Exome analyses: variant prioritizing

Virginie Bernard, David J. Arenillas, Dimas Yusuf and Wyeth W. Wasserman

virginie@cmmt.ubc.ca

Centre for Molecular Medicine and Therapeutics, Child and Family Research Institute, Department of Medical Genetics - University of British Columbia, Vancouver BC, Canada

Keywords: Exome, Variant calling, Prioritizing, Transcription Factor Binding Site, Splice sites, Diseases

The convergence of high-throughput technologies for sequencing individual exome and genomes and rapid advances in genome annotation are driving a neo-revolution in human genetics. This wave of family-based genetics analysis is revealing causal mutations responsible for striking phenotypes. By mapping this high-throughput data to the human reference genome and calling variants, a list of small nucleotide and structural variations is obtained. Filters are required to remove the common variations, and prioritize the most likely causal ones. There are widely used software tools to evaluate the putative impact of substitution on protein structure and function. One limitation of these existing tools is that only variants within the protein-encoding exons are scored. However, researchers frequently discover cases in which no causal mutation is apparent in protein-encoding exons. In such cases, examining the promoters, enhancers and splice-regulating sites may lead to the discovery of causal variants. This project attempts to address this issue by prioritizing variants which (i) occur within regulatory elements, (ii) lead to new splice-regulating sites or (iii) are observed within a gene well known to be involved in some disease of interest.