Abstract/Session Information for Program Number 1769W

Session Information

Session Title: Statistical Genetics and Genetic Epidemiology **Session Type:** Poster

Session Location: Exhibit Hall, Level 1, Convention Center **Session Time:** Wed 10:00AM-4:30PM

Abstract Information

Program Number: 1769W **Presentation Time:** Wed, Oct 23, 2013, 10:30AM-11:30AM

Keywords: Statistical Genetics and Genetic Epidemiology, KW080 - genome-wide association, KW083 - genotype-phenotype correlations, KW159 - SNP analysis/discovery, KW008 - bioinformatics

Abstract Content

PODKAT: a non-burden test for associating complex traits with rare and private variants. *U. Bodenhofer, S. Hochreiter* Institute of Bioinformatics, Johannes Kepler University, Linz, Austria.

High-throughput sequencing technologies have facilitated the identification of large numbers of single-nucleotide variations (SNVs), many of which have already been proven to be associated with diseases or other complex traits. Since association tests considering individual SNVs independently are known to be underpowered, different collapsing strategies have been proposed to consider multiple SNVs occurring in a region simultaneously. Such strategies can be classified into burden tests and non-burden tests, an important representative of which is the acclaimed Sequence Kernel Association Test (SKAT) by Wu et al. Several large sequencing studies, such as, the 1000 Genomes Project, the UK10K project, or the NHLBI-Exome Sequencing Project, have consistently reported a large proportion of private SNVs, that is, variants that are unique to a family or even a single individual. Non-burden tests like SKAT are typically utilizing correlations between SNVs to increase statistical power - a strategy that is not applicable to private SNVs, since singular events are generally uncorrelated. Burden tests are potentially able to deal with private SNVs, but only if the number of private SNVs occurring in a region is correlated with the trait under consideration. Moreover, burden tests have a disadvantage if deleterious and protective SNVs occur together in the same region. We propose the Position-Dependent Kernel Association Test (PODKAT). By means of a position-dependent kernel approach, PODKAT can potentially detect associations of rare and private SNVs with the trait under consideration even if the burden scores are not correlated with the trait. PODKAT assumes that, the closer two SNVs are on the genome, the more likely they have similar effects on the trait under consideration. This assumption is fulfilled as long as deleterious, neutral, and protective variants are grouped sufficiently well along the genome. PODKAT can be applied to testing focused regions as well as to whole-exome and whole-genome association testing. We evaluated PODKAT on simulated genomic data with simulated traits (both quantitative and dichotomous) and real sequencing data with simulated and real traits to illustrate its potential for association testing involving rare and private variants.

You may contact the first author (during and after the meeting) at bodenhofer@bioinf.jku.at