# Extended Homozygosity Score Test to Detect Positive Selection in Genome-wide Scans

**Ruzong Fan, NICHD/NIH, February 2012**

## 1 Overview

This document describes a C++ package to implement the models for Extended Homozygosity Score Test to Detect Positive Selection in Genome-wide Scans. Section 2 briefly describes the installation and compile the program. Section 3 explains how to run the program for population data with one example of HapMap data.

The theoretical basis for this program is given in our research papers in **Reference**. Please refer to the reference if you use in any published work. In case of suggestions and questions and/or problems, you can contact us via e-mail (fanr@mail.nih.gov).

## 2 Installation and Compile the Program

The package is good in Linux. First, download the package "EGHH.cc" and "EGHH.h" from EHST.zip. Use the following steps on Linux to install and compile the package:

- Put EGHH.cc and EGHH.h in a directory
- Compile the package by

  **g++ -O3 EGHH.cc -o EGHH**

## 3 How to Run the Program for HapMap Data

This program requires a pair of matched data files: "genotypes_chr2_YRI_r22_nr.b36_fwd_legend.txt" and "genotypes_chr2_YRI_r22_nr.b36_fwd.phase", for instance. The first one is SNP legend file, and the second is the haplotype phase file. To run analysis, simply type

EGHH genotypes_chr2_YRI_r22_nr.b36_fwd_legend.txt genotypes_chr2_YRI_r22_nr.b36_fwd.phase YRI_chr2

Then, three output files are generated: "YRI_chr2_egh.out", "YRI_chr2_ehh.out", and "YRI_chr2_hmm.out".

The file "YRI_chr2_egh.out"contains the results of extended genotype-based homozygosity score test (EGHST); the file "YRI_chr2_hmm.out" contains the results of hidden Markov model score test (HMMST); and "YRI_chr2_ehh.out" contains results of extended haplotype-based homozygosity score test (EHHST). For the theoretical description of the three score tests, refer to Zhong et al. (2010).

# 4 References


1. Zhong M, Lange K, Papp JC, and Fan RZ (2010) Extended homozygosity score tests to detect positive selection in genome-wide scans. *European Journal of Human Genetics* 18:1148-1159.