

Functionating the Proteome of *Aquifex aeolicus*

M. J. Thompson^{1,*}, M. Pellegrini¹, E. M. Marcotte^{1,2}

T. O. Yeates^{1,2}, & D. Eisenberg¹,

* mjt@proteinpathways.com

¹Protein Pathways, Inc., 1145 Gayley Ave., #304, Los Angeles, CA 90024

²University of California Los Angeles, Los Angeles, CA 90095-1570

Motivation

Assigning functions to the vast number of proteins comprising the proteomes of organisms whose genomes have been completely sequenced is one of the foremost challenges for computational biology today. Traditionally, most computational techniques for addressing this problem have relied on inference-by-homology. Recently, new techniques, including the Rosetta-Stone method and the Phylogenetic Profile method, have been developed that can identify the functional role of a protein independently of homology [1, 2, 3, 4]. Rather, the functional role of the protein of interest can be inferred through its functional couplings to other proteins whose cellular roles (*e.g.* pathway or complex) are known.

Results

Here we illustrate the power of these new approaches through application to the proteome of *Aquifex aeolicus* [5]. The functional couplings predicted by this analysis are presented as a global linkage map for the open reading frames (ORFs) of this proteome. We see the correspondence between sets of linkages and known pathways and protein complexes. In the accompanying tables, we present a statistical summary of the estimated true and false positive prediction rates. An overall evaluation of the performance of these approaches is given as the average functional keyword recovery obtained for linked pairs of ORFs with characterized function and this is compared against the keyword recovery for randomly linked proteins, following Marcotte, *et al.* [3]. We also provide estimates for the fractional coverage of hypothetical ORFs and conserved hypothetical ORFs for which functional information can be gleaned from their predicted functional couplings.

References

- [1] Pellegrini, M., Marcotte, E. M., Thompson, M.J., Eisenberg, D., & Yeates, T.O., *Proc. Natl. Acad. Sci.* **96**, 4285 (1999).
- [2] Overbeek, R., Fonstein, M., D'Souza, M., Pusch, G.D., & Maltsev, N. *Proc. Natl. Acad. Sci. USA* **96**, 4285-4288 (1999).
- [3] Marcotte, E.M., Pellegrini, M., Ng, H.-L., Rice, D.W., Yeates, T.O., & Eisenberg, D., *Science* **285**, 751-753 (1999).
- [4] Marcotte, E.M., Pellegrini, M., Thompson, M.J., Yeates, T.O., & Eisenberg, D., *Nature* **402**, 83-86 (1999).
- [5] Deckert, G., Warren, P.V., Gaasterland, T., *et al.* *Nature* **392**, 353-358 (1998).