Identifying functional SNVs that map to non-coding regions of the genome and alter RNA Structure.

Alain Laederach

Department of Biology, University of North Carolina at Chapel Hill (EEUU)

alain@unc.edu

Abstract. Genome-wide association studies (GWAS) often identify disease-associated muta-tions in intergenic and non-coding regions of the genome. Given the high percentage of the genome that is transcribed, we postulate that for some observed associations the disease phenotype is caused by a structural rearrangement in a regulatory region of the RNA transcript. To identify such mutations we have performed a genome wide analysis of all known disease-associated Single Nucleotide Variants (SNVs) from the Human Gene Mutation Database (HGMD) that map to the untranslated regions (UTRs) of a gene. Rather than using minimum free energy approaches (e.g. mFold), we use a partition function calculation that takes in to consideration the ensemble of possible RNA conformations for a given sequence. For six disease-states (Hyperferritinaemia Cataract Syndrome, β-Thalassemia, Cartilage-Hair Hypoplasia, Retinoblastoma, Chronic Obstructive Pulmonary Disease (COPD), and Hypertension) we identified multiple SNVs in UTRs that alter the mRNA structural ensemble of the associated genes. Using a Boltzmann sampling procedure for sub-optimal RNA structures, we are able to characterize and visualize the nature of the conformational changes induced by the disease-associated mutations in the structural ensemble. We observe experimentally using chemical probing that SNV induced conformational changes analogous to those in bacterial regulatory Riboswitches when specific ligands bind occur readily in the human genome. We propose that the UTR and SNV combinations we identify constitute a "RiboSNitch," that is a regulatory RNA in which a specific SNV has a structural consequence that results in a disease phenotype. We find that there are specific structural features, in particular whether the SNV increases or de-creases the RiboSNitch ensemble entropy that are good predictors of function.

Keywords: SNP; SNV; ncRNA; RiboSNitch