**Y window系统下本地blast+安装与使用教程**

<http://blog.sina.com.cn/s/blog_567b43d00101maqx.html>

一、blast的下载与安装

1.程序下载：访问blast本地软件包链接 blast\_latest 下载适合自己系统的blast版本，这里我选择 ncbi-blast-2.2.28+-win64.exe。<ftp://ftp.ncbi.nlm.nih.gov/blast/executables/LATEST>

2.安装流程：下载完毕后，双击安装到C:\Blast，生成bin和doc两个子目录，其中bin是程序目录， doc是文档目录，这样就安装完成。

3.用户环境变量设置：右键点击“我的电脑”-属性，然后“高级系统设置”选项-“环境变量”，在用户变量下方点击“新建”-变量名：BLASTDB，变量值：C:\Blast\db(即数据库路径)。在系统变量下方“Path”添加变量值：C:\Blast\bin。

4.查看程序版本信息:点击window的"开始"菜单，在运行中输入cmd，调出MS-DOS命令行，转到blast安装目录，输入命令“ blastn -version"即可查看版本。

## Databases

BLAST databases are updated daily and may be downloaded via FTP from [ftp://ftp.ncbi.nlm.nih.gov/blast/db/](https://ftp.ncbi.nlm.nih.gov/blast/db/).

二、blast本地数据库的构建

1.数据的获取

1.1 直接从NCBI或者其他数据库网站下载所需序列做成数据库，或者自己已有的测序数据（格式必须是fasta,名字可以自己随便命名）。

1.2 从NCBI中的ftp库下载所需要的某一个库或几个库（其链接为ftp://ftp.ncbi.nlm.nih.gov/blast/db/FASTA/）其中nr.gz为非冗余的数据库,nt.gz为核酸数据库,month.nt.gz为最近一个月的核酸序列数据。

1.3 利用新版blast自带的update\_blastdb.pl进行下载，这需要安装perl程序。

上述三种方法各有优缺点：前两种下载速度较快，但是检索前都需要对数据库进行格式化（转化成二进制数据），第三种方法下载速度较慢，但是是NCBI中已经格式化好的，在进行本地检索时不需再进行格式化，直接用即可。

2.数据的格式化

以xk001.fasta作为查询序列，以nr.fasta作为数据库文件为例。首先将nr.fasta放到C:\Blast\db文件夹下，然后调出MS-DOS命令行，转到C:\Blast\db文件夹下运行格式化命令。

格式化nr.fasta命令：

C:\Blast\db>makeblastdb -in nr.fasta -dbtype prot -title "nr" -out NR

Building a new DB, current time: 08/28/2013 08:59:17

New DB name: T

New DB title: nr

Sequence type: Protein

Keep Linkouts: T

Keep MBits: T

Maximum file size: 1000000000B

Adding sequences from FASTA; added 32044604 sequences in 3134.69 seconds.

因此，本地数据库已经建立完毕。

三、blast的使用方法

以xk001.fasta作为查询序列，以nr.fasta作为数据库文件为例。

blsat运行命令：

C:\Blast>blastx.exe -db NR -query xk001.fasta –out xk001.out -evalue 0.00001 -max\_target\_seqs 5 -num\_threads 4 -outfmt 11

注：1、将核酸序列比对到蛋白库需要用blastx,outfmt的格式选择11，11为ans.格式，此格式可以转化为其他格式文件。其他参数可以通过-h选择。

如果一切运行良好的话，待运行完毕，你将在blast文件下看到一个xk001.out的结果。

2、查询序列应放在blast目录下，不能放在数据库的目录下。放在数据库的目录下，程序找不到。