

BLAST 程序的参数

徐桂霞 国春策

中国科学院植物研究所

2009.04.24

Windows 系统下BLAST程序的安裝



BLAST程序的的下载:

<ftp://ftp.ncbi.nih.gov/blast/executables/release/2.2.20/>

A screenshot of the NCBI website. On the left is a blue navigation sidebar with several menu items. The "FTP site" item is highlighted with a red border. The main content area shows "NCBI News" with a list of links. On the right, a list of resources is displayed with blue arrow icons.

Software engineering
Tools, R&D, and databases

Education
Teaching resources and on-line tutorials

FTP site
Download data and software

Contact information
How to reach us

Internet connection.

NCBI News

NCBI News available online
[April issue](#)

- [Featured Resource: PubChem Offers 3-D Small Molecule Structures and a New Conformer Viewer \(Pc3D\)](#)
- [New Databases and Tools](#)
- [GenBank News](#)
- [Updates and Enhancements](#)
- [Announce Lists and RSS Feeds](#)

sequence project

- ▶ Short Read Archive
- ▶ SKY/CGH database
- ▶ dbSNP
- ▶ Trace Archives
- ▶ VecScreen
- ▶ Viral Genotyping Tool
- ▶ Web (RSS) Feeds

Windows系统下BLAST程序的安装



BLAST程序的的下载：进入NCBI的FTP站点



The screenshot shows the NCBI FTP site page. At the top, there is a navigation bar with links for PubMed, Entrez, BLAST, OMIM, Books, TaxBrowser, and Structure. Below this is a search bar with a dropdown menu set to "All Databases" and a "Go" button. The main content area is titled "Major resources available by ftp (ftp.ncbi.nih.gov):". A red box highlights the link for "BLAST Basic Local Alignment Search Tool". Below this link, there is a description: "Download the BLAST database and stand-alone sequence comparison software." Another link, "CDD Data", is visible below the highlighted link, with a description: "Download data from the Conserved Domain Database." On the left side of the page, there is a sidebar with links for "NCBI", "SITE MAP", "Guide to NCBI resources", and "About NCBI".

NCBI **FTP site**

PubMed Entrez **BLAST** OMIM Books TaxBrowser Structure

Search All Databases for Go

NCBI

SITE MAP
Guide to NCBI resources

About NCBI
The science behind our resources. An introduction for researchers, educators and the

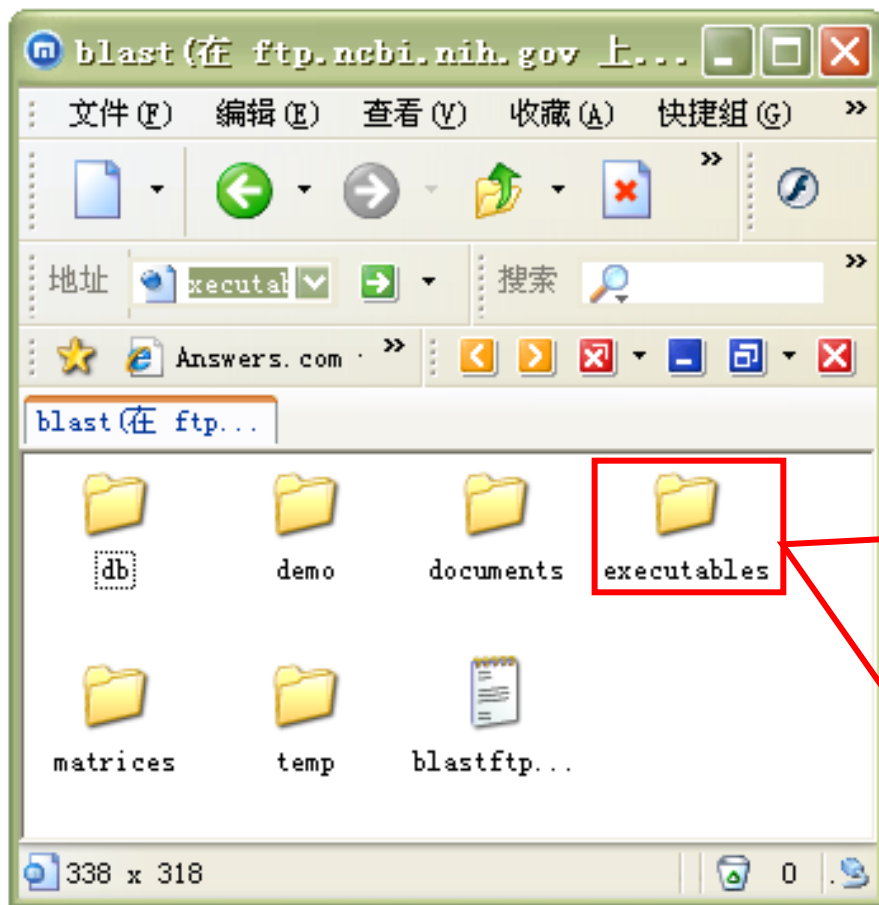
Major resources available by ftp (ftp.ncbi.nih.gov):

- ▶ [BLAST Basic Local Alignment Search Tool](#)
Download the BLAST database and stand-alone sequence comparison software.
- ▶ [CDD Data](#)
Download data from the Conserved Domain Database.

Windows系统下BLAST程序的安裝



BLAST程序的的下载: blast/executables/LATEST

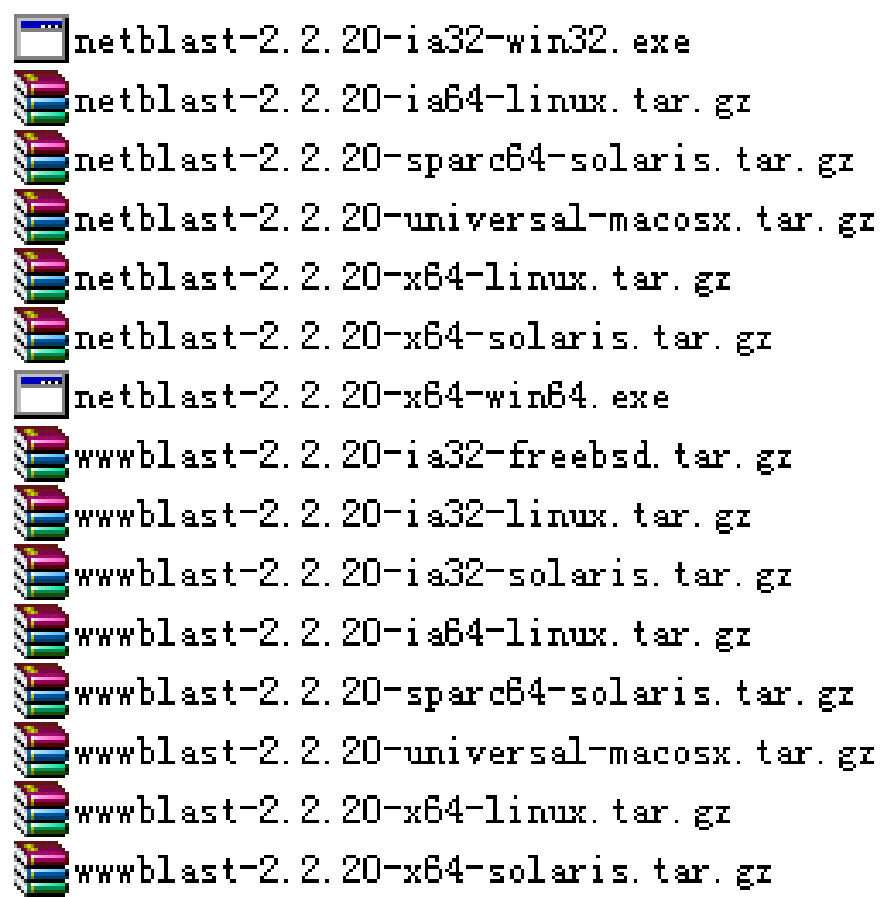
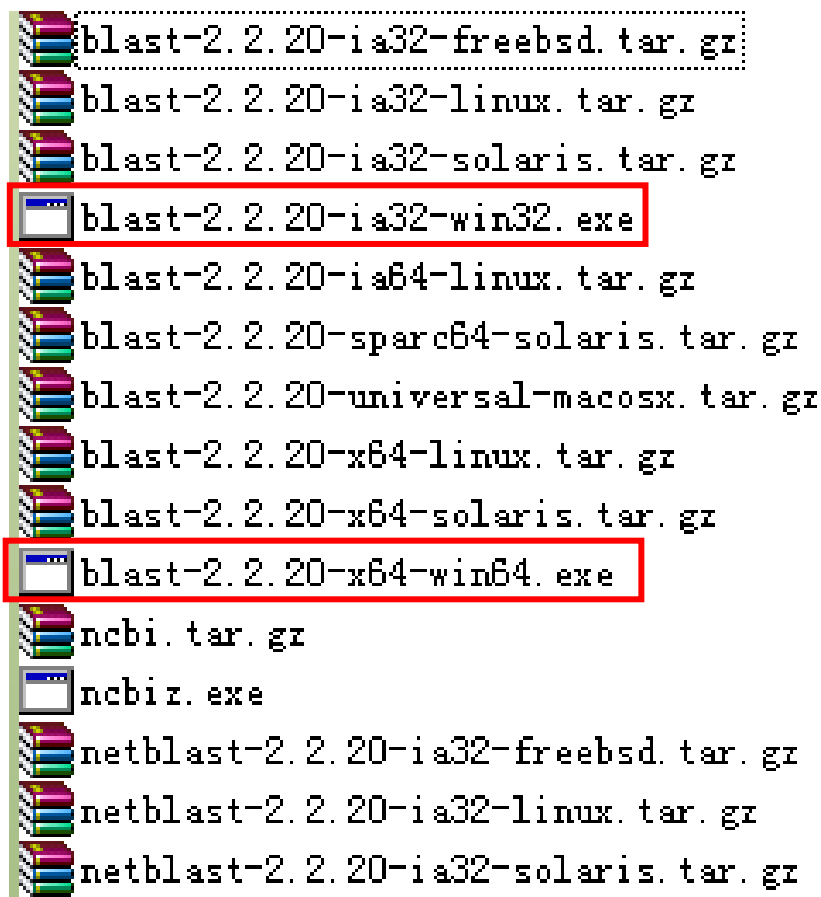


Windows系统下BLAST程序的安裝



BLAST程序的的下载：根据电脑系统的不同选择

blast-2.2.20-ia32-win32.exe 或 *blast-2.2.20-ia32-win64.exe*

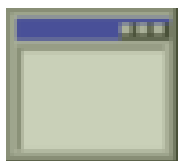


win32 - windows 32位操作系统； win64 - windows 64位操作系统，目前多数是win32的

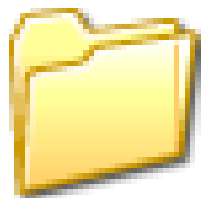
Windows系统下BLAST程序的安装



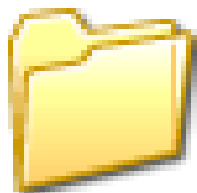
BLAST程序的的安装：解压缩到某个文件夹
自动生成三个文件夹—bin、data和doc



```
blast-2.2.20-ia3...
```



bin



data



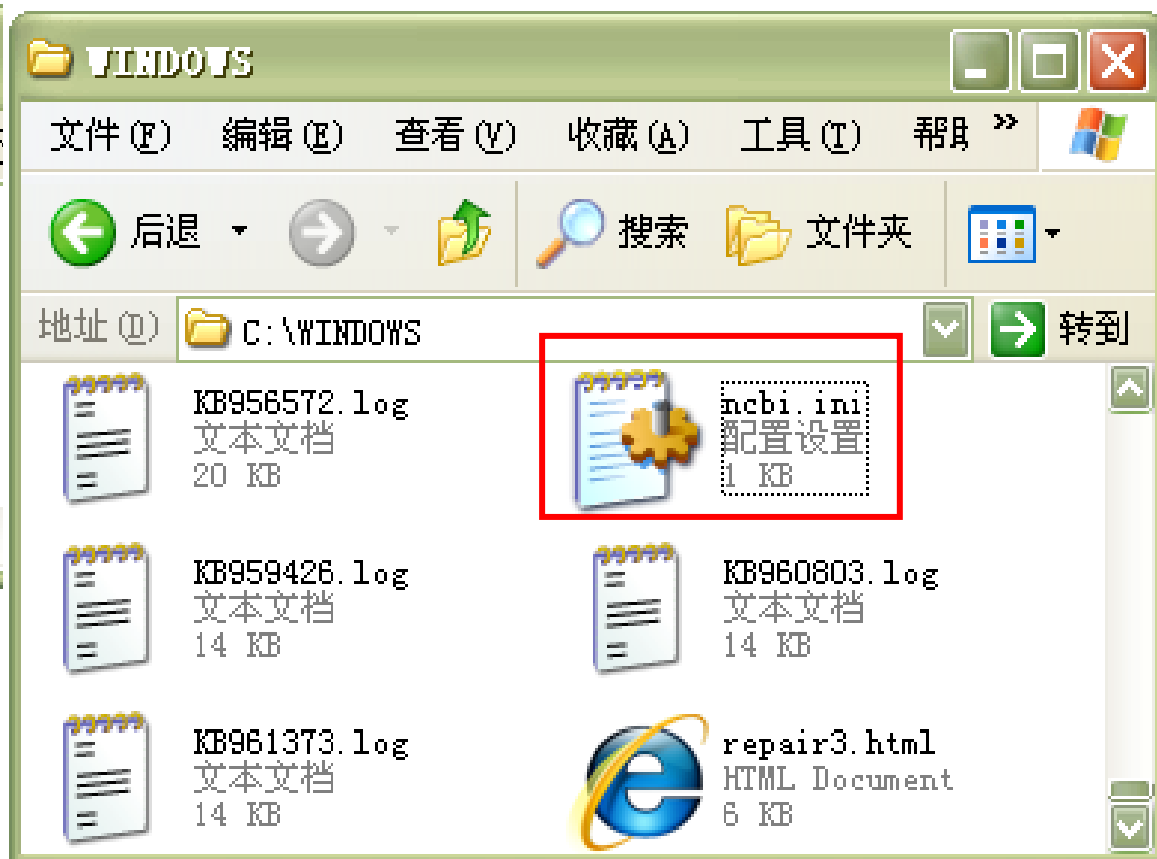
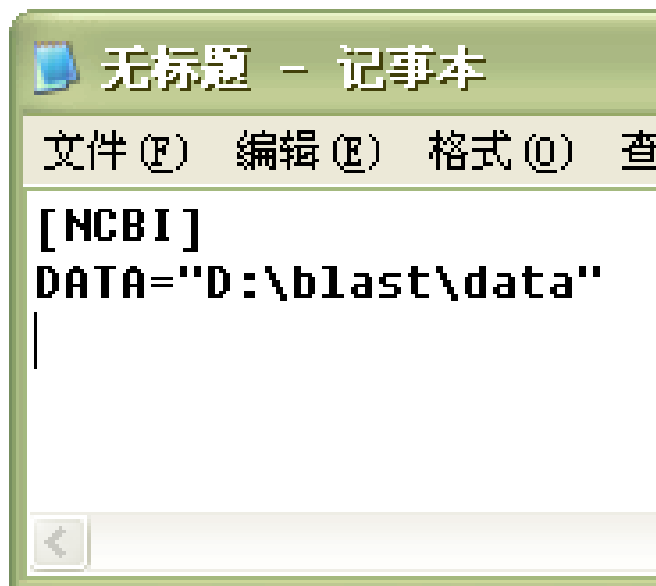
doc

Windows系统下BLAST程序的安装



[NCBI]

DATA="D:\blast\data" → ^{保存为} ncbi.ini → ^{保存位置} C:\WINDOWS



Windows 系统下BLAST程序的运行

A screenshot of a Windows XP Command Prompt window. The title bar reads "命令提示符" (Command Prompt). The window content shows the following text:

```
Microsoft Windows XP [版本 5.1.2600]
(C) 版权所有 1985-2001 Microsoft Corp.

C:\Documents and Settings\Administrator>d:

D:\>cd blast\data

D:\blast\data>
```


BLAST程序的参数



1、 -p Program Name [String]

该参数p代表的是“program”，用来选择程序。其包含五个选项：
blastp、blastn、blastx、tblastn和tblastx。

- (1) -p blastp: 用蛋白质序列搜索蛋白质序列库
- (2) -p blastn: 用核酸序列搜索核酸库
- (3) -p blastx: 核酸序列对蛋白质库的比对，核酸序列在比对之前自动按照六个读码框翻译成蛋白质序列
- (4) -p tblastn: 蛋白质序列对核酸库的比对，核酸库中的序列按照六个读码框翻译后与蛋白质序列进行比对搜索
- (5) -p tblastx: 核酸序列对核酸库在蛋白质质级别的比对，两者都在搜索之前翻译成为蛋白质质进行比对

BLAST程序的参数



2、 **-i Query File [File In] (default = stdin)**

用于搜索数据库的查询 (query) 文件，默认的是名为stdin的文件

3、 **-d Database name [String]**

选择待搜索的数据库，可以选择多个数据库

例如： -d “nr.fasta est.fasta”

4、 **-o output file name [File Out] (default = stdout)**

输出文件的名称，默认值为stdout

BLAST程序的参数

5、 -e Expected value [Real] (default = 10.0)

期望值，这一参数控制搜索的灵敏度 (search sensitivity)，可以输入整数 (如100)，分数 (如1/100)，小数 (如0.001) 或是指数 (如5e-5)，默认值是10.0。

例：分别使用e值为0.1和0.01的两种情况下，用拟南芥的AP3基因的核酸序列在水稻MADS-box基因核酸序列库中搜索同源基因。

Sequences producing significant alignments:	Score (bits)	E Value
OsMADS2	44	2e-006
OsMADS4	38	1e-004
OsMADS29	38	1e-004
OsMADS63	34	0.002
OsMADS62	34	0.002
OsMADS32	34	0.002
OsMADS23	32	0.007
OsMADS31	30	0.029
OsMADS50	30	0.029

9 hits

Sequences producing significant alignments:	Score (bits)	E Value
OsMADS2	44	2e-006
OsMADS4	38	1e-004
OsMADS29	38	1e-004
OsMADS63	34	0.002
OsMADS62	34	0.002
OsMADS32	34	0.002
OsMADS23	32	0.007

7 hits



BLAST程序的参数

6、 -m Specifies alignment view (default = 0)

设定搜索结果的显示格式，m参数的选项有12个

-m 0: 默认参数，显示query和subject两两比对的信息

```
Sequences producing significant alignments:                                Score   E
                                                                           (bits) Value

OsMADS6                                                                    1487   0.0
OsMADS17                                                                    262   5e-072
OsMADS45                                                                    192   4e-051

>OsMADS6
    Length = 750

    Score = 1487 bits (750), Expect = 0.0
    Identities = 750/750 (100%)
    Strand = Plus / Plus

Query: 1   atggggaggggaagagttgagctgaagcgcacgagaacaagatcaacaggcaggtcacc 60
          |||
Sbjct: 1   atggggaggggaagagttgagctgaagcgcacgagaacaagatcaacaggcaggtcacc 60

Query: 61  ttctccaagcgccgcaacggcctcctcaagaaggcctacgagctgtccgttctctcgac 120
          |||
Sbjct: 61  ttctccaagcgccgcaacggcctcctcaagaaggcctacgagctgtccgttctctcgac 120

Query: 121  cccgcctccgctctctctctctctctctctctctctctctctctctctctctctctct 180
```



BLAST程序的参数

-m 1: 显示query在所有subjects上的定位信息，并显示一致性比对信息，subject之间不同的碱基/氨基酸会被标出

例：用水稻的*OsMADS6*基因的核酸序列在水稻MADS-box基因核酸序列库中搜索同源基因。

命令 “blastall -p blastn -i OsMADS6_D.txt -d OsMADS_D1.txt -o OsMADS6_blastn_Ff_e-47_m1.out -F F -e 1e-47 -m 1”

```
1_0 1 atggggaggggaagagttgagctgaagcgcatcgagaacaagatcaacaggcaggtcacc 60
0 1 ..... 60
1 16 .....g.....c..... 75
25 1 .....tc.g..g.....a.g.....c.....g..g 60

1_0 61 ttctccaagcgccgcaacggcctcctcaagaaggcctacgagctgtccgttctctgcgac 120
0 61 ..... 120
1 77 .....g.....c.....g..... 136
25 61 .....g.....a.g..t....g.....g.....c.....c..... 120

1_0 121 gccgaggtcgcgctcatcatcttctccagccgcggaagctctacgagttcggcagcgcc 180
0 121 ..... 180
1 137 ..g....g.....t.....g....a....t..... 196
25 121 .....c.....a.....t..... 177

1_0 181 ggcataacaaagactttagaaaggtaccaacattggtgctacaatgctcaagattccaac 240
0 181 ..... 240
1 197 ..a.... 203
```

BLAST程序的参数

-m 2: 显示query在所有subjects上的定位信息但是不显示一致性比对信息，subject之间不同的碱基/氨基酸会被标出

例：用水稻的*OsMADS6*基因的核酸序列在水稻MADS-box基因核酸序列库中搜索同源基因。

命令 “blastall -p blastn -i OsMADS6_D.txt -d OsMADS_D1.txt -o OsMADS6_blastn_Ff_e-47_m1.out -F F -e 1e-47 -m 2”

```

1_0 1   atggggaggggaagagttgagctgaagcgcacgagaacaagatcaacaggcaggtcacc 60
0      1   atggggaggggaagagttgagctgaagcgcacgagaacaagatcaacaggcaggtcacc 60
1     16   atggggaggggaaggggtgagctgaagcgcacgagaacaagatcaaccggcaggtcacc 75
25    1   atggggaggggtcgggtggagctgaagaggatcgagaacaagatcaaccggcaggtgacg 60

1_0 61   ttctccaagcgcgcgaacggcctcctcaagaaggcctacgagctgtccgttctctgcgac 120
0     61   ttctccaagcgcgcgaacggcctcctcaagaaggcctacgagctgtccgttctctgcgac 120
1     77   ttctccaagcgcgcgaacggcctcctcaagaaggcgtacgagctctccgtgctctgcgac 136
      \
      |
      g

25   61   ttcgccaagcgcagggaatggcctgctcaagaaggcgtacgagctctccgtcctctgcgac 120

1_0 121  gccgaggtcgcgctcatcatcttctccagccgcggaagctctacgagttcggcagcgcc 180
0    121  gccgaggtcgcgctcatcatcttctccagccgcggaagctctacgagttcggcagcgcc 180
1    137  gcggaggtggcgctcatcatcttctccagccgtggcaagctgtacgaattcggtagcgcc 196
25   121  gccgaggtcgccctcatcatcttctccaaccgcggaagctctacgagttctgcagc 177

1_0 181  ggcataacaaagactttagaaaggtaccaacattggtgctacaatgctcaagattccaac 240
0    181  ggcataacaaagactttagaaaggtaccaacattggtgctacaatgctcaagattccaac 240
1    197  ggaataa 203
  
```



BLAST程序的参数

-m 3: 显示query在所有subjects的定位和一致性比对信息，不显示subjects之间的差异

例：用水稻的*OsMADS6*基因的核酸序列在水稻MADS-box基因核酸序列库中搜索同源基因。

命令 “blastall -p blastn -i OsMADS6_D.txt -d OsMADS_D1.txt -o OsMADS6_blastn_Ff_e-47_m1.out -F F -e 1e-47 -m 3”

```
1_0 1   atggggaggggaagagttgagctgaagcgcatcgagaacaagatcaacaggcaggtcacc 60
0   1   ..... 60
1   16  .....g.....c..... 75
25  1   .....tc.g.g.....a.g.....c.....g..g 60

1_0 61  -ttctccaagcgccgcaacggcctcctcaagaaggcctacgagctgtccgttctctgcca 119
0   61  -..... 119
1   76  g.....g.....c.....g..... 135
25  61  -..g.....a.g..t....g.....g.....c.....c..... 119

1_0 120 cgccgaggtcgcgctcatcatcttctccagccgcggaagctctacgagttcgggcagcgc 179
0   120 ..... 179
1   136 ...g.....g.....t.....g.....a....t..... 195
25  120 .....c.....a.....t..... 177

1_0 180 cggcataaacaagactttagaaaggtaccaacattggtgctacaatgctcaagattccaa 239
0   180 ..... 239
1   196 ...a.... 203
```



BLAST程序的参数

-m 4: 显示query在所有subjects上的定位信息但是不显示一致性比对信息，不显示subjects之间的差异

例：用水稻的*OsMADS6*基因的核酸序列在水稻MADS-box基因核酸序列库中搜索同源基因。

命令 “blastall -p blastn -i OsMADS6_D.txt -d OsMADS_D1.txt -o OsMADS6_blastn_Ff_e-47_m1.out -F F -e 1e-47 -m 4”

```
1_0 1 atggggaggggaagagttgagctgaagcgcatcgagaacaagatcaacaggcaggtcacc 60
0 1 atggggaggggaagagttgagctgaagcgcatcgagaacaagatcaacaggcaggtcacc 60
1 16 atggggaggggaagggttgagctgaagcgcatcgagaacaagatcaaccggcaggtcacc 75
25 1 atggggaggggtcgggtggagctgaagaggatcgagaacaagatcaaccggcaggtgacg 60

1_0 61 -ttctccaagcgccgcaacggcctcctcaagaaggcctacgagctgtccgttctctgcga 119
0 61 -ttctccaagcgccgcaacggcctcctcaagaaggcctacgagctgtccgttctctgcga 119
1 76 gttctccaagcgccgcaacggcctcctcaagaaggcgtacgagctctccgtgctctgcga 135
25 61 -ttcgccaagcgccaggaatggcctgctcaagaaggcgtacgagctctccgtcctctgcga 119

1_0 120 cgccgaggtcgcgctcatcatcttctccagccgcggaagctctacgagttcggtcagcgc 179
0 120 cgccgaggtcgcgctcatcatcttctccagccgcggaagctctacgagttcggtcagcgc 179
1 136 cgccgaggtggcgctcatcatcttctccagccggtggcaagctgtacgaattcggtagcgc 195
25 120 cgccgaggtcgccctcatcatcttctccaaccgcggaagctctacgagttctgcagc 177

1_0 180 cggcataacaaagactttagaaaggtaccaacattgttgctacaatgctcaagattccaa 239
0 180 cggcataacaaagactttagaaaggtaccaacattgttgctacaatgctcaagattccaa 239
1 196 cggaataa 203
```




BLAST程序的参数

-m 5: 显示query在所有subjects上的定位信息但是不显示每个碱基/氨基酸的比对信息，补充“-”比对区域，subjects之间不同的碱基/氨基酸会被标出

例：用水稻的*OsMADS6*基因的核酸序列在水稻MADS-box基因核酸序列库中搜索同源基因。

命令 “blastall -p blastn -i OsMADS6_D.txt -d OsMADS_D1.txt -o OsMADS6_blastn_Ff_e-47_m1.out -F F -e 1e-47 -m 5”

```
1_0 1 atggggaggggaagagttgagctgaagcgcatcgagaacaagatcaacaggcaggtcacc 60
0 1 atggggaggggaagagttgagctgaagcgcatcgagaacaagatcaacaggcaggtcacc 60
1 16 atggggaggggaagggttgagctgaagcgcatcgagaacaagatcaaccggcaggtcacc 75
25 1 atggggaggggtcgggtgagctgaagaggatcgagaacaagatcaaccggcaggtgacg 60

1_0 61 ttctccaagcgccgcaacggcctcctcaagaaggcctacgagctgtccgttctctgcgac 120
0 61 ttctccaagcgccgcaacggcctcctcaagaaggcctacgagctgtccgttctctgcgac 120
1 77 ttctccaagcgccgcaacggcctcctcaagaaggcgtacgagctctccgtgctctgcgac 136
25 61 ttctccaagcgccgcaacggcctcctcaagaaggcgtacgagctctccgtcctctgcgac 120

1_0 121 gccgaggtcgcgctcatcatcttctccagccgcggaagctctacgagttcggcagcgcc 180
0 121 gccgaggtcgcgctcatcatcttctccagccgcggaagctctacgagttcggcagcgcc 180
1 137 gccgaggtggcgctcatcatcttctccagccggtggcaagctgtacgaattcggtagcgcc 196
25 121 gccgaggtcgccctcatcatcttctccaaccgcggaagctctacgagttctgcagc--- 177

1_0 181 ggcataaacaagactttagaaaggtaccaacattggtgctacaatgctcaagattccaac 240
0 181 ggcataaacaagactttagaaaggtaccaacattggtgctacaatgctcaagattccaac 240
1 197 ggaataa----- 203
25 176 ----- 177
```



BLAST程序的参数

-m 6: 显示query在所有subjects上的定位信息但是不显示每个碱基/氨基酸的比对信息，补充“-”对齐比对区域，不显示subjects之间的差异

例：用水稻的*OsMADS6*基因的核酸序列在水稻MADS-box基因核酸序列库中搜索同源基因。

命令 “blastall -p blastn -i OsMADS6_D.txt -d OsMADS_D1.txt -o OsMADS6_blastn_Ff_e-47_m1.out -F F -e 1e-47 -m 6”

```
1_0 1   atggggaggggaagagttgagctgaagcgcatcgagaacaagatcaacaggcaggtcacc 60
0      1   atggggaggggaagagttgagctgaagcgcatcgagaacaagatcaacaggcaggtcacc 60
1     16   atggggaggggaagggttgagctgaagcgcatcgagaacaagatcaaccggcaggtcacc 75
25    1   atggggaggggtcgggtggagctgaagaggatcgagaacaagatcaaccggcaggtgacg 60

1_0 61   -ttctccaagcgccgcaacggcctcctcaagaaggcctacgagctgtccgttctctgcga 119
0      61   -ttctccaagcgccgcaacggcctcctcaagaaggcctacgagctgtccgttctctgcga 119
1     76   gttctccaagcgccgcaacggcctcctcaagaaggcgtacgagctctccgtgctctgcga 135
25    61   -ttcgccaagcgccaggaatggcctgctcaagaaggcgtacgagctctccgtcctctgcga 119

1_0 120  cgccgaggtcgcgctcatcatcttctccagccgcggaagctctacgagttcggcagcgc 179
0      120  cgccgaggtcgcgctcatcatcttctccagccgcggaagctctacgagttcggcagcgc 179
1     136  cgccgaggtggcgctcatcatcttctccagccgtggcaagctgtacgaattcggtagcgc 195
25    120  cgccgaggtcgccctcatcatcttctccaaccgcggaagctctacgagttctgcagc-- 177

1_0 180  cggcataacaaagactttagaaaggtaccaacattgttgctacaatgctcaagattccaa 239
0      180  cggcataacaaagactttagaaaggtaccaacattgttgctacaatgctcaagattccaa 239
1     196  cggaataa----- 203
25    176  ----- 177
```

BLAST程序的参数



-m 7: 输出XML格式的blast结果

例：用水稻的*OsMADS6*基因的核酸序列在水稻MADS-box基因核酸序列库中搜索同源基因。

命令 “blastall -p blastn -i OsMADS6_D.txt -d OsMADS_D1.txt -o OsMADS6_blastn_Ff_e-47_m1.out -F F -e 1e-47 -m 7”

```
ersion="1.0"??>
  BlastOutput PUBLIC "-//NCBI//NCBI BlastOutput/EN" "http://www.ncbi.nlm.nih.gov/dtd/NCBI_BlastOutput.dtd">
  put>
  utput_program>blastn</BlastOutput_program>
  utput_version>blastn 2.2.20 [Feb-08-2009]</BlastOutput_version>
  utput_reference>~Reference: Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schaffer, ~Jinghui Zhang, Zheng Zhang, Webb Miller, and David
  utput_db>OsMADS_D1.txt</BlastOutput_db>
  utput_query-ID>lcl|1_0</BlastOutput_query-ID>
  utput_query-def>OsMADS6</BlastOutput_query-def>
  utput_query-len>753</BlastOutput_query-len>
  utput_param>
  meters>
  rameters_expect>1e-047</Parameters_expect>
  rameters_sc-match>1</Parameters_sc-match>
  rameters_sc-mismatch>-3</Parameters_sc-mismatch>
  rameters_gap-open>5</Parameters_gap-open>
  rameters_gap-extend>2</Parameters_gap-extend>
  rameters_filter>F</Parameters_filter>
  ameters>
  Output_param>
  utput_iterations>
  ation>
  eration_iter-num>1</Iteration_iter-num>
  eration_query-ID>lcl|1_0</Iteration_query-ID>
  eration_query-def>OsMADS6</Iteration_query-def>
  eration_query-len>753</Iteration_query-len>
  eration_hits>
```



BLAST程序的参数

-m 8: 用列表格式显示比对结果。从左到右各列的意义依次是: query名, subject名, 一致性百分数, 比对长度, 错配数, 空位数, query比对起始位点和终止位点, subject比对起始位点和终止位点, 期望值, 比对得分

例: 用水稻的*OsMADS6*基因的核酸序列在水稻MADS-box基因核酸序列库中搜索同源基因。

命令 “blastall -p blastn -i OsMADS6_D.txt -d OsMADS_D1.txt -o OsMADS6_blastn_Ff_e-47_m1.out -F F -e 1e-47 -m 8”

OsMADS6	OsMADS6	100.00	750	0	0	1	750	1	750	0.0	1487
OsMADS6	OsMADS17	93.09	188	12	1	1	187	16	203	5e-072	262
OsMADS6	OsMADS45	88.70	177	20	0	1	177	1	177	4e-051	192

-m 9: 用带有注释行的列表格式显示比对结果, 格式与-m 8一样, 只是在每个query的必读结果前面加了注释行用于说明列表中各列的意义

例: 用水稻的*OsMADS6*基因的核酸序列在水稻MADS-box基因核酸序列库中搜索同源基因。

命令 “blastall -p blastn -i OsMADS6_D.txt -d OsMADS_D1.txt -o OsMADS6_blastn_Ff_e-47_m1.out -F F -e 1e-47 -m 9”

```
# BLASTN 2.2.20 [Feb-08-2009]
# Query: OsMADS6
# Database: OsMADS_D1.txt
# Fields: Query id, Subject id, % identity, alignment length, mismatches, gap openings, q. start, q. end, s. start, s. end, e-value, bit score
OsMADS6 OsMADS6 100.00 750 0 0 1 750 1 750 0.0 1487
OsMADS6 OsMADS17 93.09 188 12 1 1 187 16 203 5e-072 262
OsMADS6 OsMADS45 88.70 177 20 0 1 177 1 177 4e-051 192
```

BLAST程序的参数

-m 10: 输出文件为ASN格式的文本文件

例：用水稻的*OsMADS6*基因的核酸序列在水稻MADS-box基因核酸序列库中搜索同源基因。

命令 “blastall -p blastn -i OsMADS6_D.txt -d OsMADS_D1.txt -o OsMADS6_blastn_Ff_e-47_m1.out -F F -e 1e-47 -m 10”

```
Seq-annot ::= {
  desc {
    user {
      type
      str "Hist Seqalign" ,
      data {
        {
          label
          str "Hist Seqalign" ,
          data
          bool TRUE } } } ,
    user {
      type
      str "Blast Type" ,
      data {
        {
          label
          str "BLASTN" ,
          data
          int 1 } } } } ,
  data
```

-m 11: 输出文件为二进制文件

BLAST程序的参数



7、-F Filter query sequence (DUST with blastn, SEG with others) [String] (default = T)

用来屏蔽简单重复和低复杂度序列的参数，有T和F两个选项，选择“T”，则程序在比对过程中会屏蔽掉query序列中的简单重复和低复杂度序列；选择“F”则不会屏蔽。默认值是“T”

8、-I Show GI's in deflines [T/F] default = F

提示行显示GI number, 默认不显示

例：

-I F (default): ref|NP_001005339.1| Regulator of G-protein ...

-I T: [gi/52694755](#)/ref|NP_001005339.1| Regulator of G-protein ...



BLAST程序的参数

9、-G Cost to open a gap (zero invokes default behavior) [Integer] (default = 0)

起始空位罚分

例：分别设定G为12和10用拟南芥的AP3基因的蛋白质序列在水稻MADS-box基因蛋白质序列库中搜索同源基因

命令：

- 1) “blastall -p blastp -i AtAP3_P.txt -d OsMADS_P.txt -o OsAP3_blastp_5e-30_G12.out -e 5e-30 -G 12”
- 2) “blastall -p blastp -i AtAP3_P.txt -d OsMADS_P.txt -o OsAP3_blastp_5e-30_G12.out -e 5e-30 -G 10”

Sequences producing significant alignments:	Score (bits)	E Value
OsMADS16	203	9e-056
OsMADS29	132	2e-034
OsMADS6	131	3e-034
OsMADS2	126	1e-032
OsMADS4	123	9e-032
OsMADS18	122	2e-031
OsMADS50	121	4e-031
OsMADS1	121	4e-031
OsMADS45	120	8e-031
OsMADS17	118	3e-030

10 hits

Sequences producing significant alignments:	Score (bits)	E Value
OsMADS16	177	5e-048
OsMADS29	117	3e-030

2 hits

BLAST程序的参数

10、-E Cost to extend a gap [Integer] (default = 0)

空位延伸罚分

例：分别设定E为1和2用拟南芥的AP3基因的蛋白质序列在水稻MADS-box基因蛋白质序列库中搜索同源基因

命令：

- 1) “blastall -p blastp -i AtAP3_P.txt -d OsMADS_P.txt -o OsAP3_blastp_5e-30_G12.out -e 5e-30 -E 1”
- 2) “blastall -p blastp -i AtAP3_P.txt -d OsMADS_P.txt -o OsAP3_blastp_5e-30_G12.out -e 5e-30 -E 2”

Sequences producing significant alignments:	Score	E
	(bits)	Value
OsMADS16	192	1e-052
OsMADS29	126	9e-033
OsMADS6	126	9e-033
OsMADS2	120	8e-031
OsMADS18	117	4e-030

5 hits

Sequences producing significant alignments:	Score	E
	(bits)	Value
OsMADS16	212	1e-058
OsMADS29	133	7e-035
OsMADS2	132	2e-034
OsMADS4	128	2e-033
OsMADS6	128	3e-033
OsMADS1	124	5e-032
OsMADS45	122	1e-031
OsMADS18	120	7e-031
OsMADS17	119	1e-030
OsMADS27	118	2e-030
OsMADS50	118	3e-030

11 hits

BLAST程序的参数

11、-q Penalty for a nucleotide mismatch (blastn only) [Integer] (default = -3)

在用blastn程序搜索序列时，设定核酸错配的罚分，默认值是-3

例：分别设定q为-4和-2用拟南芥的AP3基因的核酸序列在水稻MADS-box基因核酸序列库中搜索同源基因

命令：

1) “blastall -p blastn -i AtAP3_D.txt -d OsMADS_D.txt -o OsAP3_blastn_e0.1_q-4_Ff.out -e 0.1 -F F -q -4”

2) “blastall -p blastn -i AtAP3_D.txt -d OsMADS_D.txt -o OsAP3_blastn_e0.1_q-2_Ff.out -e 0.1 -F F -q -2”

Sequences producing significant alignments:	Score (bits)	E Value
OsMADS2	40	3e-005
OsMADS4	34	0.002
OsMADS29	34	0.002
OsMADS63	32	0.007
OsMADS32	32	0.007

5 hits

Sequences producing significant alignments:	Score (bits)	E Value
OsMADS15	53	5e-009
OsMADS29	51	2e-008
OsMADS62	49	8e-008
OsMADS4	49	8e-008
OsMADS2	47	3e-007
OsMADS23	45	1e-006
OsMADS63	39	6e-005
OsMADS50	37	2e-004
OsMADS32	35	9e-004
OsMADS25	35	9e-004
OsMADS31	31	0.012
OsMADS22	30	0.047

12 hits

BLAST程序的参数

12、-r Reward for a nucleotide match (blastn only) [Integer] (default = 1)

在用blastn程序搜索序列时，比对上的每个核酸位点的得分，默认值是1

例：分别设定r为2和3用拟南芥的AP3基因的核酸序列在水稻MADS-box基因核酸序列库中搜索同源基因

命令：

- 1) “blastall -p blastn -i AtAP3_D.txt -d OsMADS_D.txt -o OsAP3_blastn_e0.1_r2_Ff.out -e 0.1 -F F -r 2”
- 2) “blastall -p blastn -i AtAP3_D.txt -d OsMADS_D.txt -o OsAP3_blastn_e0.1_r2_Ff.out -e 0.1 -F F -r 3”

Sequences producing significant alignments:	Score (bits)	E Value
OsMADS16	77	2e-016
OsMADS15	73	3e-015
OsMADS62	63	5e-012
OsMADS2	61	2e-011
OsMADS29	61	2e-011
OsMADS4	57	2e-010
OsMADS63	55	8e-010
OsMADS23	55	8e-010
OsMADS31	55	8e-010
OsMADS25	45	1e-006
OsMADS50	43	5e-006
OsMADS32	37	2e-004

12 hits

Sequences producing significant alignments:	Score (bits)	E Value
OsMADS16	164	8e-043
OsMADS29	109	4e-026
OsMADS2	100	3e-023
OsMADS15	97	3e-022
OsMADS4	97	3e-022
OsMADS31	92	7e-021
OsMADS23	90	2e-020
OsMADS62	82	5e-018
OsMADS50	78	1e-016
OsMADS63	73	4e-015
OsMADS55	67	3e-013
OsMADS25	59	8e-011
OsMADS32	57	2e-010
OsMADS22	55	7e-010

14 hits

BLAST程序的参数

13、-v Number of database sequences to show one-line descriptions for (V) [Integer] (default = 500)

设定输出文件中，匹配列表最多显示多少个subject，默认值为500

例：分别设定v为5和2用拟南芥的AP3基因的核酸序列在水稻MADS-box基因核酸序列库中搜索同源基因

命令：

1) “blastall -p blastn -i AtAP3_D.txt -d OsMADS_D.txt -o OsAP3_blastn_e0.1_v5_Ff.out -e 0.1 -F F -v 5”

2) “blastall -p blastn -i AtAP3_D.txt -d OsMADS_D.txt -o OsAP3_blastn_e0.1_v5_Ff.out -e 0.1 -F F -v 2”

Sequences producing significant alignments:	Score (bits)	E Value
OsMADS2	44	2e-006
OsMADS4	38	1e-004
OsMADS29	38	1e-004
OsMADS63	34	0.002
OsMADS62	34	0.002

显示5个

Sequences producing significant alignments:	Score (bits)	E Value
OsMADS2	44	2e-006
OsMADS4	38	1e-004

显示2个

BLAST程序的参数

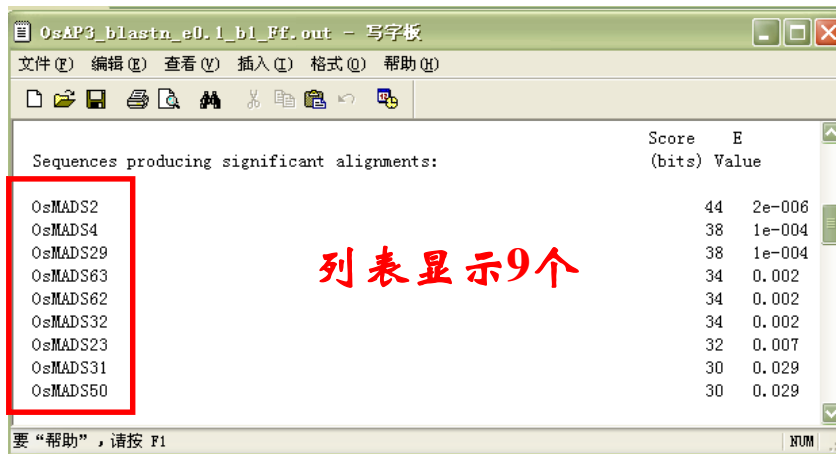
14、-b Number of database sequence to show alignments for (B) [Integer] (default = 250)

设定输出文件中，最多显示多少个query-subject两两比对文本描述，默认值为250

例：分别设定b为1和3用拟南芥的AP3基因的核酸序列在水稻MADS-box基因核酸序列库中搜索同源基因

命令：

1) “blastall -p blastn -i AtAP3_D.txt -d OsMADS_D.txt -o OsAP3_blastn_e0.1_b3_Ff.out -e 0.1 -F F -b 1”

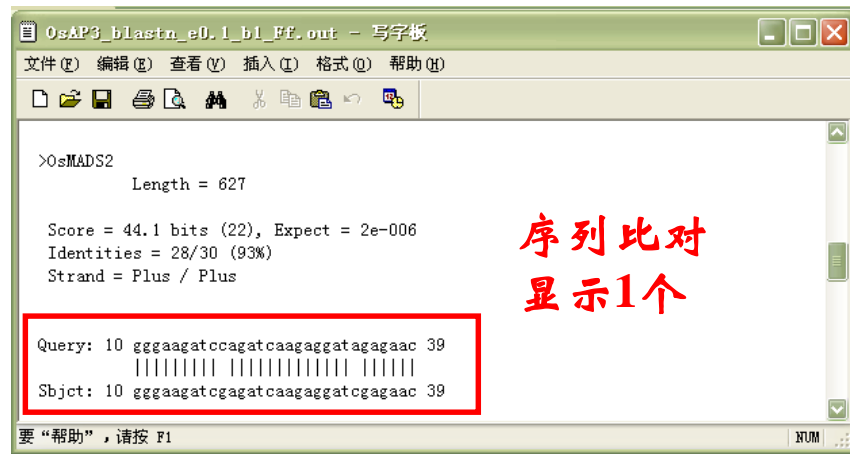


Sequences producing significant alignments:

	Score	E
	(bits)	Value
OsMADS2	44	2e-006
OsMADS4	38	1e-004
OsMADS29	38	1e-004
OsMADS63	34	0.002
OsMADS62	34	0.002
OsMADS32	34	0.002
OsMADS23	32	0.007
OsMADS31	30	0.029
OsMADS50	30	0.029

列表显示9个

A



```

>OsMADS2
    Length = 627

Score = 44.1 bits (22), Expect = 2e-006
Identities = 28/30 (93%)
Strand = Plus / Plus

Query: 10 gggaagatccagatcaagaggatagagaac 39
      |||
Sbjct: 10 gggaagatcgagatcaagaggatcgagaac 39
  
```

序列比对显示1个

B



BLAST程序的参数

15、-f Threshold for extending word hits, blastp 11, blastx 12, tblastn 13, tblastx 13, [Integer] default = 0

设定延伸字长 (word) 的阈值

例：分别设定f为13和14用拟南芥的AP3基因的核酸序列在水稻MADS-box基因核酸序列库中搜索同源基因

命令：

1) “blastall -p blastn -i AtAP3_D.txt -o OsAP3_blastn_Ff_W13.out -d OsMADS_D.txt -F F -W 13”

2) “blastall -p blastn -i AtAP3_D.txt -o OsAP3_blastn_Ff_W13.out -d OsMADS_D.txt -F F -W 14”

Sequences producing significant alignments:	Score (bits)	E Value	N
OsMADS16	220	3e-060	1
OsMADS2	144	2e-037	1
OsMADS4	138	2e-035	1
OsMADS29	121	5e-035	2
OsMADS14	100	7e-034	2
OsMADS15	103	1e-033	2
OsMADS6	100	2e-033	2
OsMADS45	102	4e-033	2
OsMADS50	99	5e-033	4
OsMADS17	99	3e-032	2
OsMADS1	93	7e-032	2

11 hits

Sequences producing significant alignments:	Score (bits)	E Value	N
OsMADS16	220	3e-060	1
OsMADS2	144	2e-037	1
OsMADS4	138	2e-035	1

3 hits

BLAST程序的参数

16、-g Perform gapped alignment (not available with tblastx) [T/F] (default = T)

执行空位比对 (tblastx程序不能使用此参数), 默认值为T

例: 分别设定g为T和F用拟南芥的AP3基因的蛋白质序列在水稻MADS-box基因蛋白质序列库中搜索同源基因

命令:

1) “blastall -p blastp -i AtAP3_P.txt -d OsMADS_P.txt -o OsAP3_blastp_5e-30__Ff.out -e 5e-30 -F F -g T”

2) “blastall -p blastp -i AtAP3_P.txt -d OsMADS_P.txt -o OsAP3_blastp_5e-30__Ff.out -e 5e-30 -F F -g F”

Sequences producing significant alignments:	Score (bits)	E Value
OsMADS16	192	1e-052
OsMADS29	126	9e-033
OsMADS6	126	9e-033
OsMADS2	120	8e-031
OsMADS18	117	4e-030

5hits

Sequences producing significant alignments:	Score (bits)	E Value	N
OsMADS16	220	1e-060	1
OsMADS2	144	9e-038	1
OsMADS4	138	7e-036	1
OsMADS14	100	3e-034	2
OsMADS15	103	6e-034	2
OsMADS6	100	7e-034	2
OsMADS45	102	2e-033	2
OsMADS50	99	4e-033	3
OsMADS17	99	1e-032	2
OsMADS23	98	3e-031	2
OsMADS21	102	6e-031	2
OsMADS29	121	9e-031	1
OsMADS24	100	3e-030	2

13hits

BLAST程序的参数



17、 -Q Genetic code for translation of the query sequence [Integer] (default = 0)

运用blastx和tblastx程序时，设定进行翻译query序列的遗传密码类型，默认值是0

18、 -D Genetic code for translation of the database [Integer] (default = 0)

运用tblastx和tblastn程序时，设定用于翻译数据库中核酸序列的遗传密码类型，默认值是0

BLAST程序的参数



19、-a Number of processors to use [Integer] (default = 1)

设定搜索过程中要用到的处理器的数目，取决于可使用的CPU或是处理器的数目，默认值为1

20、-J Believe the query defline, [T/F] default = F

显示query的defline

21、-O Saves SeqAlign output to specified file [File out]

输出SeqAlign文件，该文件可以利用*blast_demo*程序将搜索结果转换为多种格式

BLAST程序的参数

22、-M Matrix [String] (default = BLOSUM62)

指定蛋白质比对的打分矩阵，默认为BLOSUM62

例：分别设定M为BLOSUM80和PAM30用拟南芥的AP3基因的蛋白质序列在水稻MADS-box基因蛋白质序列库中搜索同源基因

命令：

- 1) “blastall -p blastp -i AtAP3_P.txt -d OsMADS_P.txt -o OsAP3_blastp_5e-30_Ff_blosum80.out -e 5e-30 -F F -M BLOSUM80”
- 2) “blastall -p blastp -i AtAP3_P.txt -d OsMADS_P.txt -o OsAP3_blastp_5e-30_Ff_blosum80.out -e 5e-30 -F F -M PAM30”

Sequences producing significant alignments:	Score (bits)	E Value
OsMADS16	201	4e-055
OsMADS29	132	2e-034
OsMADS6	130	9e-034
OsMADS2	124	4e-032
OsMADS1	122	2e-031
OsMADS18	122	2e-031
OsMADS4	121	4e-031
OsMADS50	121	5e-031
OsMADS45	119	1e-030

9 hits

Sequences producing significant alignments:	Score (bits)	E Value
OsMADS16	200	8e-055
OsMADS29	130	1e-033
OsMADS18	121	5e-031
OsMADS50	121	5e-031

4 hits



BLAST程序的参数

23、-W Word size (blastn 11, megablast 28, all others 3) [Integer] (default = 0)

设定字长大小，默认值为0，blastn为11，megablast为28，其他为3

例：分别设定W为13和14用拟南芥的AP3基因的核酸序列在水稻MADS-box基因的核酸序列库中搜索同源基因

命令：

1) “blastall -p blastn -i AtAP3_D.txt -o OsAP3_blastn_Ff_W13.out -d OsMADS_D.txt -F F -W 13”

2) “blastall -p blastn -i AtAP3_D.txt -o OsAP3_blastn_Ff_W14.out -d OsMADS_D.txt -F F -W 14”

	Score	E
Sequences producing significant alignments: (bits) Value		
OsMADS2	44	2e-006
OsMADS4	38	1e-004
OsMADS29	38	1e-004
OsMADS63	34	0.002
OsMADS32	34	0.002
OsMADS22	28	0.11
OsMADS90	26	0.45

7 hits

	Score	E
Sequences producing significant alignments: (bits) Value		
OsMADS63	34	0.002
OsMADS32	34	0.002

2 hits

BLAST程序的参数

24、-z Effective length of the database (use zero for the real size) [Real] (default = 0)

指定数据库的有效长度，默认值0代表数据库的实际长度；当输入数值时，BLAST将根据该数值计算E值

例：分别设定z为默认值和5000用拟南芥AP3基因的蛋白质序列在水稻MADS-box基因蛋白质序列库中搜索同源基因

命令：

1) “blastall -p blastp -i AtAP3_P.txt -o OsAP3_blastp_5e-30_Ff_z5000.out -d OsMADS_P.txt -e 5e-30 -F F”

2) “blastall -p blastp -i AtAP3_P.txt -o OsAP3_blastp_5e-30_Ff_z5000.out -d OsMADS_P.txt -e 5e-30 -F F -z 5000”

Sequences producing significant alignments:	Score (bits)	E Value
OsMADS16	192	1e-052
OsMADS29	126	9e-033
OsMADS6	126	9e-033
OsMADS2	120	8e-031
OsMADS18	117	4e-030

5 hits

Sequences producing significant alignments:	Score (bits)	E Value
OsMADS16	192	8e-053
OsMADS29	126	7e-033
OsMADS6	126	7e-033
OsMADS2	120	6e-031
OsMADS18	117	3e-030
OsMADS1	117	4e-030
OsMADS4	117	5e-030

7 hits



BLAST程序的参数

25、 -K Number of best hits from a region to keep. Off by default. If used a value of 100 is recommended. Very high values of -v or -b is also suggested [Integer]

指定query的一段区域所需保持最佳匹配的数目

例：分别设定K为100和2，用拟南芥的AP3基因的核酸序列借助tblastx子程序在水稻MADS-box基因核酸序列库中搜索同源基因

命令：

- 1) “blastall -p tblastx -i AtAP3_D.txt -d OsMADS_D.txt -o OsAP3_tblastx_Ff_e-29_K100.out -F F -e 1e-29 -K 100”
- 2) “blastall -p tblastx -i AtAP3_D.txt -d OsMADS_D.txt -o OsAP3_tblastx_Ff_e-29_K100.out -F F -e 1e-29 -K 2”

Sequences producing significant alignments:	Score (bits)	E Value	N
OsMADS16	220	3e-060	1
OsMADS2	144	2e-037	1
OsMADS4	138	2e-035	1
OsMADS14	100	7e-034	2
OsMADS15	103	1e-033	2
OsMADS6	100	2e-033	2
OsMADS50	99	9e-031	2
OsMADS29	121	2e-030	1

8 hits

Sequences producing significant alignments:	Score (bits)	E Value	N
OsMADS16	220	3e-060	1
OsMADS2	144	2e-037	1
OsMADS14	53	7e-034	2
OsMADS15	49	1e-033	2
OsMADS6	53	2e-033	2

5 hits

BLAST程序的参数

26、-P 0 for multiple hit, 1 for single hit (does not apply to blastn) [Integer] (default = 0)

单字长匹配与多字长匹配的转换，默认为多字长匹配，1为单字长匹配

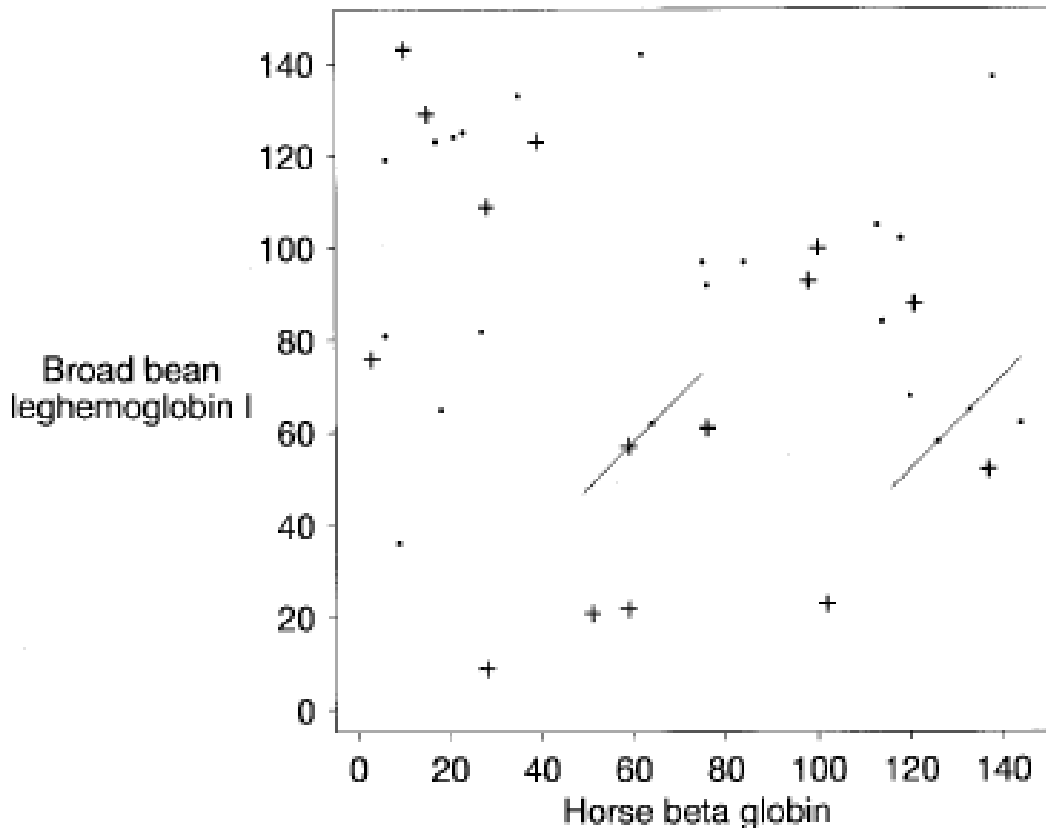


Figure 2. The BLAST comparison of broad bean leghemoglobin I (87) (SWISS-PROT accession no. P02232) and horse b-globin (88) (SWISS-PROT accession no. P02062). The 15 hits with score at least 13 are indicated by plus signs. An additional 22 non-overlapping hits with score at least 11 are indicated by dots. Of these 37 hits, only the two indicated pairs are on the same diagonal and within distance 40 of one another. Thus the two-hit heuristic with $T = 11$ triggers two extensions, in place of the 15 extensions invoked by the one-hit heuristic with $T = 13$. Because this is just one example, the relative numbers of hits and extensions at the various settings of T correspond only roughly to the ratios found in a full database search.

BLAST程序的参数

26、-P 0 for multiple hit, 1 for single hit (does not apply to blastn) [Integer] (default = 0)

单字长匹配与多字长匹配的转换，默认为多字长匹配，1为单字长匹配

例：分别设定P为0和1，用拟南芥AP3基因的核酸序列借助tblastx子程序在水稻MADS-box基因核酸序列库中搜索同源基因

命令：

- 1) “blastall -p tblastx -i AtAP3_D.txt -d OsMADS_D.txt -o OsAP3_tblastx_Ff_e-29_P0.out -F F -e 1e-29 -P 0”
- 2) “blastall -p tblastx -i AtAP3_D.txt -d OsMADS_D.txt -o OsAP3_tblastx_Ff_e-29_P0.out -F F -e 1e-29 -P 1”

Sequences producing significant alignments:	Score	E	
(bits)	Value	N	
OsMADS16	220	3e-060	1
OsMADS2	144	2e-037	1
OsMADS4	138	2e-035	1
OsMADS14	100	7e-034	2
OsMADS15	103	1e-033	2
OsMADS6	100	2e-033	2
OsMADS50	99	9e-031	2
OsMADS29	121	2e-030	1

8 hits

Sequences producing significant alignments:	Score	E	
(bits)	Value	N	
OsMADS16	220	3e-060	1
OsMADS2	144	2e-037	1
OsMADS4	138	2e-035	1
OsMADS14	100	7e-034	2
OsMADS15	103	1e-033	2
OsMADS6	100	2e-033	2
OsMADS45	102	4e-033	2
OsMADS50	99	5e-033	4
OsMADS17	99	3e-032	2
OsMADS1	93	7e-032	2
OsMADS18	98	2e-031	2
OsMADS23	98	7e-031	2
OsMADS21	102	1e-030	2
OsMADS13	102	1e-030	2
OsMADS29	121	2e-030	1
OsMADS24	100	6e-030	2
OsMADS5	93	8e-030	2

17 hits

BLAST程序的参数

27、-Y Effective length of the search space (use zero for the real size) [Real] (default = 0)

搜索空间的有效长度，默认值0代表数据库实际的有效长度，为有效query长度与有效数据库长度之积。

例：分别设定g为Y为 10^6 和 10^4 ，用拟南芥AP3基因的核酸序列借助tblastx子程序在水稻MADS-box基因核酸序列库中搜索同源基因

命令：

- 1) “blastall -p tblastx -i AtAP3_D.txt -d OsMADS_D.txt -o OsAP3_tblastx_Ff_e-29_Y106.out -F F -Y 1000000 -e 1e-29”
- 2) “blastall -p tblastx -i AtAP3_D.txt -d OsMADS_D.txt -o OsAP3_tblastx_Ff_e-29_Y106.out -F F -Y 10000 -e 1e-29”

Sequences producing significant alignments:	Score	E	N
	(bits)	Value	
OsMADS16	220	8e-061	1
OsMADS2	144	6e-038	1
OsMADS4	138	5e-036	1
OsMADS14	100	2e-034	2
OsMADS15	103	4e-034	2
OsMADS6	100	5e-034	2
OsMADS50	99	3e-031	2
OsMADS29	121	6e-031	1

8 hits

Sequences producing significant alignments:	Score	E	N
	(bits)	Value	
OsMADS16	220	8e-063	2
OsMADS2	144	6e-040	1
OsMADS4	138	5e-038	1
OsMADS14	100	2e-036	2
OsMADS15	103	4e-036	2
OsMADS6	100	5e-036	2
OsMADS50	99	3e-033	2
OsMADS29	121	6e-033	1
OsMADS56	95	2e-030	2

9 hits



BLAST程序的参数

28、-S Query strands to search against database (for blast[nx], and tblastx) 3 is both, 1 is top, 2 is bottom [Integer] (default = 3)

指定用query的哪一条核酸链进行搜索，只用于blastn, blastx, and tblastx, 1表示输入的序列, 2表示反向互补序列, 3表示两种都包括

例：分别设定S为1和2用拟南芥的AP3基因的核酸序列在水稻MADS-box基因核酸序列库中搜索同源基因

命令：

1) “blastall -p blastn -i AtAP3_D.txt -o OsAP3_blastn_Ff_S1.out -d OsMADS_D.txt -F F -S 1”

2) “blastall -p blastn -i AtAP3_D.txt -o OsAP3_blastn_Ff_S1.out -d OsMADS_D.txt -F F -S 2”

Sequences producing significant alignments:	Score (bits)	E Value
OsMADS2	44	2e-006
OsMADS4	38	1e-004
OsMADS29	38	1e-004
OsMADS63	34	0.002
OsMADS62	34	0.002
OsMADS32	34	0.002
OsMADS23	32	0.007
OsMADS31	30	0.029
OsMADS50	30	0.029
OsMADS22	28	0.11
OsMADS90	26	0.45
OsMADS15	26	0.45
OsMADS25	26	0.45
OsMADS89	24	1.8
OsMADS95	24	1.8
OsMADS47	24	1.8
OsMADS18	22	7.1
OsMADS16	22	7.1
OsMADS57	22	7.1
OsMADS27	22	7.1
OsMADS55	22	7.1
OsMADS17	22	7.1

22 hits

Sequences producing significant alignments:	Score (bits)	E Value
OsMADS88	24	1.8
OsMADS47	24	1.8
OsMADS98	22	7.1
OsMADS91	22	7.1
OsMADS76	22	7.1
OsMADS73	22	7.1
OsMADS71	22	7.1
OsMADS70	22	7.1
OsMADS63	22	7.1
OsMADS25	22	7.1
OsMADS29	22	7.1
OsMADS55	22	7.1

12 hits

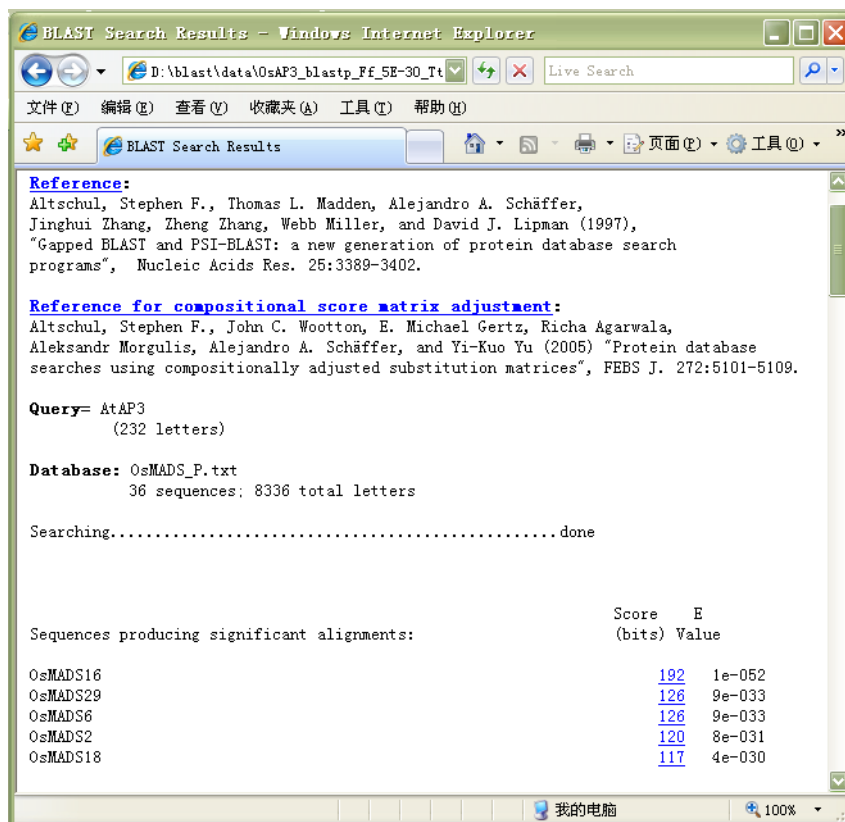
BLAST程序的参数

29、-T Produce HTML output [T/F] (default = F)

输出HTML 格式文件

例：用拟南芥的AP3基因的蛋白质序列在水稻MADS-box基因蛋白质序列库中搜索同源基因，设定参数T为T，使输出结果为HTML格式。

命令：“blastall -p blastn -i AtAP3_D.txt -o OsAP3_blastn_Ff_S2_TT.out -dOsMADS_D.txt -F F -S 2 -T T”



Reference:
Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402.

Reference for compositional score matrix adjustment:
Altschul, Stephen F., John C. Wootton, E. Michael Gertz, Richa Agarwala, Aleksandr Morgulis, Alejandro A. Schäffer, and Yi-Kuo Yu (2005) "Protein database searches using compositionally adjusted substitution matrices", FEBS J. 272:5101-5109.

Query= AtAP3
(232 letters)

Database: OsMADS_P.txt
36 sequences: 8336 total letters

Searching.....done

Sequences producing significant alignments:	Score (bits)	E Value
OsMADS16	192	1e-052
OsMADS29	126	9e-033
OsMADS6	126	9e-033
OsMADS2	120	8e-031
OsMADS18	117	4e-030

BLAST程序的参数



30、-l Restrict search of database to list of GI's [String]

将搜索时使用的数据库限制在输入文件中GIs所限制的子集内，仅用于对NCBI数据库的搜索

31、-U Use lower case filtering of FASTA sequence [T/F]

将query中小写字母格式的序列屏蔽，该功能可以滤除FASTA格式序列中的小写字母(在蛋白质序列中，一些不准确区域通常用小写字母表示)

BLAST程序的参数

32、 -y X dropoff value for ungapped extensions in bits, blastn 20, megablast 10, all others 7 [Real] (default = 0)

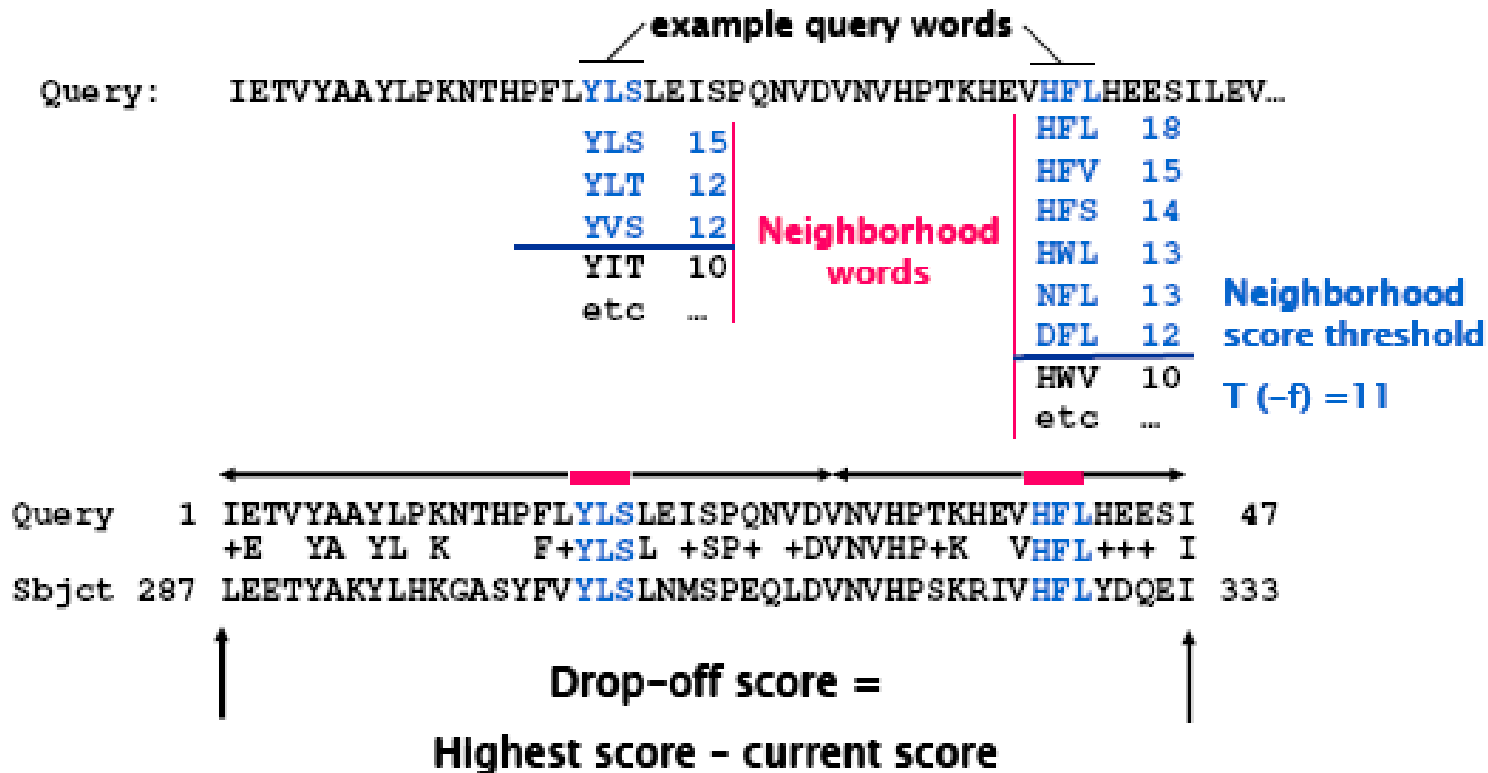
非空位延伸下降的域值，y值越小，则延伸越短

Drop-off score = Highest score – current score

QQGPHUIQEGQQGKEEDPP PKLMMQQGKQEGM	Blosum(QQG,QQG) = 16	
QQGPHUIQEGQQGKEEDPP PKLMMQQGKQEGM	Blosum(QQGK,QQGK) = 21	
QQGPHUIQEGQQGKEEDPP PKLMMQQGKQEGM	Blosum(QQGKE,QQGKQ) = 23	
QQGPHUIQEGQQGKEEDPP PKLMMQQGKQEGM	Blosum(QQGKEE,QQGKQE) = 28	
QQGPHUIQEGQQGKEEDPP PKLMMQQGKQEGM	Blosum(QQGKEED,QQGKQEG) = 27	下降1分
QQGPHUIQEGQQGKEEDPP PKLMMQQGKQEGM	Blosum(QQGKEEDP,QQGKQEGM) = 25	下降3分

BLAST程序的参数

33、 -X Specifies X dropoff value for gapped alignment (in bits)
 [Integer] (default = 0)



-X X dropoff value for gapped alignment (in bits)
 blastn 30, megablast 20, tblastx 0, all others 15

BLAST程序的参数

33、 -X Specifies X dropoff value for gapped alignment (in bits)
[Integer] (default = 0)

```

Query 1 IETVYAAAYLPKNTHPFLYLSLEISPQNVDVNVHPTKHEVHFLHEESI 47
      +E YA YL K F+YLSL +SP+ +DVNVHP+K VHFL+++ I
Sbjct 287 LEETYAKYLHKGASYFVYLSLNMSPEQLDVNVHPSKRIVHFLYDQEI 333
  
```

High-scoring pair (HSP)



Gapped extension with trace back

```

Query 1 IETVYAAAYLPKNTHPFLYLSLEISPQNVDVNVHPTKHEVHFLHEESI-LEV... 50
      +E YA YL K F+YLSL +SP+ +DVNVHP+K VHFL+++ I + +
Sbjct 287 LEETYAKYLHKGASYFVYLSLNMSPEQLDVNVHPSKRIVHFLYDQEIATSI... 337
  
```

Final HSP



BLAST程序的参数

33、-X Specifies X dropoff value for gapped alignment (in bits) [Integer] (default = 0)

设定空位比对中的下降阈值，X值越小，延伸越短

例：分别设定X为1和10用拟南芥的AP3基因的蛋白质序列借助tblastn子程序在水稻MADS-box基因核酸序列库中搜索同源基因

命令：

1) “-p tblastn -i AtAP3_P.txt -d OsMADS_D.txt -o OsAP3_tblastn_Ff_e-29_X1.out -F F -e 1e-29 -X 1”

2) “-p tblastn -i AtAP3_P.txt -d OsMADS_D.txt -o OsAP3_tblastn_Ff_e-29_X1.out -F F -e 1e-29 -X 10”

	Score	E
Sequences producing significant alignments:	(bits)	Value
OsMADS16	192	2e-052
OsMADS2	120	2e-030

2 hits

	Score	E
Sequences producing significant alignments:	(bits)	Value
OsMADS16	192	2e-052
OsMADS29	126	2e-032
OsMADS6	126	2e-032
OsMADS2	120	2e-030
OsMADS18	117	9e-030

5 hits

BLAST程序的参数



34、 -Z X dropoff value for final gapped alignment in bits, blastn/megablast 100, tblastx 0, all others 25 [Integer] (default = 0)

设定最终空位比对的下限值

35、 -R PSI-TBLASTN checkpoint file [File In], Uses a PSI-TBLASTN checkpoint file as the scoring matrix

使用PSI-TBLASTN checkpoint文件为打分矩阵

BLAST程序的参数

36、-n MegaBlast search [T/F] (default = F)

激活blastn中的MEGABLAST算法

例：分别设定n为F和T用水稻*OsMADS6*基因的核酸序列在水稻MADS-box基因核酸序列库中搜索同源基因

命令：

1) “blastall -p blastn -i OsMADS6_D.txt -d OsMADS_D.txt -o OsMADS6_blastn_Ff_e-39_nF.out -e 1e-39 -F F”

2) “blastall -p blastn -i OsMADS6_D.txt -d OsMADS_D.txt -o OsMADS6_blastn_Ff_e-39_nT.out -e 1e-39 -F F -n T”

Sequences producing significant alignments:	Score (bits)	E Value
<i>OsMADS6</i>	1487	0.0
<i>OsMADS17</i>	276	3e-076
<i>OsMADS45</i>	192	4e-051
<i>OsMADS24</i>	178	6e-047
<i>OsMADS50</i>	176	2e-046
<i>OsMADS1</i>	167	2e-043

6 hits

Sequences producing significant alignments:	Score (bits)	E Value
<i>OsMADS6</i>	1487	0.0
<i>OsMADS17</i>	276	3e-076
<i>OsMADS50</i>	176	2e-046

3 hits

BLAST程序的参数

37、-L Location on query sequence [String]

指定用query的某一段序列进行搜索

例：分别设定L为” 91-300”和” 181-300”用拟南芥的AP3基因的蛋白质序列在水稻MADS-box基因蛋白质序列库中搜索同源基因

命令：

- 1) “blastall -p blastp -i AtAP3_P.txt -d OsMADS_P.txt -o OsAP3_blastp_Ff_L181-300.out -F F -L “181,300””
- 2) “blastall -p blastp -i AtAP3_P.txt -d OsMADS_P.txt -o OsAP3_blastp_Ff_L91-300.out -F F -L “91,300””

Sequences producing significant alignments:	Score (bits)	E Value
2 hits		
OsMADS34	17	2.0
OsMADS56	15	8.3

Sequences producing significant alignments:	Score (bits)	E Value
20 hits		
OsMADS16	74	5e-017
OsMADS1	46	1e-008
OsMADS6	45	2e-008
OsMADS17	38	3e-006
OsMADS45	35	3e-005
OsMADS50	33	9e-005
OsMADS23	32	2e-004
OsMADS34	31	3e-004
OsMADS27	30	4e-004
OsMADS56	30	9e-004
OsMADS57	29	0.001
OsMADS21	28	0.002
OsMADS24	28	0.003
OsMADS2	24	0.031
OsMADS4	24	0.034
OsMADS14	24	0.048
OsMADS58	23	0.10
OsMADS25	22	0.16
OsMADS31	17	3.9
OsMADS15	17	4.3

BLAST程序的参数

38、-A Multiple Hits window size, (blastn/megablast 0, all others 40 [Integer] (default = 0))

指定多字长匹配的窗口大小，窗口越大，BLAST搜索的灵敏度越高

例：分别设定A为2和20用水稻*OsMADS6*基因的核酸序列在水稻MADS-box基因核酸序列库中搜索同源基因命令：

1) “blastall -p blastn -i OsMADS6_D.txt -d OsMADS_D.txt -o OsMADS6_blastn_Ff_nF.out -F F -A 2”

2) “blastall -p blastn -i OsMADS6_D.txt -d OsMADS_D.txt -o OsMADS6_blastn_Ff_nF.out -F F -A 20”

Sequences producing significant alignments:	Score (bits)	E Value
OsMADS6	1487	0.0
OsMADS17	276	3e-076
OsMADS45	192	4e-051
OsMADS24	178	6e-047
OsMADS50	176	2e-046
OsMADS14	139	5e-035
OsMADS22	107	2e-025
OsMADS63	86	6e-019
OsMADS15	84	2e-018
OsMADS75	78	2e-016
OsMADS76	60	4e-011

11 hits

Sequences producing significant alignments:	Score (bits)	E Value
OsMADS6	1487	0.0
OsMADS17	276	3e-076
OsMADS45	192	4e-051
OsMADS24	178	6e-047
OsMADS50	176	2e-046
OsMADS1	167	2e-043
OsMADS3	153	3e-039
OsMADS56	141	1e-035
OsMADS14	139	5e-035
OsMADS34	123	3e-030
OsMADS5	123	3e-030
OsMADS18	121	1e-029
OsMADS22	107	2e-025
OsMADS58	105	7e-025
OsMADS63	86	6e-019
OsMADS15	84	2e-018
OsMADS75	78	2e-016
OsMADS37	76	6e-016
OsMADS21	76	6e-016
OsMADS13	76	6e-016
OsMADS2	72	9e-015
OsMADS27	66	6e-013
OsMADS26	64	2e-012
OsMADS76	60	4e-011
OsMADS4	60	4e-011
OsMADS80	54	2e-009

26 hits

BLAST程序的参数

39、-w Frame shift penalty (OOF algorithm for blastx) [Integer] (default = 0)

设定移码的罚分

例：分别设定w为6和10用拟南芥的AP3基因的核酸序列，借助blastx子程序在水稻MADS-box基因蛋白质序列库中搜索同源基因

命令：

- 1) “blastall -p blastx -i AtAP3_D.txt -d OsMADS_P.txt -o OsAP3_blastx_Ff_e-30_w6.out -F F -e 1e-30 -w 6”
- 2) “blastall -p blastx -i AtAP3_D.txt -d OsMADS_P.txt -o OsAP3_blastx_Ff_e-30_w6.out -F F -e 1e-30 -w 10”

Sequences producing significant alignments:	Score (bits)	E Value
OsMADS16	198	2e-054
OsMADS1	134	4e-035
OsMADS6	134	6e-035
OsMADS29	133	8e-035
OsMADS18	133	1e-034
OsMADS2	131	4e-034
OsMADS45	130	5e-034
OsMADS4	130	5e-034
OsMADS15	129	2e-033
OsMADS14	128	3e-033
OsMADS17	125	2e-032
OsMADS34	125	3e-032
OsMADS24	123	8e-032
OsMADS57	122	2e-031
OsMADS50	121	3e-031
OsMADS3	120	5e-031
OsMADS27	120	9e-031

17 hits

Sequences producing significant alignments:	Score (bits)	E Value
OsMADS16	195	2e-053
OsMADS6	128	3e-033
OsMADS29	128	3e-033
OsMADS4	124	4e-032
OsMADS2	124	5e-032
OsMADS1	123	8e-032
OsMADS45	122	2e-031
OsMADS18	121	4e-031
OsMADS15	120	5e-031
OsMADS14	120	7e-031

10 hits

BLAST程序的参数



40、 -t Length of the largest intron allowed in a translated nucleotide sequence when linking multiple distinct alignments [Integer] (default = 0)

设定在tblastn过程中，连接多个不同比对片段时所允许的内含子长度的最大值

41、 -B Number of concatenated queries, for blastn and tblastn [Integer] (default = 0)

指定搜索过程中，query文件中相连接的序列数

BLAST程序的参数

42、-V Force use of the legacy BLAST engine [T/F] (default = F)

强制使用旧版搜索引擎，在使用参数B时，V参数必须设为F

例：分别设定V为T和F用拟南芥AP3基因的核酸序列在水稻MADS-box基因核酸序列库中搜索同源基因

命令：

- 1) “blastall -p blastn -i AtAP3_D.txt -d OsMADS_D.txt -o OsAP3_blastn_Ff_Vt.out -F F -V T”
- 2) “blastall -p blastn -i AtAP3_D.txt -d OsMADS_D.txt -o OsAP3_blastn_Ff_Vf.out -F F -V F”

```
Matrix: blastn matrix:1 -3
Gap Penalties: Existence: 5, Extension: 2
Number of Hits to DB: 35
Number of Sequences: 72
Number of extensions: 35
Number of successful extensions: 35
Number of sequences better than 10.0: 29
Number of HSP's better than 10.0 without gapping: 29
Number of HSP's successfully gapped in prelim test: 0
Number of HSP's that attempted gapping in prelim test: 0
Number of HSP's gapped (non-prelim): 35
length of query: 696
length of database: 54,516
effective HSP length: 13
effective length of query: 683
effective length of database: 53,580
effective search space: 36595140
effective search space used: 36595140
T: 0
A: 0
X1: 11 (21.8 bits)
X2: 15 (29.7 bits)
S1: 12 (24.3 bits)
S2: 11 (22.3 bits)
```

```
Matrix: blastn matrix:1 -3
Gap Penalties: Existence: 5, Extension: 2
Number of Sequences: 72
Number of Hits to DB: 468
Number of extensions: 35
Number of successful extensions: 35
Number of sequences better than 10.0: 29
Number of HSP's gapped: 35
Number of HSP's successfully gapped: 34
Length of query: 696
Length of database: 54,516
Length adjustment: 13
Effective length of query: 683
Effective length of database: 53,580
Effective search space: 36595140
Effective search space used: 36595140
X1: 11 (21.8 bits)
X2: 15 (29.7 bits)
X3: 50 (99.1 bits)
S1: 11 (22.3 bits)
S2: 11 (22.3 bits)
```

BLAST程序的参数

43、-C Uses composition-based statistics for tblastn

在运用tblastn子程序时，采用基于成分的统计分析方法

例：分别设定C为1和2用拟南芥的AP3基因的蛋白质序列，借助tblastn子程序在水稻MADS-box基因核酸序列库中搜索同源基因

命令：

1) “blastall -p tblastn -i AtAP3_P.txt -d OsMADS_D.txt -o OsAP3_tblastn_Ff_e-24_X1.out -F F -e 1e-24 -C 1”

2) “blastall -p tblastn -i AtAP3_P.txt -d OsMADS_D.txt -o OsAP3_tblastn_Ff_e-24_X1.out -F F -e 1e-24 -C 2”

Sequences producing significant alignments:	Score (bits)	E Value
OsMADS16	179	2e-048
OsMADS29	122	4e-031
OsMADS6	117	9e-030
OsMADS2	113	3e-028
OsMADS45	110	1e-027
OsMADS17	110	1e-027
OsMADS18	110	2e-027
OsMADS1	108	6e-027
OsMADS4	107	1e-026
OsMADS24	106	3e-026
OsMADS21	105	6e-026
OsMADS13	105	6e-026
OsMADS50	103	1e-025
OsMADS27	102	3e-025
OsMADS31	102	3e-025

15 hits

Sequences producing significant alignments:	Score (bits)	E Value
OsMADS16	192	2e-052
OsMADS29	126	2e-032
OsMADS6	126	2e-032
OsMADS2	120	2e-030
OsMADS18	117	9e-030
OsMADS1	117	1e-029
OsMADS4	117	1e-029
OsMADS50	117	2e-029
OsMADS45	115	5e-029
OsMADS17	114	2e-028
OsMADS27	111	6e-028
OsMADS21	110	1e-027
OsMADS13	110	1e-027
OsMADS23	110	1e-027
OsMADS24	108	4e-027
OsMADS57	108	5e-027
OsMADS3	108	7e-027
OsMADS31	108	7e-027
OsMADS56	104	9e-026
OsMADS58	103	1e-025
OsMADS34	102	3e-025
OsMADS5	102	4e-025

22 hits



BLAST程序的参数

44、-s Compute locally optimal Smith-Waterman alignments (This option is only available for gapped tblastn.) [T/F] (default = F)

计算局部最佳的Smith-Waterman比对

例：分别设定s为F和T，用拟南芥AP3基因蛋白质序列的61-180区段，借助tblastn子程序在水稻MADS-box基因核酸序列库中搜索同源基因

命令：

- 1) “blastall -p tblastn -i AtAP3_P.txt -d OsMADS_D.txt -o AtAP3_tblastn_Ff_sF_L61-180.out -F F -L "61,180"”
- 2) “blastall -p tblastn -i AtAP3_P.txt -d OsMADS_D.txt -o AtAP3_tblastn_Ff_sF_L61-180.out -F F -L "61,180" -s T”

Sequences producing significant alignments:	Score (bits)	E Value
OsMADS16	92	3e-022
OsMADS6	44	9e-008
OsMADS50	39	2e-006
OsMADS56	32	4e-004
OsMADS14	24	0.074
OsMADS15	19	1.8
OsMADS69	18	4.7
OsMADS37	17	8.9

8 hits

Sequences producing significant alignments:	Score (bits)	E Value
OsMADS16	92	3e-022
OsMADS6	44	9e-008
OsMADS50	39	2e-006
OsMADS56	32	4e-004
OsMADS14	24	0.074
OsMADS15	19	1.8
OsMADS69	18	4.7
OsMADS65	18	5.9
OsMADS37	17	8.9

9 hits

谢谢！