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

K. Anton Feenstra[1], Anita de Ruiter[1], Bas E. Dutilh[2], Vera van Noort[2], Giacomo Bastianelli[3], Martijn Huijnen[2] and Jaap Heringa[1]

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

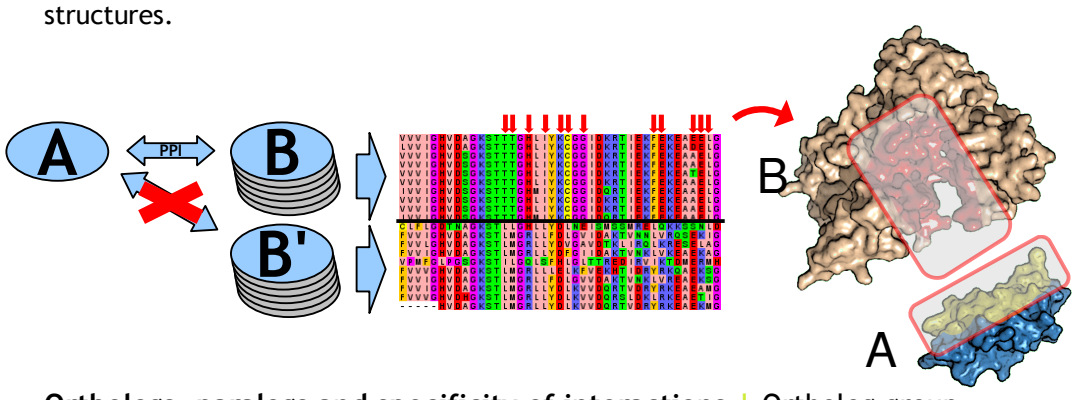
Contact: feenstra@few.vu.nl www.ibi.vu.nl Download this poster at: www.few.vu.nl/~feenstra/posters.html

Orthologs, Paralogs and Interactions

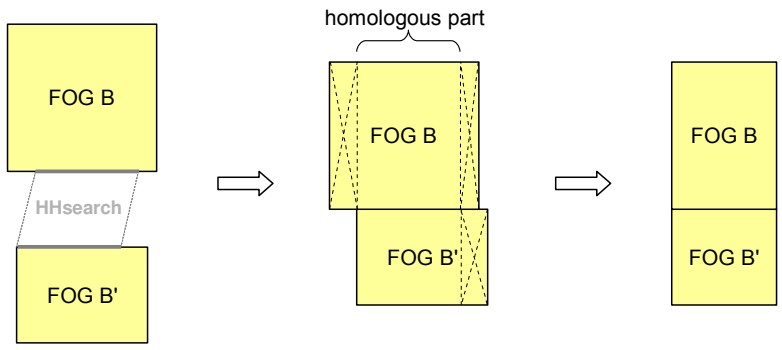
Many protein families contain sub-families that interact with different protein binding partners. Specificity in these interactions is often critical to the function of the proteins involved, therefore this specificity may be used to pinpoint protein-protein interaction (PPI) sites. We have applied the 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.



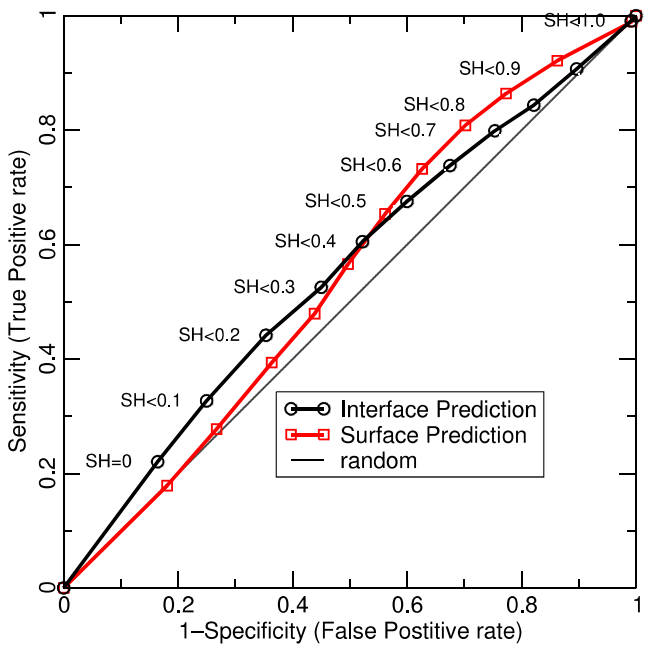
Orthologs, paralogs and specificity of interactions | Ortholog group B interacts with protein A, and is matched to its paralog group B' that does not interact with A. The specificity signal from the alignment between B and B' correlates with the interface of the complex between proteins B (member of ortholog group B) and A.



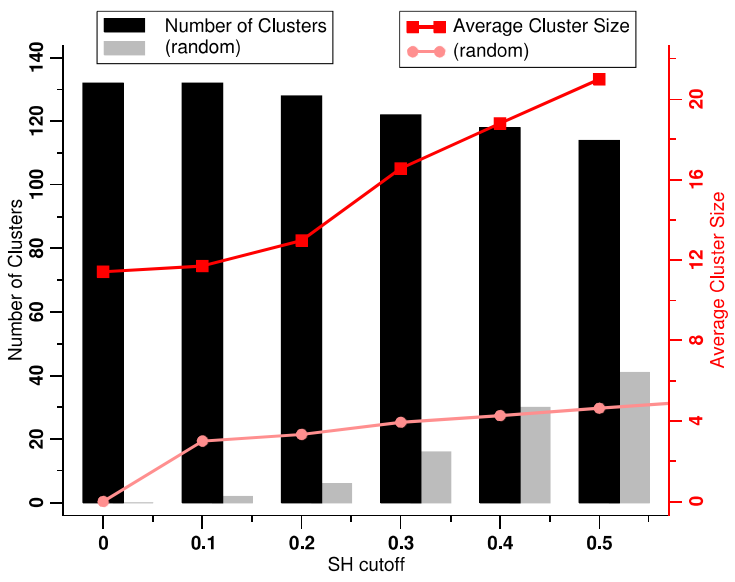
Scheme for aligning Ortholog groups | Ortholog group B is aligned with its paralog group B', guided by the information from HHsearch between homolgous regions of FOG B and FOG B'. In the alignment parts of FOG B and B' that are not homologous are removed, so that the alignment only contains the homologous regions.

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). For 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.



Prediction of Interface and Surface Residues | The performance for predicting interface or surface residues by using the Sequence Harmony specificity signal, measured by ROC plots. All predictions score slightly, but significantly, above random. Most specific predictions for interface residues occur at lower SH scores, thus at higher specificity, than predictions for surface residues.



Clustering of Predicted Interface Residues | Selected specificity sites, with low Sequence Harmony scores, are clustered. With more restrictive (lower) SH cutoff, the number of clusters remains relatively the same, while the average size of clusters decreases. The trends expected for a random distribution (grey plot and bars) of selected sites is very different.

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

[1] Feenstra, KA, W Pirovano, K Krab & J Heringa. Sequence Harmony: Detecting Functional Specificity from Alignments, *Nucl. Acid. Res.* **35** W495 2007 www.ibi.vu.nl/programs/seqharmwww
[2] Pirovano, W, KA Feenstra & J Heringa Sequence comparison by sequence harmony identifies subtype specific functional sites. *Nucl. Acids Res.* **34** 6540-6548 2006
[3] Dutilh, BE, V van Noort, RTJM van der Heijden, T Boekhout, B Snel & MA Huynen. Phylogenetics Assessment of phylogenomic and orthology approaches for phylogenetic inference. *Bioinformatics* **23** 815-824 2007
[4] Van Noort, V, and B Snel & M Huynen Exploration of the omics evidence landscape: adding qualitative labels to predicted protein-protein interactions. *Genome Biol.* **8** R197 2007
[5] Feenstra, KA, G Bastianelli & J Heringa. Predicting Protein Interactions from Functional Specificity. *in:* From Computational Biophysics to Systems Biology - *NIC Series* **40** 89-92 2008

CENTRE FOR INTEGRATIVE BIOINFORMATICS VU

nbic Netherlands Bioinformatics Centre