GMR Genetics and Molecular Research
Online Journal
ISSN:1676-5680

A novel method for identifying SNP disease association based on maximal information coefficient

H.M. Liu[1,2], N. Rao[1], D. Yang[2], L. Yang[1], Y. Li[1] and F. Ou[1]

[1]School of Life Science and Technology,
University of Electronic Science and Technology of China,
Chengdu, Sichuan, China
[2]School of Mathematics and Computer Science,
Gannan Normal University, Ganzhou, Jiangxi, China

Corresponding author: N. Rao
E-mail: raonn@uestc.edu.cn

Genet. Mol. Res. 13 (4): 10863-10877 (2014)
Received January 16, 2014
Accepted July 7, 2014
Published December 19, 2014
DOI http://dx.doi.org/10.4238/2014.December.19.7

ABSTRACT. To improve single-nucleotide polymorphism (SNP) association studies, we developed a method referred to as maximal information coefficient (MIC)-based SNP searching (MICSNPs) by employing a novel statistical approach known as the MIC to identify SNP disease associations. MIC values varied with minor allele frequencies of SNPs and the odds ratios for disease. We used a Monte Carlo-based permutation test to eliminate the effects of fluctuating MIC values and included a sliding-window-based binary search whose time-cost was 0.58% that of a sequential search to save time. The experiments examining both simulation and actual data demonstrated that our method is computationally and statistically feasible after reducing the resampling count to 4 times the number of markers and applying a

sliding-window-based binary search to the method. We found that our method outperforms existing approaches.

Key words: Sliding-windows; Maximal information coefficient (MIC); Fluctuation of MIC values; Monte Carlo-based permutation test; Binary search; SNP disease association studies

