Sequence clustering for genetic mapping of binary traits

Charith Karunarathna
Dalhousie University

Abstract

Sequence relatedness has potential application to fine-mapping genetic variants contributing to inherited traits. In this talk, I present the utility of genealogical tree-based approaches to fine map causal variants in three different projects. In the first project, through coalescent simulation, we compare the ability of several popular methods of association mapping to localize causal variants in a sub-region of a candidate genomic region. We consider four broad classes of association methods, which we describe as single-variant, pooled-variant, joint-modelling and tree-based, under an additive genetic-risk model. We also investigate whether differentiating case sequences based on their carrier status for a causal variant can improve fine-mapping. Our results lend support to the potential of tree-based methods for genetic fine-mapping of disease. In the second project, we develop an R package to dynamically cluster a set of single-nucleotide variant sequences. The resulting partition structures provide important insight into the sequence relatedness. In the third project, we investigate the ability of methods based on sequence relatedness to fine-map rare causal variants and compare it to genotypic association methods. Since the true gene genealogy is unknown in reality, we apply the methods developed in the second project to estimate the sequence relatedness. We also pursue the idea of reclassifying case sequences into their carrier status using the idea of genealogical nearest neighbors. We find that method based on sequence relatedness is competitive for fine-mapping rare causal variants. We propose some general recommendations for fine-mapping rare variants in case-control association studies.

Keywords: Fine mapping; gene genealogy; association methods; sequence relatedness; disease association; causal variants