Susana Vinga (INESC-ID)
Instituto de Sistemas e Computadores
Investigação e Desenvolvimento
Rua Alves Redol, 9
1000-29 Lisboa, Portugal
e-mail svinga@vinci.inesc-id.pt

Análise de sequências biológicas por funções vectoriais: classificações de AND e proteínas através de métodos sem alinhamento

Biological sequence analysis by vector-valued functions: alignment-free methodologies for DNA and protein classification

Abstract:

Biological sequence analysis is at the core of most Bioinformatics applications. In particular, the recent development of alignment-free methodologies provides an elegant and natural approach to study DNA and proteins. They are based on mapping sequences onto n-dimensional spaces and on subsequently exploring the properties of the obtained points. One example is the estimation and comparison of the relative frequency of short segments, e.g. triplet content. The vector of frequencies defines a genomic signature, which characterizes the original sequence and the corresponding organism. Iterative function systems (IFS) can further generalize this approach by considering all scales or resolutions simultaneously. One application of IFS is the estimation of DNA Rényi continuous entropy, a quantity that measures sequence randomness.

Biographical note:

Susana Vinga is a researcher at INESC-ID and an invited assistant professor at Faculdade de Ciências Médicas (FCM/UNL). Her major areas of interest are bioinformatics and biomathematics. She got a PhD in Biological Sciences from Instituto de Tecnologia Química e Biológica (ITQB/UNL) in 2005, a Post-Graduation in Applied Mathematics - Probability and Statistics from IST in 2002 and a BSc in Mechanical Engineering - Automation and Robotics from IST in 1999. She attended Biomedical Engineering for one year at Politecnico di Milano (Italy) and finished the Music/Piano degree at Instituto Gregoriano de Lisboa. She co-authors a number of articles in the area of biological sequence analysis.