9th International Conference on

STRUCTURAL BIOLOGY

September 18-20, 2017   Zurich, Switzerland

Functional divergence in protein families: A co-variation analysis

Marie Chabbert
University of Angers, France

Statement of the Problem: Co-variations between positions in a multiple sequence alignment may reflect structural, functional, and/or phylogenetic constraints. Numerous co-variation methods have been developed and may yield a wide variety of results. However, few studies have been undertaken to determine co-variations methods adequate to gain information on functional divergence within a protein family.

Methodology & Theoretical Orientation: We explore co- variation methods for their capability to mine co-varying positions related to the functional divergence in a protein family. To reach this objective, we compare several methods on a model system that consists of three nested sets of about 300, 100, and 20 paralogous sequences of a protein family, the class A of G protein-coupled receptors. The co- variation methods analyzed are based on chi2 scores, mutual information, substitution matrices, or perturbation methods. We analyze the dependence of the co-variation scores on residue conservation, measured by sequence entropy, and the networking structure of the top pairs.

Findings: Out of the four methods that privilege top pairs with intermediate entropy, two favor individual pairs, whereas the other two methods, OMES (Observed minus Expected Squared) and ELSC (Explicit Likelihood of Subset Covariation), favor a network structure with a central residue involved in several high scoring pairs. This network structure is observed for the three sequence sets, making a hierarchical analysis possible. In each case, the central residue corresponds to a residue known to be crucial for the evolution of the protein family and the sub-family specificity. Positions co-varying with this central residue form a few clusters in the receptor 3D structure.

Conclusion & Significance: The central residues obtained with the OMES or ELSC methods can be viewed as evolutionary hubs, in relation with an epistasis-based mechanism of functional divergence within a protein family.
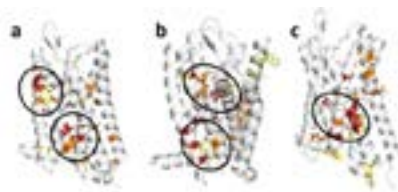

Figure1: *Visualization of co-varying residues in set 1 (a), 2 (b) and 3 (c) on the 3D structure of typical G protein coupled receptors. The color code, from yellow to red, is indicative of the co-variation score with the central residue of the network.*

Biography

Marie Chabbert is a Scientist from the French CNRS (Centre National de la Recherche Scientifique). She has her expertise in molecular modeling and bioinformatics approaches to the structure-function relationship of proteins. She has special interest in deciphering the mechanisms that drove protein evolution and in using evolutionary data to gain structural and functional information on protein families. She is presently working on the G protein-coupled receptors, especially chemotaxis and vasoactive peptide receptors.

marie.chabbert@univ-angers.fr

Notes:

J Proteomics Bioinform, an open access journal
ISSN: 0974-276X

Structural Biology 2017
September 18-20, 2017

Volume 10, Issue 8 (Suppl)

Page 132