Protein-protein interaction networks and codon bias: the case of Escherichia Coli.

Maddalena Dilucca[1], Giulio Cimini [2], Andrea Semmoloni [1], Antonio Deiana [1] and Andrea Giansanti [1,3].
[1] Dipartimento di fisica, Università Sapienza, Rome, Italy
[2] Istituto di sistemi complessi (ISC-CNR) UoS Università Sapienza, Rome, Italy
[3] INFN Roma1 unit, Rome, Italy
e-mail: maddalena.dilucca@gmail.com

Keywords: (tRNA, codon bias, E. Coli, PIN).

Synonymous codons, DNA nucleotide triplets coding for the same amino acid, are differently used in different organisms, a phenomenon known as codon-bias. In order to measure this bias, several indices have been proposed, such as the codon adaptation index (CAI), the tRNA adaptation index (tAI) and the effective number of codons (Nc). Here we propose a new index (CompAI) that does not make reference to the codon usage bias of highly expressed genes but is based on the competition between cognate and near-cognate tRNAs, in the codon-anticodon recognition process on the ribosome. We perform an extensive genome-scale comparison of the various codon-bias indices for Escherichia coli (E.coli), focusing, in particular, on their distribution over the communities of the E.coli protein-protein interaction network (PIN). The codon-bias indices that are based on tRNA abundance, tAI and CompAI, besides being correlated with gene expression levels, correlate with the conservation of a gene among several species and its internal essentiality. More importantly, high codon-bias of the gene is systematically associated to high connectivity (i.e. number of interactions) of the corresponding protein, and the most densely connected communities of the PIN share a common level of codon bias, indicating that functional subnets of protein interactions co-evolve with the number of tRNA genes and the usage of codons in the DNA regions coding for proteins. A PCA-based cluster analysis of the above mentioned indices confirms that tAI and CompAI account for most of the variability in codon usage bias over the whole E.coli genome and is consistent with the independent partition of the PIN into highly connected components or communities.
This paper clearly shows that the communities of the protein-protein interaction network are also subsets of functionally related proteins that coherently share, as a consequence of adaptive processes, their codon bias. The CompAI index, here introduced, is the most consistent, among other indices, in its coherent distribution over the communities of the PIN. As a conclusion we can say that a small difference in codon bias of the corresponding genes is, statistically, a prerequisite for a couple of proteins to interact.

[1] Fluitt A, Pienaar E, Viljoen H (2007) Ribosome Kinetics and aa-tRNA Competition determine Rate and Fidelity of Peptide Synthesis. Comput. Biol Chem. 31(5-6): 335–346.
[2] Dos Reis M., Savva R., Wernisch L. Solving the riddle of codon usage preferences: a test for translational selection. Nucleic Acids Res, 32(17):5036-5044, 2004.
[3] S. Y. Gerdes, M. D. Scholle, J. W. Campbell, G. Balzsi, E. Ravasz, M. D. Daugherty, A. L. Somera, N. C. Kyrpides, I. Anderson, M. S. Gelfand, A. Bhattacharya, V. Kapatral, M. D'Souza, M. V.Baev, Y. Grechkin, F. Mseeh, M. Y. Fonstein, R. Overbeek, A-L. Barabsi, Z. N. Oltvai, and A. L.Osterman. Experimental determination and system level analysis of essential genes in Escherichia coli mg1655. J Bacteriol, 185(19):5673{5684, Oct 2003.

.

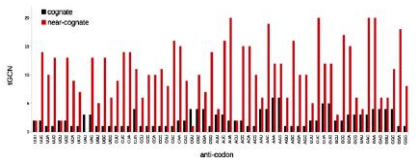

Figure 1. Abundance of tGCN cognate and near-cognate for each anti-codon.

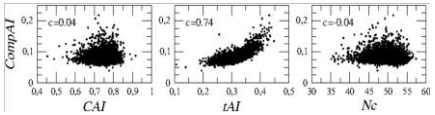

Figure 2. Correlation between various codon bias indices and compAI.

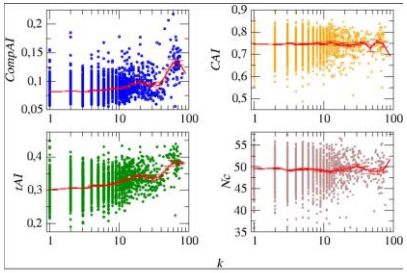

Figure 3. Relation between the various codon bias indices and protein degree in PIN.

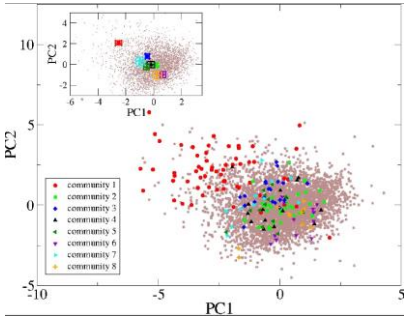

Figure 4. The set of expressed sequences represented in the first two PCA components.