

Help Document

DeMix : Decovolution for Mixed cancer genomes using raw measured data.

Updated date : 04/01/2013

Version : v1.0.0

Jaeil Ahn :email to (jahn1@mdanderson.org)

1. How to use :

To use demix, you need to move all files after unzipping Demix.zip to the designated folder where the dataset locates. Then, you need to have 'R' on your machine. When you open R, follow the instruction below. Please note that there are two versions of DeMix: one for Mac and other for linux (for PC users). When you download the DeMix software, you must download appropriate version depending on what kind of operation system you have.

- a. First load a function by using `'source("TumorHv1.R");'`
- b. The function name is '`Rdemix`'. To use this function, you need to input parameters listed below
 - i. Inputmat1 : Matrix of gene expressions after the RMA normalization. The G*S matrix where G is the number of genes and S is the number of samples.
 - ii. Groupid : Vector of indicators. Element 0 for normal and 1 for mixed samples. For 6 sample setup, we might have `c(0,1,0,1,0,1)`.
 - iii. Cit : Cit=64 tells the machine is built based on 64 bit. It can be 32 bit for old machines.
 - iv. Matched : For matched sample, we need to specify `matched=1`, otherwise `matched=0`. For the matched case, the Groupid needs to specified in the form of `(0,1,0,1, ..., 0,1)`, i.e., the tumor samples need to be aligned next to the corresponding normal sample.
 - v. Nref : If you have reference genes, set `Nref=1` otherwise `Nref=0`
 - vi. RefMat : If `Nref =1`, we need to define the matrix with reference gene expressions with variance (See Example).
 - vii. Nstart : Number of initial values for tumor proportion phi. The running time is proportional to 'nstart'.
 - viii. Nhuman : If the data are collected from human sample, specify `Nhuman=1`, otherwise put 0.

2. Return values

DeMix will return two types of outcomes.

- a. phis : Mixed proportions of the second cell types.
- b. Decov : Deconvolved expressions after log-transformed corresponding to mixed samples.

3. Example when there is no reference

Please see the following example.

```
rm(list=ls(all=TRUE))
setwd('/Users/jaeilahn/CNV/CNV/Grid/R_pack/');

# Load maqc3 data
rnadata<-as.matrix(read.table("maqcdata3.txt", header=FALSE ))

# Load R function
source("TumorH.R");

# Define group: 0 for normal 1 for mixed
cnvgroup<-c(rep(0, 3), rep(1, 12)); # maqc

# Define other parameter
cbit=64;
matched=0;
nstart=2
ref=0
refmatrix=matrix(0, nrow(rnadata), 2)
nhuman=1

testr<-Rdemix(rnadata, cnvgroup, cbit, matched, ref, refmatrix,
nstart, nhuman)

# Estimated proportions (vector with 12 elements)
> testr$phi
[1] 0.3175708 0.3087230 0.3008485 0.7205728 0.7477822 0.7293976
0.6478953 0.6695369 0.6420604 0.6438080 0.6358323 0.6279505

# Estimated deconvolved expressions (log2-transformed)
> testr$decov[1:5,1:5]
      [,1]      [,2]      [,3]      [,4]      [,5]
[1,] 10.692062 10.768543 10.673647 10.156418  9.799552
[2,]  9.638798  9.708867  9.610924  9.014184  8.988028
[3,]  8.979478  9.068793  9.058247  9.493976  9.408624
[4,] 10.077569 10.069213 10.052203 10.515880 10.455641
[5,] 10.191303 10.203609 10.187710 10.928037 10.826842
```

1. Example when there is reference
Please see the following example.

```
rm(list=ls(all=TRUE))
setwd('/Users/jaeilahn/CNV/CNV/Grid/R_pack/');

# Example maqc1 data
rnadata<-as.matrix(read.table("Rmaqcdata1.txt", header=FALSE ))
rnadata2<-as.matrix(read.table("Rmaqcdata1ori.txt", header=FALSE
))

nabsv<-(abs(apply(rnadata[,1:5], 1, mean)-apply(rnadata2[,1:5],
1, mean) )>3 & abs(2^apply(rnadata[,1:5], 1, mean)-
2^apply(rnadata2[,1:5], 1, mean) )<5000)

rnadata=rnadata[nabsv,]
rnadataT<-rnadata2[nabsv,]
refdata2<-cbind((apply(rnadataT, 1, mean)), (apply(rnadata, 1 ,
sd))^2);
source("TumorHv1.R");

# define group 0 for normal 1 for mixed
cnvgroup<-c(rep(0, 5), rep(1, 10)); # Maqc

cbit=64;
matched=0;
nstart=2
inputmat1<-rnadata

testr<-Rdemix(inputmat1, cnvgroup, cbit, matched, 1, refdata2,
nstart, 1)

# Estimated proportions (vector with 12 elements)
> testr$phi
[1] 0.11095734 0.08985242 0.13702647 0.12959593 0.13757911
0.58713162 0.59679857 0.58952894 0.58964421

# Estimated deconvolved expressions (log2-transformed)
> testr$decov[1:5,1:5]
[,1]      [,2]      [,3]      [,4]      [,5]
[1,]  9.283528  9.326418  9.329793  9.352642  9.283690
[2,] 10.398591 10.362051 10.408206 10.385461 10.286755
[3,]  9.142444  9.256425  9.251844  9.245567  9.155768
[4,]  9.329873  9.291780  9.356418  9.335080  9.336665
[5,]  9.253079  9.316123  9.288251  9.227155  9.267332
```