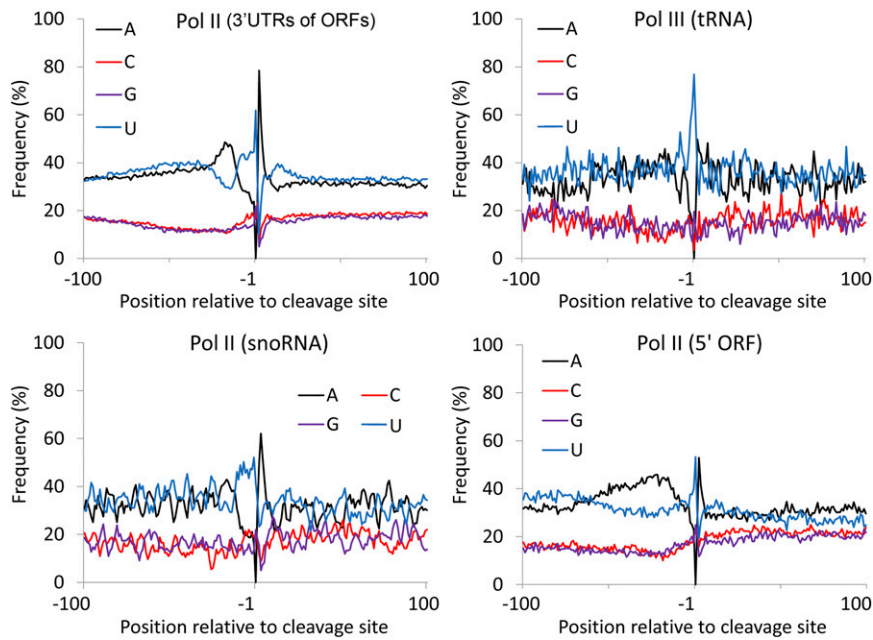
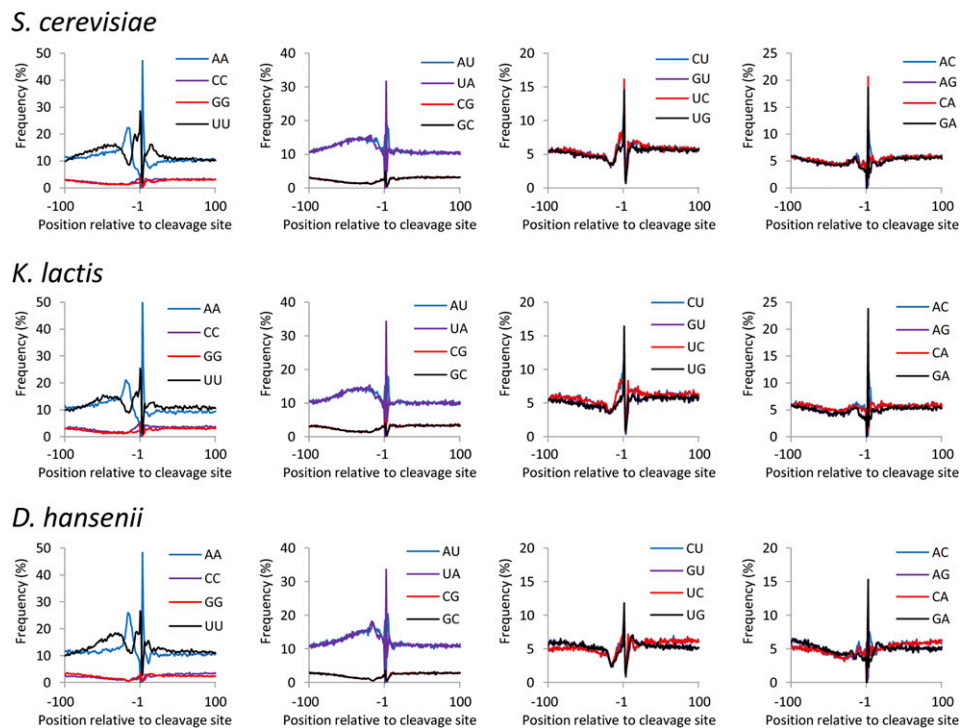


# Supporting Information

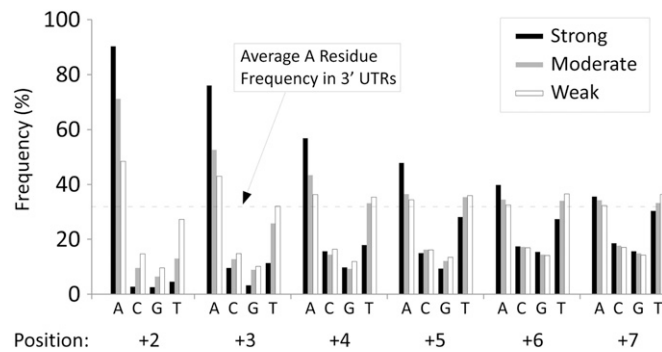
Moqtaderi et al. 10.1073/pnas.1309384110



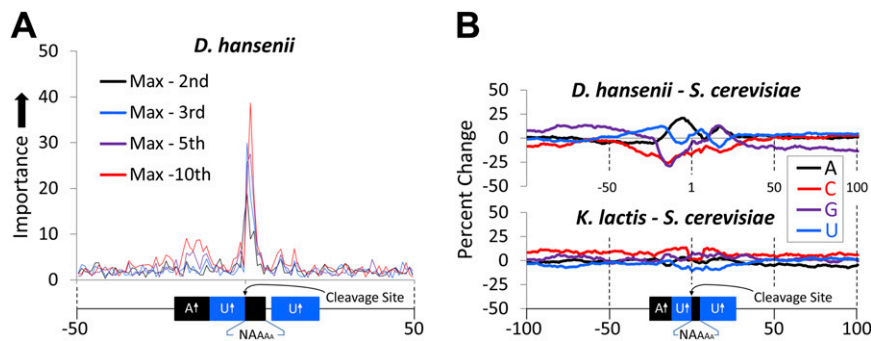
**Fig. S1.** Nucleotide frequency patterns at tRNA cleavage sites are highly dissimilar to those at cleavage sites of protein-coding genes and snoRNA genes. Cleavage/polyadenylation sites occurring early in the ORF possess a nucleotide frequency pattern reminiscent of that seen in 3'UTRs. Baseline variation is greater in the tRNA, snoRNA, and early ORF plots, in accord with the smaller number of genes analyzed.



**Fig. S2.** Frequencies of all possible dinucleotides near preferred poly(A) sites in *Saccharomyces cerevisiae* (Top), *Kluyveromyces lactis* (Middle), and *Debaryomyces hansenii* (Bottom).



**Fig. S3.** Decrease in A-residue frequencies immediately downstream of mRNA 3' ends correlated with a drop in signal strength. Averaged nucleotide frequencies for strong, moderate, and weak sites (>40%, 10%, and 2% of overall 3'UTR signal strength, respectively) show a large decrease in A-residue composition, especially at positions +2 to +5 relative to the cleavage site.



**Fig. S4.** Nucleotide patterns near preferred cleavage sites in *K. lactis* and *D. hansenii*. (A) Position-specific difference plots showing the Euclidean distance (*Materials and Methods*) between the most highly preferred cleavage site and the indicated less preferred (second-best, third-best, etc.) cleavage sites for every nucleotide shown. The extent of change is an indicator of the importance of the position. (B) Interspecies difference plots illustrate the distinctions in nucleotide frequencies between preferred sites in *S. cerevisiae* and those in the more distantly related *D. hansenii* (Upper) and the more closely related *K. lactis* (Lower).

Position Weight Matrix					
Position within 3' end	Position	A	C	G	T
-26	1	39	13	11	37
-25	2	38	13	13	37
-24	3	41	13	11	35
-23	4	41	12	12	36
-22	5	43	12	11	34
-21	6	45	11	11	33
-20	7	45	11	11	33
-19	8	48	11	10	31
-18	9	47	11	12	30
-17	10	47	12	12	30
-16	11	47	11	12	29
-15	12	46	13	12	30
-14	13	42	14	12	32
-13	14	40	14	14	33
-12	15	35	15	12	37
-11	16	32	15	13	40
-10	17	29	15	13	43
-9	18	28	15	13	44
-8	19	27	15	13	45
-7	20	27	16	14	43
-6	21	26	19	14	41
-5	22	27	17	13	43
-4	23	24	18	14	44
-3	24	22	19	14	46
-2	25	24	17	15	44
-1	26	25	25	25	25
1	27	14	25	22	38
2	28	83	5	3	9
3	29	57	13	8	22
4	30	47	14	8	31
5	31	35	16	11	37
6	32	32	19	15	34
7	33	31	19	15	35
8	34	31	17	16	35
9	35	29	18	17	36
10	36	28	18	17	37
11	37	28	17	17	37
12	38	27	18	16	40
13	39	27	17	15	41
14	40	29	17	15	40
15	41	28	17	14	41
16	42	29	17	15	39
17	43	29	17	16	38
18	44	30	19	14	37
19	45	30	18	15	37
20	46	29	19	16	37
21	47	30	18	16	36
22	48	31	19	16	35
23	49	31	17	15	36
24	50	31	18	17	34
25	51	31	18	17	34

**Fig. S5.** Position weight matrix used in assessing a region's likelihood to act as a cleavage site. The matrix was generated based on observed nucleotide frequencies at locations within 100 nt of nearly 5,000 dominant cleavage positions.