Editorial

Novel Statistical Approaches for High-dimensional Gene-gene and Gene-environment Interaction Analyses

With the advance of high-throughput technologies, large-scale genetic studies have been popularly used to uncover novel genes predisposing to complex traits. During the past decade, numerous genes have been identified through genome-wide association, exome-sequencing, and whole-genome-sequencing studies, some with compelling biological plausibility for a role in complex traits. Despite such success, for most complex traits, identified genes so far have only explained a small fraction of total phenotype variation.

Qing Lu

Gene-gene/gene-environment (G-G/G-E) interactions may play an important role in complex traits. However, they have not been taken fully into account by current analyses, which have commonly adopted a single-locus approach or a SNP-set approach. The identification of G-G/G-E interactions should explain additional phenotypic variation and may also lead to novel insights into the mechanisms underlying complex traits. The yield from these G-G/G-E interaction studies can be enhanced by adopting computationally efficient and powerful approaches.

In this thematic issue, a group of researchers developed and applied innovative approaches to address several issues relate to the G-G/G-E interaction research. Liu *et al.* developed a G-G-E triple interaction model, where the environmental moderation effect was assessed via a varying-coefficient approach. Li *et al.* proposed a stepwise U-test for detecting gene-gene interaction associated with multiple related traits. To facilitate the G-G/G-E interaction analysis, Xu *et al.* develop GMDR software that can handle various traits (e.g., continuous and dichotomous traits), different types of data (e.g., unrelated case-control and family-based data), and covariates. Luo *et al.* applied a nonparametric varying-coefficient model to a genome-wide association dataset, and identified genetic variants that were significantly modulated by baby's cord glucose level through either linearly or nonlinearly interactions. Peng *et al.* developed a soft thresholding approach to construct networks of functional modules using gene expression datasets, in which nodes were co-expressed genes in the same subcellular location and edges represented inter-module connections.

The high-throughput genotyping and sequencing technologies take the G-G/G-E interactions research into a new level. The comprehensive evaluation of massive genetic data from the new technologies holds great promise of identifying new G-G/G-E interaction associated with complex traits, yet the complex relationship between genetic, environmental and phenotypic data and the large amounts of data bring additional statistical and computational challenges. The development of innovative statistical and computational approaches helps address these challenges, and facilitates discovering new G-G/G-E interactions of possible functional importance relating to complex traits. The five articles published in this thematic issue present the examples of using new statistical and computational approaches to tackle important topics in the G-G/G-E interaction research. We believed that the innovative ways of modeling and analyzing genetic and environmental data will play a more important role in answering various biological questions and unraveling the mechanisms underlying complex traits.

Qing Lu

Guest Editor
Department of Epidemiology and Biostatistics
Michigan State University, East Lansing
MI 48824
U.S.A