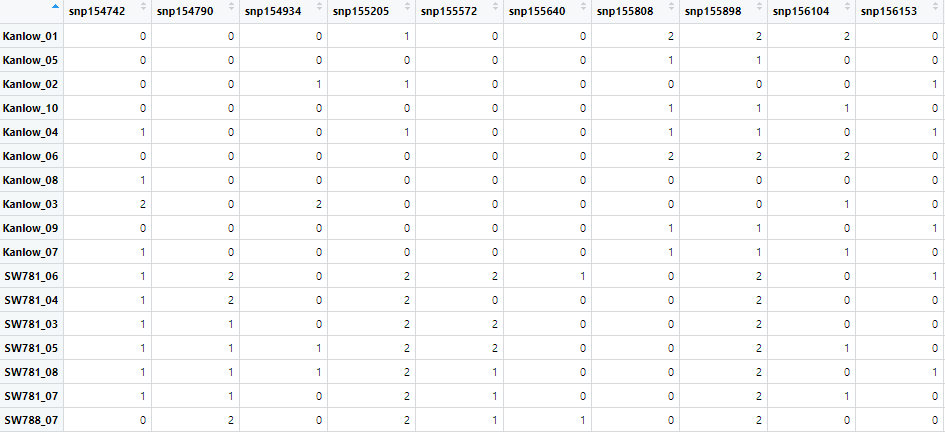
**TetraplodLD**

1. **Introduction**

At present, the **TetraplodLD** package is used to linkage disequilibrium analysis in the nature population of allotetraploid. This guide gives some brief instructions on how to perform the tasks of linkage disequilibrium analysis by this package. The outline of this guide is as follows:

1. **Data format**



This table indicates dosage-unknown markers, each marker contain three genotypes (aaaa=0, A\_\_\_=1, AAAA=2).

1. **Work example**

#Dosage-unknown marker

#read genotype file

chr03a\_m <- read.csv("./data/chr03am.csv")[,-1]

#extract snp ID

snpn <- colnames(chr03a\_m)

#load functions

source("util\_PC\_m.R")

source("util\_PC.R")

# Linkage disequilibrium analysis

LD <- work\_test1\_mix(M=chr03a\_m,mn=snpn)

#M is a matrix; mn is a character vector giving the snp ID.

#work\_test1\_mix produces a matrix with some or all of the following elements in :

**LR** The log-likelihood ratio between haplotype and diplotype models.

**Pv** P-value is calculated through chi-square distribution based on LR.

**m1\_pA** The estimated allele frequency of A by haplotype model.

**m1\_pB** The estimated allele frequency of B by haplotype model.

**m1\_D** The estimate of the LD coefficient by haplotype model.

**m2\_pA** The estimated allele frequency of A by diplotype model.

**m1\_pB** The estimated allele frequency of B by diplotype model.

**m1\_DA** The estimate of the LD coefficient at the locus A.

**m1\_DB** The estimate of the LD coefficient at the locus B.

**m1\_Deab** The sum of the estimate of the LD coefficient between two nonalleles at different loci on the same haplotype and the estimate of the LD coefficient between two nonalleles on different haplotypes.

**m1\_DAb** The estimate of the LD coefficient between two alleles from SNP **A** and one allele from SNP **B**.

**m1\_DaB** The estimate of the LD coefficient between two alleles from SNP **A** and one allele from SNP **B**.

**m1\_DAB** between two alleles from SNP **A** and two alleles from SNP **B**.

**m2\_DA\_n, m2\_DB\_n, m2\_Deab\_n, m2\_DAb\_n, m2\_DaB\_n, m2\_DAB\_n** The estimate of the standardized LD coefficient.

1. **Computer simulation**

# Simulation dosage-unknown markers based on diplotype model

#load functions

source("util\_FC.R")

# The true value of LD coefficient in computer simulation

pall <- c(0.6,0.45,0.12,0.12,0.02,0.02,0.02,0.01)

# 1000 simulation replicates

allres <- c()

for(i in 1:1000){

geno <- sim\_geno\_auto(allpar=pall,n=500) #n represents the sample size

res <- geno\_est\_auto(geno=geno)

allres <- rbind(allres,res)

}

#Calculate the mean value of the estimation paraeters

colMeans(allres)

#Calculate the standard deviation of the estimation paraeters

apply(allres,2,sd)

# Simulation dosage-known markers based on diplotype model

#load functions

source("util\_PC.R")

# The true value of LD coefficient in computer simulation

pall <- c(0.6,0.45,0.12,0.12,0.02,0.02,0.02,0.01)

# 1000 simulation replicates

allres <- c()

for(i in 1:1000){

geno <- sim\_geno\_autoP(allpar=pall,n=500) #n represents the sample size

res <- geno\_est\_autoP(geno=geno)

allres <- rbind(allres,res)

}

#Calculate the mean value of the estimation paraeters

colMeans(allres)

#Calculate the standard deviation of the estimation paraeters

apply(allres,2,sd)