Supplementary materials for:

Translation of poly(A) tails leads to precise mRNA cleavage
Guydosh and Green, 2017

Supplemental Figure Legends

Figure S1. Translation of the poly(A) tail leads to widespread upstream mRNA cleavage and ribosome stalling at the 3' ends of the resulting truncated fragments.

A) Examples of mapped footprints on genes (*GAT2* and *SMX2*) with UAA stop codons and no in-frame stop codons in the 3'UTR. Data reveal enrichment of long reads near poly(A) and short reads over a longer region upstream when *DOM34* is knocked out and the suppressor tRNA is present. In the case of *SMX2*, long reads are somewhat enriched in the absence of the suppressor tRNA (as shown previously (Guydosh and Green, 2014), presumably due to the accumulation of likely non-translating ribosomes. However, short reads do not enrich upstream in this case. Approximate region where polyadenylation takes place is indicated with a dotted arrow.

B) Scatter plot describing effect of knocking out *DOM34* or *SKI2* on ribosome occupancy in 3'UTRs in the presence or absence of *SUP4-**o***. Enrichment ratios of short (15-18 nt) reads are plotted as a function of the ratio of long (25-34 nt) reads. Ratios are computed from data on genes with UAA stop codons and no in-frame stop codons in 3'UTRs (*SUP4-**o*** strains shown in red, dark yellow, and blue) or for all genes (strain where *SUP4-**o*** is absent shown in gray). Note that when *SUP4-**o*** is absent (gray dots), the long reads (but not short reads) are enriched in the *dom34*Δ strain as previously observed for likely non-translating ribosomes (Guydosh and Green, 2014).

Figure S2. (Short (<10 aa) runs of polybasic residues stall the ribosome but do not trigger significant upstream mRNA cleavage.

A) Same as Figure 4 for subsets of KKK motifs. Note that the peak in the AAA data for short reads is present in both footprint and mRNA-Seq data, implying it is arises due to an artifact and is not reflective of mRNA cleavage.

B) Same as Figure 4 when *DOM34* is present (*ski2*Δ background).

C) Same as Figure 4 for polyarginine residues. Unlike the pausing observed with lysine, pausing with arginine does not suddenly appear when 3 codons are repeated but instead gets larger with the addition of each arginine to the sequence.

Figure S3. Translation of poly(A) tails cannot be observed while ORF-internal poly(A) is readily translated.

A) Histogram of consecutive A count on the 3' ends of footprint reads before any alignments were performed.

B) Ribosome footprints mapped to a reporter construct with 9 consecutive lysines (green box) that are encoded by either all AAA (top) or AAG (bottom) codons. Regions where either short (yellow) or long (brown) reads cannot map due to low complexity of the sequence are indicated above the plots. All reads normalized to the long-read ribosome occupancy of the *URA3* gene on the reporter plasmid (long and short reads sequenced together from a single library preparation). Data from the AAG reporter was then scaled by a factor of 2.2 so that long-read density was the same as for the AAA reporter. Finally, short reads for both reporters were scaled by a factor of 2 to facilitate easier comparison.

Strains Used

Strain ID	**Genotype**	**Source**
yNG100	*MAT*α *his3*Δ1 *leu2*Δ0 lys2Δ0 *ura3*Δ0 (*ski2*::*KanMX4*)	Thermo Scientific
yNG103a	*MAT***a** *his3*Δ1 *leu2*Δ0 *ura3*Δ0 (*ski2*::*KanMX4*) (*dom34*::*KanMX4*)	Guydosh and Green, 2014
yNG103b	*MAT***a** *his3*Δ1 *leu2*Δ0 met15Δ0 lys2Δ0 *ura3*Δ0 (*ski2*::*KanMX4*) (*dom34*::*KanMX4*)	Guydosh and Green, 2014
yNG104	*MAT*α *his3*Δ1 *leu2*Δ0 lys2Δ0 *ura3*Δ0 (*ski2*::*KanMX4*) p[*SUP4*-**o**, *HIS3*]	Guydosh and Green, 2014
yNG107	*MAT***a** *his3*Δ1 *leu2*Δ0 met15Δ0 lys2Δ0 *ura3*Δ0 (*ski2*::*KanMX4*) (*dom34*::*KanMX4*) p[SUP4-o, HIS3]	This study
yNG108	*MAT***a** *his3*Δ1 *leu2*Δ0 *met15*Δ0 *ura3*Δ0 (*dom34*::*KanMX4*) p[SUP4-o, HIS3]	This study
yNG109	*MAT***a** *his3*Δ1 *leu2*Δ0 *met15*Δ0 *ura3*Δ0 p[SUP4-o, HIS3]	This study
yNG110	*MAT***a** *his3*Δ1 *leu2*Δ0 met15Δ0 lys2Δ0 *ura3*Δ0 (*ski2*::*KanMX4*) (*dom34*::*KanMX4*) p[Thrdx-HA-(AAA)9-mCherry, URA3]	This study
yNG111	*MAT***a** *his3*Δ1 *leu2*Δ0 met15Δ0 lys2Δ0 *ura3*Δ0 (*ski2*::*KanMX4*) (*dom34*::*KanMX4*) p[Thrdx-HA-(AAG)9-mCherry, URA3]	This study

Sequence of the (AAA)$_9$ reporter

```
GACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGTATGATCCAATATCAAAGGAAATGATAGCATT
GAAGGATGAGACTAATCCAATTGAGGAGTGGCAGCATATAGAACAGCTAAAGGGTAGTGCTGAAGGAAGCATACGATACCCCGCATGGAATGGGATA
ATATCACAGGAGGTACTAGACTACCTTTCATCCTACATAAATAGACGCATATAAGTACGCATTTAAGCATAAACACGCACTATGCCGTTCTTCTCAT
GTATATATATATACAGGCAACACGCAGATATAGGTGCGACGTGAACAGTGAGCTGTATGTGCGCAGCTCGCGTTGCATTTTCGGAAGCGCTCGTTTT
CGGAAACGCTTTGAAGTTCCTATTCCGAAGTTCCTATTCTCTAGAAAGTATAGGAACTTCAGAGCGCTTTTGAAAACCAAAAGCGCTCTGAAGACGC
ACTTTCAAAAAACCAAAAACGCACCGGACTGTAACGAGCTACTAAAATATTGCGAATACCGCTTCCACAAACATTGCTCAAAAGTATCTCTTTGCTA
TATATCTCTGTGCTATATCCCTATATAACCTACCCATCCACCTTTCGCTCCTTGAACTTGCATCTAAACTCGACCTCTACATTTTTTATGTTTATCT
CTAGTATTACTCTTTAGACAAAAAAATTGTAGTAAGAACTATTCATAGAGTGAATCGAAAACAATACGAAAATGTAAACATTTCCTATACGTAGTAT
ATAGAGACAAAATAGAAGAAACCGTTCATAATTTTCTGACCAATGAAGAATCATCAACGCTATCACTTTCTGTTCACAAAGTATGCGCAATCCACAT
CGGTATAGAATATAATCGGGGATGCCTTTATCTTGAAAAAATGCACCCGCAGCTTCGCTAGTAATCAGTAAACGCGGGAAGTGGAGTCAGGCTTTTT
TTATGGAAGAGAAAATAGACACCAAAGTAGCCTTCTTCTAACCTTAACGGACCTACAGTGCAAAAAGTTATCAAGAGACTGCATTATAGAGCGCACA
AAGGAGAAAAAAAGTAATCTAAGATGCTTTGTTAGAAAAATAGCGCTCTCGGGATGCATTTTTGTAGAACAAAAAAGAAGTATAGATTCTTTGTTGG
TAAAATAGCGCTCTCGCGTTGCATTTCTGTTCTGTAAAAATGCAGCTCAGATTCTTTGTTTGAAAAATTAGCGCTCTCGCGTTGCATTTTTGTTTTA
CAAAAATGAAGCACAGATTCTTCGTTGGTAAAATAGCGCTTTCGCGTTGCATTTCTGTTCTGTAAAAATGCAGCTCAGATTCTTTGTTTGAAAAATT
AGCGCTCTCGCGTTGCATTTTTGTTCTACAAAATGAAGCACAGATGCTTCGTTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGT
TTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTC
AACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGA
TCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGC
ACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGG
TTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGC
CAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCG
GAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTA
CTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGC
TGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACG
GGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCAT
ATATACTTTAGATTGATTTAAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGA
GTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAA
AAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATA
CTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGC
TGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACA
CAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACA
GGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCA
CCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTT
TGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCA
GCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATG
CAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTACCTCACTCATTAGGCACCCCAGGCTTTACAC
TTTATGCTTCCGGCTCCTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAACAGCTATGACCATGATTACGCCAAGCGCGCAATTA
ACCCTCACTAAAGGGAACAAAAGCTGGAGCTCTTCGAGTTTATCATTATCAATACTCGCCATTTCAAAGAATACGTAAATAATTAATAGTAGTGATT
TTCCTAACTTTATTTAGTCAAAAAATTAGCCTTTTAATTCTGCTGTAACCCGTACATGCCCAAAATAGGGGGCGGGTTACACAGAATATATAACACC
GAAGGTGCCTGGGTGAACAGGTTTATTCCTGGCATCCACTAAATATAATGGAGCCCGCTTTTTAAGCTGGCATCCAGAAAAAAAAAGAATCCCAGCA
CCAAAATATTGTTTTCTTCACCAACCATCAGTTCATAGGTCCATTCTCTTAGCGCAACTACAGAGAACAGGGGCACAAACAGGCAAAAAACGGGCAC
AACCTCAATGGAGTGATGCAACCTGCCTGGAGTAAATGATGACACAAGGCAATTGACCCACGCATGTATCTATCTCATTTTCTTACACCTTCTATTA
CCTTCTGCTCTCTCTGATTTGGAAAAAGCTGAAAAAAAAGGTTGAAACCAGTTCCCTGAAATTATTCCCCTACTTGACTAATAAGTATATAAAGACG
GTAGGTATTGATTGTAATTCTGTAAATCTATTTCTTAAACTTCTTAAATTCTACTTTTATAGTTAGTCTTTTTTTTAGTTTTAAAACACCAGAACTT
AGTTTCGACGGATTCTAGAACTAGTGGATCCCCCATCACAAGTTTGTACAAAAAAGCAGGCTCCGCGGCCGCCCCCTTCACCAAAAAAATGGGATCT
GATAAAATTATTCATCTGACTGATGATTCTTTTGATACTGATGTACTTAAGGCAGATGGTGCAATCCTGGTTGATTTCTGGGCACACTGGTGCGGTC
CGTGCAAAATGATCGCTCCGATTCTGGATGAAATCGCTGACGAATATCAGGGCAAACTGACCGTTGCAAAACTGAACATCGATCACAACCCGGGCAC
TGCGCCGAAATATGGCATCCGTGGTATCCCGACTCTGCTGCTGTTCAAAAACGGTGAAGTGGCGGCAACCAAAGTGGGTGCACTGTCTAAAGGTCAG
TTGAAAGAGTTCCTCGACGCTAACCTGGCCGGCTCTGGATCCGGTGATGACGATGACAAGCTGGGAATTATCACAAGTTTGTACAAAAAAGCAGGCT
CCGCGGCCGCCCCCTTCACCATGGGCTACCCATACGATGTTCCAGATTACGCTGCGGCCTACCCGTATGACGTACCGGATTATGCAAAAAAAAAAAA
AAAAAAAAAAAAAAAAGTGAGCAAGGGCGAGGAGGATAACATGGCCATCATCAAGGAGTTCATGCGCTTCAAGGTGCACATGGAGGGCTCCGTGAAC
GGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCGCCCCTACGAGGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGGTGGCCCCCTGCCCT
TCGCCTGGGACATCCTGTCCCCTCAGTTCATGTACGGCTCCAAGGCCTACGTGAAGCACCCCGCCGACATCCCCGACTACTTGAAGCTGTCCTTCCC
CGAGGGCTTCAAGTGGGAGCGCGTGATGAACTTCGAGGACGGCGGCGTGGTGACCGTGACCCAGGACTCCTCCCTGCAGGACGGCGAGTTCATCTAC
AAGGTGAAGCTGCGCGGCACCAACTTCCCCTCCGACGGCCCCGTAATGCAGAAGAAGACCATGGGCTGGGAGGCCTCCTCCGAGCGGATGTACCCCG
AGGACGGCGCCCTGAAGGGCGAGATCAAGCAGAGGCTGAAGCTGAAGGACGGCGGCCACTACGACGCTGAGGTCAAGACCACCTACAAGGCCAAGAA
GCCCGTGCAGCTGCCCGGCGCCTACAACGTCAACATCAAGTTGGACATCACCTCCCACAACGAGGACTACACCATCGTGGAACAGTACGAACGCGCC
GAGGGCCGCCACTCCACCGGCGGCATGGACGAGCTGTACAAGTAAAGGGTGGGCGCGCCGACCCAGCTTTCTTGTACAAAGTGGTGATGGGCTGCA
GGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGTCATGTAATTAGTTATGTCACGCTTACATTCACGCCCTCCCCCCACATCCGCTCTAA
CCGAAAAGGAAGGAGTTAGACAACCTGAAGTCTAGGTCCCTATTTATTTTTTTATAGTTATGTTAGTATTAAGAACGTTATTTATATTTCAAATTTT
TCTTTTTTTTCTGTACAGACGCGTGTACGCATGTAACATTATACTGAAAACCTTGCTTGAGAAGGTTTTGGGACGCTCGAAGGCTTTAATTTGCGGC
CGGTACCCAATTCGCCCTATAGTGAGTCGTATTACGCGCGCTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACT
TAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGC
GAATGGCGCGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCG
CTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGC
TTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAG
TCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTT
CGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGTTTACAATTTCCTGATGCGGTATTTTCTC
CTTACGCATCTGTGCGGTATTTCACACCGCATAGGGTAATAACTGATATAATTAAATTGAAGCTCTAATTTGTGAGTTTAGTATACATGCATTTACT
```

```
TATAATACAGTTTTTTAGTTTTGCTGGCCGCATCTTCTCAAATATGCTTCCCAGCCTGCTTTTCTGTAACGTTCACCCTCTACCTTAGCATCCCTTC
CCTTTGCAAATAGTCCTCTTCCAACAATAATAATGTCAGATCCTGTAGAGACCACATCATCCACGGTTCTATACTGTTGACCCAATGCGTCTCCCTT
GTCATCTAAACCCACACCGGGTGTCATAATCAACCAATCGTAACCTTCATCTCTTCCACCCATGTCTCTTTGAGCAATAAAGCCGATAACAAAATCT
TTGTCGCTCTTCGCAATGTCAACAGTACCCTTAGTATATTCTCCAGTAGATAGGGAGCCCTTGCATGACAATTCTGCTAACATCAAAAGGCCTCTAG
GTTCCTTTGTTACTTCTTCTGCCGCCTGCTTCAAACCGCTAACAATACCTGGGCCCACCACACCGTGTGCATTCGTAATGTCTGCCCATTCTGCTAT
TCTGTATACACCCGCAGAGTACTGCAATTTGACTGTATTACCAATGTCAGCAAATTTTCTGTCTTCGAAGAGTAAAAAATTGTACTTGGCGGATAAT
GCCTTTAGCGGCTTAACTGTGCCCTCCATGGAAAAATCAGTCAAGATATCCACATGTGTTTTTAGTAAACAAATTTTGGGACCTAATGCTTCAACTA
ACTCCAGTAATTCCTTGGTGGTACGAACATCCAATGAAGCACACAAGTTTGTTTGCTTTTCGTGCATGATATTAAATAGCTTGGCAGCAACAGGACT
AGGATGAGTAGCAGCACGTTCCTTATATGTAGCTTTCGACATGATTTATCTTCGTTTCCTGCAGGTTTTTGTTCTGTGCAGTTGGGTTAAGAATACT
GGGCAATTTCATGTTTCTTCAACACTACATATGCGTATATATACCAATCTAAGTCTGTGCTCCTTCCTTCGTTCTTCCTTCTGTTCGGAGATTACCG
AATCAAAAAAATTTCAAAGAAACCGAAATCAAAAAAAAGAATAAAAAAAAAATGATGAATTGAATTGAAAAGCTGTGGTATGGTGCACTCTCAGTAC
AATCTGCTCTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTAC
AGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGCGCGA
```

Fig. S1

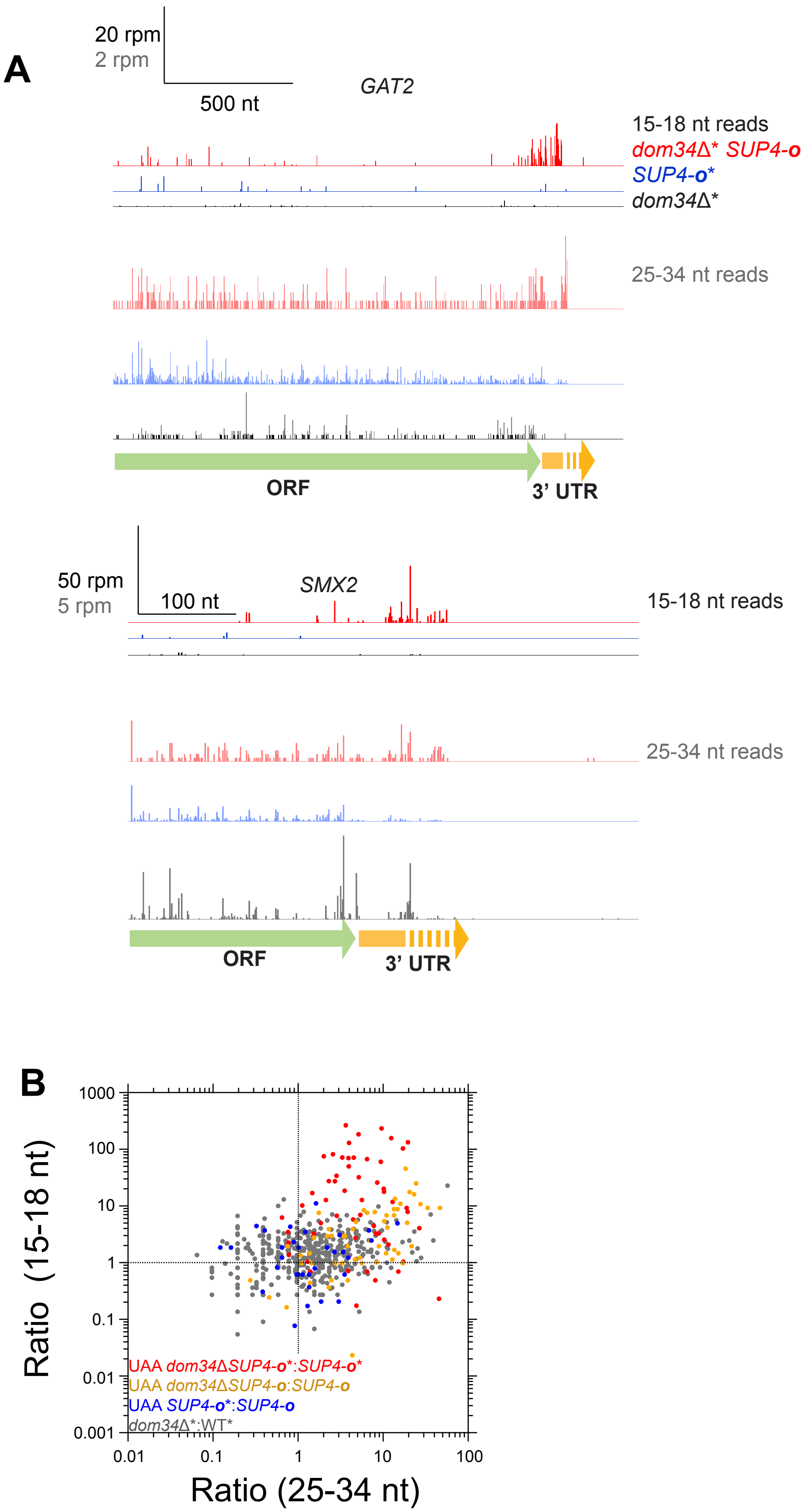

A
20 rpm
2 rpm
500 nt
GAT2
15-18 nt reads
dom34Δ* SUP4-o
SUP4-o*
dom34Δ*
25-34 nt reads
ORF
3' UTR
50 rpm
5 rpm
100 nt
SMX2
15-18 nt reads
25-34 nt reads
ORF
3' UTR
B
Ratio (15-18 nt)
1000
100
10
1
0.1
0.01
0.001
0.01
0.1
1
10
100
Ratio (25-34 nt)
UAA dom34ΔSUP4-o*:SUP4-o*
UAA dom34ΔSUP4-o:SUP4-o
UAA SUP4-o*:SUP4-o
dom34Δ*:WT*

Fig. S2

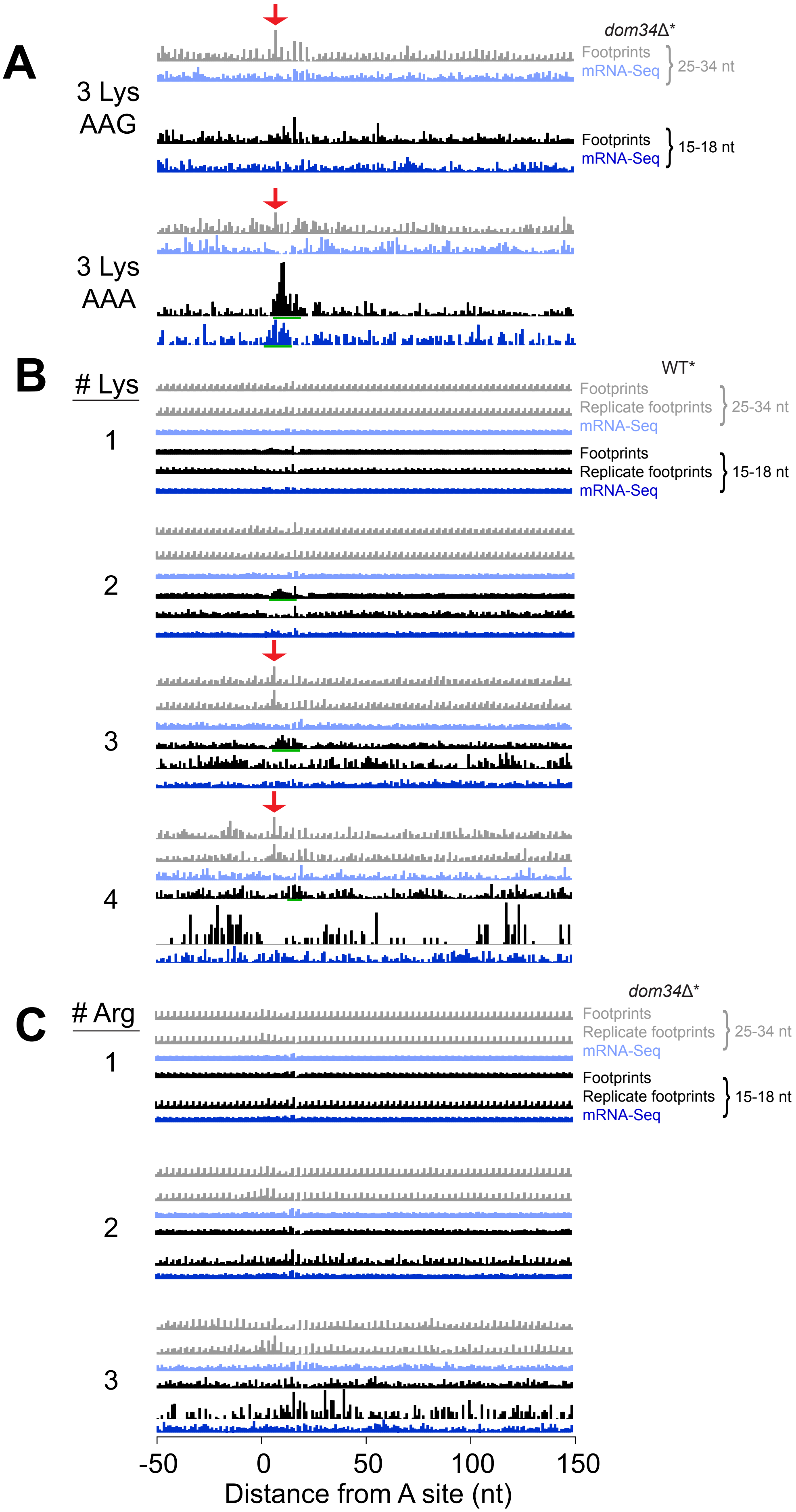
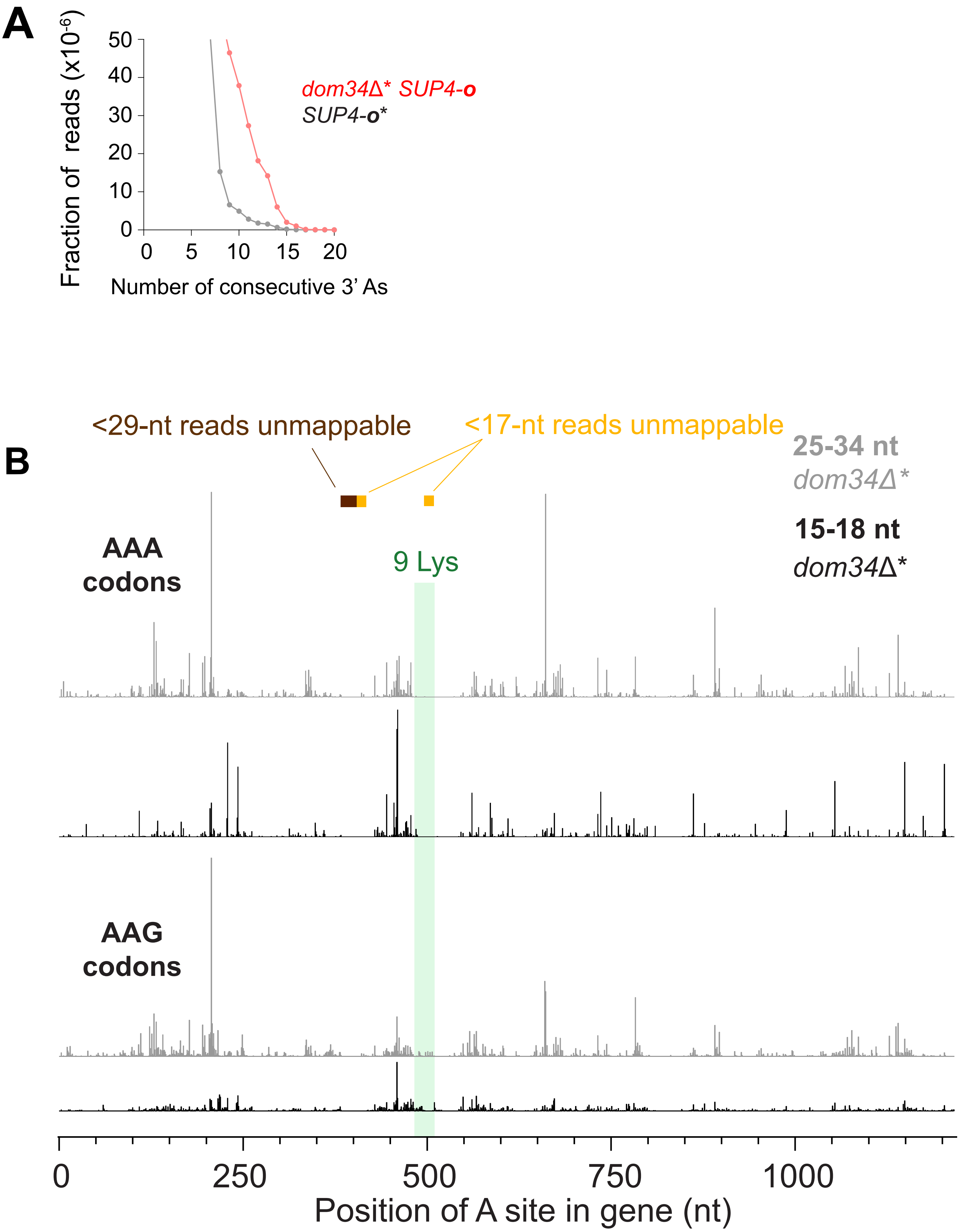

A
dom34Δ*
Footprints
mRNA-Seq
25-34 nt
3 Lys
AAG
Footprints
mRNA-Seq
15-18 nt
3 Lys
AAA
B
Lys
WT*
Footprints
Replicate footprints
mRNA-Seq
25-34 nt
1
Footprints
Replicate footprints
mRNA-Seq
15-18 nt
2
3
4
C
Arg
dom34Δ*
Footprints
Replicate footprints
mRNA-Seq
25-34 nt
1
Footprints
Replicate footprints
mRNA-Seq
15-18 nt
2
3
-50
0
50
100
150
Distance from A site (nt)

Fig. S3


A
Fraction of reads (x10-6)
50
40
30
20
10
0
0
5
10
15
20
Number of consecutive 3' As
dom34Δ* SUP4-o
SUP4-o*
B
<29-nt reads unmappable
<17-nt reads unmappable
25-34 nt
dom34Δ*
15-18 nt
dom34Δ*
AAA
codons
9 Lys
AAG
codons
0
250
500
750
1000
Position of A site in gene (nt)