The Roles of Short Linear Motifs in Human Diseases

Bora Uyar and Toby Gibson

European Molecular Biology Laboratory
Meyerhoffstrasse 1, 69126 Heidelberg, Germany
{bora.uyar,toby.gibson}@embl.de

Abstract. Intrinsically disordered regions (IDRs) are natively unfolded segments of proteins, which lack tertiary structure under physiological conditions and often contain short linear motifs (SLiMs). These are small sequence elements of typically 2-10 amino residues length, which are involved in key cellular processes including signaling and regulation. IDRs are found in many proteins associated with human disease genes and mutations in IDRs are thought to be causal in mis-signalling/regulation based diseases. It has been postulated that mutations in IDRs may damage SLiMs (Tompa, 2009) leading to disease, as even a single amino acid substitution may impair the function of any functional motif. In fact, several diseases are known where the disease-relevant mutation(s) hit SLiMs. However the link between SLiMs and human diseases has not been studied at a proteome-wide scale.
In this study, proteome-wide comparison of the distribution of missense mutations revealed that, in disordered segments of proteins, disease-related substitutions (somatic cancer mutations (COSMIC) and inherited Mendelian disease mutations (OMIM)) are more likely to disrupt experimentally defined SLiMs than neutral missense mutations (1000GP). Similar results were reproduced on a set of predicted SLiMs detected in disordered and relatively conserved regions. In order to find out which biological processes suffer most due to damaged SLiMs, a sub-group analysis was carried out. Biological processes such as regulation of transcription, cell migration, proliferation were among those that had highest frequency of cancer mutations but lowest frequency of neutral substitutions in SLiMs. Finally, by constructing predicted SLiM mediated interaction networks, we observed that in disordered regions, cancer mutations tend to impact higher number of interactions than neutral missense substitutions.

Keywords: short linear motifs, disease mutations, protein interactions, intrinsically disordered proteins