

Influence of Protein Network Dynamics on Protein Evolution

Brian Mannakee¹ and Ryan Gutenkunst²

Summary

While intuition and theory predict that the importance of a protein's function should affect its rate of evolution, previous studies have found only a weak connection. We couple biochemically detailed protein network simulations with molecular evolution methods to more deeply explore the link between protein structure, function, and evolution. Using this novel approach, we show that protein function does indeed affect evolutionary rate. In addition to contributing to a deeper understanding of evolution, this work also informs researchers seeking to understand the roles of network dynamics and mutations in cancer and evolved drug resistance.

How do proteins evolve?

Proteins are strings of amino acids transcribed from genes, and they typically fold into a particular shape to perform some function in a living system. Proteins evolve when one amino acid is substituted for another. As we see in **Figure 1**, amino acid substitutions occur frequently as species evolve from a common ancestor. In 1965 Emil Zuckerkandl and Linus Pauling observed that the rate of change in the amino acid sequence of hemoglobin was linear in time across species. This led to the idea of a molecular clock, or constant underlying rate of change that characterizes the evolution of a protein. Surprisingly, this rate is not the same for all proteins and in fact varies dramatically. The cause of this variability is among the fundamental questions of evolutionary biology, and although many factors are found to contribute to this variability, none explains more than a small portion of it. In particular, protein function has been found to have only a weak correlation with evolutionary rate.

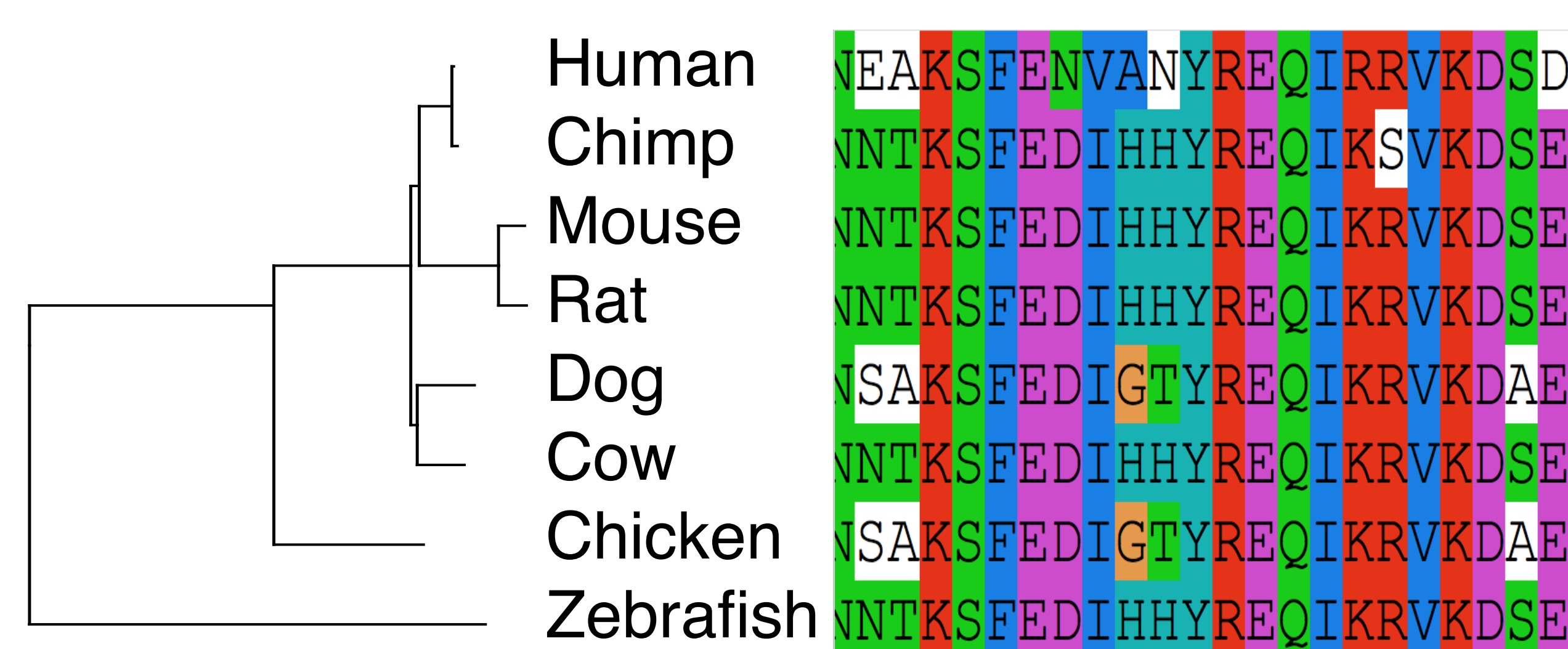


Figure 1. Amino acid sequences change as vertebrates evolve from their most recent common ancestor.

¹Chemistry and Biochemistry
²Molecular and Cellular Biology

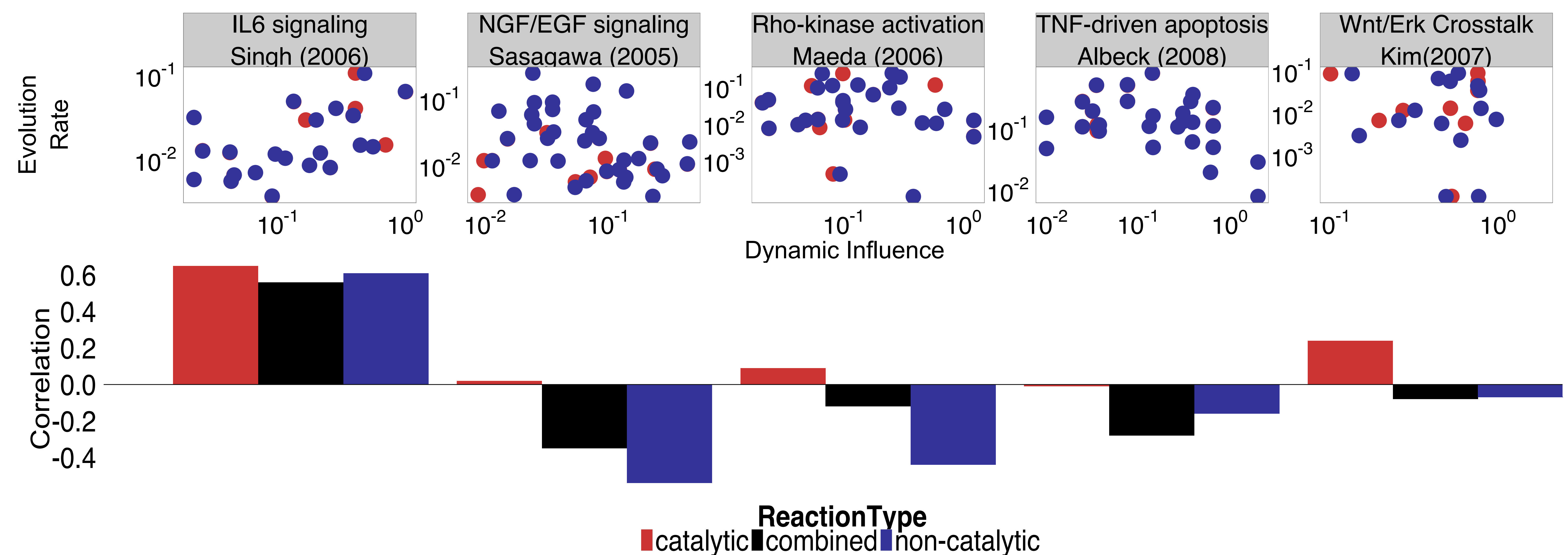


Figure 2. Correlation between protein domain influence on network dynamics and evolutionary rate for five cell signaling networks.

Network functional importance affects evolutionary rate at the domain level

Previous studies relied on protein knockout assays to assign importance to a protein and analyzed evolutionary rate at the whole protein level. We hypothesize that this approach masks subtle changes in protein interactions that contribute to organismal fitness. Using computational methods coupled with detailed biochemical models, we take a novel approach to measuring functional importance that offers a finer view than knockouts.

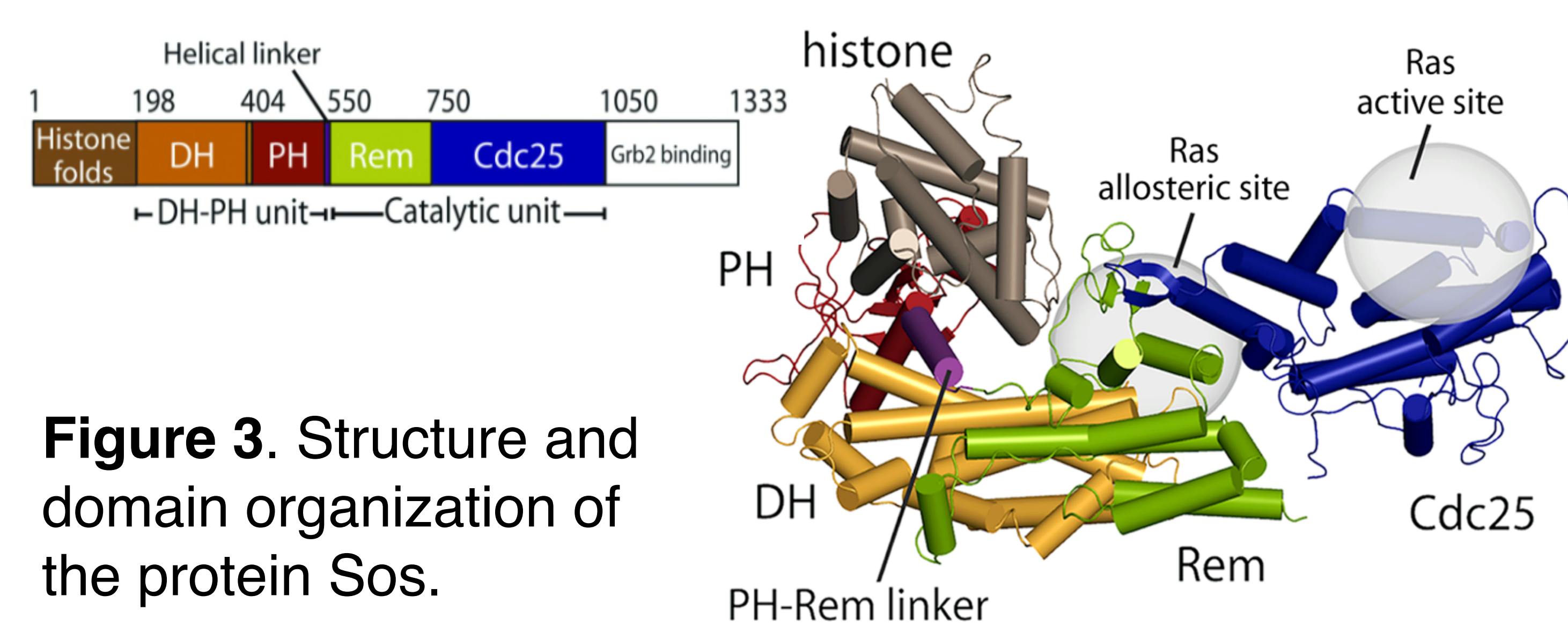


Figure 3. Structure and domain organization of the protein Sos.

Many proteins are organized into functional units called domains, as exemplified by the protein shown in **Figure 3**. Protein domains are evolutionarily conserved and evolve independently of one another. We classify domains into “catalytic” and “non-catalytic”, depending on the reactions they participate in. In **Figure 2** we analyze functional importance and evolutionary rate at the domain level, finding that not only does the correlation between function and evolutionary rate increase at this finer level of detail, but interesting patterns emerge when results are compared by reaction type. In general, more important catalytic domains evolve faster, while more important non-catalytic domains evolve more slowly. This suggests that networks typically evolve under stabilizing selection. Interestingly, the IL-6 signaling system, which is involved in the immune response, is an exception in which functionally more important domains evolve faster.

Our Method

We create computer simulations of protein interaction networks using models obtained from the BioModels database. For each protein in the network we review the biochemical literature to assign amino acids to domains, and then we use phylogenetic analysis to estimate the rate of evolution for these domains along the tree pictured in **Figure 1**.

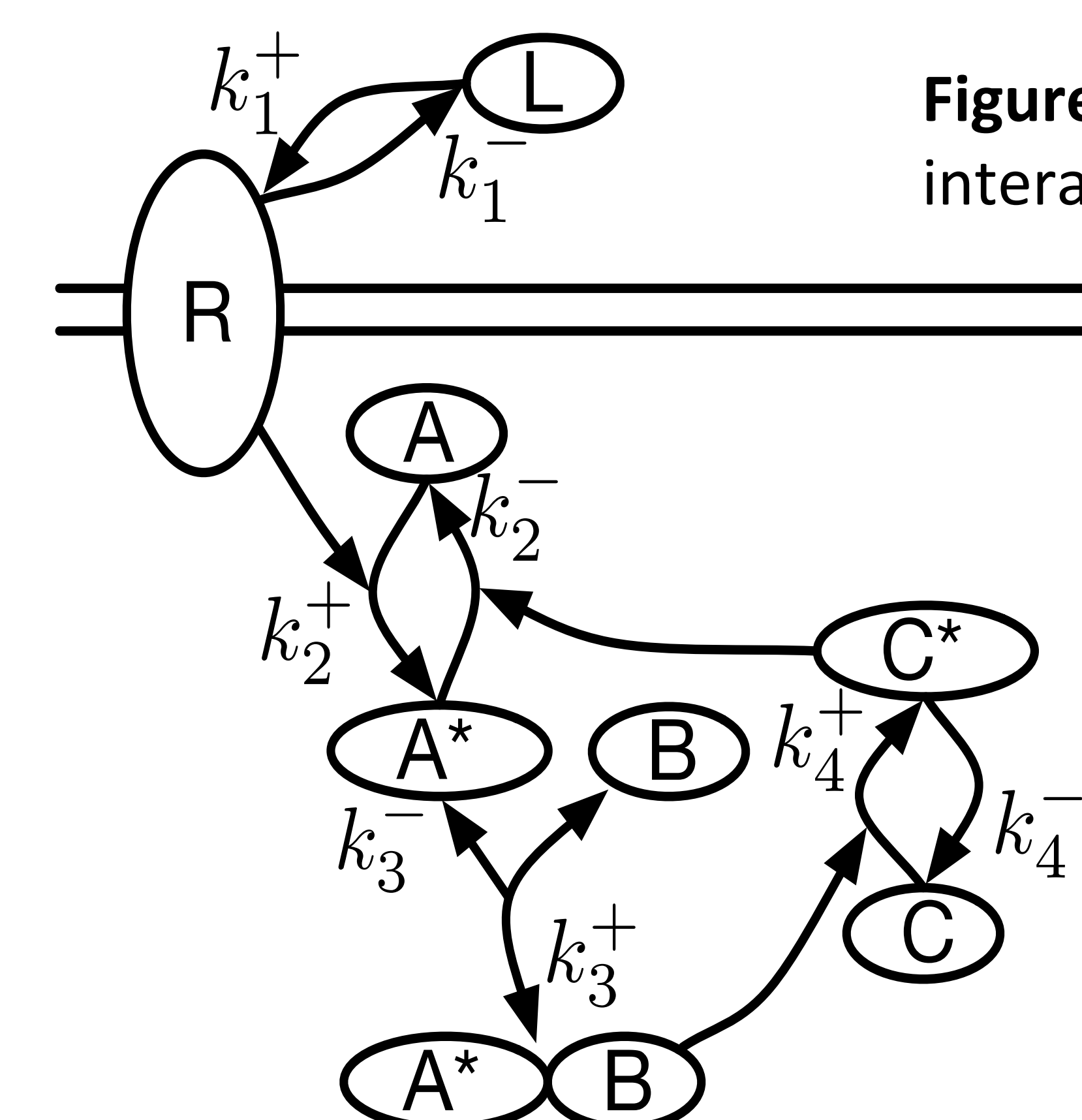


Figure 4. A small protein interaction network

We measure the functional importance of each protein domain in a network by calculating the influence that domain exerts on the dynamics of the network. **Figure 4** shows a small protein interaction network that transmits a signal from outside the cell and causes a chain of reactions on the inside. The interactions among these proteins are governed by rate parameters denoted by k 's. The dynamical influence of a parameter k is calculated by finding the influence on every protein in the network of small changes in the value of that parameter. The dynamical influence of a protein domain is then the average dynamical influence of all the rate constants for all the reactions that domain participates in.