

# Document for LDMgen software

Shuanglin Zhang & Hongyu Zhao

Department of Epidemiology and public Health  
Yale University School of Medicine  
60 College street  
New Haven, CT 06520, USA

## INTRODUCTION

The program *LDMhap* implements a fine-mapping method based the paper by Zhang and Zhao titled "Linkage Disequilibrium Mapping by Using Genotype Data" that can be find in our website.

We strongly encourage users to send me a brief email at [shuanglin.zhang@yale.edu](mailto:shuanglin.zhang@yale.edu) or [hongyu.zhao@yale.edu](mailto:hongyu.zhao@yale.edu) , saying that you have download the program and what sort of the questions the program applied to. This will help us to assess the amount of interest in *LDMhap*, and how much effort we should put into maintaining and improving it.

## GETTING STARTED

You will need an executable file: *LDMhap.exe*, one file of parameters : *parameters.txt* and a data file (the name of the data file given in the file *parameters.txt*). We strongly suggest that you put that executable file, parameters file and data file in a same folder. To analyze your own data, you need to prepare a data file in the appropriate format.

## FORMAT FOR INPUT DATA FILE

In the parameter file *parameters.txt* (see example file), there are description lines and number lines. For each term, only has one description line, the number following the description line can take one line or more than one line.

The format of the data files are as follows. The data file *disease\_p.txt* for disease individual's genotypes. Each line for one individual. First and second rows are the two alleles of the individual at the first marker, and third and fourth rows are the two alleles of the individual at the second marker, et al. The format of the data file containing normal individual's genotypes is same as that of *disease\_p.txt*.

## **FORMAT FOR OUT DATA FILE**

There are two out put files: *loca\_likeli.txt* and *result.out*. The estimations of location, the age of disease mutation, ancestral haplotype, mutation rate and the heterogeneity parameter  $p$  are given in the file *result.out*. Fixing other parameters, the values of location versus log-likelihood are given in the file *loca\_likeli.txt*. You can use the file *location.txt* to get the confidence interval (detail see the paper).