方法1 SCeQTL

Zero-inflated generalized linear model

零膨胀负二项分布模型

<https://github.com/XuegongLab/SCeQTL/>

图形用户界面, 文本, 应用程序

描述已自动生成

SCeQTL

SCeQTL is an R package that uses zero-inflated negative binomial regression to do eQTL analysis on single-cell data. lt can distinguish two type of gene-expression differences among different genotype groups. lt's more suitable to use SCeOTL to identify eQTLs from single-cell data. lt can also be used for finding gene expression variations associated with other grouping factors like cell lineages.

SCeQTL 是一个 R 软件包，它使用零膨胀负二项式回归对单细胞数据进行 eQTL 分析。它可以区分不同基因型组之间的两种基因表达差异。使用SCeQTL从单细胞数据中识别eQTL更合适。它还可用于查找与其他分组因素（如细胞谱系）相关的基因表达变异。

<https://github.com/XuegongLab/SCeQTL/blob/master/README.md>

**installation：**

if(!require(devtools)) install.packages("devtools")

devtools::install\_github("XuegongLab/SCeQTL")

**Input:**

**SCeQTL** takes two inputs: gene and snp.

输入是SNP 和gene

The input gene is a scRNA-seq **raw read counts matrix**. The rows of the matrix are genes and columns are cells.

行是基因列是细胞

The other input snp is a **genotype matrix**, where each element should be 0/1/2 indicating three types of genotypes. The order of columns should be the same with gene.

基因型：0/1/2

列的基因序列应该与表达矩阵相同

Usage :

#an example to run **SCeQTL** with read counts matrix input

# Load library and the test data for SCeQTL

library(SCeQTL)

data(test)

# check whether the non-zero part could fit negative binomial well

# This function may fail since It's possible that the random picked gene can't be fit to negative bionomial distribution, all zero value for example

checkdist(gene)

# normalize gene matrix using DEseq method

normalize(gene)

# Detecting the SCeQTL

result <- cal.pvalue(gene, snp)

# Picking one sample to visualize

check.sample(gene[1,], snp[1,])

Output:

SCeQTL will output the p-value of each gene-snp pair. You could also get q-value by running cal.qvalue(gene, snp) instead of cal.pvalue(gene, snp).

得到每一对gene-snp间的P值或q值

#这里的snp 应该是以阿拉伯数字为索引的

表格

描述已自动生成

SCeQTL::cal.qvalue(gene[1:10,], snp)

#Then you may interested in the significant gene-snp pairs. You could stat and visualize it by check.sample. Here take the most significant pair for example.

SCeQTL::check.sample(gene["CDC5L",], snp[10, ])

可以得到这对pair 的统计分析结果和箱型图

#Before run cal.qvalue, you may want to check whether the distribution fit the bionomial negative assumption well by checkdist

检测分布是否满足负二项分布

checkdist(gene)

是随机挑选一些基因的，可以看看分布情况