

Supplementary Material

Shemesh, R., *et al.* 2010. Follow the leader: preference for specific amino acids directly following the initial methionine in proteins of different organisms. *Genomics Proteomics Bioinformatics* 8(3): 180-189.

DOI: 10.1016/S1672-0229(10)60020-4

Table S1 Distribution of the second amino acid in proteins of different phylogenetic groups

Amino acid	A	C	D	E	F	G	H	I	K	L	M	N	P	Q	R	S	T	V	W	Y	n
Species																					
Human	22.9	0.8	6.2	9.1	1.8	7.5	0.9	1.5	4.6	5.5	1.6	3.6	4.9	2.4	5.0	11.4	4.4	3.8	1.2	0.9	10,790
Dog	17.5	1.2	7.3	8.5	1.8	9.4	1.2	2.1	6.9	5.4	3.0	4.2	5.1	3.0	4.2	7.9	5.4	3.3	2.1	0.3	331
Rat	22.4	0.5	5.3	9.1	1.7	6.8	0.8	2.4	5.6	4.8	1.7	3.8	4.6	2.3	4.8	11.6	5.1	4.6	0.8	1.3	8,420
Mouse	24.1	0.7	5.8	9.4	2.0	6.7	0.9	1.4	4.9	5.2	1.5	3.1	5.0	2.4	4.5	12.2	4.3	3.5	1.2	1.0	7,764
Chicken	21.8	1.3	6.0	9.0	1.9	5.9	1.7	1.2	4.7	4.4	1.7	3.7	5.4	2.4	6.6	10.2	5.0	4.7	0.6	1.9	835
Frog	17.1	1.1	6.7	9.5	3.8	5.5	1.8	1.4	4.2	4.8	2.5	4.6	5.7	2.7	3.6	13.7	5.7	2.9	0.0	2.7	715
Zebrafish	13.8	0.6	7.9	6.5	2.8	3.1	2.5	2.5	3.7	4.8	3.7	8.7	5.1	1.7	3.1	15.5	9.0	2.8	0.3	2.0	355
Fruitfly	14.0	1.2	6.6	5.7	4.0	5.8	1.3	2.3	5.9	6.8	2.3	3.7	4.9	3.2	3.6	16.3	4.7	5.0	1.0	1.7	2,104
Worm	10.0	1.1	5.7	4.5	3.5	5.4	1.6	3.8	5.0	6.2	1.1	5.3	4.9	2.7	5.3	19.9	7.9	4.3	0.5	1.4	2,566
Trypanosoma	12.4	0.0	4.7	3.1	7.8	7.8	0.8	2.3	5.4	8.5	0.8	2.3	3.9	2.3	1.6	21.7	10.1	4.7	0.0	0.0	129
<i>S. cerevisiae</i>	7.9	0.7	5.3	4.3	3.7	4.6	1.4	3.2	5.9	6.6	1.6	5.4	4.4	2.6	2.9	23.7	7.6	6.0	0.6	1.5	4,819
<i>S. pombe</i>	10.9	0.3	6.9	6.3	3.7	5.1	0.9	2.8	5.9	5.2	0.7	5.7	4.3	2.4	3.5	23.3	6.4	4.2	0.6	0.9	2,672
<i>C. albicans</i>	8.9	0.0	2.8	1.4	5.7	2.8	0.7	4.3	6.4	7.8	1.4	4.6	5.7	2.5	3.9	28.5	7.1	3.9	0.0	1.4	281
<i>P. aeruginosa</i>	11.3	0.0	3.7	2.1	1.3	1.8	1.4	4.2	9.4	5.2	1.4	6.4	5.9	3.8	8.0	21.3	10.2	1.8	0.3	0.7	1,002
<i>N. crassa</i>	24.9	0.6	4.5	3.4	3.6	5.0	1.4	1.7	3.6	4.5	0.8	3.4	8.4	0.8	3.6	19.0	4.5	4.5	0.0	1.7	357
<i>Arabidopsis</i>	29.6	0.2	5.5	10.0	1.4	8.3	0.3	2.6	5.6	2.4	1.8	3.3	2.4	1.6	3.3	10.8	4.5	5.1	0.3	1.0	3,056
Maize	40.9	0.3	4.5	6.3	1.0	6.8	0.3	2.5	3.0	3.3	1.0	1.5	2.8	2.0	5.3	9.1	4.0	4.0	0.3	1.0	396
<i>E. coli</i>	7.1	0.4	3.5	4.5	3.0	2.0	1.5	5.7	16.7	6.2	2.1	8.9	2.9	4.0	6.2	13.6	8.1	2.1	0.4	1.2	4,474
<i>B. subtilis</i>	6.2	0.3	4.1	6.8	3.4	3.3	1.1	5.1	19.7	6.2	1.8	11.1	2.0	2.6	5.1	9.7	6.7	2.7	0.4	1.7	2,651
<i>H. influenzae</i>	7.5	0.4	2.7	4.2	3.1	1.7	0.9	6.4	18.6	8.0	2.1	8.9	2.3	5.7	4.0	10.3	9.1	2.1	0.4	1.6	1,706
<i>S. pneumonia</i>	11.2	0.0	4.7	4.0	2.1	2.1	0.5	5.1	21.9	4.0	0.9	6.5	2.1	1.9	4.4	16.0	9.3	2.3	0.2	0.9	430
<i>Shige lla.</i>	7.7	0.7	4.6	4.6	2.8	1.7	1.8	6.1	13.6	5.7	2.4	7.0	1.7	3.5	3.9	18.8	10.5	1.5	0.0	1.5	543
<i>Salmo nella</i>	7.4	0.0	4.0	4.5	2.6	1.4	1.4	5.4	13.4	6.3	2.0	7.0	2.1	4.5	5.1	16.1	13.0	2.4	0.0	1.3	759

Note: The numbers are presented in percents, “n” represents the total number of proteins analyzed. The most two abundant amino acids in the second position of protein sequences are presented in bold.