Evolutionary couplings from sequences: prediction of 3D structure and fitness

Debora S. Marks[1] and Chris Sander
[1]Harvard Medical School, Systems Biology Department.
debbie@hms.harvard.edu

Genomic sequences contain rich evolutionary information about functional constraints on macromolecules such as proteins. This information can be efficiently mined to detect evolutionary couplings between residues in proteins and address the long-standing challenge to compute protein three-dimensional structures from amino acid sequences.

Substantial progress on this problem has become possible because of the explosive growth in available sequences and the application of statistical methods that use a global probability model. In addition to three-dimensional structure of single proteins, this statistical analysis of evolutionary constraints identifies functional residues involved in ligand binding, protein-interactions and the hard-to-see structures of disordered proteins.

We expect computation of co-variation patterns to complement experimental structural biology in elucidating the full spectrum of protein structures, their functional interactions, evolutionary dynamics and the design of novel proteins and interactions.

See evfold.org