

SNPs on a signaling network

Jieming Chen¹, Ekta Khurana^{1,2} and Mark Gerstein^{1,2,3}

¹Program in Computational Biology and Bioinformatics, ²Department of Molecular Biochemistry and Biophysics, ³Department of Computer Science, Yale University, New Haven, CT

The 1000 Genomes Project has provided the scientific community with an unprecedentedly dense resource of human genomic variation data. In order to elucidate their functional roles, it will require additional dimensions of information. To this end, we map the SNP variation data in the CEU population onto protein-coding genes whose products are involved in seven major signal transduction pathways curated from the *SignaLink* database. Simply by dissecting the data in multiple ways, we are able to investigate the variations at the SNP level, protein level and their interactions in the context of the signal transduction network. Specifically, we see that as proteins become involved in more interactions, they also tend to be under more selective constraints. We also proceeded to look at proteins that are involved in multiple signaling pathways and found evidence to support that these proteins do generally have more interactions in individual pathways. However, it is also interesting to observe that not all of them are highly-interactive hubs. This represents a preliminary work in analyzing variations, proteins and protein interaction networks.