Limma是基于R平台的常用来对芯片数据进行常见分析的程序包。本代码是使用limma包对GEO公用数据库上的GSE56015数据进行分析，筛选其中经过石蒜碱处理的和没有经过石蒜碱处理的HSPC细胞中差异表达的基因。简要说明下该组芯片中共包含18组数据：共有三种细胞类型，每种细胞类型都分为经过石蒜碱处理的实验组以及没有经过石蒜碱处理的对照组且每组做了三次生物学重复。具体的实验设计等详细信息可以登录GEO数据库搜索GSE56015获得。在此提供已经下载完毕的芯片矩阵文本文件GSE56015\_series\_matrix.txt通过以下代码，即可筛选出经过石蒜碱处理的在HSPC中差异表达的基因，得到的dif文件既包含了差异表达基因的全部信息。

代码如下:

#设置工作目录,确保该目录下有存放芯片文本文件以及注释文件

setwd(“G:/R”)

#下载安装脚本

source("http://bioconductor.org/biocLite.R")

#安装所有 Bioconductor 核心包

biocLite()

#在bioconductor网站下载limma包

biocLite("limma")

#加载limma包

library(limma)

#读取芯片文本文件

date<-read.table("GSE56015\_series\_matrix.txt", header = TRUE, sep = "",comment.char = "!",row.names="ID\_REF")

#显示基因表达矩阵前六行

head(date)

#构建实验设计矩阵

condition=factor(c(1,2,3,1,2,3,1,2,3,4,5,6,4,5,6,4,5,6))

design <- model.matrix(~-1+condition)

# 构 建 对比 模 型，比较两个实验条件下表达数据，其中condition2代表只用DMSO处理过的HSPC细胞，condition5代表用石蒜碱处理过的HSPC

contrast.matrix <- makeContrasts (contrasts = "condition2 - condition5", levels = design)

#线性模型拟合

fit <- lmFit(date, design)

#根据比对模型进行差值计算#

fit1 <- contrasts.fit(fit, contrast.matrix)

# 贝 叶 斯 检验

fit2 <- eBayes(fit1)

# 生成所有 基因的检验结果报表

dif <- topTable(fit2, coef = "condition2 - condition5", n = nrow(fit2), lfc = log2(1.5))

# 根据P.Value对结果进行筛选，得到全部差异表达基因

dif <- dif[dif[, "P.Value"] < 0.01, ]

#查看结果的前6行

head(dif)

#将结果保存在dif文件中

write.table(dif,"dif.txt",row.names=T, sep="\t",quote=F)