1. **所需軟體**
2. **RSEM(v1.3.0) : (quantifying transcript abundances from RNA-Seq data)**

<https://deweylab.github.io/RSEM/>

# wget <https://github.com/deweylab/RSEM/archive/v1.3.0.tar.gz>

# tar zxvf v1.3.0.tar.gz

# cd ~/RSEM/RSEM-1.3.0

# make

# PATH="$HOME/bio\_tool/RSEM/RSEM-1.3.0:$PATH"

# source ~/.profile

1. **bowtie(v1.0.0) : (align to reference)**

**(RSEM1.3.0與bowtie1.2在reads 為paired-end時不相容,故不安裝最新版)**

[https://sourceforge.net/projects/bowtie-bio/files/bowtie/1.0.0/](https://sourceforge.net/projects/bowtie-bio/files/bowtie/1.0.0/%20)

# unzip

# PATH="$HOME/bio\_tool/bowtie/bowtie-1.0.0:$PATH"

# source ~/.profile

要使用bowtie需有libtbb套件

# sudo apt-get update

# sudo apt-get install libtbb-dev (or #sudo apt-get install libtbb2)

1. **cutadapt(v1.13) : (修剪fastq軟體)**

<http://cutadapt.readthedocs.io/en/stable/installation.html>

# sudo pip install cutadapt (請管理員安裝)

1. **sratoolkit(2.8.2) : (sra 轉 fastq , 下載sra 資料)**

<https://trace.ncbi.nlm.nih.gov/Traces/sra/sra.cgi?view=software>

# wget <https://ftp-trace.ncbi.nlm.nih.gov/sra/sdk/2.8.2-1/sratoolkit.2.8.2-1-ubuntu64.tar.gz>

# tar zxf sratoolkit.2.8.2-1-ubuntu64.tar.gz

# PATH="$HOME/bio\_tool/sratoolkit/sratoolkit.2.8.2-1-ubuntu64/bin:$PATH"

# source ~/.profile

**使用 :**

# prefetch SRR975551 SRR975552 SRR975553 (download sra file)

# fastq-dump file.sra (sra to fastq) (single-end)

# fastq-dump –split-3 file.sra (sra to \_1.fastq and \_2.fastq) (paired-end)

1. **FastQC(v0.11.5) : (quality control checks on raw sequence data)**

<http://www.bioinformatics.babraham.ac.uk/projects/download.html#fastqc>

# wget <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/fastqc_v0.11.5.zip>

# unzip fastqc\_v0.11.5.zip

# chmod 755 fastqc

# PATH="$HOME/bio\_tool/FastQC/FastQC:$PATH"

# source ~/.profile

1. **資料前處裡**

**當拿到reads資料(.sra)時，須將sra轉檔為fastq (sample.sra 🡺 sample.fastq)**

若此reads為single-end(SE) : fastq-dump sample.sra

**一次轉多個檔**

# bash dump\_SE\_multi.sh ~/ncbi/public/sra/ #sra檔所在資料夾

若此reads為paired-end(PE) : fastq-dump --split-3 sample.sra

(切割成upstream、downstream)

**一次轉多個檔**

# bash dump\_PE\_multi.sh ~/ncbi/public/sra/ #sra檔所在資料夾

1. **計算表現量**

**RSEM.sh流程 :**

**(已將多支程式彙整在RSEM.sh 或 RSEM\_multi.sh，執行RSEM.sh即可算出表現量並匯入資料庫)**

reads

FastQC

trim

untrim

cutadapt

rsem-calculate-expression

FastQC

rsem-calculate-expression

MySQL

MySQL

**(SE or PE)**

**一次計算一個fastq**

# bash RSEM.sh \

~/RNA\_data/RNAseq/sample.fastq \

trim ( or untrim) \

single-end ( or paired-end)

(# bash RSEM.sh 序列原始檔 要不要去除adapter SE或PE)

**一次計算多個fastq**

# bash RSEM\_multi.sh \

~/RNA\_data/RNAseq/sample.fastq \

~/RNA\_data/RNAseq/RNA\_GFP.fastq \

~/RNA\_data/RNAseq/RNA\_GFPQ1.fastq \

trim (or untrim) \

single-end (or paired-end)

(# bash RSEM\_multi.sh 序列原始檔 要不要去除adapter SE或PE)