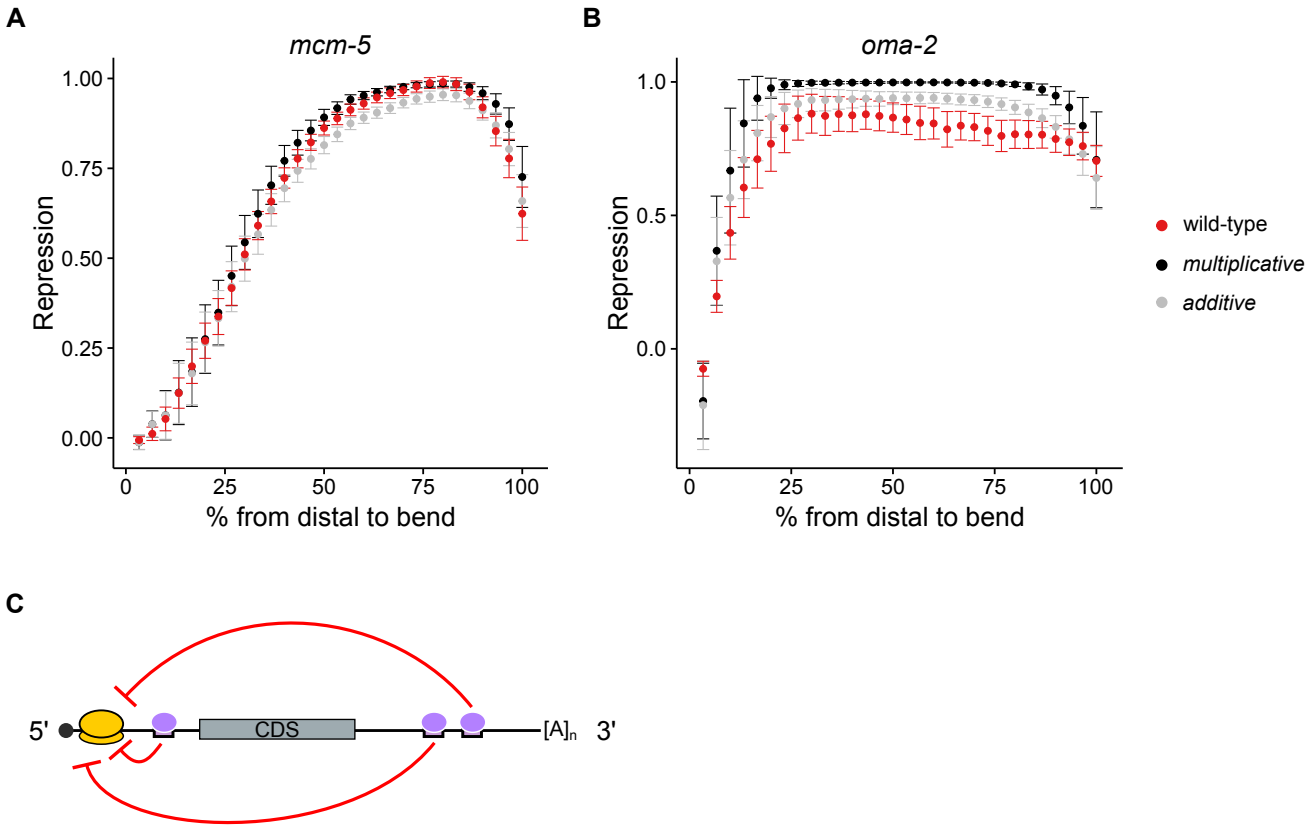


Figure S4. Modeling of repression suggests that sites act independently. Related to Figure 3 and Figure 4.
(A and B) Repression was calculated by assuming (1) complete repression at minimal GFP signal for each wild-type reporter, and (2) GLD-1-independent expression for reporters without GLD-1 sites. Wild-type: calculated experimental repression, *multiplicative*: theoretical wild-type reporter repression if repression was conferred independently by both sites, *additive*: theoretical wild-type reporter repression if binding of additional sites did not result in increased repression. For detailed explanation of calculations, see the **Supplemental Experimental Procedures**. Displayed are mean ± SEM.
(C) Sites in both 5' and 3' UTRs can potently repress translation. CDS: coding sequence.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Raw data, scripts, reporter constructs and worm strains are available upon request.

Estimation of repression

To estimate the degree of repression, we made the following assumptions: (1) RNA levels are the same along the germline for different reporter constructs of the same target (see **Figures S3C** and **S3E**). (2) repression reaches its maximum in the medial gonad of wild-type reporters, and remaining signal is due to autofluorescence (baseline (b) = complete repression, for *oma-2*: b = 0.9, for *mcm-5*: b = 0.35), and (3) the *mcm-5* double mutant and the *oma-2* quadruple mutant ("no site" mutants) are expressed independently of GLD-1. Normalization of relative GFP intensities of reporters with intact binding sites to the respective "no site" mutant yields GLD-1-dependent repression. Repression fold changes (f) along the germline were calculated for each reporter with intact GLD-1 site(s) as follows (E: relative GFP expression of respective reporter):

$$f = (E["no\ site"] - b)/(E["intact\ site(s)"] - b)$$

Based on repression fold changes of "single site" reporters, theoretical wild-type repression along the germline was calculated, assuming either multiplicative (independent repression by multiple sites) or additive repression (binding of additional sites within a transcript already bound does not lead to increased repression).

Multiplicative repression:

mcm-5:

$$f["wild\text{-}type"] = f["intact\ 5'\ site"] \times f["intact\ 3'\ site"]$$

oma-2 :

$$f["wild\text{-}type"] = f["intact\ 5'a\ site"] \times f["intact\ 5'b\ site"] \times f["intact\ 3'a\ site"] \times f["intact\ 3'b\ site"]$$

Additive repression:

mcm-5:

$$f["wild\text{-}type"] = f["intact\ 5'\ site"] + f["intact\ 3'\ site"] - 1$$

oma-2 :

$$f["wild\text{-}type"] = f["intact\ 5'a\ site"] + f["intact\ 5'b\ site"] + f["intact\ 3'a\ site"] + f["intact\ 3'b\ site"] - 3$$

Oligonucleotides for amplification, mutagenesis and sequencing

Restriction sites and introduced mutations are in capital letters.

Promoter:

Gene	Description	Sequence
gld-1	fwd	actACTAGTattgagatacacaagtgtttttta
	rev	actAGATCTtcttcgatggttaacctgtaag
	rev2[a]	actAGATCTctgtaagaaaacttagtaataatatattttc

[a]used instead of "rev" for *oma-2* reporters and for *mcm-5* uncUTR reporter; generates a 15 nt-shortened *gld-1* promoter sequence (→ *gld*-1 promoter without annotated *gld-1* 5' UTR)

3' UTRs:

Gene	Description	Sequence
R09E10.6	fwd	actCCTAGGcgctctttgcgttgctagac
	rev	actCTTAAGtggttccatttcacacttcc
gipc-1	fwd	actCCTAGGcttttactgtggaataaattattttaatc
	rev	actCTTAAGaactgggaatcgatgaatgc
mcm-5	fwd	actCCTAGGattctcaaaattttgtctttttatctg
	rev	actCTTAAGaaactagaatttgattatttttcaagg
unc-54	fwd	actCCTAGGgtccaattactcttcaacatcc
	rev	actCTTAAGaaacagttatgtttggtatattggg
oma-2	fwd	actCCTAGGataaccacctccttcccca
	rev	actCTTAAGccctgtcgcgaaactaattg

5' UTR::GFP::H2B fusions and mutagenesis (Hobert, 2002):

Gene	Description	Sequence
GFP::H2B	primer D	ttgtccggcagtgcactgatgataa
	primer D*	actCCTAGGttacttgctggaagtgtacttggtgacg
	primer C	atgagtaaaggagaagaacttttc
R09E10.6	primer A	cgtgaaagcctattcaagtttc
	primer A*	actAGATCTcagtactagaaaattttctcgacttc
	primer B	gaaaagttcttctcctttactcattgttagtaaagttagttgttgaag
	primer B with mutation	gaaaagttcttctcctttactcattgCtCgtaaagctcgttgttgaa
gipc-1	primer A	cagcgtggcactctaattt
	primer A*	actAGATCTtgattcattcatcgatttaacacc
	primer B	gaaaagttcttctcctttactcattgtggagaatgttagtagttg
	primer B with mutation	gaaaagttcttctcctttactcattgtggagaatgCtCgtagttg
mcm-5	primer A	ccgcattctgttttatttcca
	primer A*	actAGATCTatggtttaaaatttcagtaatccac
	primer B	gaaaagttcttctcctttactcatattggtgttagtggattactg
	primer B with mutation	gaaaagttcttctcctttactcatattggtgCtCgtggattactg
oma-2	primer A	ttataccgttttggcgcatt
	primer A*	actAGATCTgggaatcaattaaccgagcac
	primer B	gaaaagttcttctcctttactcattatttcgttagtagtgcagg

Site-directed mutagenesis (3' UTR):

Gene	Description	Sequence
mcm-5	fwd	ctttctcatactcaattctGaGcatattttctcattcaatttg
	rev	caaattgaatgagaaaatatgCtCagaattgagtatgagaaag
oma-2 (3'a site)	fwd	ccatacgtttttatgtcccatacGaGctataataacgcgcgcaattt
	rev	aaattgcgcgcgttattatagCtCgtatgggacataaaaacgtatgg
oma-2 (3'b site)	fwd	cgtatcctctccccacacGaGcaatcaataaactttatcgaatgtca
	rev	tgacattcgataaagtttattgattgCtCgtgtggggagaggatacg

Site-directed mutagenesis (5' UTR):

Gene	Description	Sequence
gipc-1 (mut nbs)	fwd	tcttgattcattcatcgTACtaacaccaactcacttc
	rev	gaagtgagttggtgttaGTAcgatgaatgaatcaaga
gipc-1 (shift)	fwd	acaactactaacattctccacaATCGATTTAACAatgagtaaaggagaagaacttttc
	rev	gaaaagttcttctcctttactcatTGTTAAATCGATtgtggagaatgttagtagttgt
oma-2 (5'a site)	fwd	caccccacgcctaatcacGaGccacccatttcccctgca
	rev	tgcaggggaaatgggtggCtCgtgattaggcgtggggtg
oma-2 (5'b site)	fwd	cacccatttcccctgcactacGaGcgaaataatgagtaaaggagaagaac
	rev	gttcttctcctttactcattatttcgCtCgtagtgcaggggaaatgggtg

Verification of integration on chromosome II (Frøkjær-Jensen *et al.*, 2008):

Region	Name	Sequence
Insertion site	oCF418	tctggctctgcttcttcgtt
Transgene	oCF419	caattcatcccggtttctgt

Sequencing:

Region	Description	Sequence
5' UTR, CDS, 3' UTR	fwd	gcagagtttaagcttatgtcg
	rev	cgctgtcctgtcacactc
gld-1 promoter	fwd	tacagaagaccgttacgaaac
	rev	gaattgggacaactccagtg

qPCR:

Gene	Description	Sequence
gfp (1)	fwd	ggccctgtccttttaccagac
	rev	tcgtccatgccatgtgtaatc
gfp (2)	fwd	tggaaacattcttggacacaa
	rev	tcgccaattggagtattttg
pmp-3 (control)	fwd	gttcccgtgttcatcactcat
	rev	acaccgtcgagaagctgtaga
tba-1 (control)	fwd	gtacactccactgatctctgctgacaag
	rev	ctctgtacaagaggcaaacagccatg
Y45F10D.4 (control)	fwd	gtcgcttcaaatcagttcagc
	rev	gttcttgtcaagtgatccgaca