Supplementary Information: Haenni *et al.*

Riboprobes for the NRO analysis

Probe	Accession number	Coordinates	Probe length (nt)[1]	Number of uridines[2]
ets	X03680.1	526-828	303	97
18S	X03680.1	2303-2603	301	81
its1	X03680.1	2730-3029	300	90
5.8S	X03680.1	3158-3308	151	39
its2	X03680.1	3314-3613	300	90
26S-1	X03680.1	3779-4079	301	71
26S-2	X03680.1	6832-7130	299	83
pGEM[3]			128	25
5S	X06102.1	1-119	119	30
rps-6	AL132902.4	358-660	303	54
v2	U56966	22220 - 22519	300	66
v3	U56966	26894 - 27339	446	131
its1/5.8S	X03680	3115 - 3429	315	92

[1] Length of the specific target sequence (insert in pGEM-T easy).

[2] Uridine contents of the complementary sequence (T7 antisense transcripts).

[3] Empty pGEM-T easy vector. Length of the SP6 transcript is indicated in parenthesis. All SP6 transcribed probes contain these flanking sequences from the vector.

Supplementary Figure Legends

Figure S1. Fluorescently labeled nuclei from several somatic tissues can be released. (A) Strain PK2011 with a somatic nuclear marker (*sur-5*p*::nls::gfp*). **(B)** Strain AW60 with a seam cell nuclear marker (*scm::nls::gfp*). **(C)** Strain MS604 with a neuronal nuclear marker (*unc-119*p*::nls::yfp::lacZ*), image captured using GFP filter. **(D)** Strain GFPF with a body wall muscle nuclear marker (*myo-3*p*::nls::gfp*). DIC, D, G, and M are as in Figure 1.

Figure S2. Analysis of gene expression by 3'end-seq. (A) Number of genes considered as expressed in unsorted and sorted samples using different cutoffs (RPM). The inset shows the 0 to 20 RPM region. **(B)** List of number of genes considered as expressed in unsorted and sorted samples at different gene expression cutoffs. **(C)** Distribution of 3'end-seq reads across chromosomes. Reads are shown in a UCSC genome graph. Y-axis is density of RPM. **(D)** Gene density plot showing a positive correlation of gene expression difference between genomic tiling array and 3'end-seq. Genes were distributed in a 20x20 table based on the gene expression fold change in two experiments. The number of genes in each cell of the table was normalized

to an expected number derived from randomized data. Relative gene density, $\log_2$(Obs/Exp), where Obs is observed number of genes and Exp is expected number of genes, is presented in a heat map. The color scheme is shown in the figure.

Figure S3 Comparison of expressed genes detected by Pauli *et al.*, McGhee *et al.*, Spencer *et al.*, and this study. (A) Fractions of genes detected in the intestine by Pauli *et al.*, McGhee *et al.*, Spencer *et al.*, and this study. For genes listed by Pauli *et al.*, 199 genes had been removed from the current database, and were not used in the analysis. **(B)** Comparison of expression level for the genes detected by both this study and another study, and those by this study only.

Figure S4. The 80 genes considered as not expressed in the intestine as reported by Pauli *et al.* (A) The number of genes detected in different studies. **(B)** There are 31 genes in Pauli's list which were detected in this study. Of these, 28 were found expressed by Spencer *et al.*, and 3 were not studied by their microarray.

Figure S5. Analysis of alternative polyadenylation using 3'end-seq. (A) Frequencies of different PAS types for polyA sites identified in this study. A recursive search method was used to identify top PAS hexamers (see Materials and Methods for details). The percentage for each PAS indicates the fraction of polyA sites that 1) have the PAS and 2) do not have a more prominent PAS. **(B)** Nucleotide content around polyA sites in different PAS groups.

Figure S6. Analysis of polyA sites uniquely identified in different studies. (A) Nucleotide profile around the polyA sites uniquely identified in each study. **(B)** PAS usage of the polyA sites uniquely identified in each study.

Figure S7. Specific examples of polyA sites uniquely identified in this study. The polyA sites shown were not identified by Mangone *et al.* or Jan *et al.*

Figure S8. Examples of genes with different polyA site usage in the unsorted sample vs. sorted sample. *C48E7.1* (top) has APA regulation in the 3'-most exon, and *C28C12.7* (bottom) has APA regulation in an intron.

Figure S9. APA. (A) *Trans*-splicing structure of nematode genes. **(B)** Fraction of genes with different *trans*-splicing structures that have APA regulation. Operon genes were divided into first, middle and last genes, based on the location in the operon. **(C)** Frequencies of different PAS associated with polyA sites of genes with different *trans*-splicing structures. **(D)** Genes with significant APA regulation in the 3'-most exon ($P < 0.05$, Fisher's exact test; change of usage >= 5%). We divided genes into two groups, i.e., 550 genes with more usage of promoter-distal polyA sites in the intestine (lengthened) and 636 genes with more usage of promoter-proximal pA sites in the intestine (shortened). We plotted the distribution of the distances between the two most regulated pA sites in the 3'-most exon. The red dotted line indicates 40 nt.

Figure S10: Comparison of gene expression levels for solo genes vs. other genes.

Figure S1

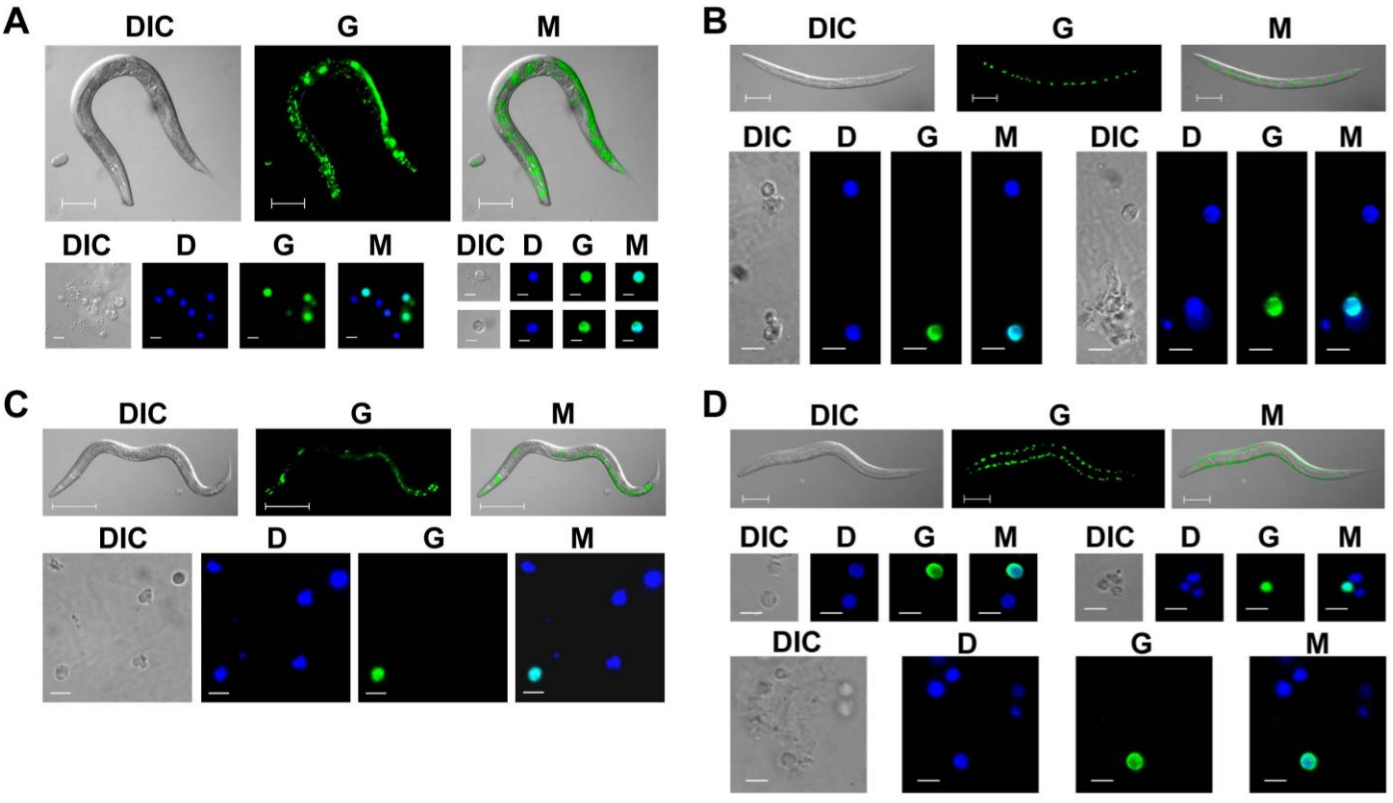
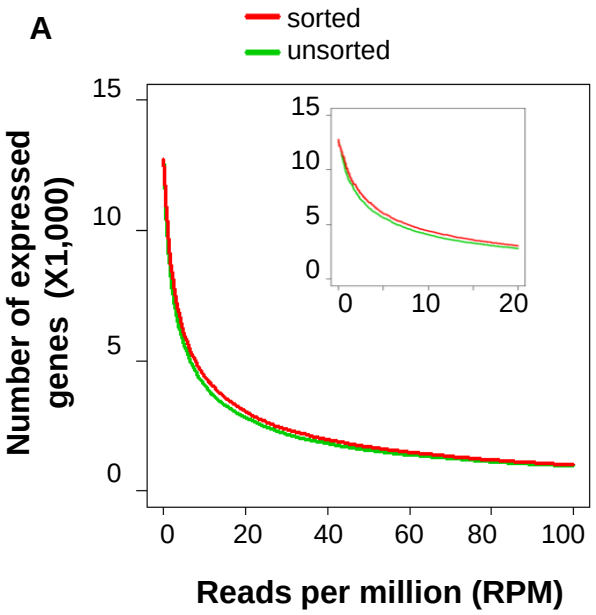

A
DIC
G
M
DIC
D
G
M
DIC
D
G
M
B
DIC
G
M
DIC
D
G
M
DIC
D
G
M
C
DIC
G
M
DIC
D
G
M
D
DIC
G
M
DIC
D
G
M
DIC
D
G
M
DIC
D
G
M

Figure S2

A

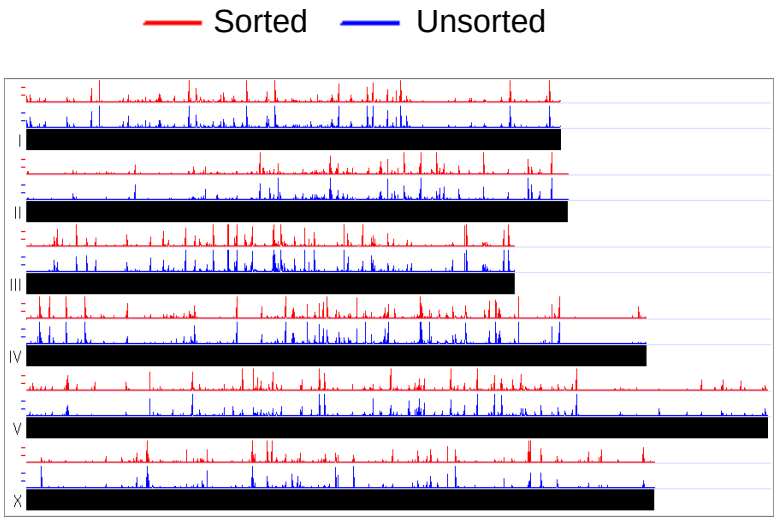


B

RPM cutoff	Number of genes considered as expressed		
	Unsorted	Sorted	Both
0	12,397	12,177	11,824
0.5	10,959	11,205	9,962
1	9,419	10,149	8,647
2	7,868	8,386	7,114
5	5,620	6,047	5,002
10	4,075	4,395	3,589
20	2,821	3,052	2,452
50	1,575	1,703	1,306
100	962	1,017	773

C

Chr.	Unsorted	Sorted
Chr I	1,841,039	829,074
Chr II	1,109,249	552,494
Chr III	2,070,821	1,165,260
Chr IV	1,917,803	832,136
Chr V	1,395,208	776,097
Chr X	886,125	547,483



D

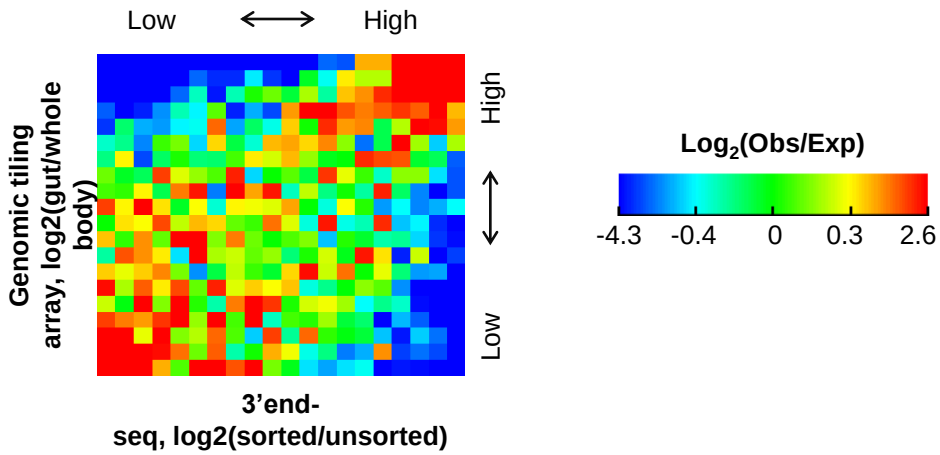


Figure S3

A

	Number of genes detected	Number of genes detected in this study (RPM>0)	Number of genes detected in this study (RPM>1)
This study		12,177	10,149
Pauli *et al.*	1,739	1,394 (80.2%)	1,296 (74.5%)
McGhee *et al.*	4,779	3,689 (77.2%)	3,375 (70.6%)
Spencer *et al.*	13,089	9,458 (72.3%)	8,181 (62.5%)

B

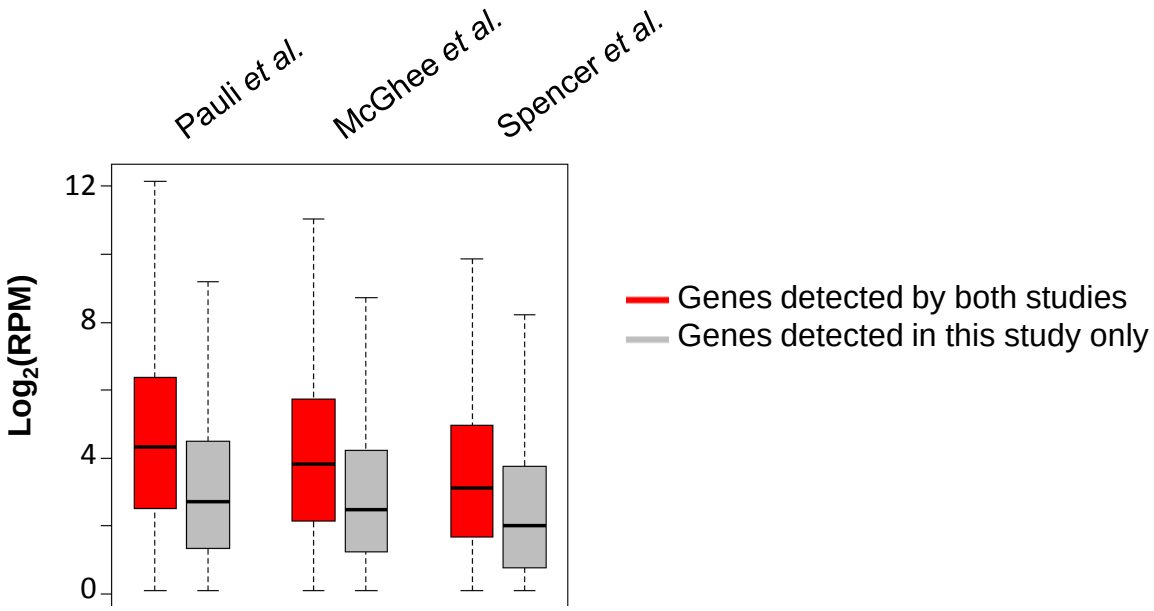


Figure S4

A

	Genes detected
This study (RPM>1)	31
McGhee *et al.*	10
Spencer *et al.*	53

B

refID	GeneID	WBGeneID	sorted	L2.Exp.FDR	McGhee (tags)
pat-10	F54C1.7	WBGene00003934	313.59	1.14E-05	NA
unc-18	F27D9.1	WBGene00006757	37.37	2.65E-05	2
unc-97	F14D12.2	WBGene00006826	3.18	5.86E-05	NA
unc-27	ZK721.2	WBGene00006764	736.94	1.14E-05	NA
apr-1	K04G2.8	WBGene00000156	1.91	5.17E-04	NA
odr-4	Y102E9.1	WBGene00003851	10.40	5.30E-05	1
sax-3	ZK377.2	WBGene00004729	25.90	9.39E-05	NA
flp-1	F23B2.5	WBGene00001444	23.14	2.65E-05	NA
unc-5	B0273.4	WBGene00006745	28.66	2.21E-04	NA
dpy-7	F46C8.6	WBGene00001069	81.10	1.53E-04	NA
mec-12	C44B11.3	WBGene00003175	156.05	7.31E-05	NA
aex-3	C02H7.3	WBGene00000086	7.86	2.64E-04	1
unc-45	F30H5.1	WBGene00006781	15.71	6.79E-04	10
gpa-7	R10H10.5	WBGene00001669	7.64	4.58E-04	NA
mec-2	F14D12.4	WBGene00003166	7.22	2.81E-04	1
unc-129	C53D6.2	WBGene00006852	79.83	8.11E-04	NA
pat-4	C29F9.7	WBGene00003931	3.82	1.30E-03	1
hlh-1	B0304.1	WBGene00001948	1.27	1.61E-03	NA
myo-2	T18D3.4	WBGene00003514	5.10	5.86E-05	NA
gon-4	K04D7.5	WBGene00001653	3.61	1.04E-03	NA
fax-1	F56E3.4	WBGene00001400	5.73	6.73E-03	NA
daf-7	B0412.2	WBGene00000903	3.82	3.94E-04	NA
let-75	R06C7.10	WBGene00002348	12.10	1.40E-04	NA
che-14	F56H1.1	WBGene00000493	2.97	2.41E-02	NA
flp-17	C52D10.11	WBGene00001460	2.12	1.88E-04	NA
ace-1	W09B12.1	WBGene00000035	1.91	1.56E-03	NA
glr-1	C06E1.4	WBGene00001612	6.16	2.20E-02	NA
deg-3	K03B8.9	WBGene00000951	2.76	6.97E-03	NA

Figure S5

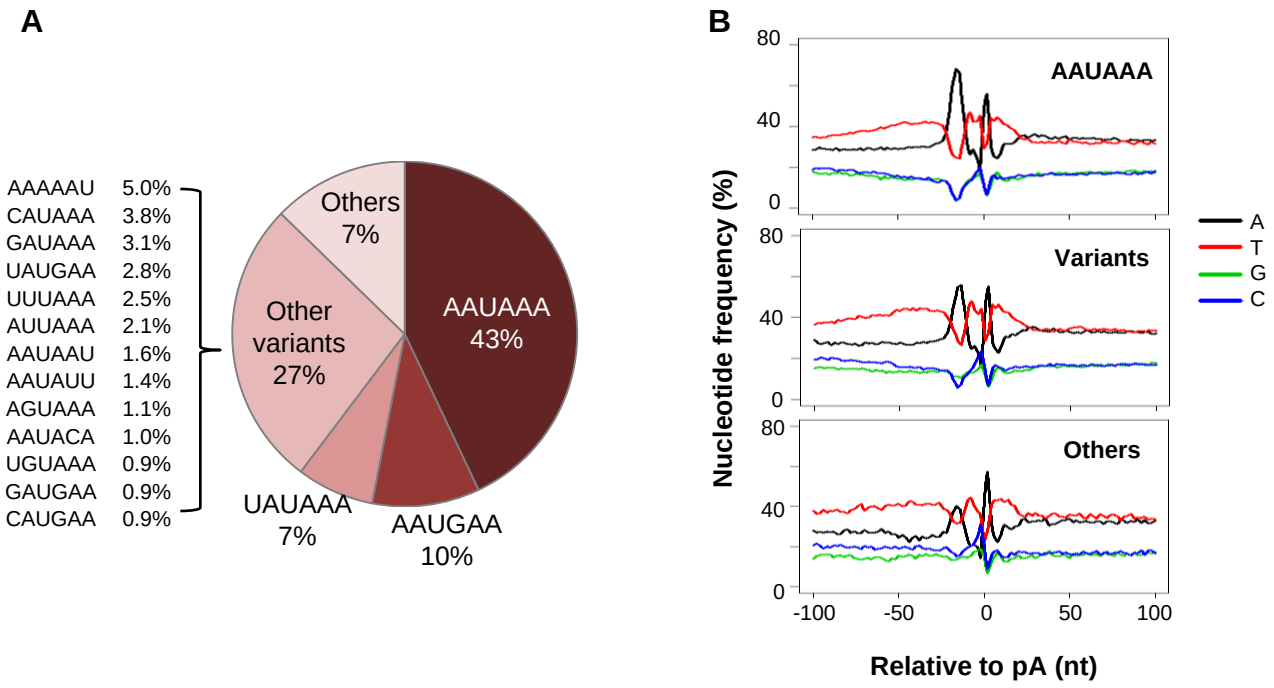
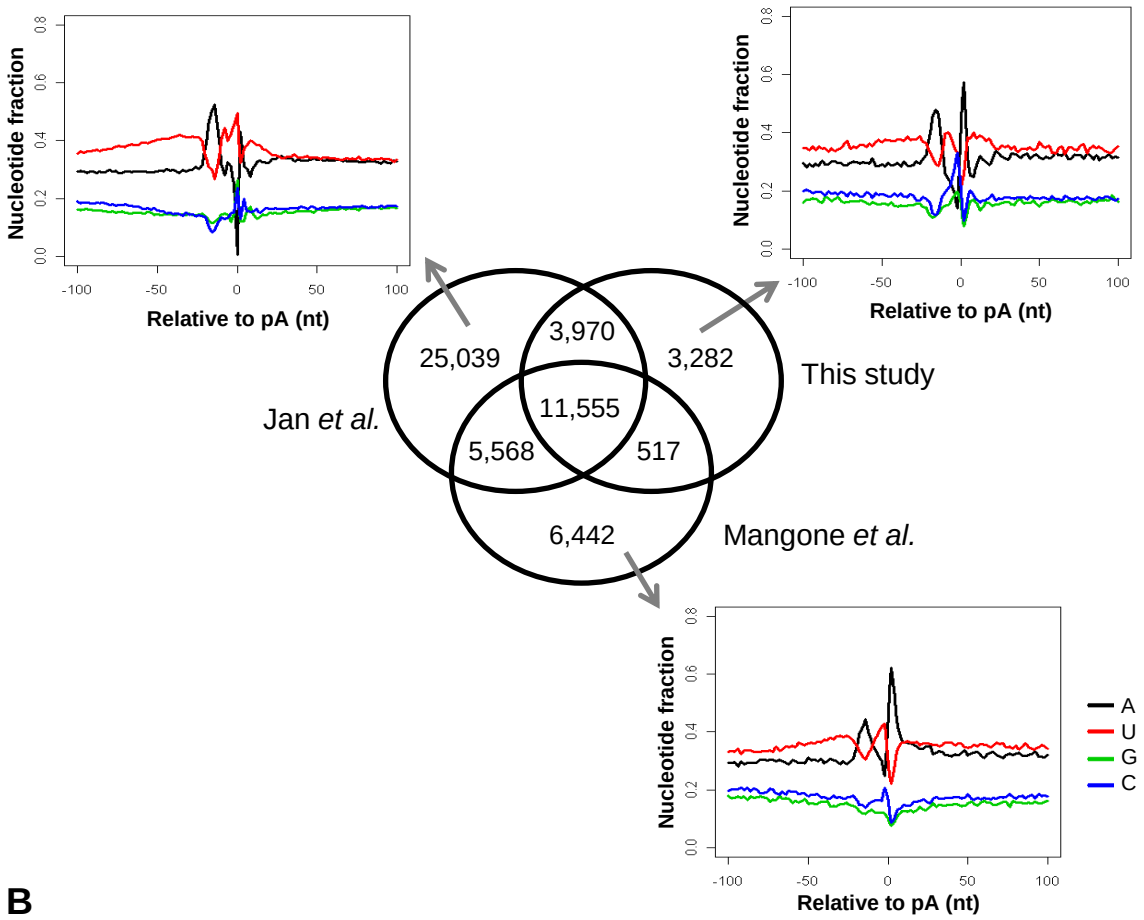

A
AAAAAU 5.0%
CAUAAA 3.8%
GAUAAA 3.1%
UAUGAA 2.8%
UUUAAA 2.5%
AUUAAA 2.1%
AAUAAU 1.6%
AAUAUU 1.4%
AGUAAA 1.1%
AAUACA 1.0%
UGUAAA 0.9%
GAUGAA 0.9%
CAUGAA 0.9%
Others 7%
Other variants 27%
AAUAAA 43%
UAUAAA 7%
AAUGAA 10%
B
AAUAAA
Variants
Others
Nucleotide frequency (%)
80
40
0
-100
-50
0
50
100
Relative to pA (nt)
A
T
G
C

Figure S6

A


Nucleotide fraction
Relative to pA (nt)
25,039
3,970
3,282
11,555
5,568
517
6,442
Jan et al.
This study
Mangone et al.
A
U
G
C

B

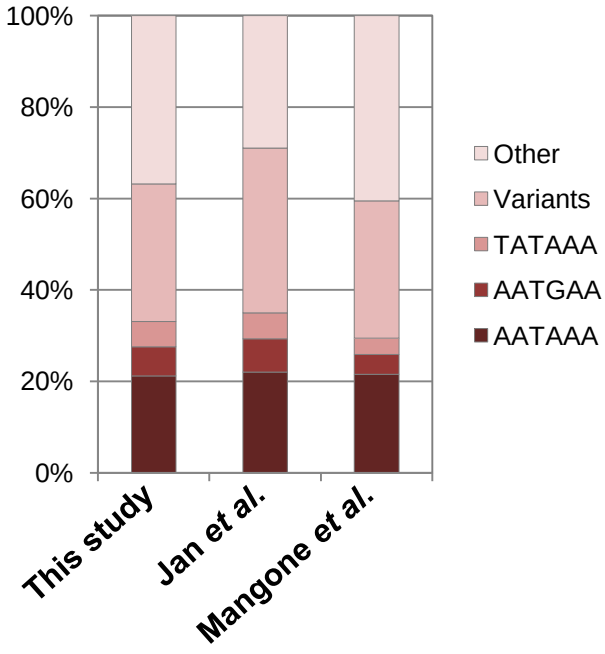

100%
80%
60%
40%
20%
0%
Other
Variants
TATAAA
AATGAA
AATAAA
This study
Jan et al.
Mangone et al.

Figure S7

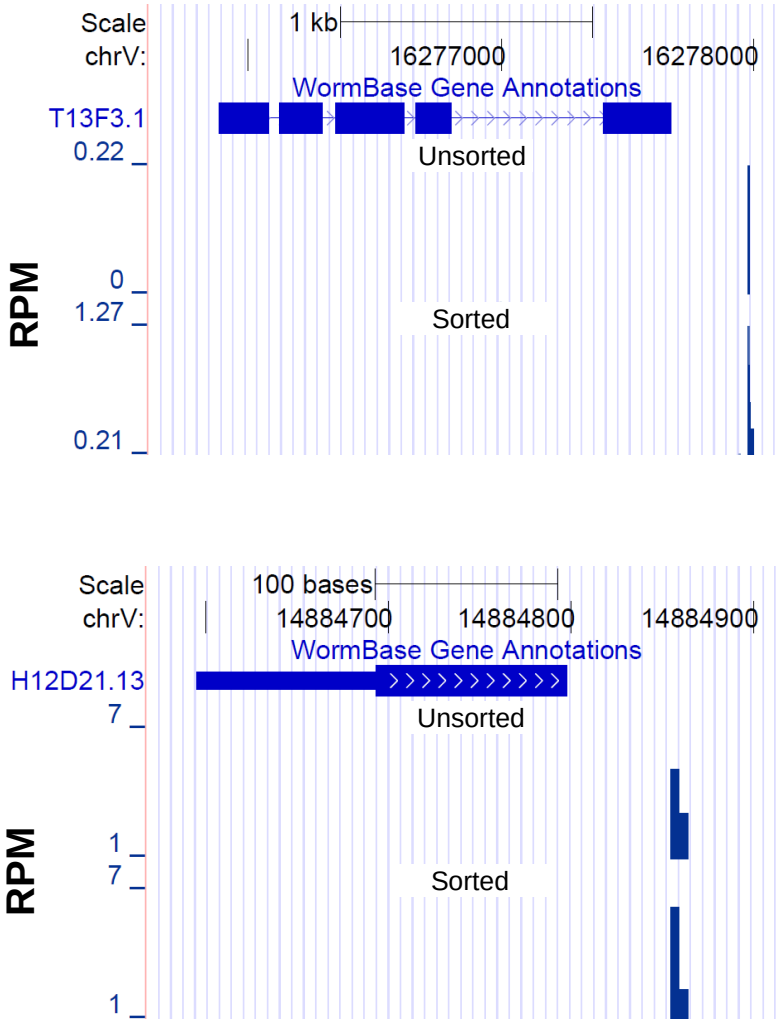
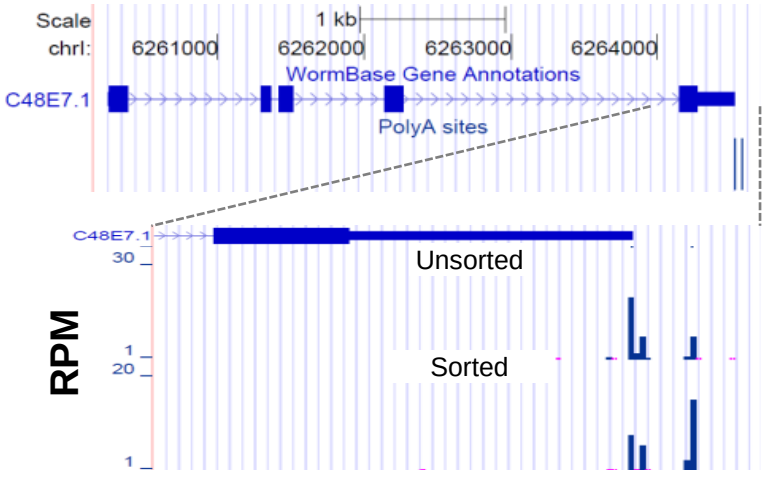
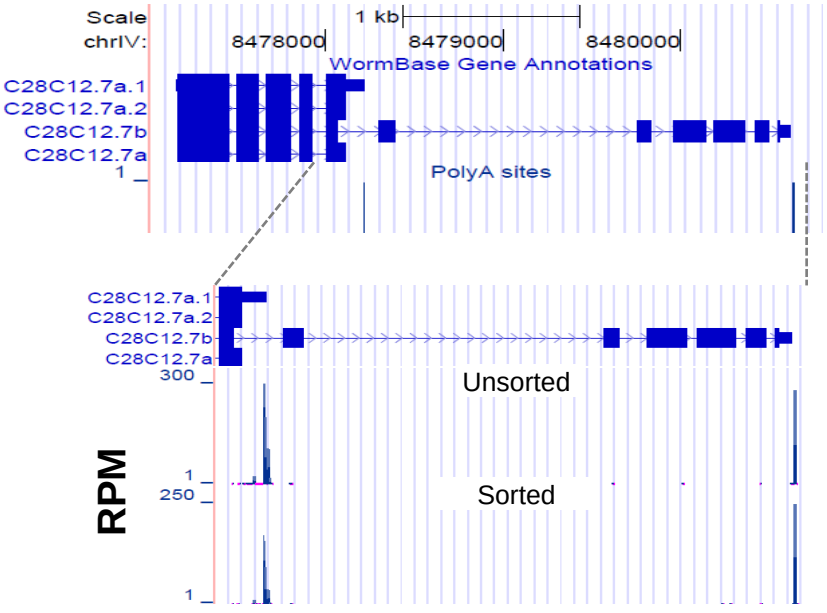

Scale
chrV:
1 kb
16277000
16278000
WormBase Gene Annotations
T13F3.1
0.22
Unsorted
RPM
0
1.27
Sorted
0.21
Scale
chrV:
100 bases
14884700
14884800
14884900
WormBase Gene Annotations
H12D21.13
7
Unsorted
RPM
1
7
Sorted
1

Figure S8

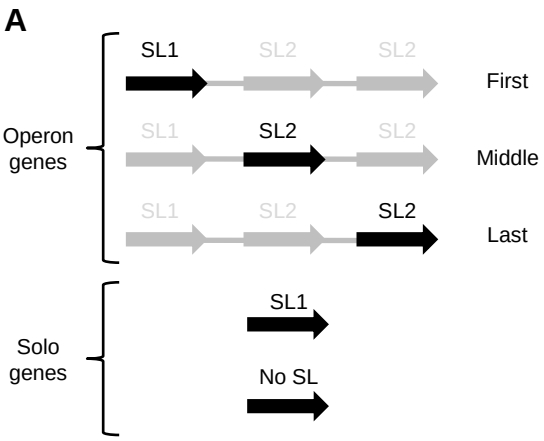


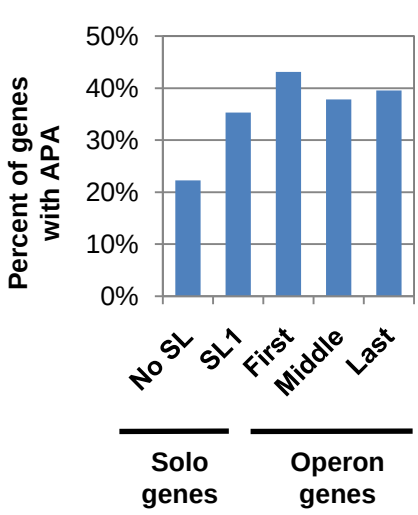
	Unsorted	Sorted
Proximal	0.78	0.48
Distal	0.22	0.52

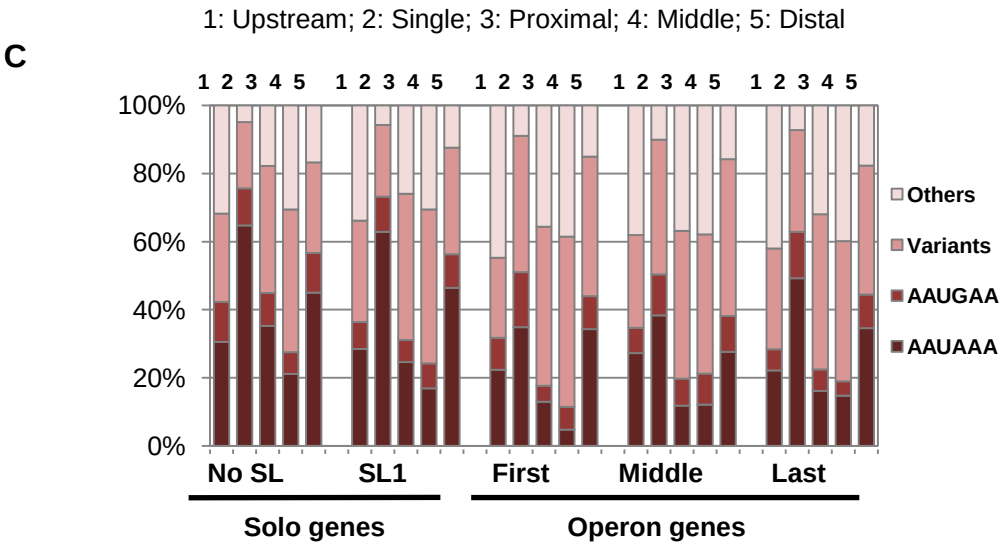


	Unsorted	Sorted
Intron	0.70	0.51
3’-most exon	0.30	0.49

Figure S9


A
SL1
SL2
SL2
First
Operon genes
SL1
SL2
SL2
Middle
SL1
SL2
SL2
Last
Solo genes
SL1
No SL


B
Percent of genes with APA
50%
40%
30%
20%
10%
0%
No SL
SL1
First
Middle
Last
Solo genes
Operon genes


1: Upstream; 2: Single; 3: Proximal; 4: Middle; 5: Distal
C
1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5
100%
80%
60%
40%
20%
0%
Others
Variants
AAUGAA
AAUAAA
No SL
SL1
First
Middle
Last
Solo genes
Operon genes

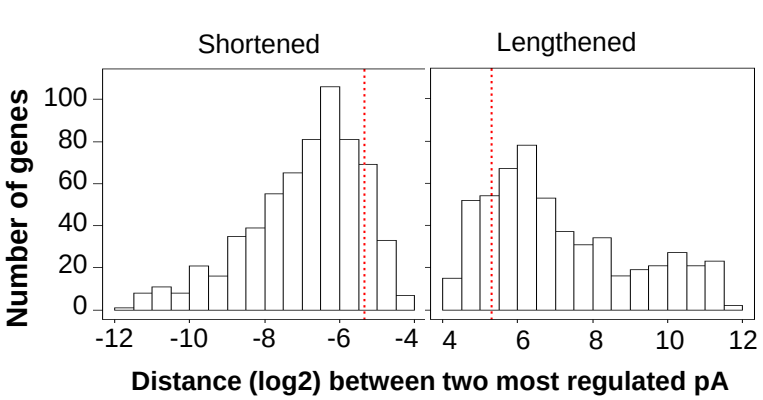
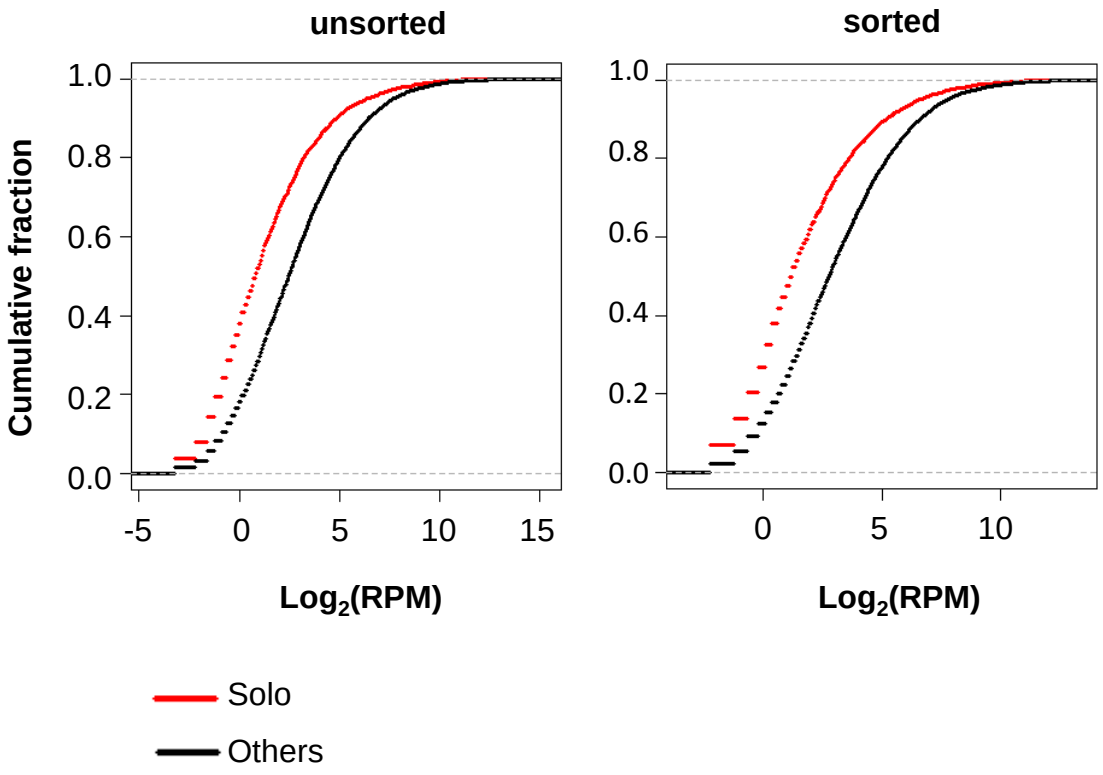

D
Shortened
Lengthened
Number of genes
100
80
60
40
20
0
-12
-10
-8
-6
-4
4
6
8
10
12
Distance (log2) between two most regulated pA

Figure S10


unsorted
sorted
Cumulative fraction
1.0
0.8
0.6
0.4
0.2
0.0
-5
0
5
10
15
Log2(RPM)
Solo
Others