**gfold软件使用说明**

**官网地址**：https://zhanglab.tongji.edu.cn/softwares/GFOLD/index.html

有重复的转录组数据：DEseq2(负二项分布)，edgeR(负二项分布)， edgeR（二项分布）。

没有重复转录组数据：推荐使用同济大学的GFOLD。

**优点**：

gfold软件特别适合当没有生物学重复的情况下的RNAseq的数据分析。该软件称尤其适合做无重复样本的差异分析，它对foldchange 的计算考虑到posterior distribution，即克服了pvalue评估显著性的缺点，同时也克服了 fold change 在评估低counts 数的gene时的缺点。

在安装Gfold之前，必须安装一款基于GNU的数值计算工具。

安装gsl软件（安装gfold之前，要保证gsl软件先安装上，否则会报错）

GSL官网：http://www.gnu.org/software/gsl/

**下载GSL：**

wget ftp://ftp.gnu.org/gnu/gsl/gsl-2.2.tar.gz

unzip gsl-2.2.tar.gz

cd gsl-2.2/

对于我们没有root权限的用户，安装必须指定路径，安装到自己的用户环境中。

./configure --prefix=/public/data/zhangpeng/resource/gsl-2.2/

make

make install

**下载gfold：**

wget http://compbio.tongji.edu.cn/~fengjx/GFOLD/gfold.V1.1.2.tar.gz

unzip gfold.V1.1.2.tar.gz

cd feeldead-gfold-1921fd6dc668

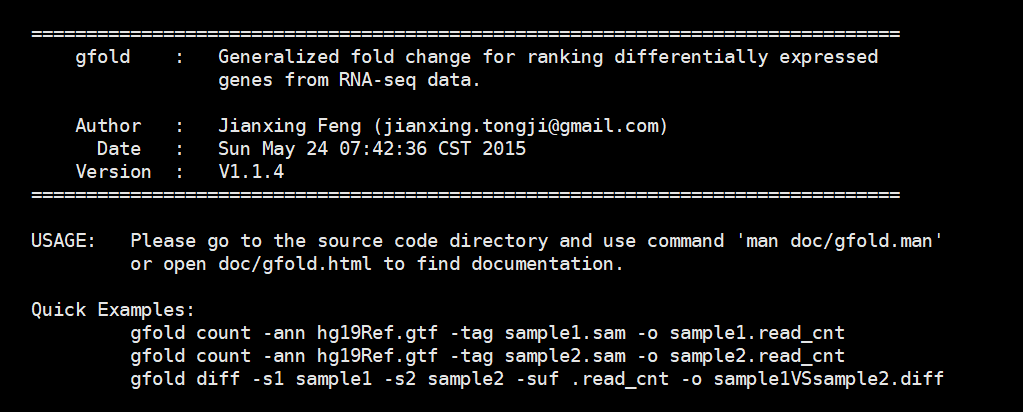
export CXXFLAGS="-g -03 -I/public/data/zhangpeng/resource/gsl-2.2/include -L/public/data/zhangpeng/resource/gsl-2.2/lib"

export LD\_LIBRARY\_PATH="/public/data/zhangpeng/resource/gsl-2.2/lib:"$LD\_LIBRARY\_PATH

g++ -O3 -Wall -g main.cc -o gfold -lgsl -lgslcblas -I/public/data/zhangpeng/resource/gsl-2.2/include -L/public/data/zhangpeng/resource/gsl-2.2/lib

**判断是否安装成功**

**在**feeldead-gfold-1921fd6dc668文件夹执行：**./gfold –h**



**功能**：

① read count计数和基因排序

gfold count -ann hg19Ref.gtf -tag sample1.sam -o sample1.read\_cnt

gfold count -ann hg19Ref.gtf -tag sample2.sam -o sample2.read\_cnt

gfold diff -s1 sample1 -s2 sample2 -suf .read\_cnt -o sample1VSsample2.diff

② read count计数

samtools view sample1.bam | gfold count -ann hg19Ref.gtf -tag stdin -o sample1.read\_cnt

③ 无重复样本的差异表达基因的识别

gfold diff -s1 sample1 -s2 sample2 -suf .read\_cnt -o sample1VSsample2.diff

④ 两组样本均含有重复样本的差异表达基因的识别

gfold diff -s1 sample1,sample2,sample3 -s2 sample4,sample5,sample6 -suf .read\_cnt -o 123VS456.diff

⑤ 只有一组样本含有重复的差异表达基因识别。

gfold diff -s1 sample1,sample2 -s2 sample3 -suf .read\_cnt -o 12VS3.diff

执行之前需要运行以下指令：

export CXXFLAGS="-g -03 -I/public/data/zhangpeng/resource/gsl-2.2/include -L/public/data/zhangpeng/resource/gsl-2.2/lib"

export LD\_LIBRARY\_PATH="/public/data/zhangpeng/resource/gsl-2.2/lib:"$LD\_LIBRARY\_PATH

**实例：**

/public/data/zhangpeng/resource/feeldead-gfold-1921fd6dc668/gfold count -ann /public/data/zhangpeng/resource/gene/Drosophila\_melanogaster.BDGP6.22.96.gtf -tag sample1.sam -o sample1.read\_cnt

/public/data/zhangpeng/resource/feeldead-gfold-1921fd6dc668/gfold diff -s1 sample1 -s2 sample2 -o sample2VSsample1.diff **# s1\_vs\_s2 的Gfold是基于s2。**

**查看.diff文件**

less -S old1\_VS\_young1.diff|column -t|head -n 50

**输出.csv文件**

less -S old1\_VS\_young1.diff | sed 's/\t/,/g;s/[[:space:]]//g' > old1\_VS\_young1.csv