**3) Functional analysis of largely noncoding SNPs**

A number of new computational tools have been developed to prioritize noncoding variants in terms of their likely functional significance. One such tool is FunSeq, developed in Gerstein Lab. Like other noncoding variant prioritization tools, FunSeq combines data types to approach the question of functional significance from multiple perspectives. It uses the ENCODE annotation to interrelate the variant with known functional elements. It identifies the evolutionary pressure on the region in which the variant occurs by both inter- and intra-species measures of evolutionary conservation. It uses the network centrality of associated genes as a proxy for how important are the genes regulated by the functional elements containing the variant. FunSeq also identifies whether the variant leads to the gain or loss of some functional motif. FunSeq assigns subscores along each of these dimensions and integrates them into a composite score representing the likely functional significance of each noncoding variant.

***Slides providing details on non-coding variant annotations in subjectZ:***[**noncoding.pptx**](http://archive.gersteinlab.org/proj/zimmerome/part05/ppt/Slides.non-coding_XL.pptx)

**a) High impact non-coding variants (FunSeq)**

Our pipeline identified thousands of rare variants affecting annotated regions of the non-coding portions of the genome, with approximately ~80K rare variants in introns, ~15K in enhancers, ~3000 in promoters, and ~42K in other annotations. These numbers are very similar to those found in the other two European individuals. Nine of the non-coding variants (which collectively affect 11 genes) are predicted to have particularly high functional impact, insofar as they i) are rare variants that occur in highly conserved regions, ii) result in altered TF binding sites, and iii) are in regions that regulate genes that serve as a hubs within a regulatory interaction network.

One illustrative example of such a variant in subjectZ is a G to C transversion in the promoter of Van-Gogh-Like Protein 2, a hub gene that helps to maintain the polarity of cells in the developing heart, nervous system, and auditory conduction system. This transversion leads to a gain of a BCL motif, which creates a new docking site for FGFR-family transcription factors, and could be expected to alter the transcription of VGLP2, although, in subjectZ’s case, apparently not by enough to cause a major developmental disorder, such as Tetralogy of Fallot.