Reconstructing ancestral protein-protein interactions of virus-host systems

Anderson F. de Brito, John W. Pinney

Theoretical Systems Biology Group, Division of Molecular Biosciences, Imperial College London

Over the last decades of advances in biomolecular research, large amounts of biological data have been made available, which now allows us to apply integrative approaches to combine information from different levels of complexity. Such approach has proven to be of particular interest to broaden our understanding on how host-pathogen systems evolve, in particular by integrating genomic, proteomic, gene ontology, structural, and taxonomy data. In our research we have been using computational tools to infer the phylogenetic history of Protein-Protein Interactions (PPIs) between viruses and their hosts. As a starting point, a reference structure of a herpesvirus-human protein complex was taken from PDB. Searches for homologous proteins in taxonomically related species were performed in order to create multiple sequence alignments (MSAs) depicting the amino acid diversity of both viruses and hosts. Such alignments were then used to create genealogies of the protein families, yielding trees whose internal nodes represent one or more ancestral states of the existing proteins included in the initial alignments. By applying in-house methods and Maximum Likelihood approaches implemented in PAML and FastML, distributions of ancestral sequences from viruses and hosts were inferred by using the genealogies and MSAs previously mentioned. In the near future homology modeling will be used to reconstruct ancestral virus-host complexes, and variations in free energy ($\Delta\Delta G$) between existing PPIs and their ancestral states will be calculated, revealing important aspects of PPI evolution. With these results we aim to expand the understanding on how mutations (substitutions and indels) determine protein affinity in similar protein pairs, which although homologous, show remarkable differences in terms of binding energy. In addition, we intend to apply this approach to predict new PPIs, what will allow us to take advantage of the knowledge obtained in widely studied systems to better understand the protein interactions in neglected virus-host pairs.

Acknowledgments: we thank Capes and Imperial College London for the financial support.