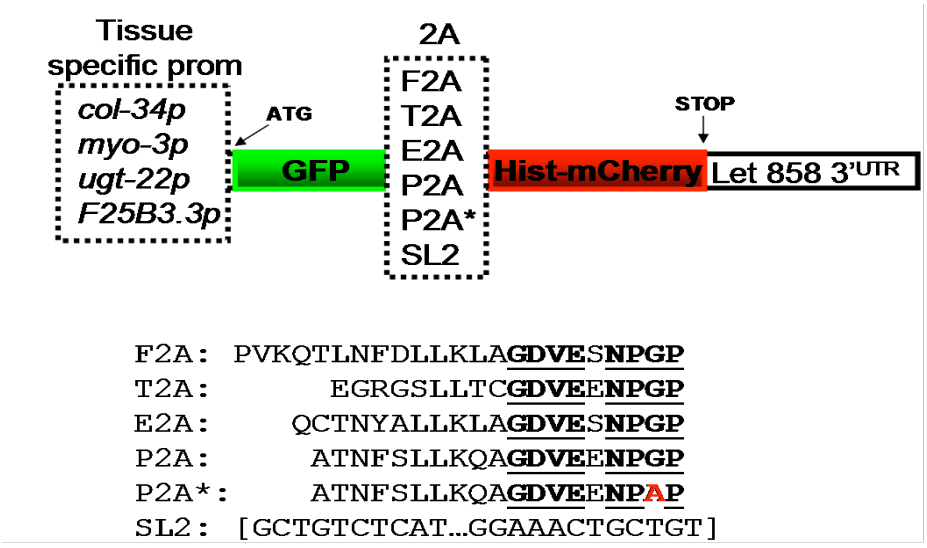
a-



b-

2a sequence	**correct localization**	2 markers in the nucleus	2 markers in the whole cell	GFP only expressed	mCherry only expressed
F2A (n=104)	98,15%±1,66%	0%±0%	1,08%±1,86%	0%±0%	0,78%±1,34%
T2A (n=115)	96,86%±2,73%	0%±0%	0%±0%	1,63%±2,82%	1,52%±2,62%
E2A (n=120)	94,35%±8,28%	0%±0%	0%±0%	0%±0%	5,65%±8,28%
P2A (n= n=144)	88,11%±3,05%	9,62%±4,29%	2,27%±2,71%	0%±0%	0%±0%
P2A* (n=96)	4,30%±7,45%	94,62%±9,31%	1,08%±1,86%	0%±0%	0%±0%
SL2 (n=150)	95,31%±1,92%	0%±0%	0%±0%	0,98%±1,70%	3,71%±1,75%

Figure S1 **a-** Schematic representation of the constructs generated to test the canonical 2A (F2A, T2A, E2A and P2A) peptides in *C. elegans*. P2A* (Hahn and Palmenberg 1996) was used as a negative control and SL2 (Macosko *et al.* 2009; Kagias *et al.* 2012) was used as a positive control. **b-** Quantification of the number of cells exhibiting a given sub-cellular localization pattern for the GFP and HISTONE::mCherry proteins in transgenic lines expressing each of the 4 canonical 2A sequences ("correct localization", GFP in the whole cell, HISTONE::mCherry in the nucleus). The very high % of cells exhibiting the expected localization pattern suggests a remarkable efficiency in *C. elegans*. In addition, note that the "self-cleavage" efficiency is not correlated with the length of the 2A peptide used. Although still very robust, we found the peptide P2A to be slightly less efficient in generating two distinct sub-cellular localizations for GFP and mCherry compared to the other 2A peptides; this however may not be due to a lesser production of 2 independent polypeptides (see Supp Fig. 2A and the western blots in Fig. 1). As expected, the "cleavage-inefficient" P2A* variant leads to the colocalization of the two fluorescent proteins in the nucleus. n= total number of cells scored.

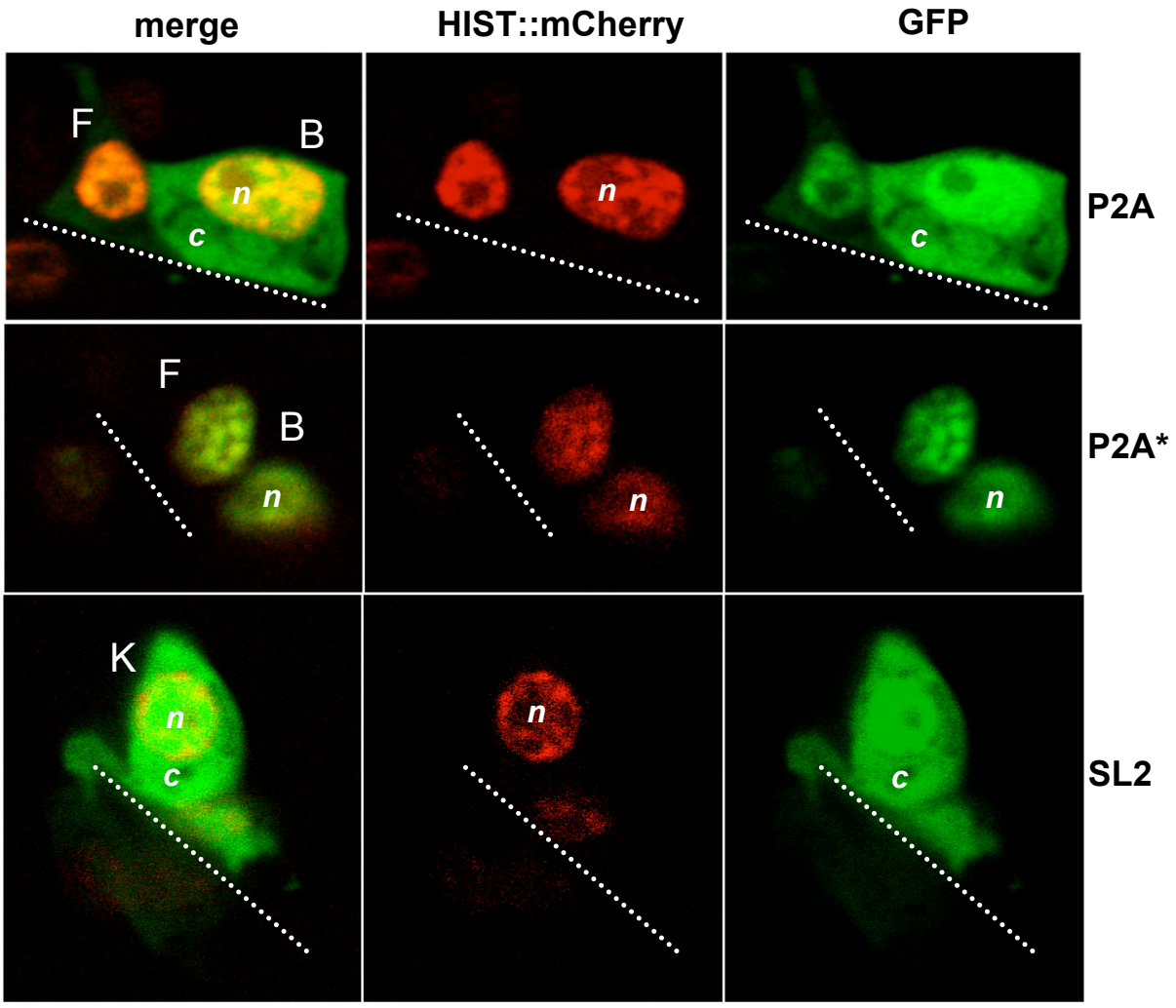


Figure S2 Differential co-localization of both fluorescent proteins, GFP and HISTONE::mCherry, expressed from the P2A vector described in Fig. S1 (with *col-34p*), is confirmed by confocal SP2 microscopy in the B, F and K rectal cells. We noted that the P2A sequence (top panels) appeared to result in slightly more GFP in the nucleus compared to the other 2A peptides instead of being uniformly spread within the whole cell. Confocal imaging shows here that some of the nuclear GFP does not co-localize with mCherry in the nucleus, suggesting that GFP products that have been split from mCherry can be retained in the nucleus. The same observations were made in transgenic lines expressing a SL2 construct (bottom panels), where substantial GFP is found in the nucleus, in a pattern that distinct from HISTONE::mChrerry. Middle panels: split-up-inactive P2A* peptide (Hahn and Palmenberg 1996). Both fluorescent proteins are exclusively co-localized in the nucleus, as expected if they form one long polypeptide. The dashed line indicates the position of the rectum; anterior is to the left and ventral to the bottom; (n), nucleus and (c) cytosol; the rectal cells are named on the merge panel.

Developmental stage	bean to 2-folds stage	L1	L4
E2A	95,62%±9,42% (n=186)	99,61%±0,88% (n=251)	100%±0% (n=131)
P2A*	0%±0% (n=327)****	0%±0% (n=200)	0%±0% (n=150)
SL2	97,34%±5,13% (n=177)	100%±0% (n=150)	100%±0% (n=117)

***** note that for ±27,5% of observed cases mCherry is barely visible; however GFP remains strictly restricted to the nucleus meaning that GFP remains fused to HIST-mCherry.

Figure S3 Quantification of the expected localization pattern (GFP in the whole cell, HISTONE::mCherry in the nucleus) observed for lines expressing *myo-3p::GFP::E2A::Histone::mCherry*, *myo-3p::GFP::P2A*::Histone::mCherry* or *myo-3p::GFP::SL2::Histone::mCherry* at different developmental stages (see also Fig. 2A); n= total number of cells scored.

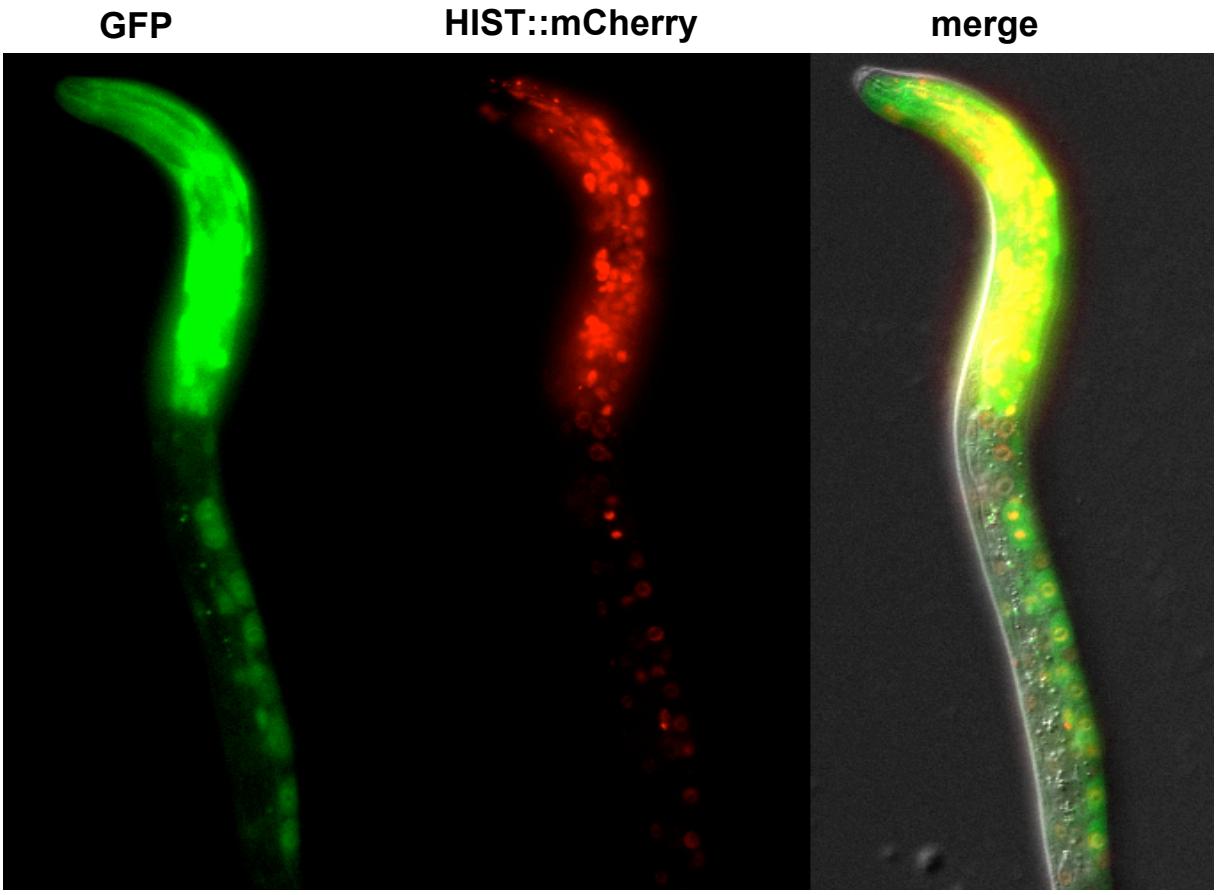
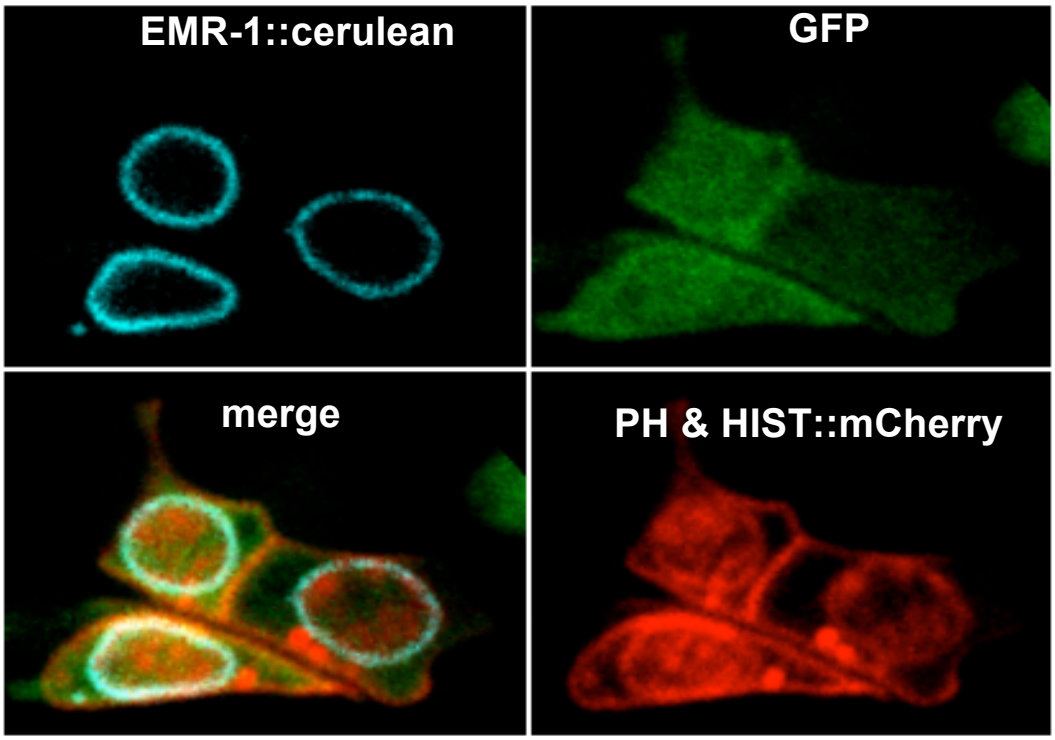


Figure S4 The expected localization of GFP and mCherry is observed in all cells analyzed in transgenic animals ubiquitously expressing a *his-72p::GFP::F2A::Histone::mCherry* construct. GFP localizes within the whole cell whereas HISTONE::mCherry remains tightly restricted to the nucleus. An L1 worm, vertically oriented and tail at the bottom, is shown.

a-

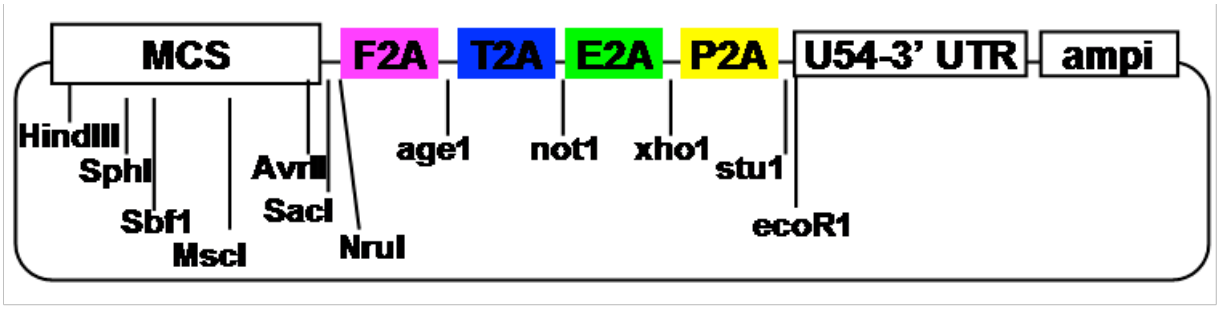


b-

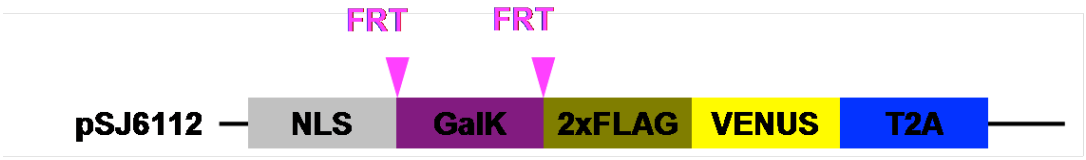
line/localization	GFP in the whole cell	mCherry::PH at the cell membrane	EMR-1::cerulean at the nuclear membrane	HISTONE::mCherry in the nucleus
IS2008 (n=267)	100%±0%	100%±0%	99,26%±1,28%	96,69% ±0,36%
IS2070 (n=207)	100%±0%	100%±0%	100%±0%	99,04% ±0,83%

Figure S5 **a-** Additional line IS2008 expressing 5 different functional products in the Y, B and F rectal cells. All the markers are addressed to the expected cellular compartments: whole cell (GFP), nucleus and plasma membrane (mCherry) and nuclear envelope (cerulean). In addition, note that the rescuing efficiencies obtained with this line or IS2070 are similar or higher to those obtained with *col-34p::sem-4a::HA::SL2::mCherry* [17,25%±2,49%; n=192 (IS1648) and 12,05%±2,5%; n=126 (IS1649)](Kagias *et al.* 2012). Anterior is to the left and ventral to the bottom. **b-** Quantification of the expected localization for each of the fluorescent protein, in the two transgenic lines IS2008 and IS2070; n= total number of cells scored.

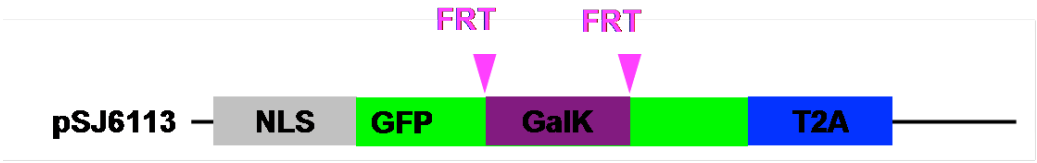
a-



b-



c-



d-

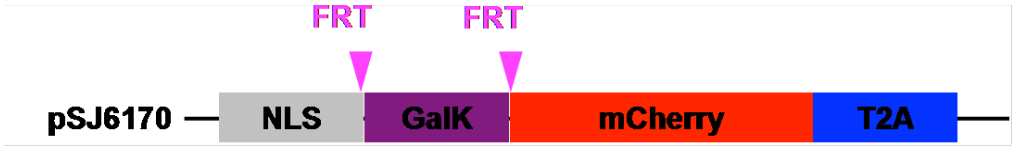


Figure S6 2A TOOLKIT for *C. elegans*. **a-** Map of pSJ6186 *[MCS ::F2A ::T2A ::E2A ::P2A ::U543' UTR].* **b-** Map of pSJ6112 *[NLS ::FRT ::galK ::FRT ::2xFLAG ::VENUS ::T2A]* [Modified from (Tursun *et al.* 2009)]. **c-** Map of pSJ6113 *[NLS::GFP(part1) ::FRT ::galK ::FRT::GFP(part2) ::T2A]* [Modified from (Tursun *et al.* 2009)]. **d-** Map of pSJ6170 *[NLS ::FRT ::galK ::FRT ::mCherry ::T2A]* [Modified from (Tursun *et al.* 2009)]. *(Sequences available on demand).*

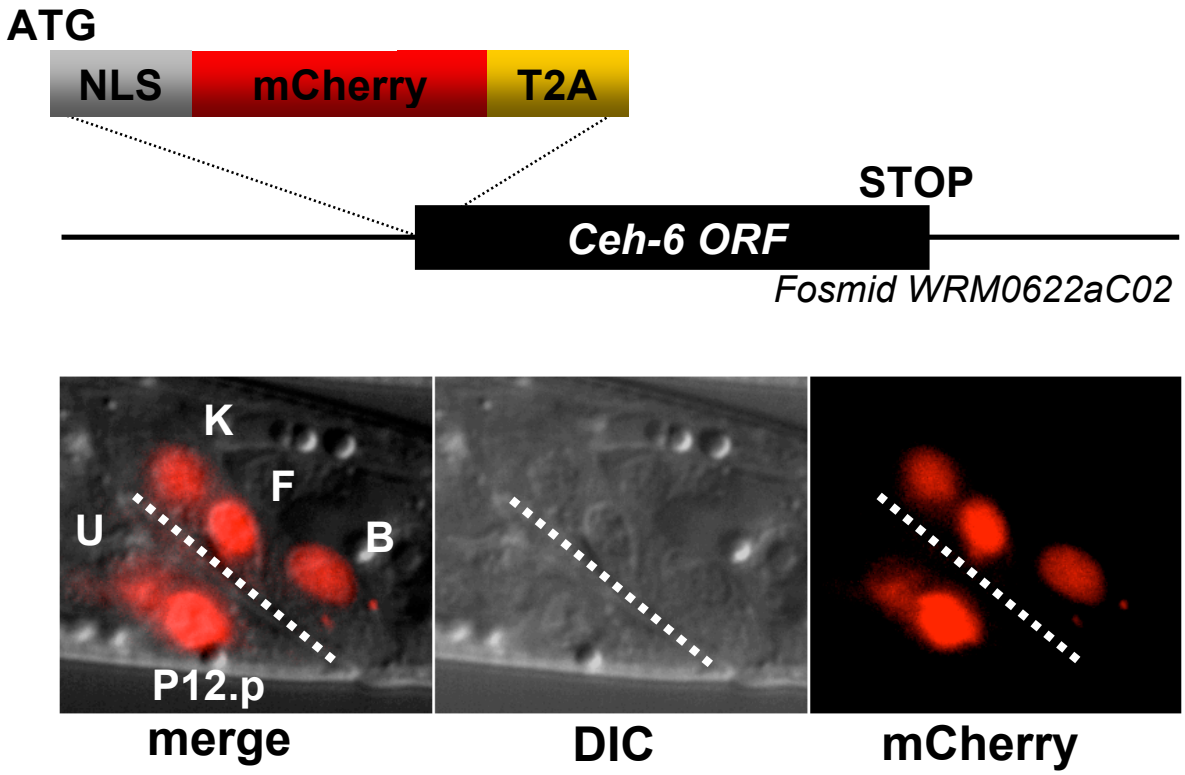


Figure S7 Schematic representation of the 2A-modified fosmid *WRM0622aC02* encompassing the *ceh-6* genomic locus and the expression it drives in rectal cells (shown are the B, P12.pa, U, F and K rectal cells) and that mimics the expression pattern observed after antibody staining for CEH-6 (Burglin and Ruvkun 2001). Anterior is to the left and ventral to the bottom.

Table S1 List of the molecular constructs generated for this study

Name	R.	Backbone	Plasmid description
pSJ6171	Ampi	pUC57	*Col-34p ::GFP ::F2A ::Hist-24 ::mCherry ::let 858 3' UTR*
pSJ6172	Ampi	pUC57	*Col-34p ::GFP ::T2A ::Hist-24 ::mCherry ::let 858 3' UTR*
pSJ6173	Ampi	pUC57	*Col-34p ::GFP ::E2A ::Hist-24 ::mCherry ::let 858 3' UTR*
pSJ6174	Ampi	pUC57	*Col-34p ::GFP ::P2A ::Hist-24 ::mCherry ::let 858 3' UTR*
pSJ6175	Ampi	pUC57	*Col-34p ::GFP ::P2A* ::Hist-24 ::mCherry ::let 858 3' UTR*
pSJ6176	Ampi	pUC57	*Col-34p ::GFP ::SL2 ::Hist-24 ::mCherry ::let 858 3' UTR*
pSJ6177	Ampi	pUC57	*Myo-3p ::GFP ::F2A ::Hist-24 ::mCherry ::let 858 3' UTR*
pSJ6178	Ampi	pUC57	*Myo-3p ::GFP ::T2A ::Hist-24 ::mCherry ::let 858 3' UTR*
pSJ6179	Ampi	pUC57	*Myo-3p ::GFP ::E2A ::Hist-24 ::mCherry ::let 858 3' UTR*
pSJ6180	Ampi	pUC57	*Myo-3p ::GFP ::P2A ::Hist-24 ::mCherry ::let 858 3' UTR*
pSJ6181	Ampi	pUC57	*Myo-3p ::GFP ::P*2A ::Hist-24 ::mCherry ::let 858 3' UTR*
pSJ6182	Ampi	pUC57	*Myo-3p ::GFP ::SL2 ::Hist-24 ::mCherry ::let 858 3' UTR*
pSJ6183	Ampi	pjet1.2™	*F25B3.3p ::GFP :: T2A ::Hist-24 ::mCherry ::let 858 3' UTR*
pSJ6187	Ampi	pjet1.2™	*Ugt-22p ::GFP :: T2A ::Hist-24 ::mCherry ::let 858 3' UTR*
pSJ6188	Ampi Kana	Strataclone™	*His-72p ::GFP :: F2A ::Hist-24 ::mCherry ::his-72 3' UTR*
pSJ6184	Ampi	pUC57	*Col-34p ::GFP ::F2A ::mCherry ::PH ::T2A ::sem-4a^{HA}::E2A ::emr-1 ::cerulean::P2A::Hist-24 ::mCherry ::let 858 3' UTR*
pSJ6185	ChloR	WRM0622aC02 fosmid	*NLS ::mCherry ::T2A ::ceh-6 gene*
pSJ6112	Ampi Kana	TOPO pcrII™	*NLS ::FRT ::galK ::FRT ::2xFLAG ::VENUS ::T2A [TOPO pcrII – Invitrogen©]*
pSJ6113	Ampi Kana	TOPO pcrII™	*NLS ::GFP(part1) ::FRT ::galK ::FRT ::GFP(part2) ::T2A [TOPO pcrII – Invitrogen©]*
pSJ6170	Ampi Kana	TOPO pcrII™	*NLS ::FRT ::galK ::FRT ::mCherry ::T2A [TOPO pcrII – Invitrogen©]*
pSJ6186	Ampi	pPD95.75	*MCS ::F2A ::T2A ::E2A ::P2A ::U543' UTR*

Table S2 List of primers used for this study

N°	name	sequence
1	acc65i gfpf	aacggtaccagaaaaaatgagtaaaggagaag
2	acc65i gfpr	ttgggtacctttgtatagttcatccatgcc
3	h1bmchlet868f	cccgggccttctagaatgtctgattccgctgttg
4	h1bmchlet868r	cagtcgacgggcccggctgggtaccgggcc
5	f2ar	gggccctgggttggactc
6	h1bf	atgtctgattccgctgttgttg
7	gfpunir	ggtacctttgtatagttcatcca
8	p2af	gccacgaacttctctctgtt
9	e2af	caatgtactaactacgctttg
10	e2ar	aggaccggggttactttca
11	t2af	gagggcagaggaagtctgct
12	t2ar	tgggccaggattctcctcg
13	p²amutf	gaagaaaaccccgcgccttctagaatgt
14	p²amutr	acattctagaaggcgcggggttttcttc
15	sl2f	ggcatggatgaactatactaagctgtctcatcctactttcac
16	sl2r	caacaacagcggaatcagacatacagcagtttccctgaattaaa
17	myo3pftepf	gtaaaacgacggccagtgaattcgatcggctataataagttcttgaataaaa
18	myo3ftepr	ggtaccgagctcgaattcactagtgattctagatggatctagtggtcgtgg
19	m13rp	caggaaacagctatgacc
20	gfpmiddlef	gatggaaacattcttggacac
21	let868spe1r	atacactagtgacgggcccgggatccgat
22	h1bmcherryr	atccccgcgggctgggtaccgggcccct
23	f25b3.3f	ggctgaaatcactcacaacgatg
24	bld204	aatttttcatcagatctagtca
25	age1mchephf	cccaccggtatggtctcaaagggtgaagaag
26	age1mchephr	ctcaccggtcttctgccgctggtccatgg
27	e2aemr-1f	cgatgttgaaagtaaccccggtcctatggacgtctcccagctgacag
28	ceremr-1r	ccagtgaaaaggtcttctcctttactcataatagtatcctccggattcgtctcg
29	cer1-p2ar	gctttaacagagagaagttcgtggcggatccacgcgtactagttttg

30	e2aemr-1f	cgatgttgaaagtaaccccggtcctatggacgtctcccagctgacag
31	ceremr-1r	ccagtgaaaaggtcttctcctttactcataatagtatcctccggattcgtctcg
32	sem-4anotf	agcggccgcaatgaatgagctgctcgccgag
33	HAnotr	tgcggccgcagcataatcaggaacatcatacgg
34	nls-f	aatgaccgctccaaagaagaaacgc
35	gfp-t2ar	tgggccaggattctcctcgacgtcaccgcatgttagcagacttcctctgccctctttgtatagttcatccatgccatgtg
36	venus-t2ar	tgggccaggattctcctcgacgtcaccgcatgttagcagacttcctctgccctccttgtacagctcgtccatgccgag
37	mcherry-t2ar	tgggccaggattctcctcgacgtcaccgcatgttagcagacttcctctgccctccttgtacagctcgtccatgccgcc
38	ceh6-nlsf	ccatcttttccacagtagtggccaccccgttgctcctggacttcccaactgatgaccgctccaaagaagaaacgc
39	ceh6-t2ar	tgctgaagcagataaagaagacggtatggatgacgacgaaggtatgagcattgggccaggattctcctcgac
40	his72T2AFus1	cctgttcgttgcaacaaattgat
41	his72T2AFus2	gaaaagttcttctcctttactcattgttgttctggaaattgagaattg
42	his72T2AFus3	tttccagaacaacaatgagtaaaggagaagaacttttcac
43	his72T2AFus5	gcaatgctttttataatgccaac
44	his72T2AFus6	gagaattggtgatggagcttactacttatacaattcatccatgccacc
45	his72T2AFus7	gaattgtataagtagtaagctccatcaccaattctcgaag
46	his72T2AFus9	gcaaacgttatagtgtggacacc
47	his72T2AFus10	cacgcaacgcgccgtaaacctac

Table S3 List of strains/array generated for this study

	col-34 promoter	*myo-3 promoter*	*ugt-22 promoter*	*F25b3. 3 promoter*
SL2	IS1942 fpEx558	IS2041 fpEx628 *W.B*		
	IS1943 fpEx559	IS2090 fpEx667		
	IS1944 fpEx560	IS2092 fpEx669		
F2A	IS1946 fpEx562	IS2003 fpEx600 *W.B*		
	IS1947 fpEx563	IS2077 fpEx658		
	IS1948 fpEx564	S2091 fpEx668		
T2A	IS1949 fpEx565	IS2071 fpEx653	IS1980 fpEx584 *W.B*	IS1982 fpEx585
	IS1950 fpEx566	IS2073 fpEx654 *W.B*	IS2075 fpEx656	IS2068 fpEx650
	IS1951 fpEx567	IS2074 fpEx655	IS2076 fpEx657	IS2069 fpEx651
		IS2078 fpEx659		
E2A	IS1952 fpEx568	IS2004 fpEx601 *W.B*		
	IS1953 fpEx569	IS2058 fpEx640		
	IS1954 fpEx570	IS2059 fpEx641		
	IS1955 fpEx571			
P2A	IS1945 fpEx561	IS2031 fpEx619		
	IS2005 fpEx602	IS2053 fpEx635 *W.B*		
	IS2006 fpEx603	IS2060 fpEx642		
		IS2064 fpEx646		
		IS2066 fpEx648		
P2A*	IS2061 fpEx64*d*	IS2028 fpEx617 *W.B*		
	IS2062 fpEx644	IS2029 fpEx618		
	IS2065 fpEx647			

W.B: strain used for Western-Blot analysis

Strains obtained by injecting pSJ6184 :

IS2008 fpEx605

IS2070 fpEx252

Strain obtained by injecting pSJ6185 :

IS2067 fpEx649

Strain obtained by injecting pSJ6188 :

IS2279 fpEx791

Acknowledgements

We thank P. Meister, O. Hobert, S. Kim, C. Bargmann and P. Sengupta for kindly providing DNA constructs. We thank D. Moerman for kindly providing fosmid reagents from the *C. elegans* Fosmid Library*.

* Jaryn Perkins[1,2], Kim Wong[3], René Warren[3], Jacquie Schein[3], Jeff Stott[3], Rob Holt[3], Steve Jones[3], Marco Marra[3], and Don Moerman[1,2]

[1]Michael Smith Laboratories, U.B.C., Vancouver, B.C., Canada V6T 1Z4

[2]Department of Zoology, University of British Columbia, B.C. Vancouver, Canada V6T 1Z4

[3]Genome Sciences Centre, BC Cancer Agency, Vancouver, B.C. Canada, V5Z 4S6

Additional references:

Burglin, T. R., and G. Ruvkun, 2001 Regulation of ectodermal and excretory function by the C. elegans POU homeobox gene ceh-6. Development 128**:** 779-790.

Hahn, H., and A. C. Palmenberg, 1996 Mutational analysis of the encephalomyocarditis virus primary cleavage. J Virol 70**:** 6870-6875.

Kagias, K., A. Ahier, N. Fischer and S. Jarriault, 2012 Members of the NODE (Nanog and Oct4-associated deacetylase) complex and SOX-2 promote the initiation of a natural cellular reprogramming event in vivo. Proc Natl Acad Sci U S A 109**:** 6596-6601.

Macosko, E. Z., N. Pokala, E. H. Feinberg, S. H. Chalasani, R. A. Butcher *et al.*, 2009 A hub-and-spoke circuit drives pheromone attraction and social behaviour in C. elegans. Nature 458**:** 1171-1175.

Tursun, B., L. Cochella, I. Carrera and O. Hobert, 2009 A toolkit and robust pipeline for the generation of fosmid-based reporter genes in C. elegans. PLoS One 4**:** e4625.