Protein-Protein Interaction: Interface, Surface & Specificity (PPI:ISS)

K. Anton Feenstra1, Anita de Ruiter1, Bas E. Dutilh2, Vera van Noort2, Giacomo Bastianelli3, Martijn Huijnen4 and Jaap Heringa1

1 IBIVU Centre for Integrative Bioinformatics, VU University, Amsterdam 2 NCMLS Nijmegen Centre for Molecular Life Sciences 3 Institut Pasteur, Paris, France

Contact: feenstra@few.vu.nl  www.ibi.vu.nl  Download this poster at: www.few.vu.nl/~feenstra/posters.html

Orthologs, Paralogs and Interactions

Many protein families contain sub-families that interact with different protein binding partners. Specificity in these interactions is often critical to the function of the proteins involved, therefore this specificity may be used to pinpoint protein-protein interaction (PPi) sites. We have applied the Sequence Harmony (SH) method [1,2] for sub-family specific site detection to detect specificity sites that determine the interaction or non-interaction between protein families.

The analysis integrates three types of information:
- Fungal Orthologous Groups of proteins, or FOGs [3];
- homology relations between the FOG groups (from HHsearch);
- genome-wide PPI data for Yeast, specifying interacting as well as non-interacting protein pairs, based on socio-affinity scores [4].

Selected specificity sites are compared to interface regions and surface residues in the protein complex as defined from the corresponding crystal structures.

Sequence Harmony & Interaction Specificity

From the alignments of the interacting FOG with the non-interacting FOG, we collect specificity sites at different cut-off values for the SH score (ranging from 0 to 1). From these selections, we construct a ROC plot by scoring the numbers of True and False positives. In the selection, True positive residues are defined as surface residues that are in contact with the binding partner in the crystal structure. In addition, we have also scored the selection of surface residues per se, as we observed earlier [2,5] that specificity sites tend to be on the surface.

Discussion & Conclusion

Our results show a clear signal of specificity between interacting and non-interacting paralogs, that allows enrichment of interface residues. For accurate prediction of interface regions, however, the signal seems to be insufficient.

Additional improvement can be expected from combining the specificity signal with, for example, conservation, and taking into account spatial patterns. The cluster analysis shows that selected sites are organized in patches, and that the centers of these patches contain the most specific sites.

References


