Quantitation of *cis*-translational control by general mRNA features in five eukaryotes

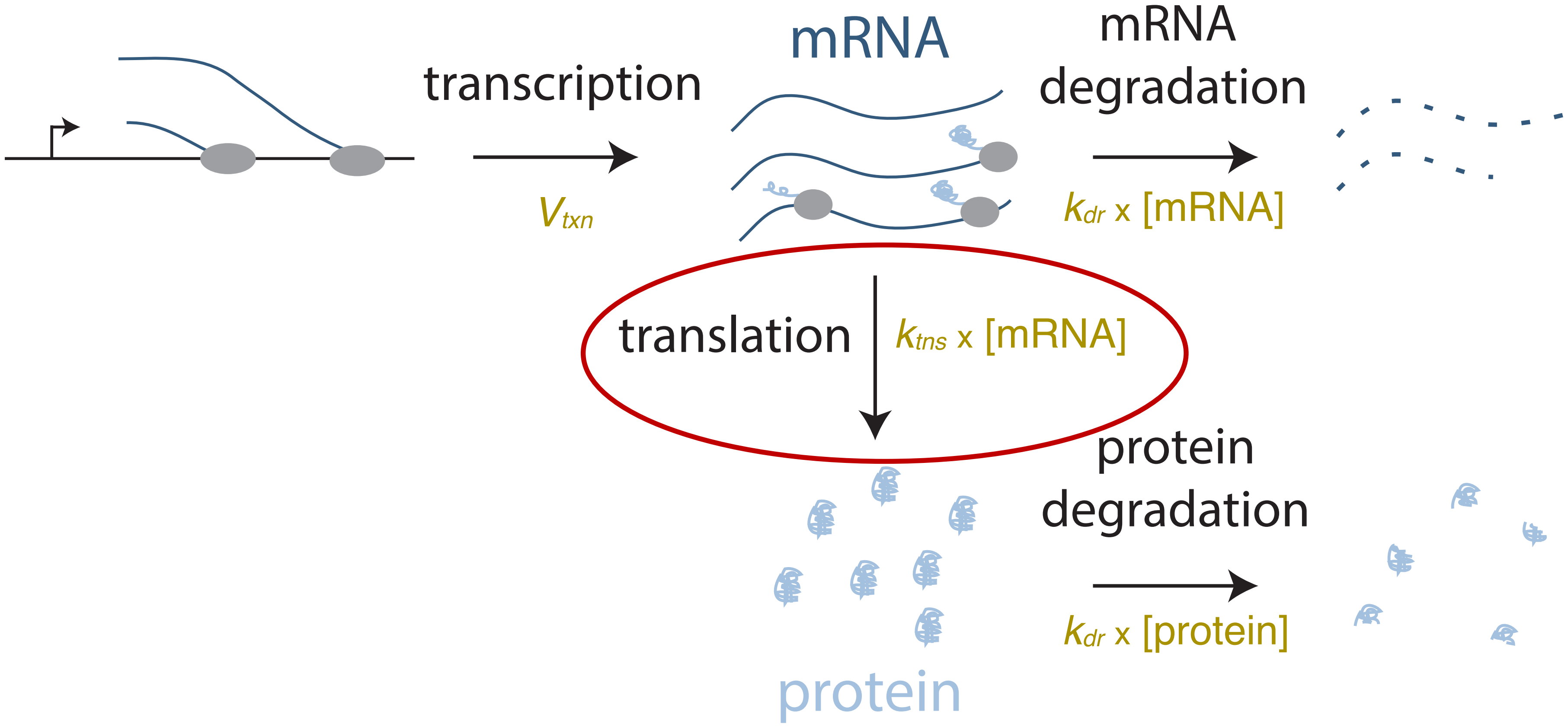
Jingyi Jessica Li

University of California, Los Angeles

Joint work with Drs. Guo-Liang Chew (Fred-Hutch) and Mark D. Biggin (LBNL)

SoCal SysBio Conference, Feb 9, 2019

Translation is one of four steps that determine protein abundances



transcription

V_{txn}

mRNA

mRNA degradation

$k_{dr} \times [mRNA]$

translation

$k_{tns} \times [mRNA]$

protein degradation

$k_{dr} \times [protein]$

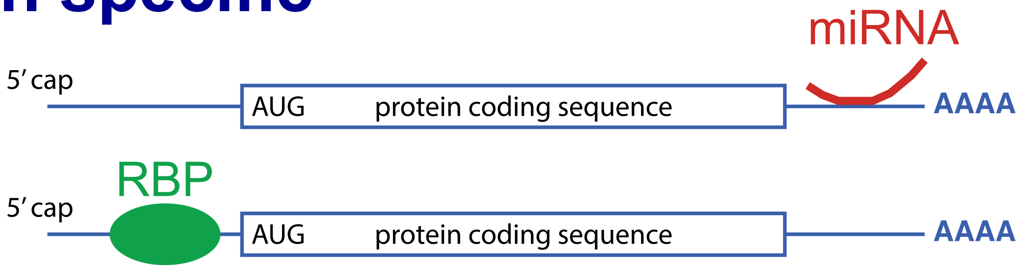
protein

Two classes of translational control

Gene and condition specific

e.g. miRNA

RBPs

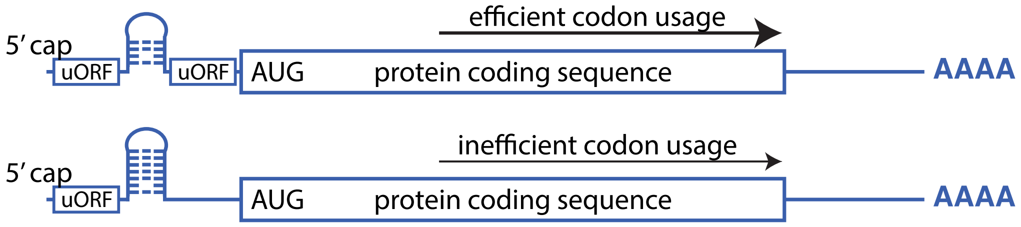


Present in a minority of mRNAs, cell types and conditions

General features

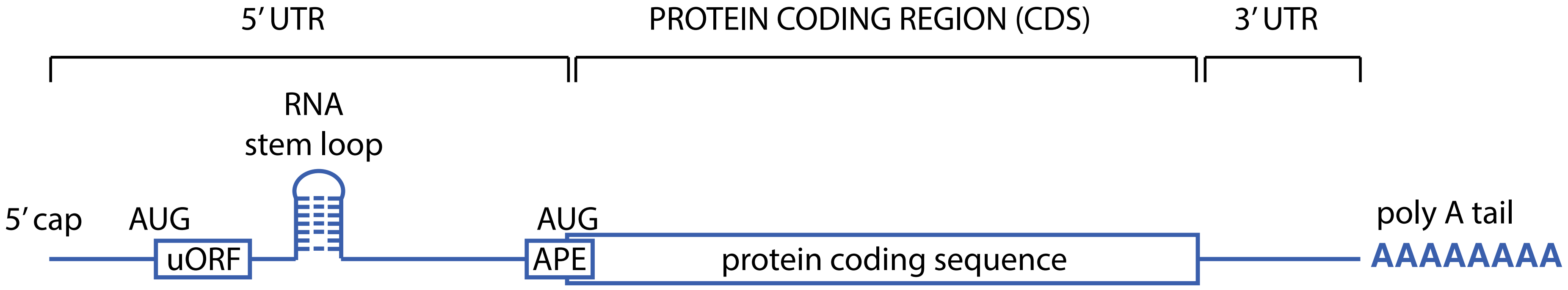
e.g. RNA structure

uORFs, codon usage



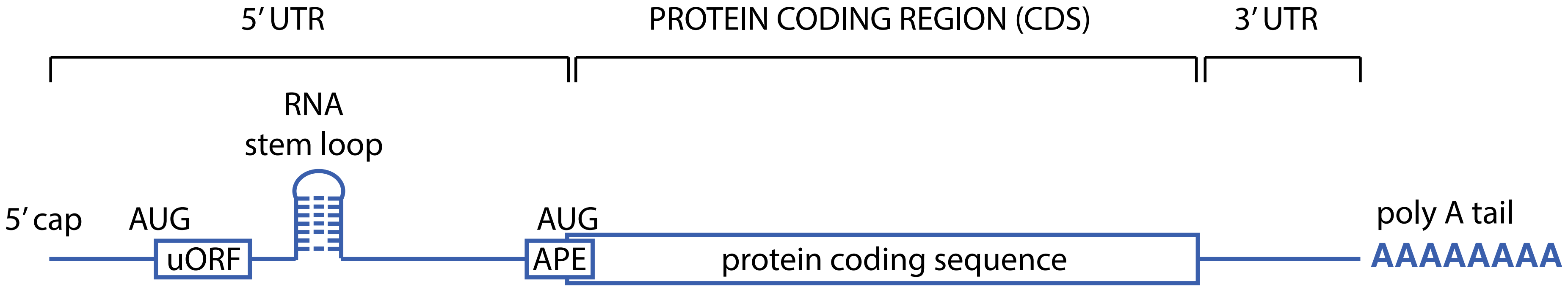
Present in most or all mRNAs, cell types and conditions

There are five general features



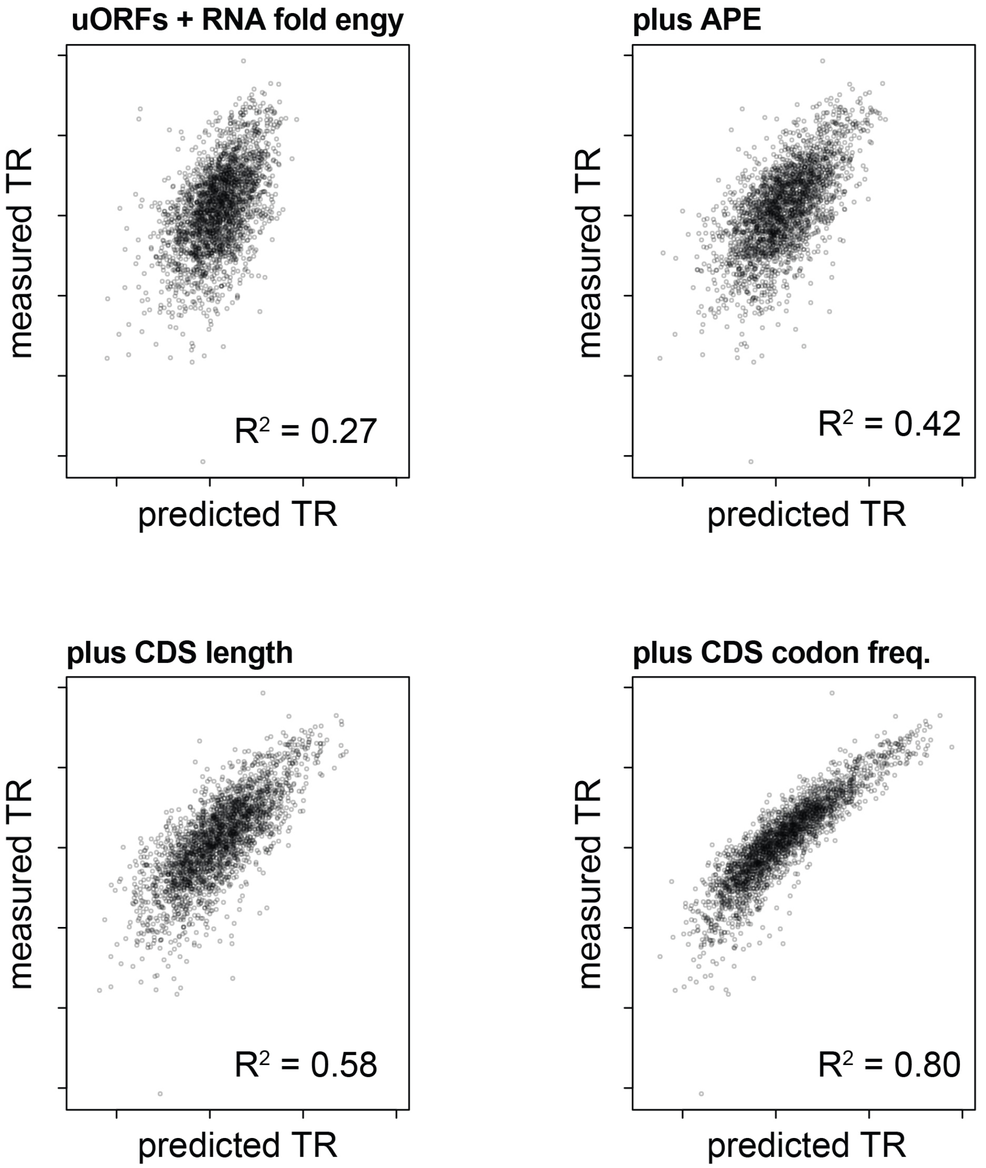
1. 5′ UTR RNA secondary structure
2. upstream open reading frames (uORFs)
3. AUG proximal element (APE)
4. protein coding sequence (CDS) length
5. CDS codon frequency

A multivariate linear model combining all five features



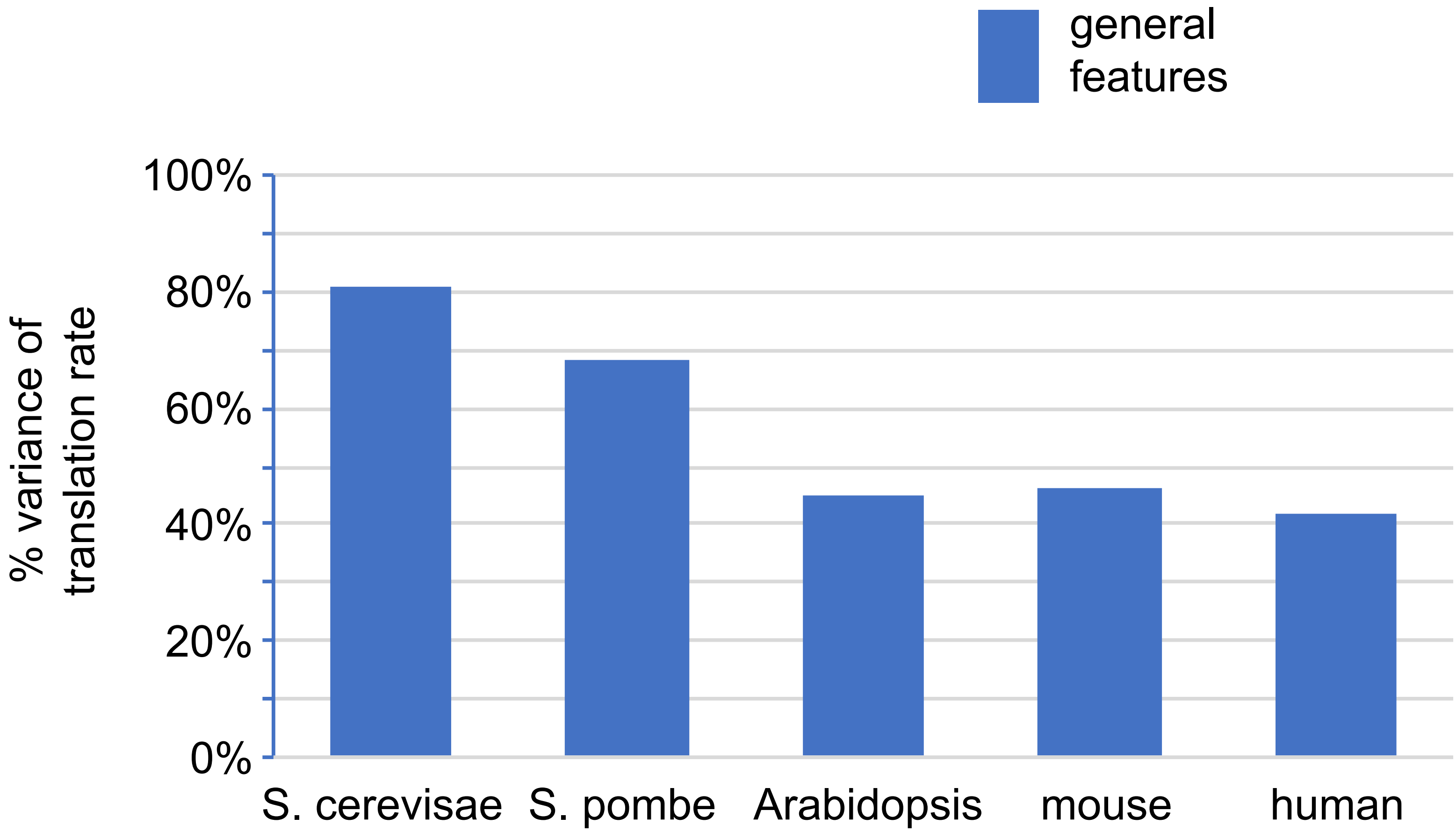
1. 5′ UTR RNA secondary structure
2. upstream open reading frames (uORFs)
3. AUG proximal element (APE)
4. protein coding sequence (CDS) length
5. CDS codon frequency

A multivariate linear model combining all five features

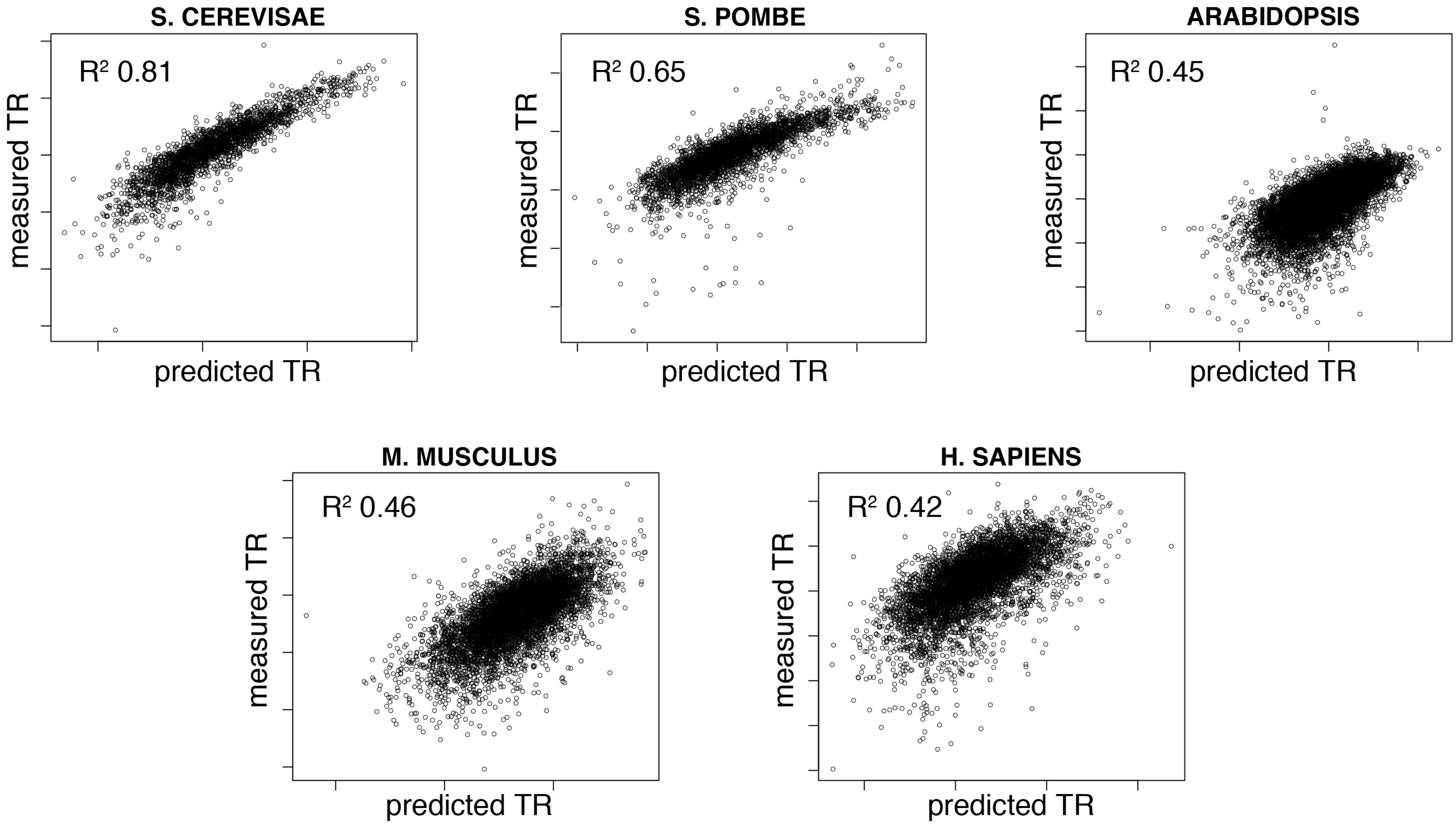


S. cerevisiae

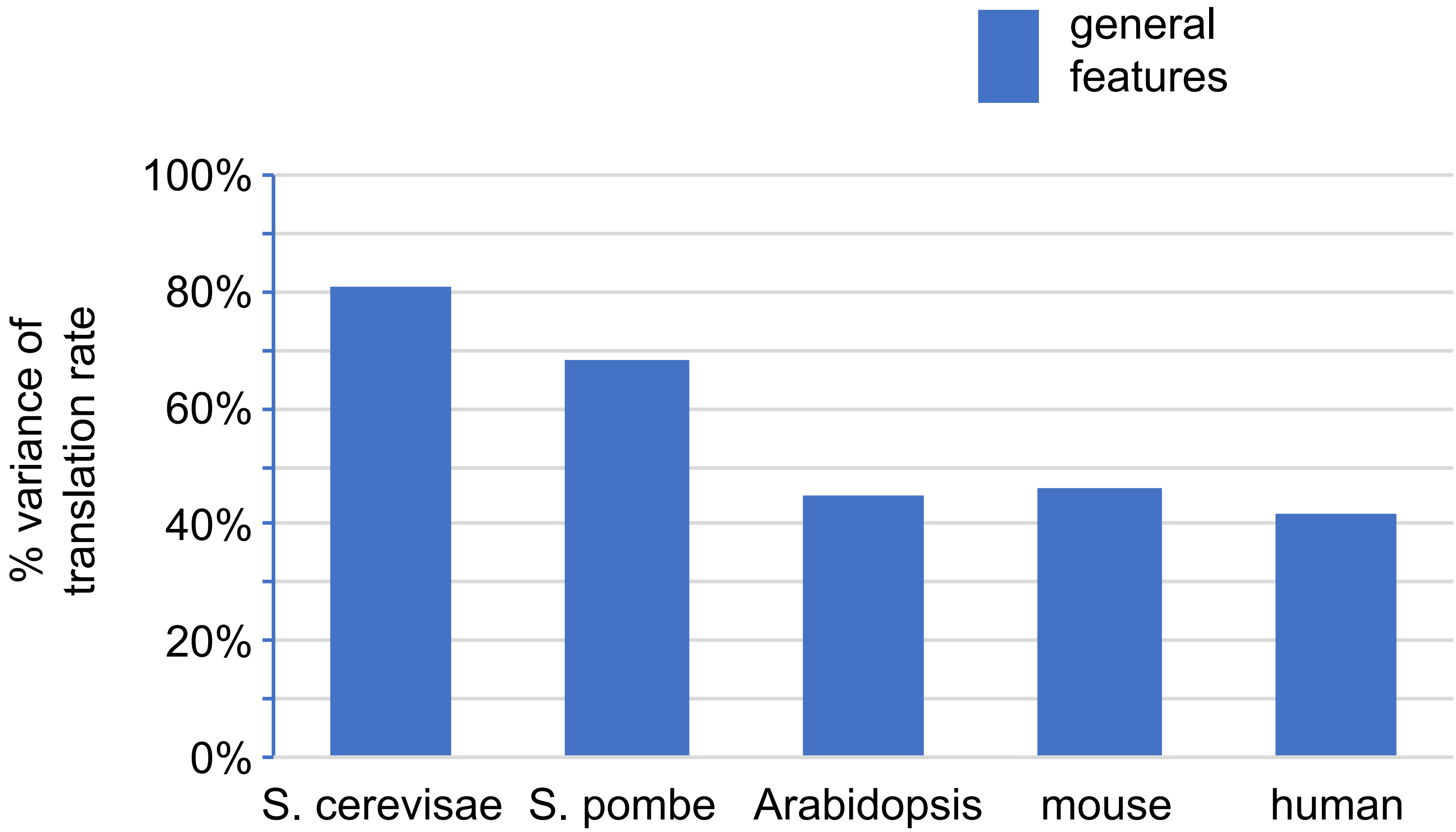
General features predict TR in five eukaryotes



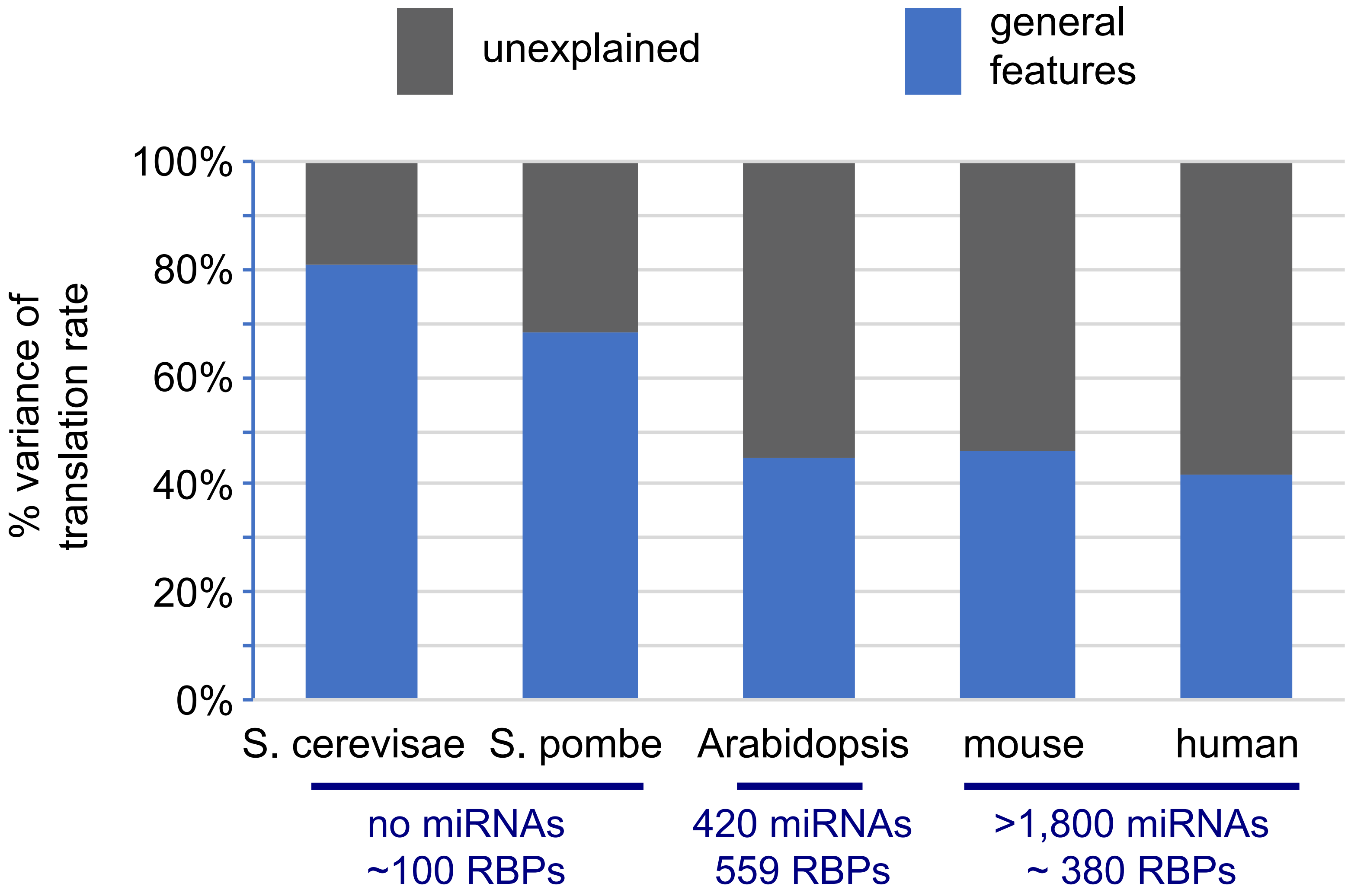
General features predict TR in five eukaryotes



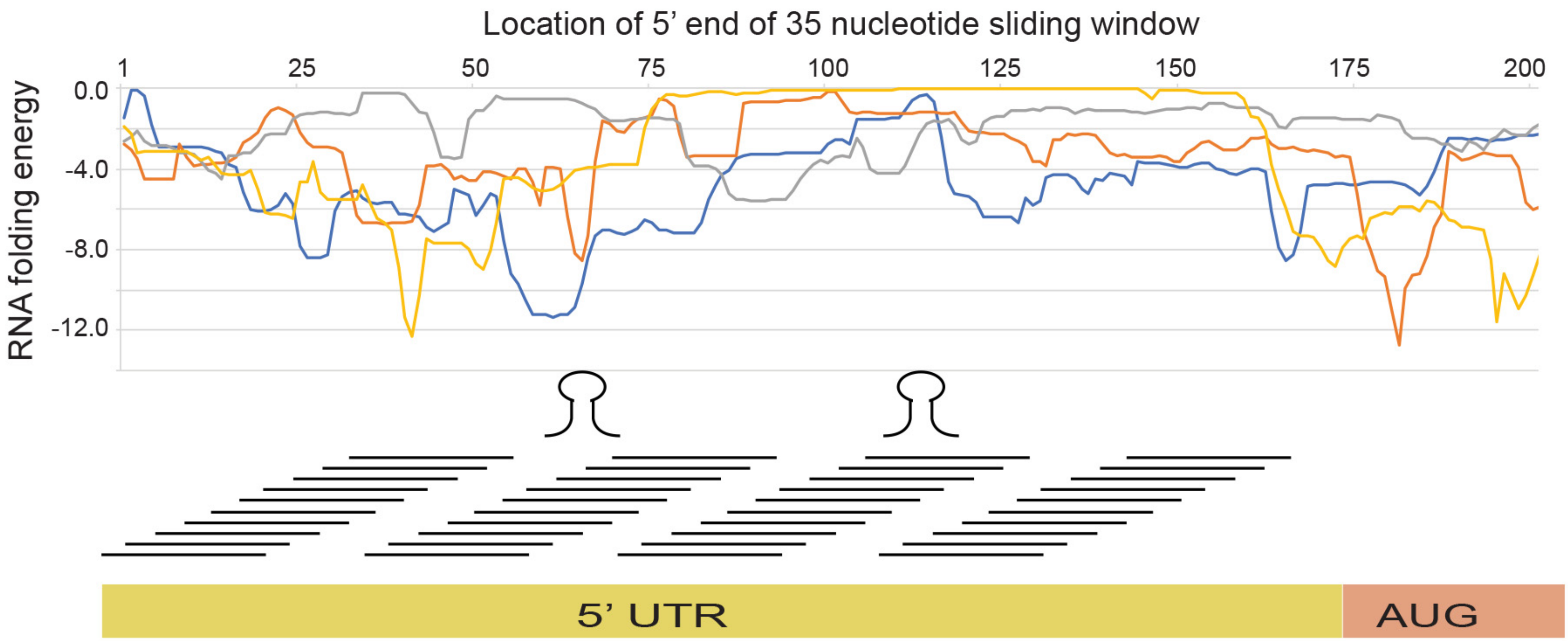
General features predict TR in five eukaryotes



Unexplained variance sets a limit on gene and condition specific control

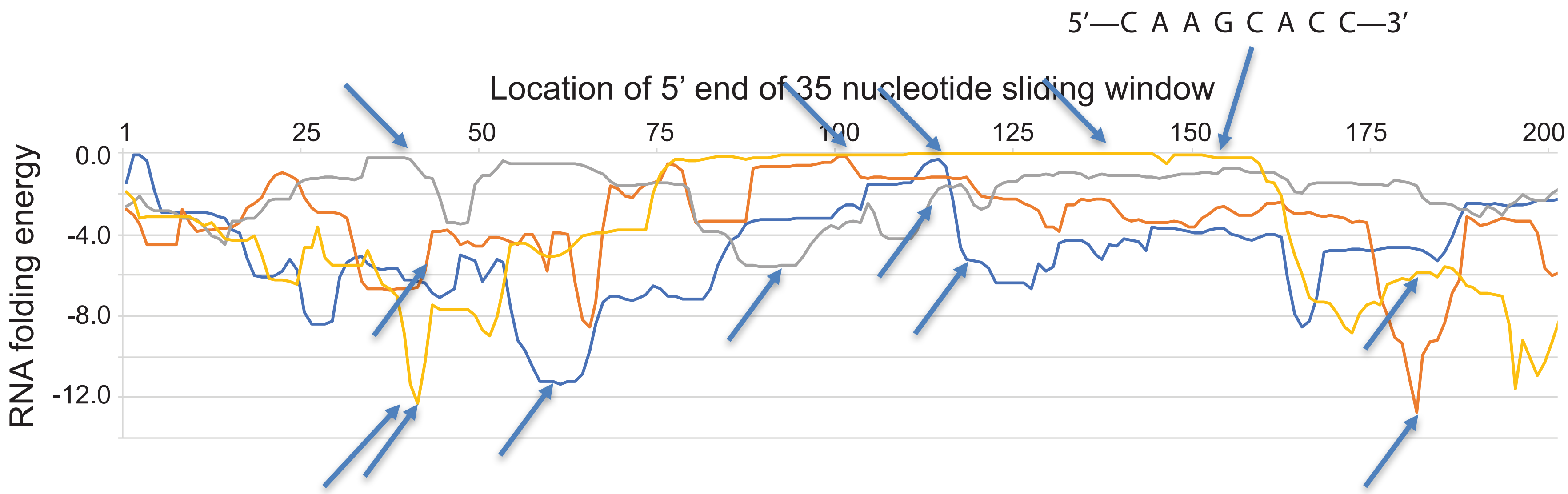


Control of translation by 5’ UTR RNA structure

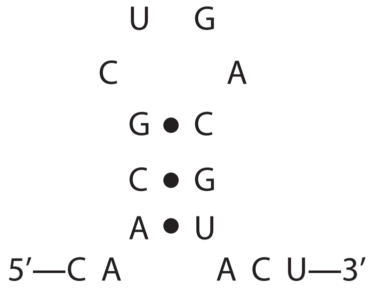


35 nucleotide sliding windows folded

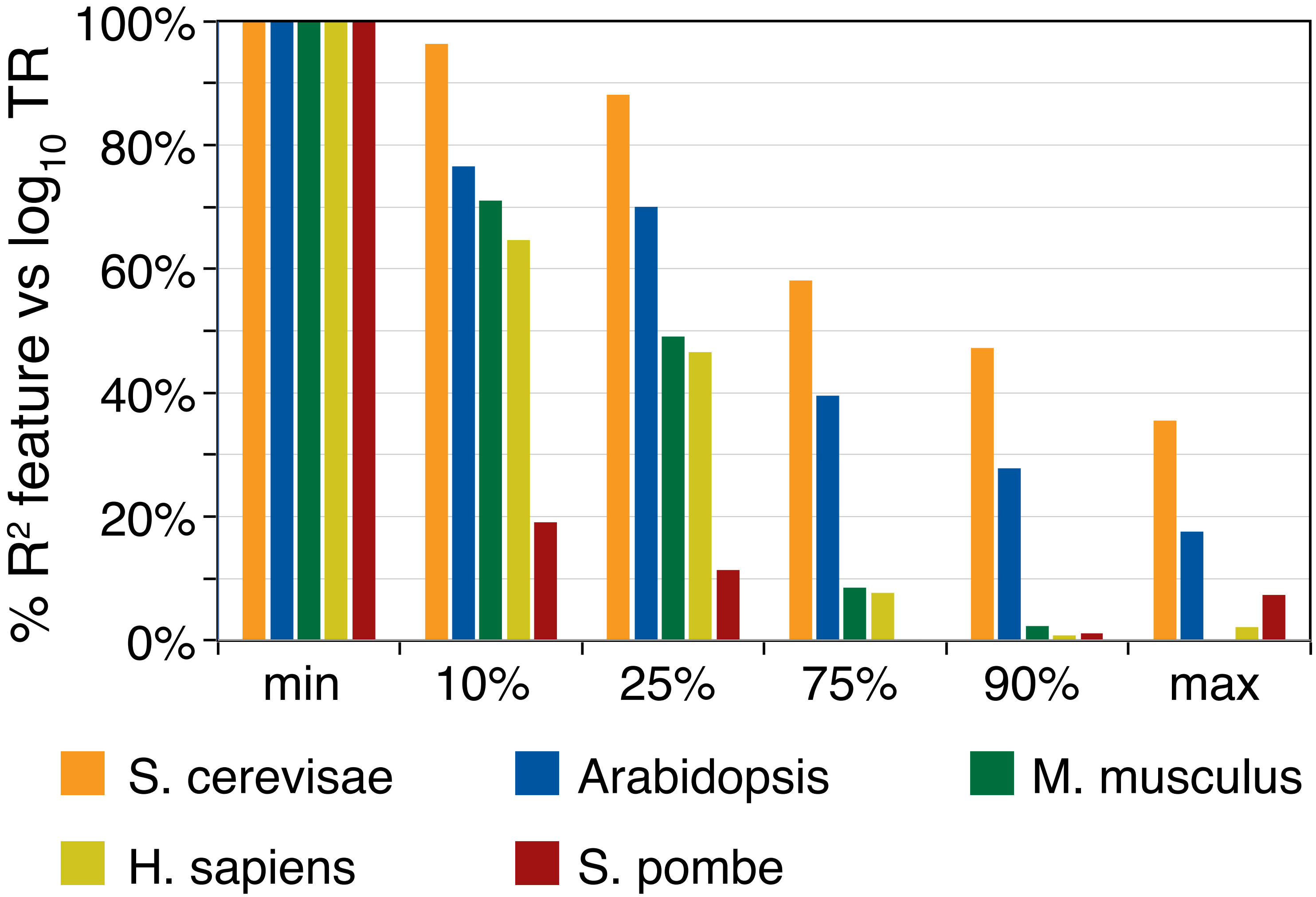
Control of translation by 5’ UTR RNA structure



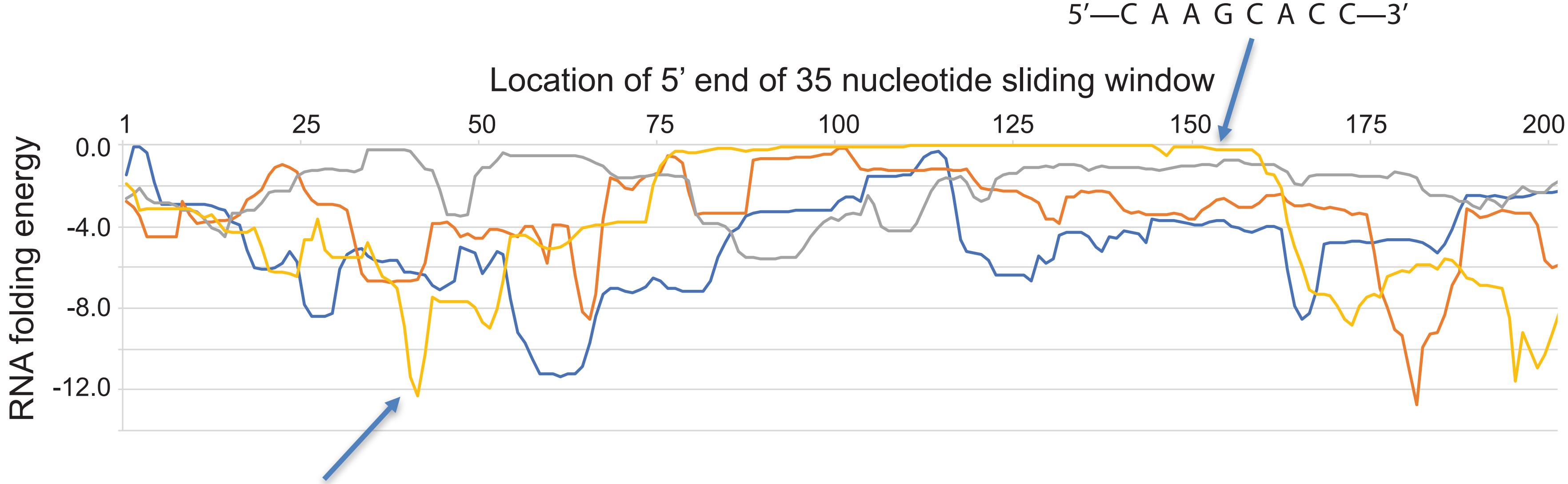
5’—C A A G C A C C—3’

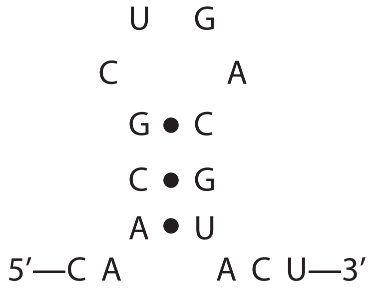


The most folded regions within 5’ UTRs determine translation rates

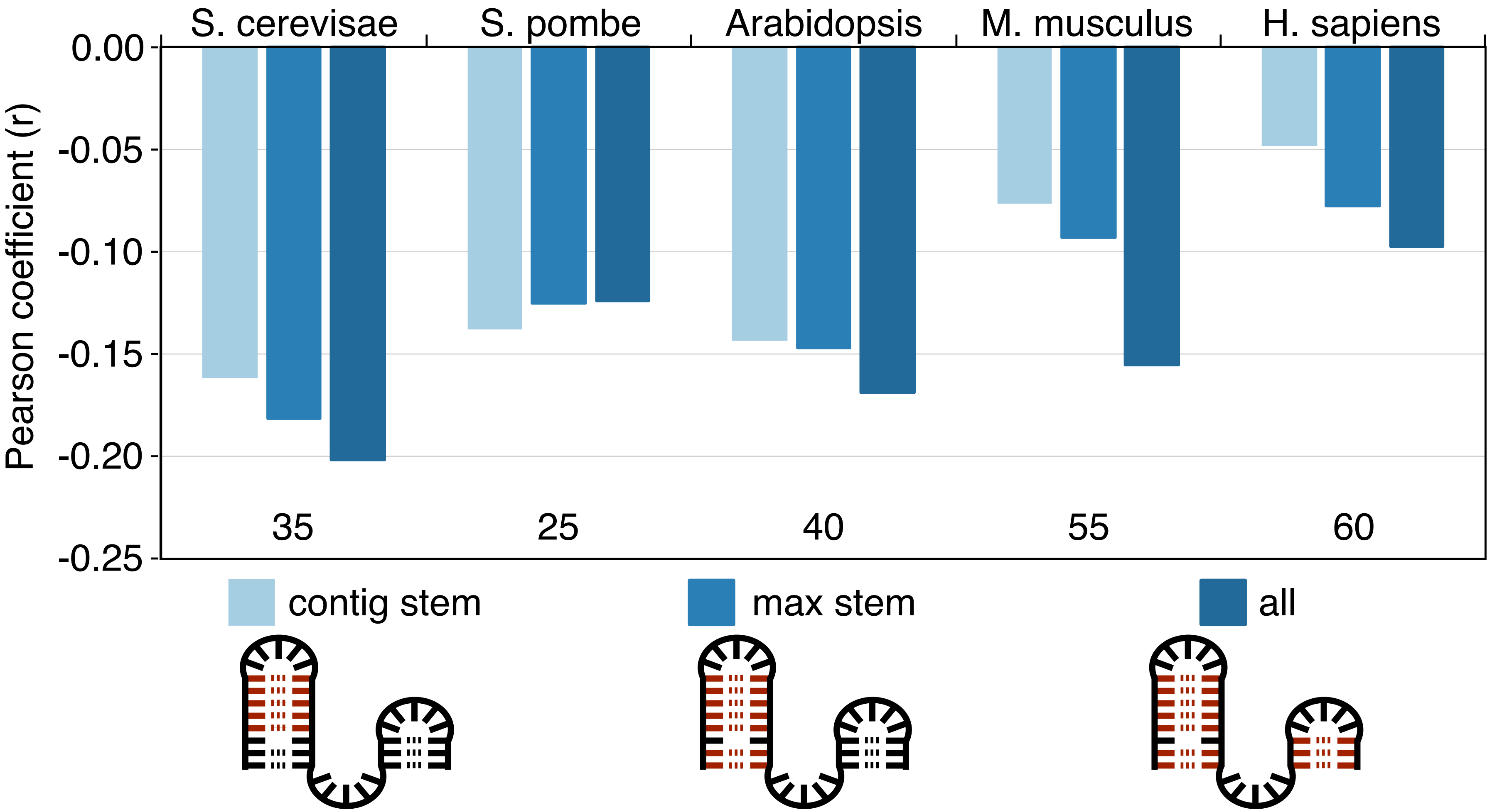


The most folded regions within 5’ UTRs determine translation rates

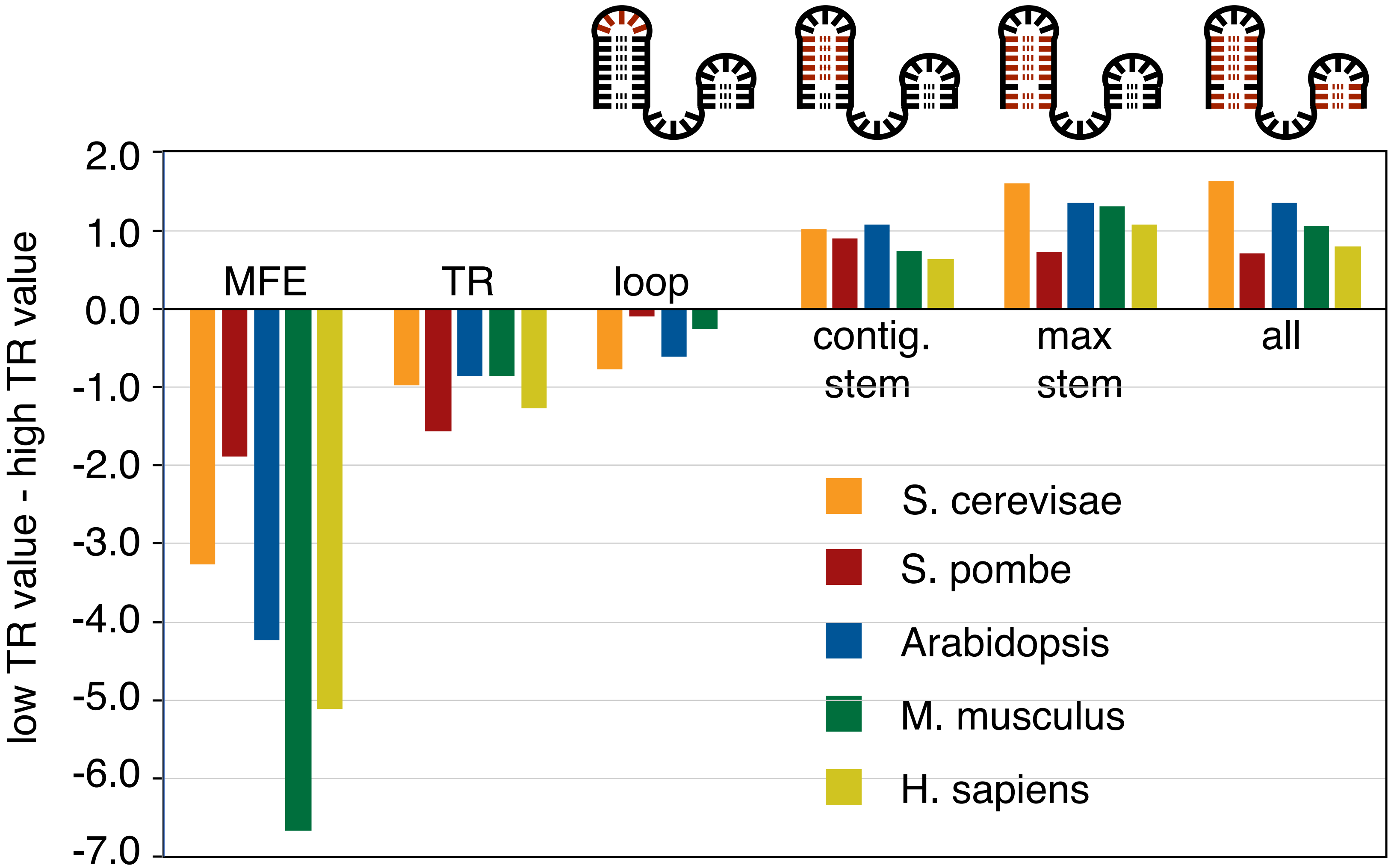




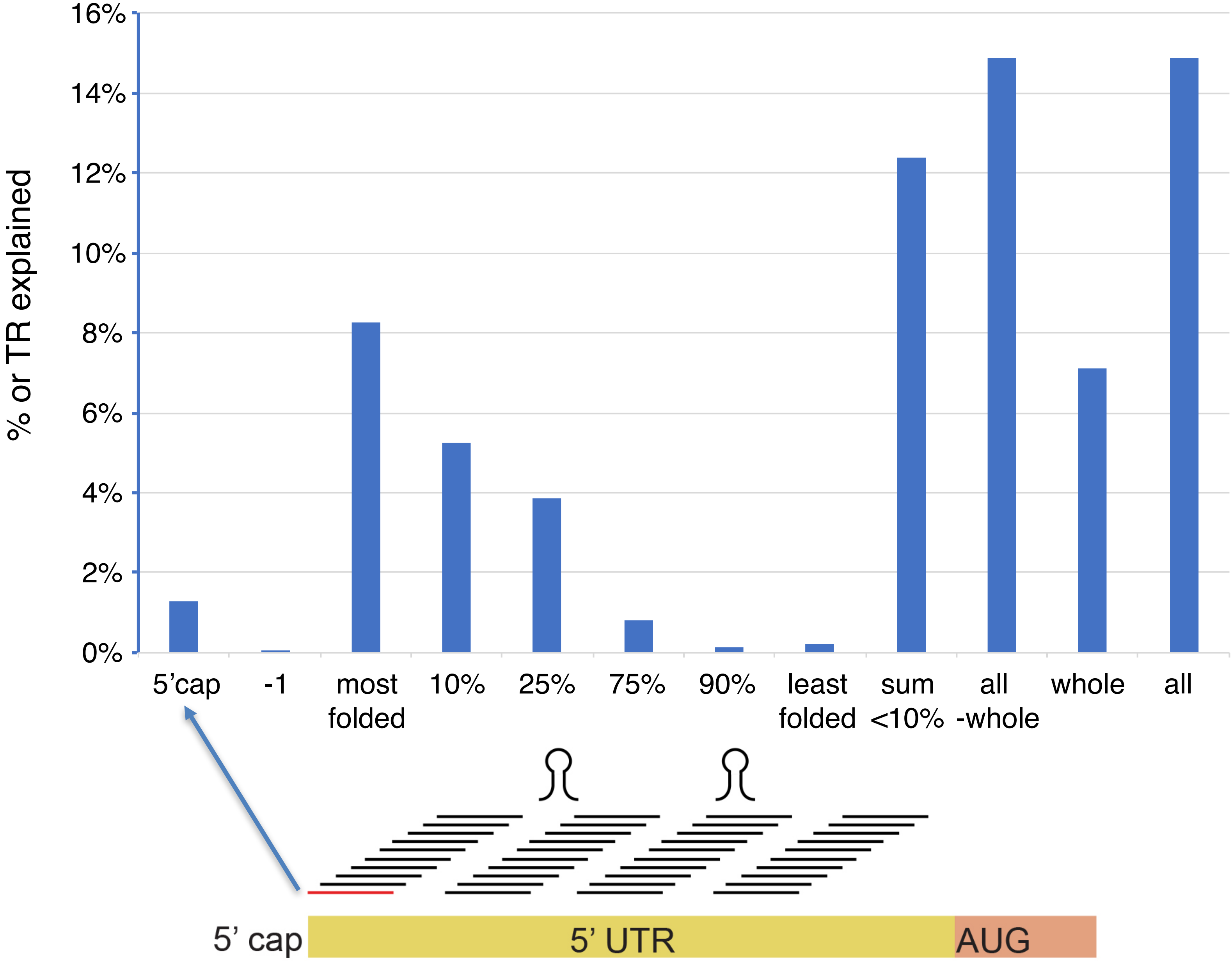
Contiguous stems are most effective, but other nucleotide pairs contribute



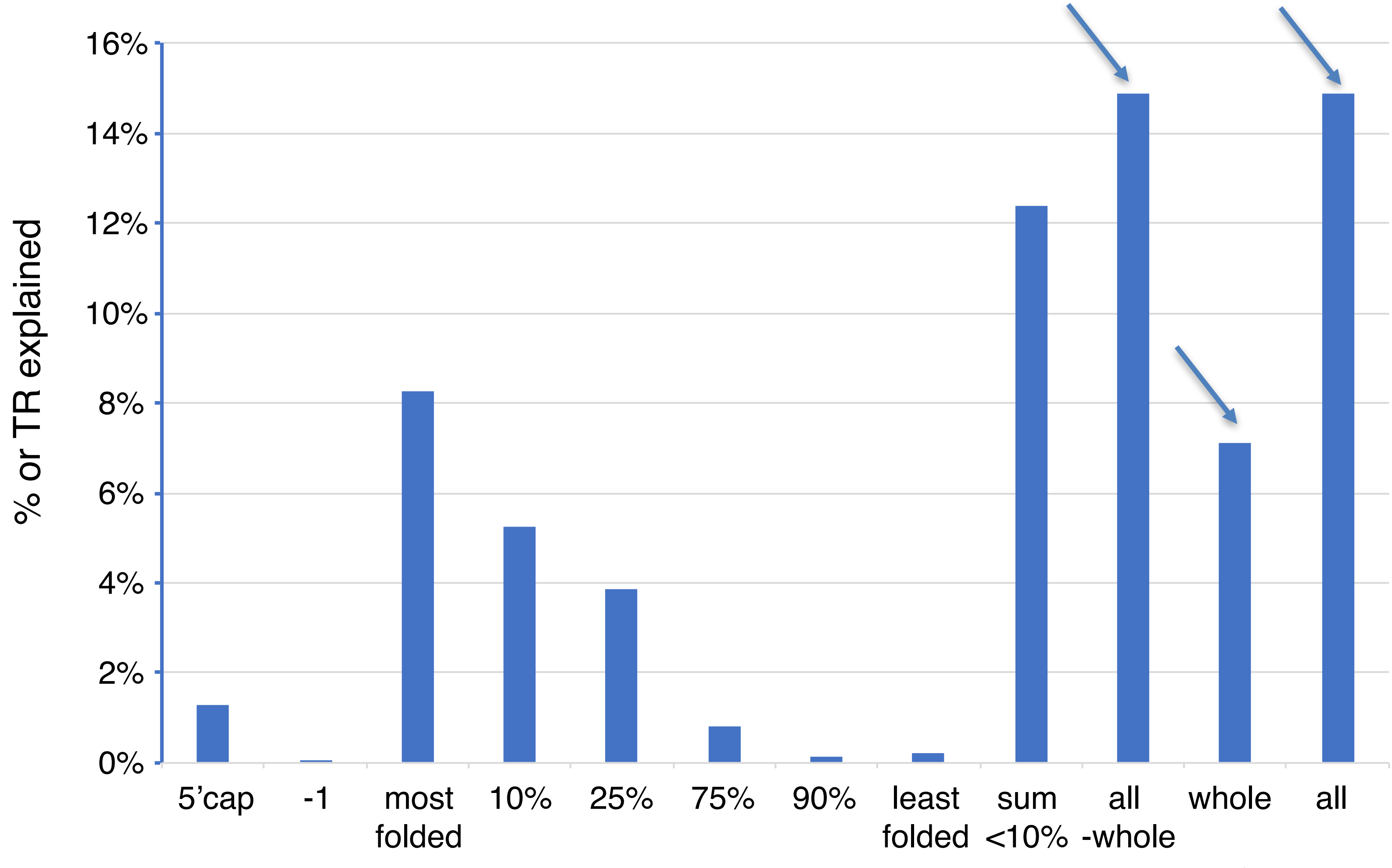
Most folded regions are similar across eukaryotes

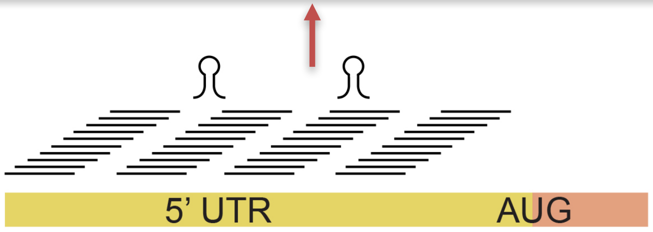


Structure at the 5’ cap contributes little

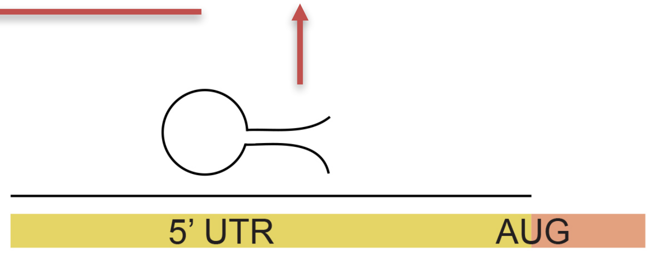


Long distance loops make little contribution



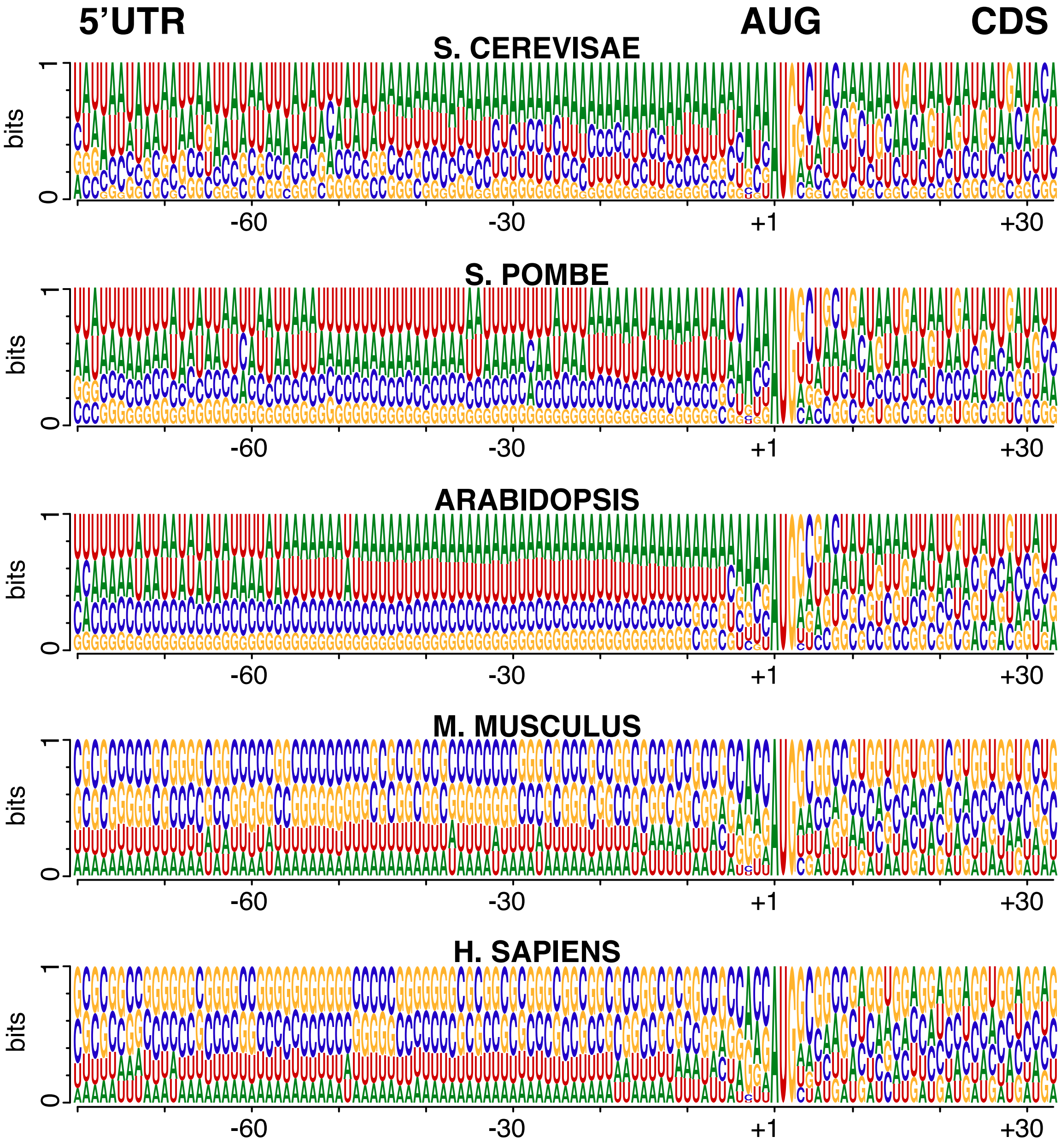


35 nucleotide sliding windows folded

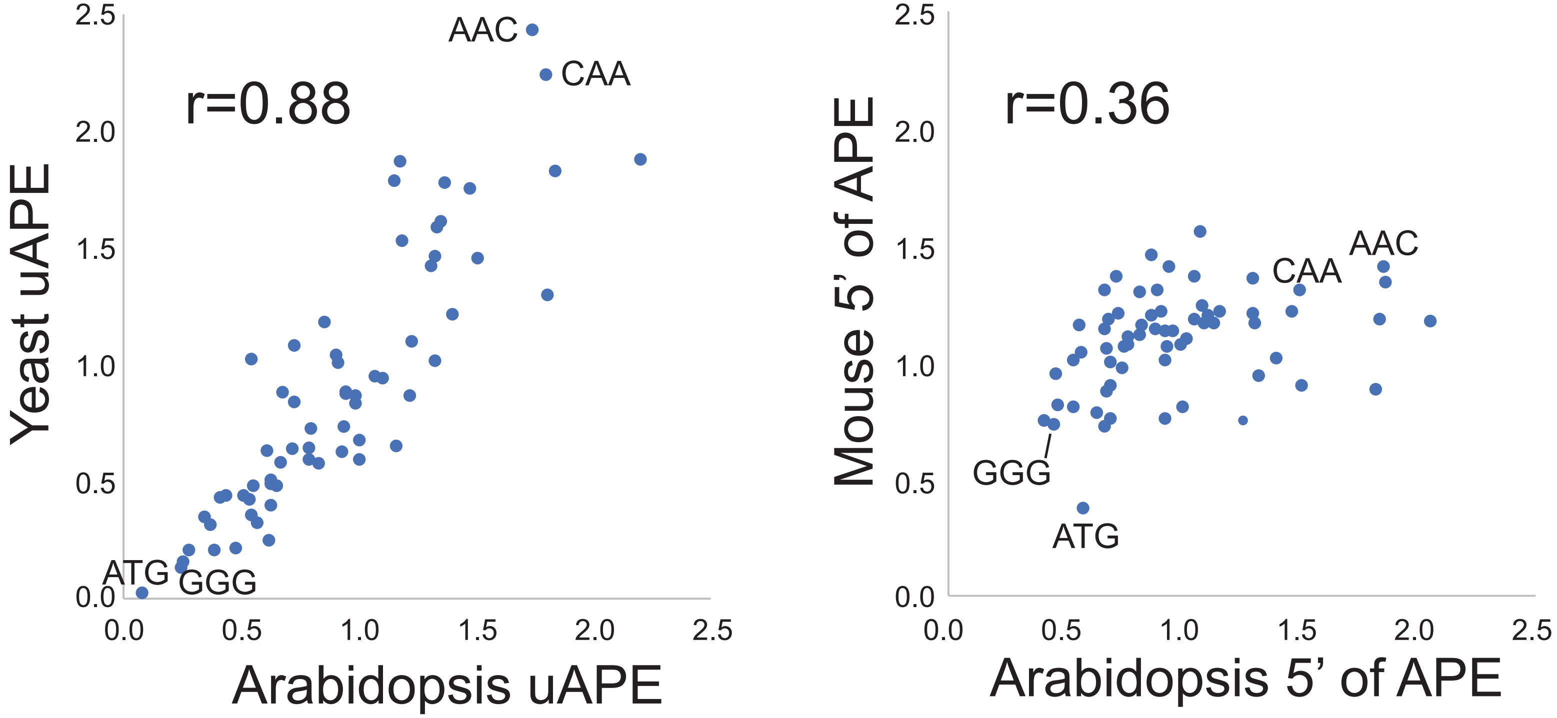


whole 5′ UTR folded

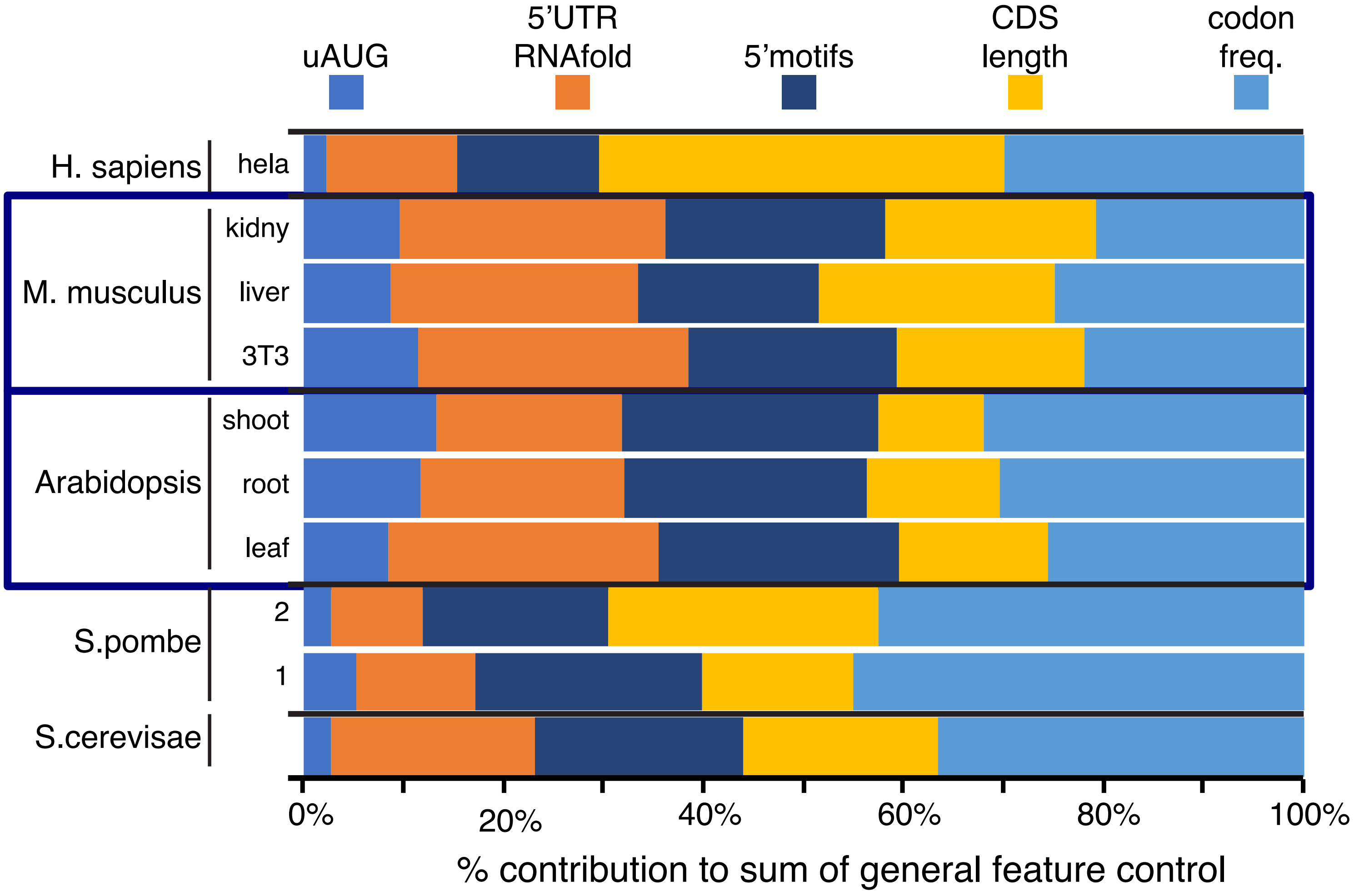
Translation sequence motifs



Similar control sequences across Eukarya



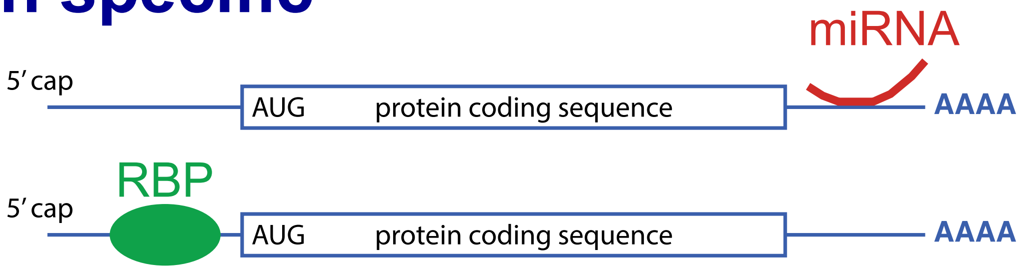
General features act similarly in different tissues, but differently between species



Gene and condition specific

e.g. miRNA

RBPs

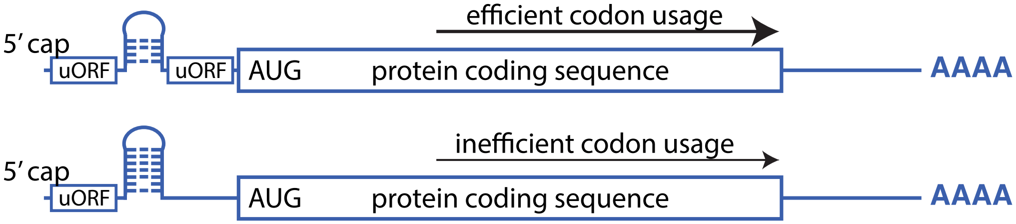


Present in a minority of mRNAs, cell types and conditions

General features

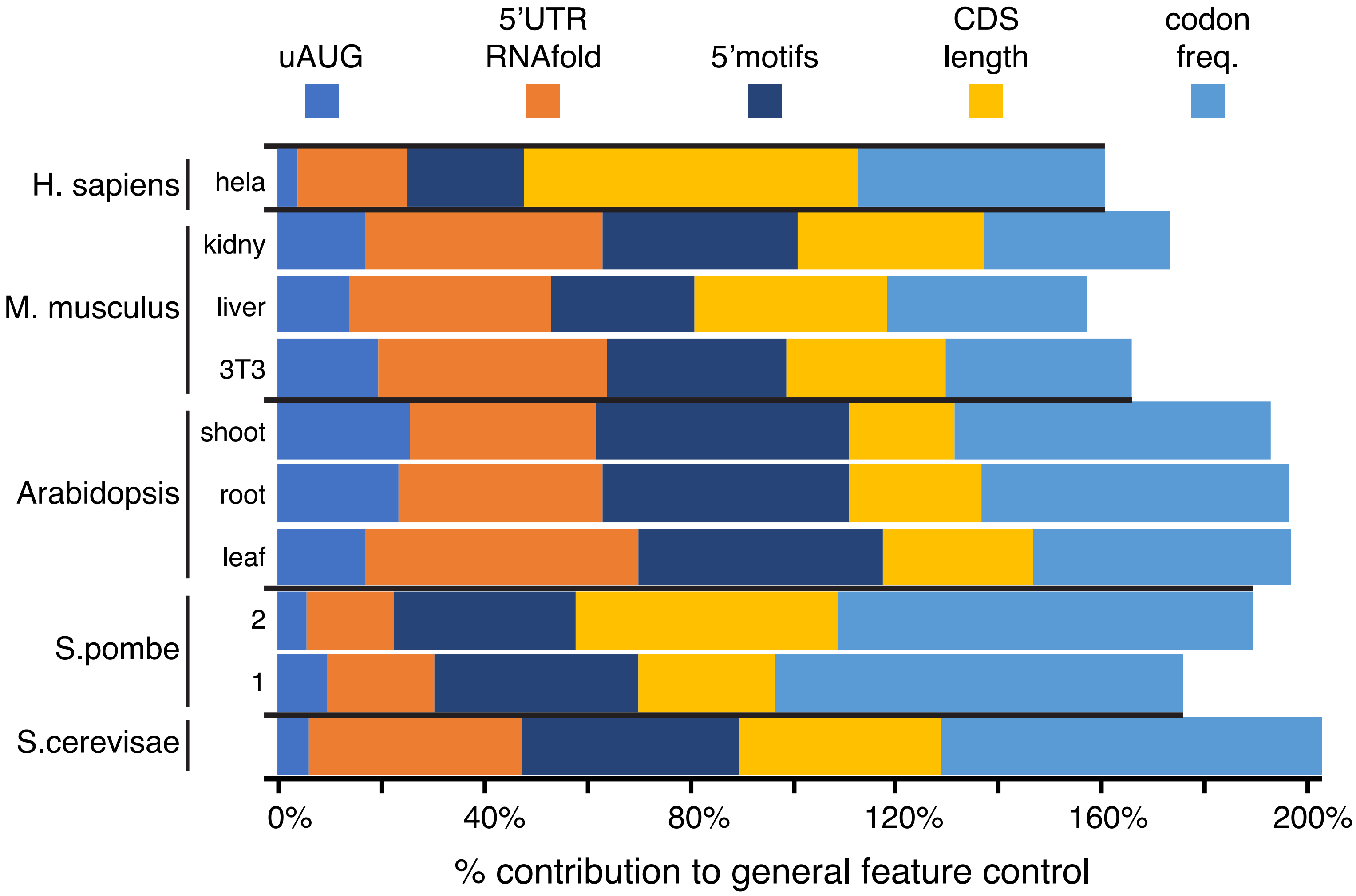
e.g. RNA structure

uORFs, codon usage



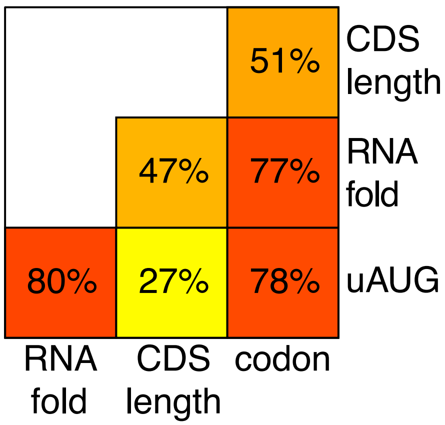
Present in most or all mRNAs, cell types and conditions

The percent contributions sum to >100%



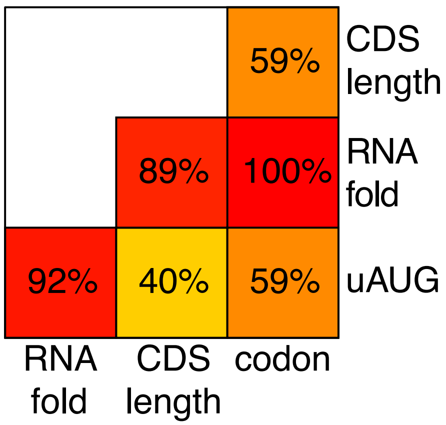
Highly correlated (collinear) control between the general features

S. CEREVISAE



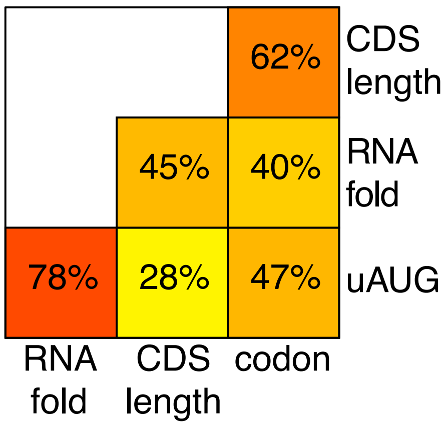
	RNA fold	CDS length	codon
CDS length			51%
RNA fold		47%	77%
uAUG	80%	27%	78%

S.POMBE



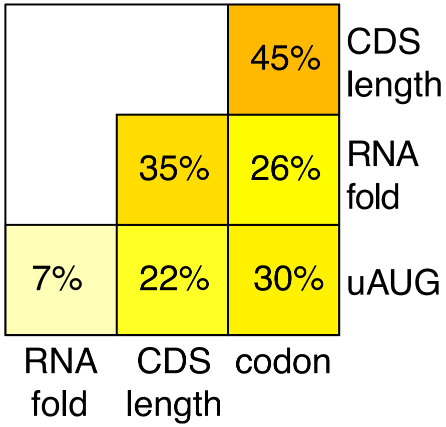
	RNA fold	CDS length	codon
CDS length			59%
RNA fold		89%	100%
uAUG	92%	40%	59%

ARABIDOPSIS



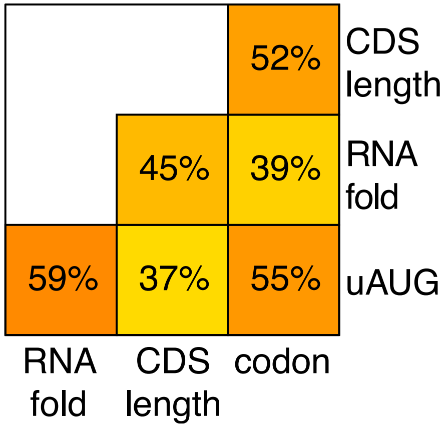
	RNA fold	CDS length	codon
CDS length			62%
RNA fold		45%	40%
uAUG	78%	28%	47%

M. MUSCULUS



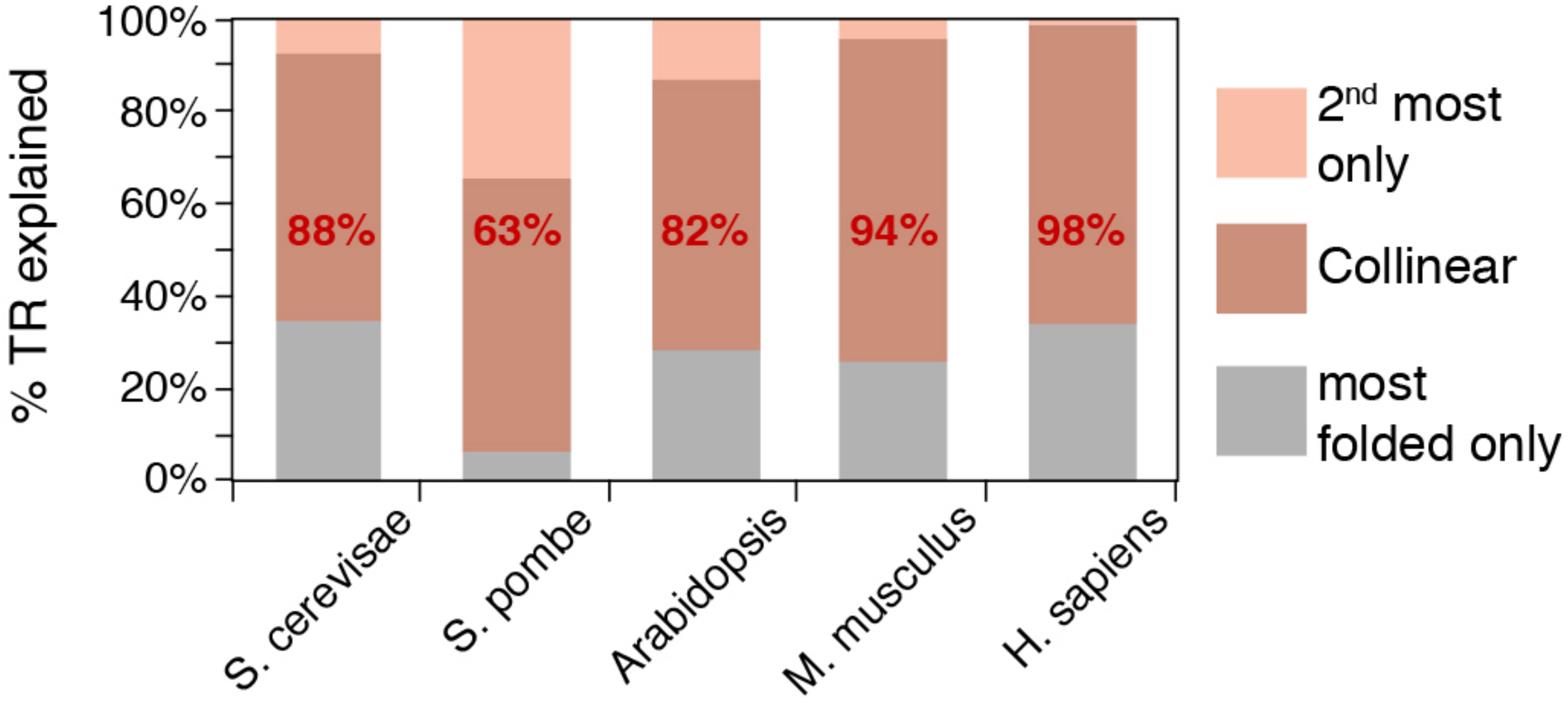
	RNA fold	CDS length	codon
CDS length			45%
RNA fold		35%	26%
uAUG	7%	22%	30%

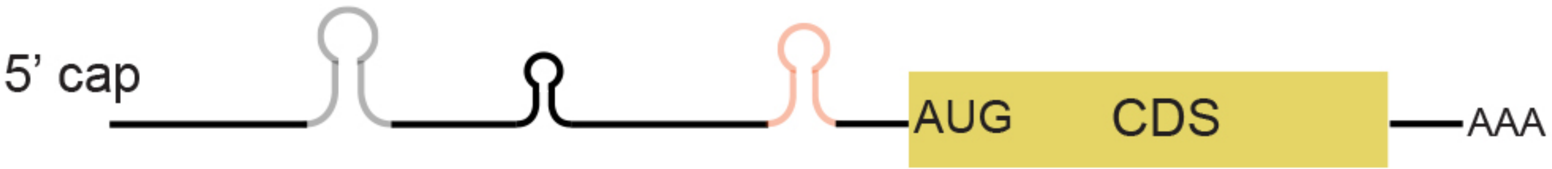
H. SAPIENS



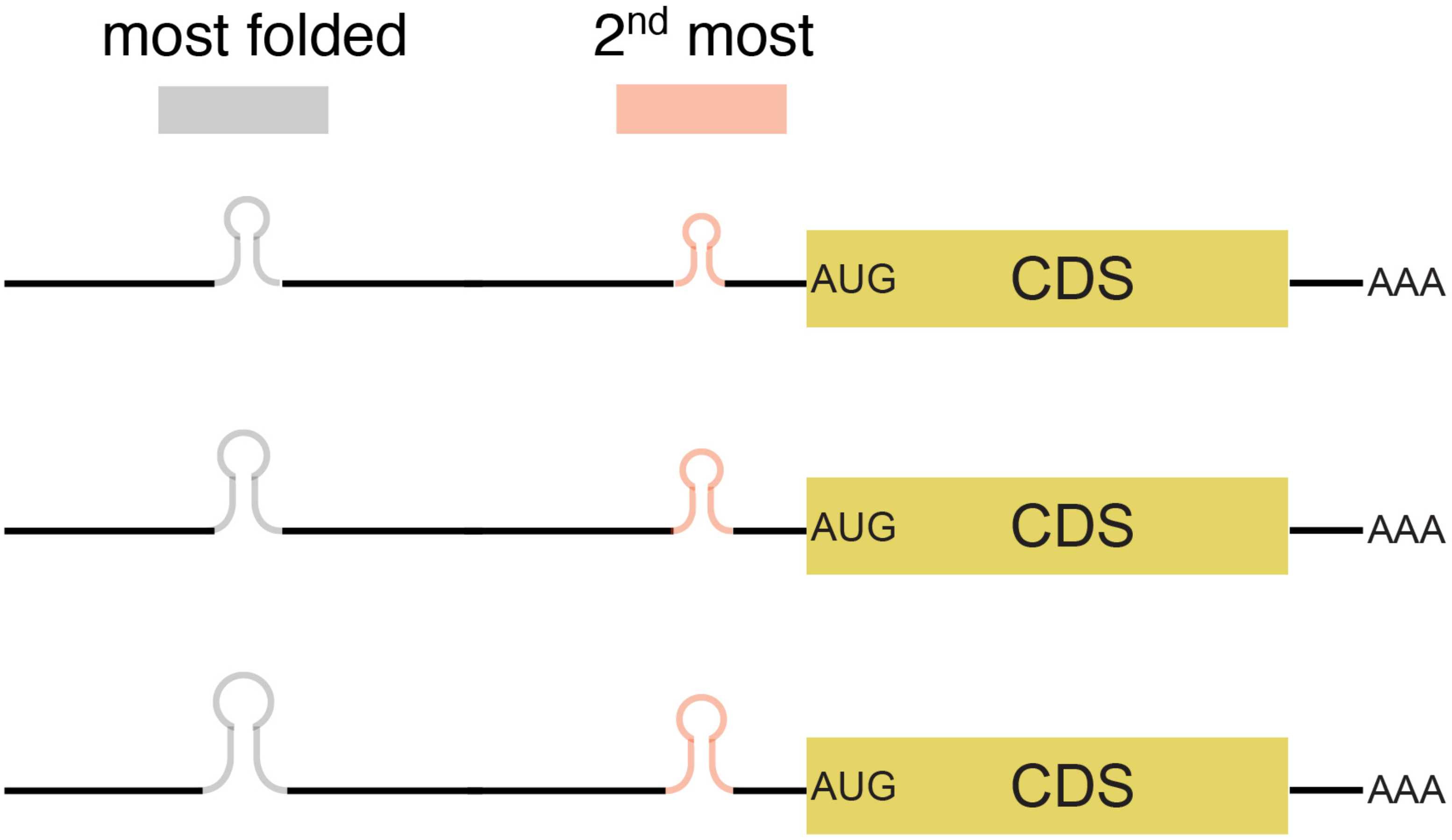
	RNA fold	CDS length	codon
CDS length			52%
RNA fold		45%	39%
uAUG	59%	37%	55%

Control by separate stem loops is highly correlated

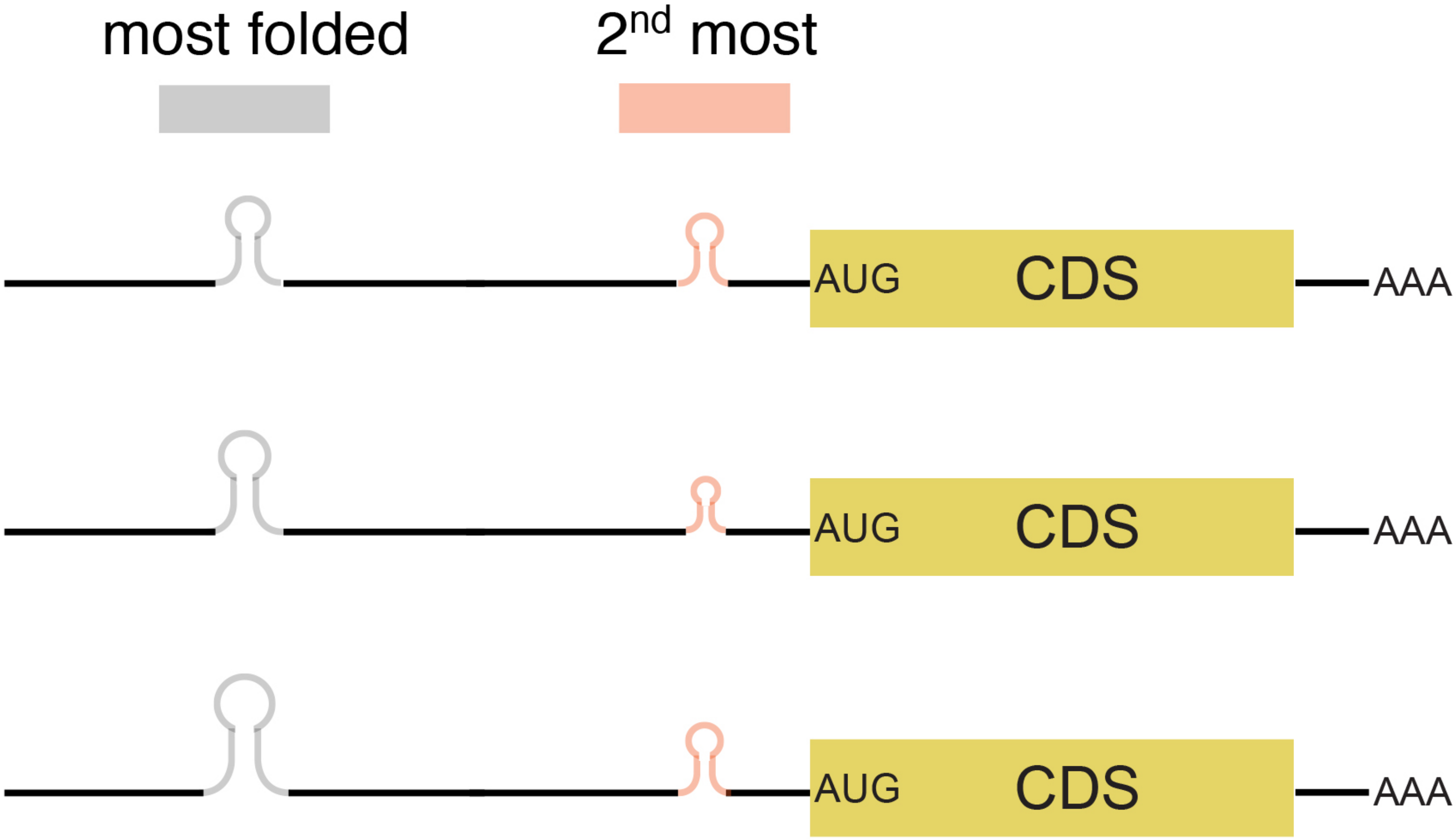




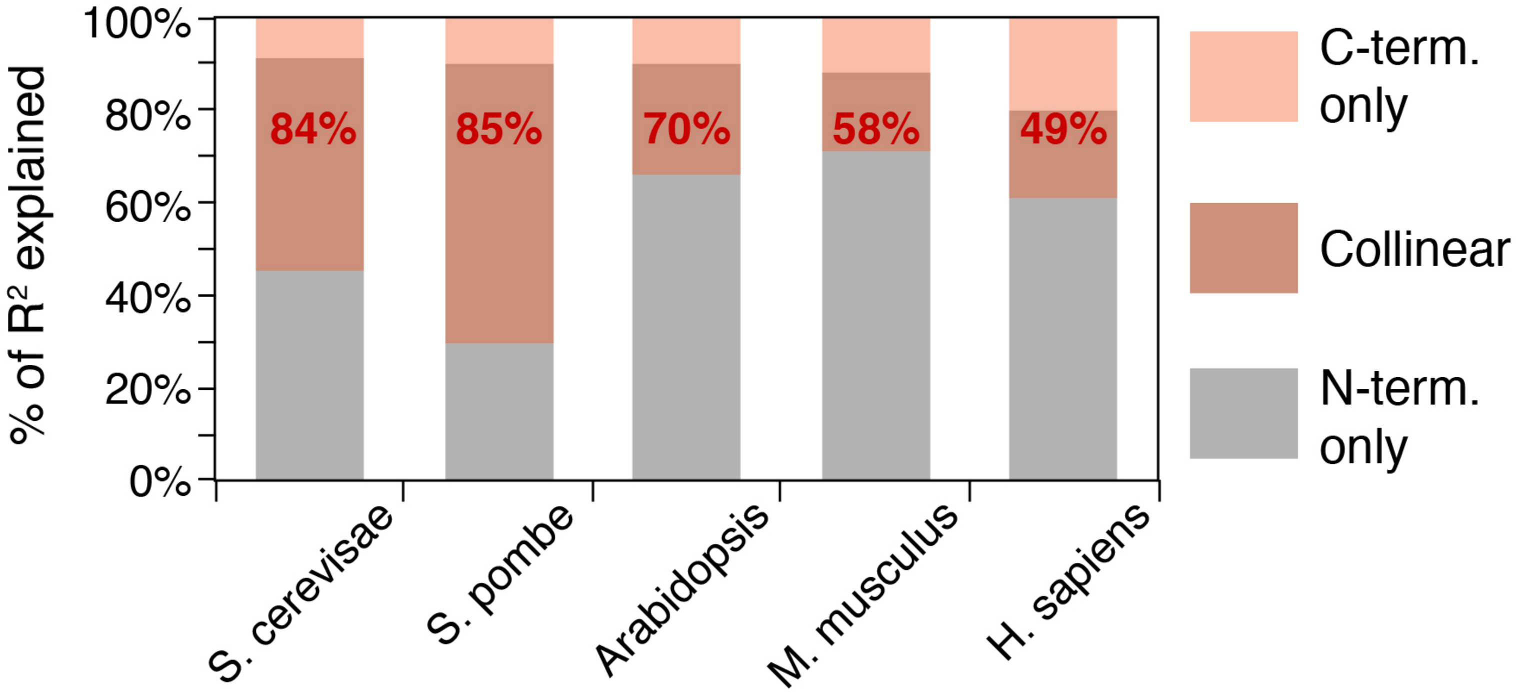
Control by separate stem loops is highly correlated

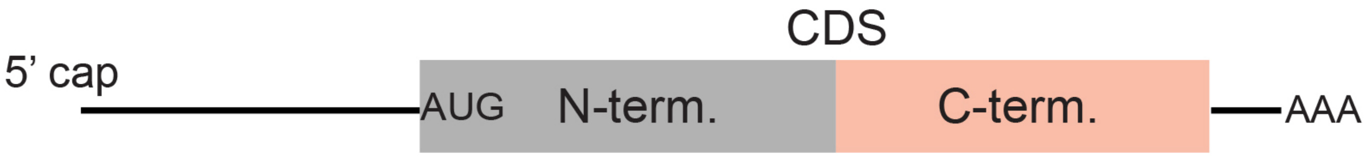


Control by separate stem loops is highly correlated

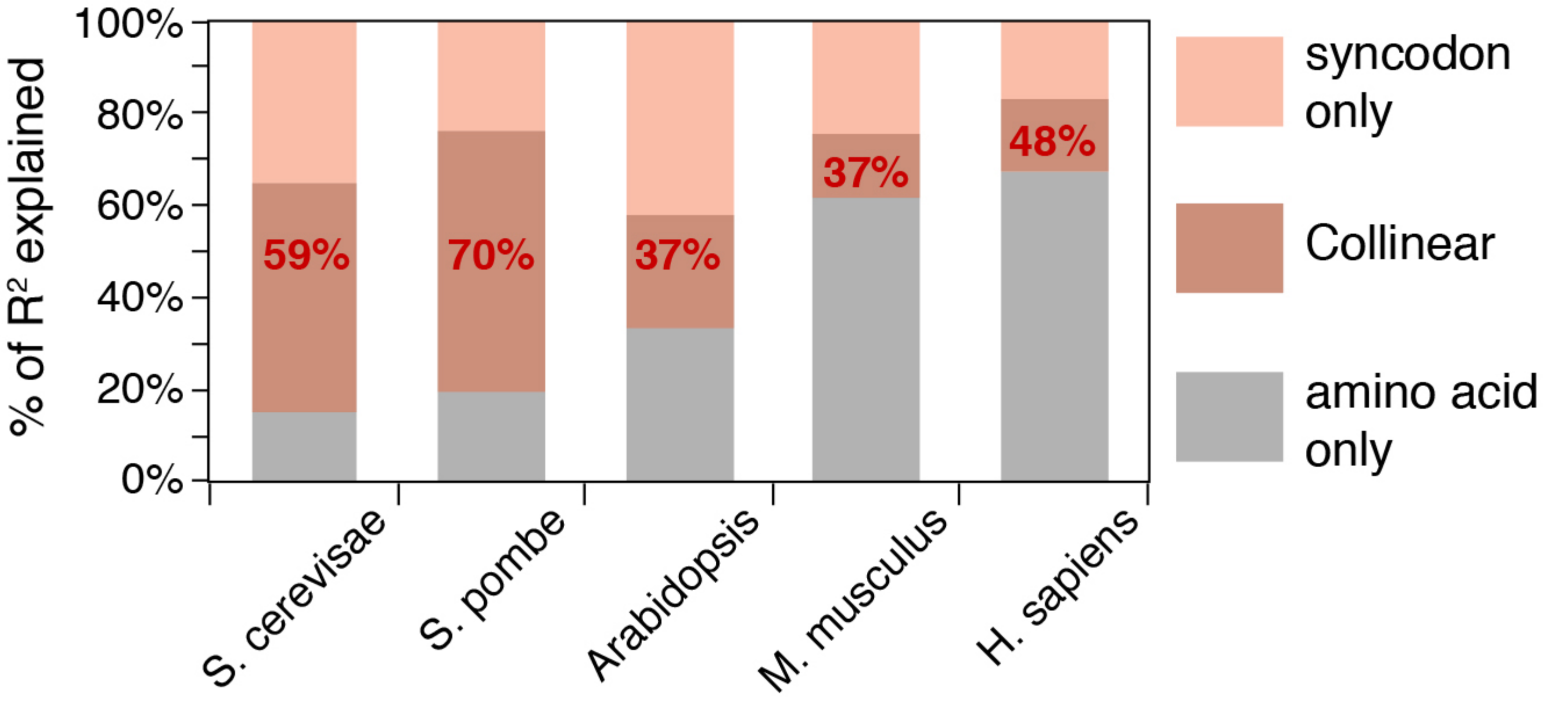


Collinear control by N- and C- terminal codon usage





Collinear control by amino acid frequency and synonymous codon preferences

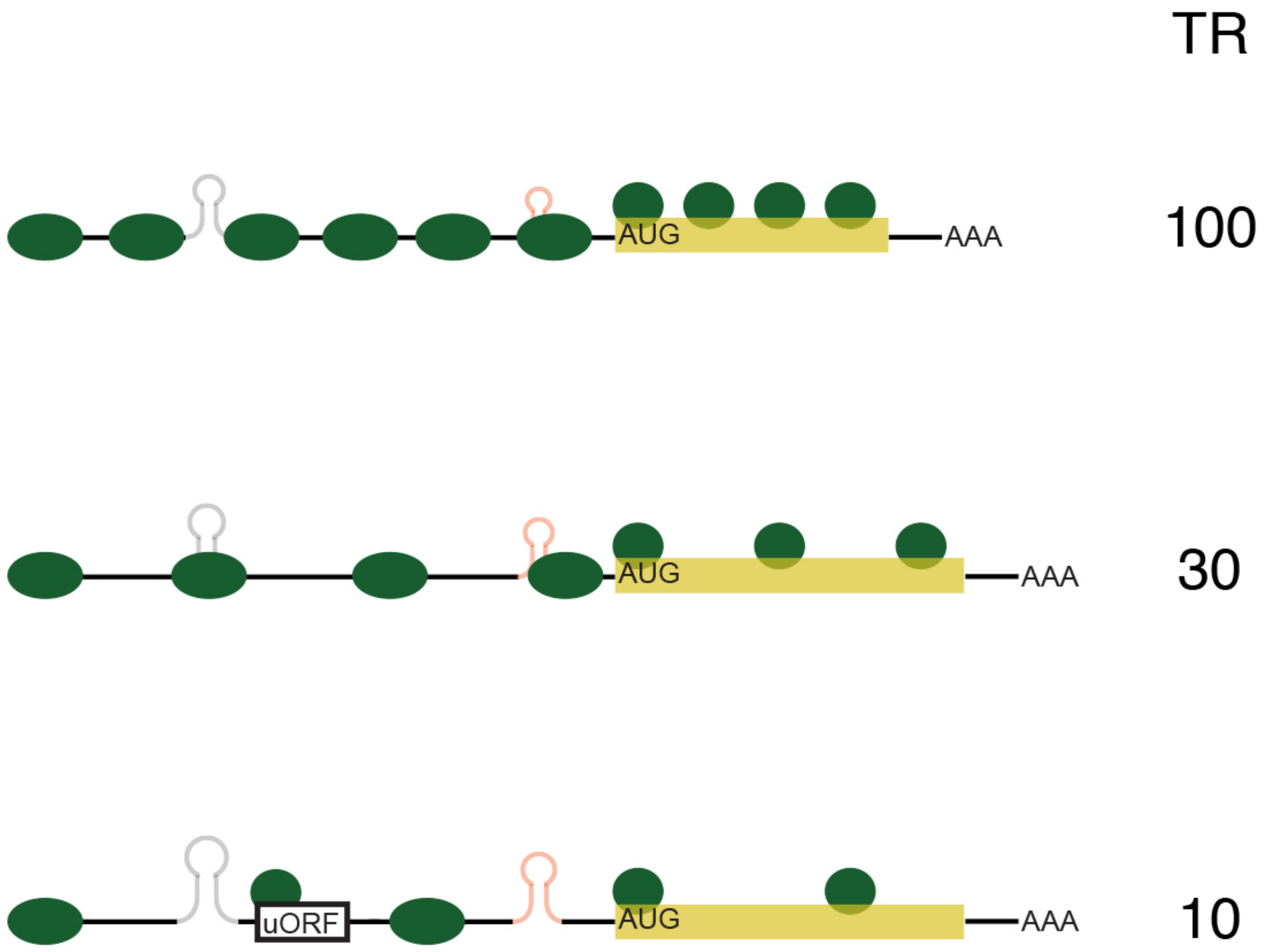


codon freq. = amino acid freq. x syn. codon pref.

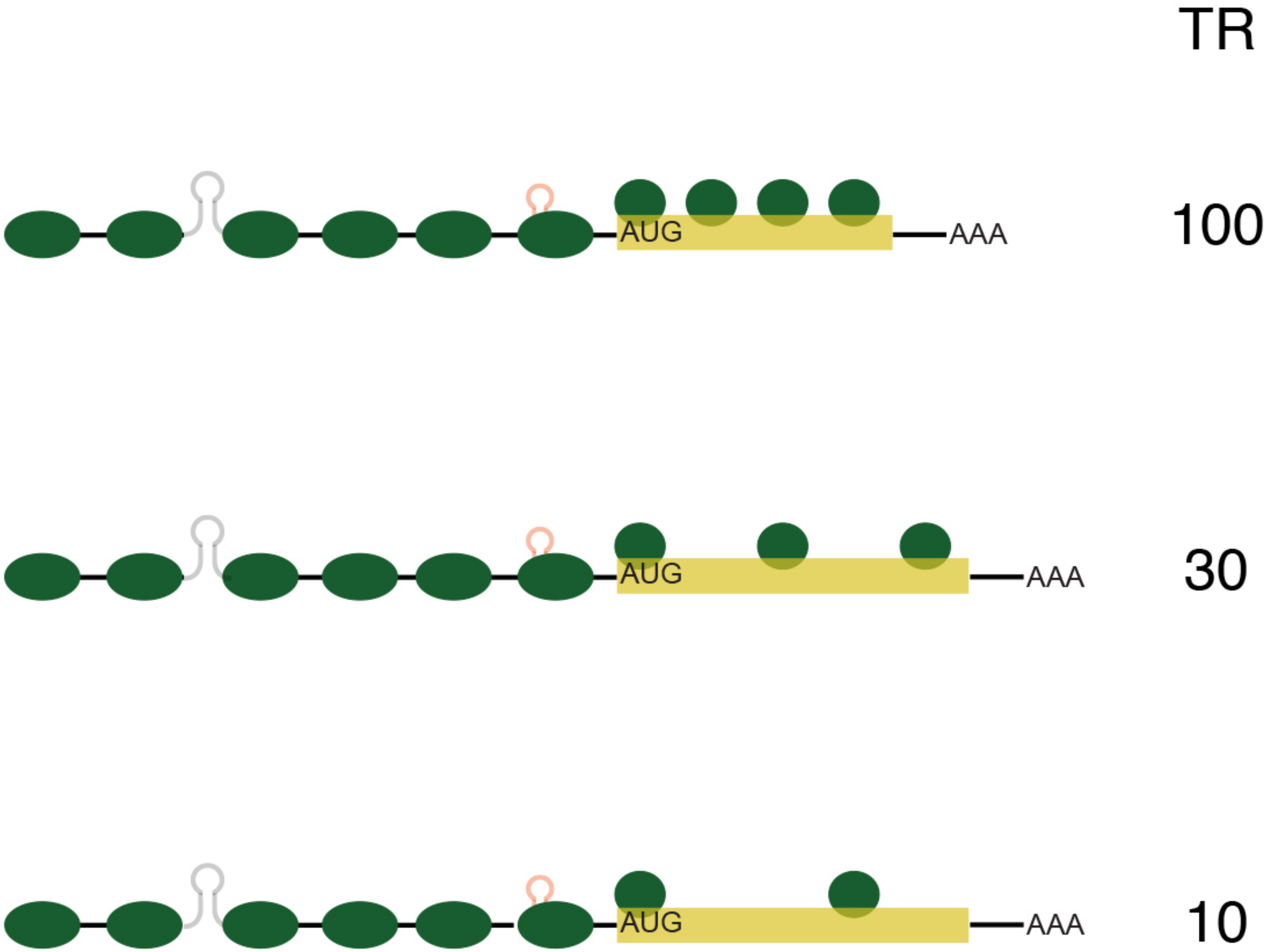
UUU freq. = Phe freq. x UUU/(UUU+UUC)

UUC freq. = Phe freq. x UUC/(UUU+UUC)

Correlated control gives even packing



Uncorrelated control gives uneven packing

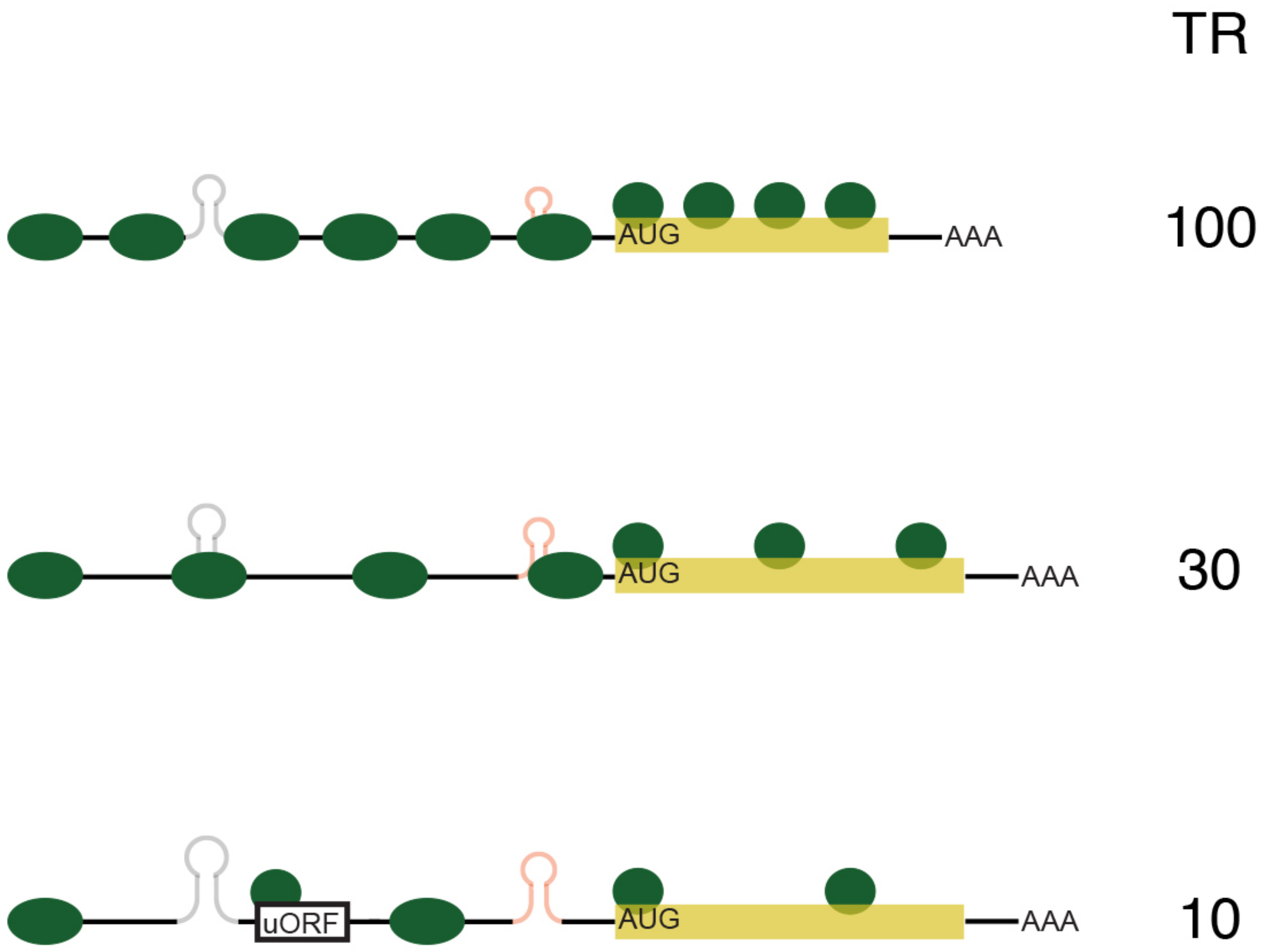


Ribosomes are limiting for cell growth

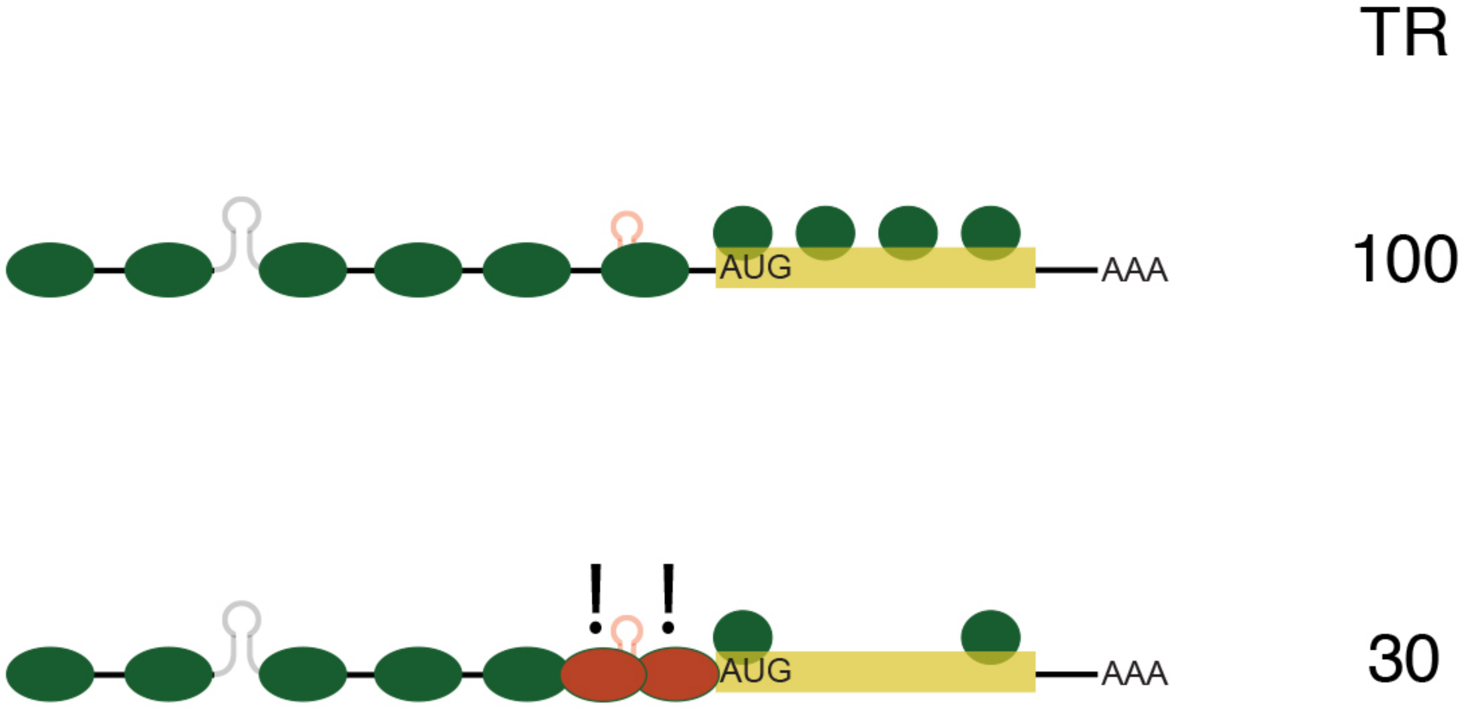
>90% ribosomes bound to mRNA

50% reductions in ribosome copy slows cell growth
(in Drosophila)

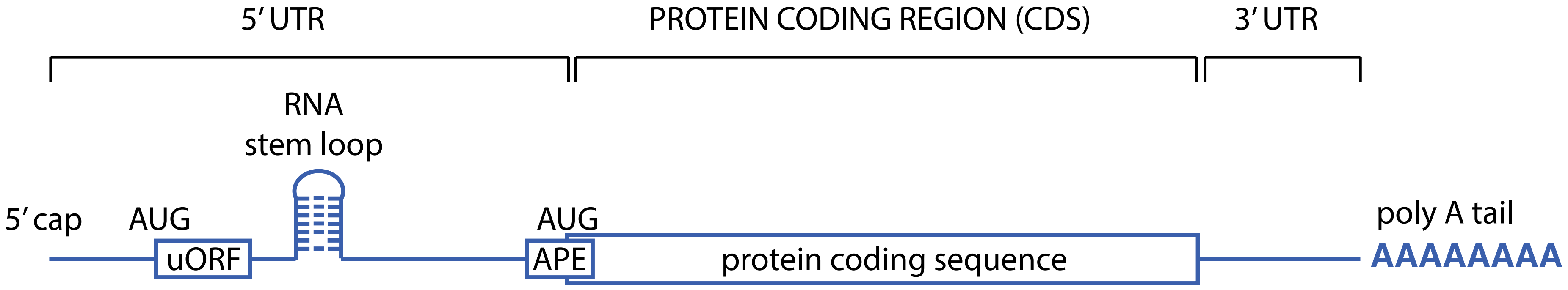
Correlated control by general features aids efficient use of the translation machinery



Uncorrelated control increases particle collisions, leading to RNA degradation



The general features have several roles



1. 5′ UTR RNA secondary structure
2. upstream open reading frames (uORFs)
3. AUG proximal element (APE)
4. protein coding sequence (CDS) length
5. CDS codon frequency

Quantitation of *cis*-translational control by general mRNA features in five eukaryotes

Mark Biggin

LBNL

Guo-Liang Chew

Fred Hutchison Center