

Differences of Translation Termination Sites Among the Three Stop Codons

Yosuke Ozawa¹²

t96214yo@sfc.keio.ac.jp

Rintaro Saito¹³

rsaito@sfc.keio.ac.jp

Satoshi Hanaoka¹²

t97780sh@sfc.keio.ac.jp

Masaru Tomita¹²

mt@sfc.keio.ac.jp

¹ Laboratory for Bioinformatics

² Department of Environmental Information

³ Graduate School of Media and Governance

Keio University Endo 5322, Fujisawa, Kanagawa 252-8520 Japan

1 Introduction

Most organisms use different mechanisms of translation termination for the three types of stop codons, i.e. UAA, UAG, and UGA [1]. Using the GenBank database, we have conducted comprehensive computer analyses of translation termination sites in order to find the difference of consensus sequence patterns around each type of stop codons. We also systematically analyzed the preference of stop codons (O/E ratio: frequency of UAA, UAG, or UGA used as stop codons divided by frequency of the same triplet over the entire genome). It is known that Release Factor 1 (RF1) is required for translation termination for the stop codon UAG, and Release Factor 2 (RF2) for the stop codon UGA. The stop codon UAA requires either RF1 or RF2, but not necessarily both. For the five species of which genes for RF1 and RF2 are annotated in the GenBank database, we computed Codon Adaptation Index (CAI) values [2] for these genes.

2 Results and Discussion

Tables 1 and 2 show the results of our analysis. For the stop codon UAA, patterns AAA and GAA are frequently observed at position -3 (from the stop codon). For the stop codon UGA, on the other hand, pattern GAG is often observed at position -12 . Consensus patterns observed for the stop codon UAG are not as strong as the other two stop codons.

Table 3 shows that the preference of each stop codon is correlated with the CAI value of its corresponding Release Factor gene; if O/E ratio of stop codon is high, the CAI value of the corresponding release factor gene is also high. Greater CAI values usually imply higher expression rates. We therefore conclude that many organisms preferably use the type of stop codon which corresponds to the most abundant Release Factor.

Acknowledgments

This work is supported in part by a Grant-in-Aid for Scientific Research on Priority Areas, “Genome Science”, from The Ministry of Education, Science, Sports and Culture in Japan.

species	cod.	pat.	pos.	Z val.	species	cod.	pat.	pos.	Z val.
<i>A.fulgidus</i>	uga	gag	-12	15.20	<i>A.fulgidus</i>	uga	cta	-3	11.07
<i>A.aeolicus</i>	uga	gag	-12	12.87	<i>H.influenzae</i>	uga	cta	-3	10.58
<i>B.subtilis</i>	uga	gag	-12	14.99	<i>E.coli</i>	uua	gag	-3	17.21
<i>E.coli</i>	uga	gag	-12	19.71	<i>M.jannaschii</i>	uua	aaa	-3	16.03
<i>M.tuberculosis</i>	uga	gag	-12	13.67	<i>E.coli</i>	uua	aaa	-3	21.65
<i>P.horikoshii</i>	uga	gag	-12	14.71	<i>H.influenza</i>	uua	aaa	-3	24.17
<i>T.maritima</i>	uga	gag	-12	15.76	<i>H.influenzae</i>	uua	gaa	-3	16.58
<i>M.genitalium</i>	uag	gaa	-12	7.19	<i>B.subtilis</i>	uua	gaa	-3	17.06
<i>H.pylori</i>	uga	gaa	-9	12.35	<i>M.jannaschii</i>	uua	gaa	-3	24.05
<i>A.fulgidus</i>	uga	gtg	-9	18.35	<i>E.coli</i>	uag	gag	-3	8.15
<i>M.jannaschii</i>	uga	gtg	-9	11.38	<i>M.pneumoniae</i>	uag	aaa	-3	9.05
<i>T.maritima</i>	uga	gtg	-9	18.86	<i>H.influenza</i>	uag	aaa	-3	8.91

Table 1: Frequent triplets before stop codons

species	cod.	pat.	pos.	Z val.	species	cod.	pat.	pos.	Z val.
<i>A.aeolicus</i>	uua	ttatg	+12	11.24	<i>B.subtilis</i>	uua	aggag	+3	13.56
<i>M.jannaschii</i>	uua	ttttt	+10	11.97	<i>A.aeolicus</i>	uua	aggag	+3	12.66
<i>A.fulgidus</i>	uua	tttaa	+7	18.89	<i>A.fulgidus</i>	uua	atddd	+3	12.65
<i>M.genitalium</i>	uag	tgcca	+6	20.48	<i>H.influenzae</i>	uua	ggaga	+3	13.63
<i>B.subtilis</i>	uua	aaaaa	+5	23.87	<i>T.maritima</i>	uua	ggagg	+3	16.72
<i>A.aeolicus</i>	uua	gaggt	+5	14.55	<i>B.subtilis</i>	uua	ggagg	+3	17.44
<i>A.fulgidus</i>	uua	tttta	+5	26.64	<i>E.coli</i>	uua	ggagt	+3	14.47
<i>P.horikoshii</i>	uua	ctddd	+4	10.78	<i>T.pallidum</i>	uua	ggggg	+3	11.28
<i>T.maritima</i>	uua	gaggt	+4	13.78	<i>M.jannaschii</i>	uua	ggtga	+3	13.70
<i>A.aeolicus</i>	uua	ggagg	+4	21.42	<i>M.jannaschii</i>	uag	gtgag	+3	16.56
<i>A.fulgidus</i>	uua	ttddd	+4	38.44	<i>T.maritima</i>	uag	gaggt	+3	14.70
<i>M.jannaschii</i>	uag	ggtga	+4	22.19	<i>B.subtilis</i>	uag	gaggt	+3	19.01
<i>A.fulgidus</i>	uga	gtggt	+4	27.10	<i>M.pneumoniae</i>	uag	ttatg	+3	21.46
<i>P.horikoshii</i>	uga	gtgat	+4	30.86	<i>T.maritima</i>	uga	ggggg	+3	31.68
<i>B.subtilis</i>	uua	aaaag	+3	19.44	<i>A.fulgidus</i>	uga	ggtgg	+3	29.10
<i>A.aeolicus</i>	uua	aataa	+3	10.75	<i>P.horikoshii</i>	uga	ggtga	+3	46.02

Table 2: Frequent pattern consist of five bases observed after UAA

Species	O/E-UAA (RF1 or RF2)	O/E-UGA (RF2)	O/E-UAG (RF1)	RF1	RF2
<i>A.aeolicus</i>	1.10	1.22	0.51	0.719	0.745
<i>B.burgdorferi</i>	1.10	0.78	0.97	0.707	0.625
<i>E.coli</i>	1.66	0.63	0.50	0.757	0.761
<i>H.pylori</i>	1.17	0.97	0.70	0.713	0.744
<i>T.pallidum</i>	0.90	0.74	1.68	0.742	0.711

Table 3: O/E ratio of stop codon and CAI value of the corresponding Release Factor

References

- [1] Brown, C.M., Stockwell, P.A., Trotman, C.N. A., and Tate, W.P., The signal for the termination of protein synthesis in procaryotes, *Nucleic Acids Res.*, 18:2079–2086, 1990.
- [2] Sharp, P.M. and Li, W.-H., The codon adaptation index – a measure of directional synonymous codon usage bias, and its potential applications, *Nucleic Acids Res.*, 15:1281–1295, 1987.