Computational molecular biology is a new discipline, bringing together computational, statistical, experimental and technological methods, which is energizing and dramatically accelerating the discovery of new technologies and tools for molecular biology. The application of wavelets in molecular biology is a thriving field of research. It has two important and linked benefits: an improved ability for capturing hidden components from biological data and a better link between biological systems and the mathematics objects used to describe them. Nowadays is a growing interest in using wavelets in the analysis of biological sequences and in different areas of molecular biology: genome sequence, protein structure, microarray data analysis. But, why wavelets? Because a wavelet is a wavy function carefully constructed so as to have certain mathematical properties. Wavelets have already had a remarkable impact because a lot of people are now applying wavelets to a lot of situations and all seem to report favorable results.

Wavelet techniques are now used in many statistical areas such as: density estimation, nonparametric regression, time series analysis, change-point problems.

Several authors revealed that wavelets can be useful in detecting patterns in DNA sequences. Lio (2000) showed that wavelet variance decomposition of bacterial genome sequences coded as G,C=1 and A,T=0 (or -1) can reveal the location of pathogenicity islands.

Wavelets have been applied to all aspects of protein structural investigation: primary sequence evolution, secondary and tertiary structure determination, refinement of X-ray crystallography, drug design and visualization. A wavelet data-dependent threshold technique is applied to predict location and topology of helices in transmembrane proteins. The CWT (continuous wavelet transforms) are used to analyse the hydrophobicity and relative accessible surface area of a variety of repeating protein motifs such as TIM barrels, propellor blades, coiled coils and leucine repeats.

Wavelets seem more suitable for describing scaling biological structures and signals then other mathematical transforms because wavelet basis functions can have different complex shapes, each adaptable for a different class of problems.


1 Faculty of Horticulture, USAMV Cluj-Napoca
2 Faculty of Veterinary Medicine, USAMV Cluj-Napoca