Lecture 11.2 (03:40-03:55)

Using genome-wide summary statistics to identify pleiotropic genes via the composite test

June En-Yu Lai (ISSAS)

Identifying the causal relationship between genotype and phenotype allows us to expand the gene regulatory network from molecular levels to perceptible traits. From a gene-centric perspective, on the one hand, a pleiotropic gene functions as a hub in the network to influence multiple outcomes. Identifying such a phenomenon involves conducting testing procedures under a composite null hypothesis where a gene is associated with at most one trait. To this end, conventional methods, such as meta-analyses of top-hit p-values and sequential testing of multiple traits, have been proposed. However, these methods fail to consider the background of genome-wide signals. Huang proposes the composite test that produces uniform p-values for the genome-wide variants under the composite null. We propose gene-level pleiotropy tests by integrating Huang's method and multivariate non-mediation tests such as Hotelling's T-square test, minimum p-value method, generalized Berk--Jones test, generalized higher criticism test, and aggregated Cauchy association test. From a trait-centric view, on the other hand, a polygenic trait involves multiple genes of different functions to co-regulate the mechanisms. Genotype and phenotype intertwine with each other as we obtain a deeper understanding of their relationship. Herein, we provide a comprehensive map by constructing the phenome--genome functional modules using the results from the proposed pleiotropy tests. Our analysis suite has been implemented as an R package \texttt{PGCtest}. The utility is demonstrated by an application study of the Taiwan Biobank database to investigate functional modules composed of specific genes and their co-regulated phenotypes.