Title: Coevolution of interacting proteins: What is behind the mirror tree?

Abstract: Understanding the relationships among proteins is crucial to understand the molecular machinery of the cell. Computational tools to predict domain-domain interactions provide a detailed molecular view of the protein interactions and complements expensive and laborious experimental techniques to identify such interactions. The evolutionary distances of interacting proteins often display a higher level of similarity than those of non-interacting proteins. This finding indicates that interacting proteins are subject to common evolutionary constraints and constitute the basis of a method to predict protein interactions known as mirrortree. It has been difficult, however, to identify the direct cause of the observed similarities between evolutionary trees. One possible explanation is the existence of compensatory mutations between partners’ binding sites to maintain proper binding. This explanation, however, has been recently challenged. It has been suggested that the signal of correlated evolution uncovered by the mirrortree method is unrelated to any correlated evolution between binding sites. We have addressed this controversial debate in the field by studying the contribution of binding sites to the correlation between evolutionary trees of interacting domains. We showed that binding neighborhoods of interacting proteins have, on average, higher co-evolutionary signal compared to the regions outside binding sites; although when the binding neighborhood was removed, the remaining domain sequence still contained some co-evolutionary signal. These results provide evidence of the role of compensatory mutations in protein co-evolution and contribute to our understanding of co-evolution of interacting proteins.