“Detecting copy number variations from next-generation sequencing data via a Bayesian procedure”

PI: Guan-Hua Huang (National Chiao Tung University)

Role: Principal investigator

Ministry of Science and Technology of Taiwan: 08/01/2016-07/31/2018

Copy number variations (CNVs) are genomic structural mutations with abnormal gene fragment copies. Read depth signal mirrors the variants directly from the next generation sequencing. Some tools have been published to predict CNVs by depths, but most of them just apply to a specific data type. Providing a comprehensive CNV detection algorithm that can easily make use of a variety of data types is difficult but valuable. In this project, we will develop a COpy Number variation detection tool by a BaYesian procedure, CONY, which adopts a hierarchical model and an efficient reversible jump Markov chain Monte Carlo inference algorithm for whole genome sequencing read depths data. CONY can be applied not only to an individual for estimating the absolute number of copies but also to case-control pairs for detecting patient specific variations. Real data from the 1000 Genomes Project will be analyzed. We will also evaluate the performance of CONY and compare it with competing approaches via simulations.

Keywords: Bayesian inference; copy number variations; next-generation sequencing; whole genome sequencing; 1000 Genomes Project.