

Manuscript Number:

Title: CODON EFFECTIVE NUMBER IN METHYLATING AND NON-METHYLATING ORGANISMS

Article Type: Regular Paper

Keywords: Codon bias; Codon usage; G + C content; Methylating organisms; Non-methylating organisms

Corresponding Author: Prof Chiara Scapoli, PhD

Corresponding Author's Institution: University of Ferrara

First Author: Italo Barraï, Professor

Order of Authors: Italo Barraï, Professor; Alberto Carrieri, BSc; Germano Salvatorelli, Professor; Stefano Volinia, PhD; Marco Galasso, PhD; Sonia De Lorenzi, BSc; Erika Bartolomei; Alvaro Rodriguez-Larralde, Professor; Chiara Scapoli, PhD

Abstract: The effective number of codons (ENC) is a necessary measure to estimate synonymous codon usage bias (CUB), namely the non-random usage of synonymous codons. There are several methods which lead to the estimation of the ENC, the most popular being the one introduced by Frank Wright in 1990. Here, we introduce two new estimates based on the entropy of the codon usage distribution within a coding sequence, a gene, a group of genes, an organism. We apply here four methods of estimation of the ENC to methylating and non-methylating organisms, and show that the effect of synonymy on the estimates lies in increasing the ENC, then in diminishing codon usage bias. We measure the effect of synonymy comparing, for the same table of codon usage, the ENC estimated under synonymy with the ENC estimated under the hypothesis that each codon is an independent entity irrespective of the coding group. We also observe that the estimates of the ENC in methylating organisms depend on the GC content in the third position of the triplets, while this is not observed in non-methylating organisms