Extensions to a Score Test for Genetic Linkage with Identity by Descent Data

Sandrine Dudoit and Darlene R. Goldstein

Abstract

Genetic analysis aims to determine which underlying genes affect traits, their chromosomal locations and variants, and, ultimately, their modes of action at the biochemical level. Linkage analysis is an initial step in elucidating the genetic mechanisms affecting a trait of interest. This paper reviews genetic linkage analysis, with an emphasis on the score test approach developed by Dudoit and Speed [8, 10]. Two extensions of the test under current investigation are also presented: use of the test with larger sets of relatives than pairs, and generalization to allow for missing DNA identity by descent (IBD) information.

Keywords: allele sharing; complex traits; identity by descent; linkage analysis; pedigree; score test

1 Introduction

A central problem in genetic analysis is to determine which gene(s), if any, affect particular phenotypes, the chromosomal locations of these genes, their different alleles and, ultimately, their biochemical modes of action. Linkage analysis is an initial step in elucidating the genetic mechanisms affecting a trait of interest. Its goal is to determine the chromosomal location of the gene(s) influencing the trait. Linkage analysis proceeds by tracking patterns of coinheritance of the trait of interest and other traits or genetic markers, relying on the varying degree of recombination between trait and marker loci to map the loci relative to one another.

Mendel's second law of inheritance hypothesizes that different "factors" (traits or genes) segregate to gametes (sperm or egg) independently. Actually, independent assortment of gene pairs only occurs when the genes are on different chromosomes or are so far apart on the same chromosome that there is the same chance of recombination as nonrecombination. Such pairs of genes are said to be *unlinked.* Two genes are *linked* when they do not segregate independently. A measure of the degree of linkage is the *recombination fraction*, the chance of recombination occurring between two loci, denoted almost universally in the genetics literature as θ. For unlinked genes, $\theta = 1/2$; for linked genes, $0 \leqslant \theta < 1/2$. The following gives a brief introduction to linkage analysis; more substantial detail is provided by Ott [30], McPeek [27], and Speed [32].