TITLE: Identifying novel functional linear motifs in human protein interaction network using host-viral interactions and the principle of convergent evolution

AUTHORS: Vitalii Kleshchevnikov, Evangelia Petsalaki

ABSTRACT:

Linear motifs are short amino-acid sequence motifs that mediate physical and selective protein-protein interactions. Linear motifs are usually located in the disordered regions of a protein and are usually recognised by structured globular domains.

Linear motif-mediated interactions are known to connect and direct cell signalling pathways often in a regulated manner. Linear motif-mediated interactions can evolve rapidly and help rewire cell signalling networks during speciation events, in disease or in host-pathogen interactions.

A number of linear motifs have been identified using traditional molecular biology approaches and hypothesis-driven research, however, those methods are laborious and most of the functional linear motifs are yet to be identified. Using computational motif search tools to identify linear motifs in homologous proteins tend to result in a large number of coincidentally identified motifs. A number of approaches have been shown to improve the efficiency of identifying functional motifs: incorporating protein-protein interaction data, the sequence conservation across species and filtering for motifs located in the unstructured regions.

In this study, we use host-viral protein interactions data as a way to limit the search space and identify novel linear motifs. Viral proteins mimic cellular linear motifs to interact with and modify cell signalling in a way that favours the progression of viral infection. We can use this functional relationship to identify novel cellular linear motifs as well as summarise the ways viruses tend to use to highjack cell signalling.

We use 13488 interactions between 4423 human and 787 viral proteins to identify domains in human proteins that may be mediating interaction with specific viral protein and then use the sequence of such viral proteins to search for linear motifs. This approach generates prediction of domain-linear motif interacting pairs that will be tested using phage display screen which allows high-throughput identification of domain-linear motif interactions.