

Cost Efficiency of Genetic Linkage Studies Using Mixtures of Selected Sib-pairs

Jian Han and Rudy Guerra

Abstract

Sib-pairs are relatively easy to collect and use of extreme quantitative phenotypes provide high statistical power. Thus, selected sib-pair (discordant, concordant) study designs are among the most useful in quantitative genetic linkage analysis. Dudoit and Speed [5, 6] proposed a score test for linkage that allows analysis of any sample, random or selected, by conditioning on phenotype and analyzing genotype. Selected sampling strategies have largely focused on studies collecting data on a single type of sib-pair. Using the score test statistic, we demonstrate that sampling designs based on a mixture of sib-pair types are more cost efficient than the traditional single selection scheme. In particular, there is no need to discard a large fraction of screened individuals. Cost efficient designs are based on a mixture of concordant and discordant sib-pairs, with the selection threshold of concordant sib-pairs more stringent than that of discordant pairs. General guidelines for the thresholds are given as a function of mode of inheritance, allele frequency, and residual correlation, as well as the cost ratio of phenotyping to genotyping. Since in many cases the mode of inheritance is not completely known, robustness with respect to assumed genetic models is also addressed.

Keywords: linkage; sample size; selected sample; sib-pair; study design

1 Introduction

It is well known that quantitative genetic linkage analysis based on random sampling of sib-pairs usually has low statistical power to detect non-Mendelian, quantitative, or complex disease loci. For example, Blackwelder and Elston [3] showed in simulations that even when heritability is moderate (30%) at a single locus, the Haseman-Elston [11] linkage test based on random sampling of sib-pairs is low. Significant improvement in power can be achieved when an unselected sibling is regressed on the proportion (π) of alleles shared identical-by-descent (IBD) with a selected sibling [4]. Eaves and Meyer [7] provide evidence of additional power increases depending on the types of sib-pairs selected: discordant, with the sib-pair representing both tails of the phenotypic distribution; or concordant high (low), where both siblings are selected from the upper (lower) tail of the phenotypic distribution. However, see Allison *et al.* [1] for some limitations on the general utility of selected sib-pairs. Risch and Zhang