[BIOINFO 2023]

Coevolution analysis of Intrinsically Disordered Proteins for characterizing complex functionalities

Joongyu Heo[1], Minjae Seo[1], Wookyung Yu[1,*]

[1]*Department of Brain Sciences, DGIST, Daegu, 42988, Republic of Korea*

**Corresponding author: wkyu@dgist.ac.kr*

Intrinsically Disordered Proteins and Intrinsically Disordered Regions of Proteins (IDP/IDRs) lack a tertiary structure in the physiological condition, not following the 'functionality follows structure' paradigm. However, IDP/IDRs have various functions, including Protein-Protein Interaction (PPI) and Liquid-Liquid Phase Separation (LLPS), which is counterintuitive.

Recent research suggests treating IDP/IDRs as complex dynamical, 'edge-of-chaos' systems to explain such functions. The sequence of the IDP/IDRs may be a major intrinsic regulator of their functionality, leaving the patterns of coevolving residues. However, research about the coevolutionary relationships of IDP/IDRs had been concentrated on a limited number of 'evolutionary couplings' and their steric and electrostatic properties, thereby leaving the overall pattern of the coevolution in the IDP/IDRs not sufficiently explored.

To analyze the coevolutionary relationship network in and between IDP/IDRs, we utilized a higher-order coevolution analysis tool, Statistical Coupling Analysis (SCA). We applied SCA to individual proteins' multiple sequence alignment (MSA) with eigenvector decomposition for noise reduction. Coevolution analysis shows residue relationships within IDP/IDRs, which also fits the classical functional classification of IDPs/IDRs. We present examples of these patterns with several representative IDP/IDRs fitting into corresponding functional classes, for example, p53 TAD1 motif as display sites and CREB kinase-inducible transactivation domain as assemblers. Finally, we suggest the application of coevolutionary IDP/IDRs classification to possible IDR-gated selectivity mechanism between Aminergic G-protein coupled receptors (GPCRs) and G-proteins by the third intracellular loop of GPCRs (ICL3).