

Dear editors and reviewer:

Here, we provide a complete way to generate .pssm and .hhm. For generation of .pssm files, local version (section 1.1) and online version (section 1.2) can be used. For generation of .hhm files, only local version (section 2.1) is available. In online version, you can save much time to download the huge database for Psi-blast.

On the GitHub page, we use spider3 to automatically call psi-blast and hhblits. If you also want to use spider3, please make sure you download the same version as us. The specific download and installation information is described in section 3.

## 1. Generation of PSSM profiles (local version and online version)

### 1.1 Local version

#### Step1. Download N-Blast 2.7:

<https://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/2.7.1/>

Index of /blast/executables/blast+/2.7.1			
Name	Last modified	Size	
Parent Directory		-	
ChangeLog	2017-10-18 23:07	85	
ncbi-blast-2.7.1-i386.rpm	2017-10-18 23:05	18M	
ncbi-blast-2.7.1-i386.rpm.md5	2017-10-18 23:08	62	
ncbi-blast-2.7.1-i686.rpm	2017-10-18 23:05	181M	
ncbi-blast-2.7.1-i686.rpm.md5	2017-10-18 23:08	65	
ncbi-blast-2.7.1-src.tar.gz	2017-10-18 23:08	22M	
ncbi-blast-2.7.1-src.tar.gz.md5	2017-10-18 23:08	63	
ncbi-blast-2.7.1-src.zip	2017-10-18 23:08	26M	
ncbi-blast-2.7.1-src.zip.md5	2017-10-18 23:08	60	
ncbi-blast-2.7.1-win64.exe	2017-10-18 23:07	87M	
ncbi-blast-2.7.1-win64.exe.md5	2017-10-18 23:08	62	
ncbi-blast-2.7.1-x64-linux.tar.gz	2017-10-18 23:07	223M	
ncbi-blast-2.7.1-x64-linux.tar.gz.md5	2017-10-18 23:08	69	
ncbi-blast-2.7.1-x64-macosx.tar.gz	2017-10-18 23:05	127M	
ncbi-blast-2.7.1-x64-macosx.tar.gz.md5	2017-10-18 23:08	70	
ncbi-blast-2.7.1-x64-win64.tar.gz	2017-10-18 23:08	87M	
ncbi-blast-2.7.1-x64-win64.tar.gz.md5	2017-10-18 23:08	69	
ncbi-blast-2.7.1.dmg	2017-10-18 23:04	128M	
ncbi-blast-2.7.1.dmg.md5	2017-10-18 23:09	56	

Decompress this file:

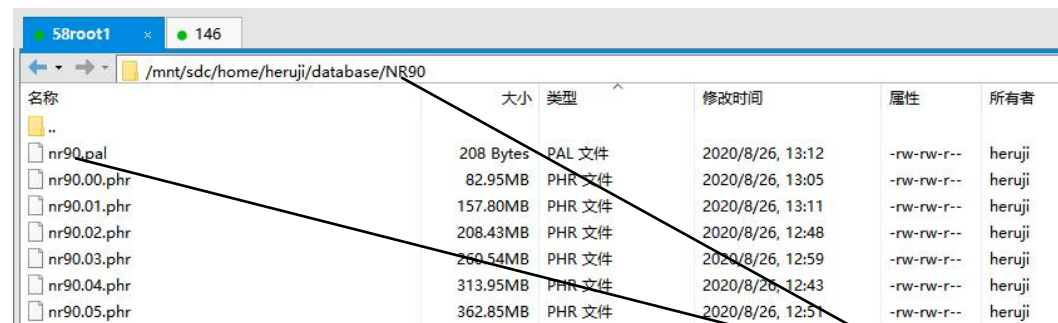
/home/teacher03/jiangyelu/ncbi-blast-2.7.1+				
名称	大小	类型	修改时间	
..				
bin		文件夹	2022/7/30, 16:06	
doc		文件夹	2022/7/30, 16:06	
ChangeLog	85 Bytes	文件	2022/7/30, 16:06	
LICENSE	27KB	文件	2022/7/30, 16:06	
ncbi_package_info	15KB	文件	2022/7/30, 16:06	
README	465 Bytes	文件	2022/7/30, 16:06	

#### Step2. Download NR90 database from NCBI: <https://ftp.ncbi.nlm.nih.gov/blast/db/>

Here, the database version used in our paper is Database NR90 (Version: Nov. 5 2015 11:07 pm), and the specific information is show in the following figure. Please check your database contain complete files. nr.00.tar.gz, nr.01.tar.gz, ..., nr.59.tar.gz, and concatenate them in your local environment. After several update the years, we found that this version of NR90 is hard to find and download. If you need this database, you can email to us. Anyway, in addition to reproducing the results of the paper, if the program is applied on a new problem, the latest version we provide

here is available.

```
lzhp@ubuntu3:/mnt/sdc/user/heruji/NR90$ blastdbcmd -db nr90 -info
Database: nr90
41,206,127 sequences; 13,918,132,356 total residues
Date: Nov 5, 2015 11:07 PM Longest sequence: 41,965 residues
Volumes:
/mnt/sdc/user/heruji/NR90/nr90.00
/mnt/sdc/user/heruji/NR90/nr90.01
/mnt/sdc/user/heruji/NR90/nr90.02
/mnt/sdc/user/heruji/NR90/nr90.03
/mnt/sdc/user/heruji/NR90/nr90.04
/mnt/sdc/user/heruji/NR90/nr90.05
/mnt/sdc/user/heruji/NR90/nr90.06
/mnt/sdc/user/heruji/NR90/nr90.07
/mnt/sdc/user/heruji/NR90/nr90.08
/mnt/sdc/user/heruji/NR90/nr90.09
/mnt/sdc/user/heruji/NR90/nr90.10
/mnt/sdc/user/heruji/NR90/nr90.11
/mnt/sdc/user/heruji/NR90/nr90.12
/mnt/sdc/user/heruji/NR90/nr90.13
lzhp@ubuntu3:/mnt/sdc/user/heruji/NR90$
```



名称	大小	类型	修改时间	属性	所有者
nr90.pal	208 Bytes	PAL 文件	2020/8/26, 13:12	-rw-rw-r--	heruji
nr90.00.phr	82.95MB	PHR 文件	2020/8/26, 13:05	-rw-rw-r--	heruji
nr90.01.phr	157.80MB	PHR 文件	2020/8/26, 13:11	-rw-rw-r--	heruji
nr90.02.phr	208.43MB	PHR 文件	2020/8/26, 12:48	-rw-rw-r--	heruji
nr90.03.phr	260.54MB	PHR 文件	2020/8/26, 12:59	-rw-rw-r--	heruji
nr90.04.phr	313.95MB	PHR 文件	2020/8/26, 12:43	-rw-rw-r--	heruji
nr90.05.phr	362.85MB	PHR 文件	2020/8/26, 12:51	-rw-rw-r--	heruji

### Step3. Run blast to generate pssm profile:

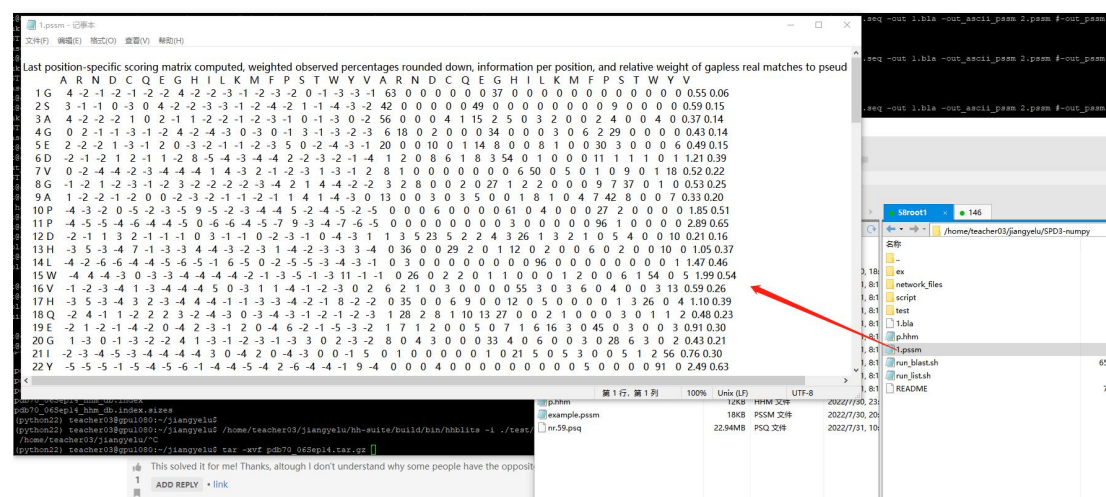
`/home/teacher03/jiangyelu/ncbi-blast-2.7.1+/bin/psiblast -db /mnt/sdc/home/heruji/database/NR90/nr90 -num_iterations 3 -num_alignments 1 -num_threads 2 -query ./test/*.seq -out 1.bla -out_ascii_pssm 1.pssm #out_pssm ./1.chk`

```
root@gu1080:/home/teacher03/jiangyelu/SPD3-numpy# /home/teacher03/jiangyelu/ncbi-blast-2.7.1+/bin/psiblast -db /mnt/sdc/home/heruji/database/NR90/nr90 -num_iterations 3 -num_alignments 1 -num_threads 2 -query ./test/*.seq -out 1.bla -out_ascii_pssm 1.pssm #out_pssm ./1.chk
```

-db :alignment database

-query : input sequence

-out\_ascii\_pssm : output .pssm file



Last position-specific scoring matrix computed, weighted observed percentages rounded down, information per position, and relative weight of gapless real matches to pseud

	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V				
1G	4.2	-1.2	-1.2	-2.4	-2.3	-1.2	-3.2	-2.0	-1.3	-3.1	63.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.55	0.06			
2S	3.1	-1.0	-3.0	4.2	-2.3	-3.1	-2.4	-2.1	-1.4	-3.2	42.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.9	0.0	0.59	0.15		
3A	4.2	-2.2	-1.0	2.1	-1.2	-2.1	-2.3	-1.0	-1.3	-2.0	56.0	0.0	0.4	1.15	2.5	0.3	2.0	0.2	4.0	0.4	0.37	0.14		
4G	0.2	-1.1	-3.1	-2.4	-2.4	-3.0	-3.0	-1.3	-1.3	-2.3	6.18	0.2	0.0	0.34	0.0	0.3	0.6	2.29	0.0	0.0	0.43	0.14		
5E	2.2	-2.1	-3.1	2.0	-3.2	-2.1	-1.2	-3.5	0.2	-4.3	-1.20	0.0	0.10	0.14	8.0	0.8	1.0	0.30	3.0	0.0	0.6	0.49	0.15	
6D	-2.1	-2.1	1.2	-1.2	-8.5	-4.3	-4.4	-2.3	-2.1	-4.1	1.2	0.8	6.1	8.3	5.4	0.1	0.0	0.11	1.1	1.0	1.1	1.21	0.39	
7V	0.2	-4.4	-2.3	-4.4	-4.1	1.4	-3.2	-1.3	-1.2	8.1	0.0	0.0	0.0	0.6	5.0	0.5	0.1	0.9	0.1	1.8	0.52	0.22		
8G	-1.2	-1.2	-3.1	-2.3	-2.2	-2.3	-4.2	1.4	-4.2	-2.3	3.28	0.0	0.27	1.2	2.2	0.0	0.9	7.37	0.1	0.0	0.53	0.25		
9A	1.2	-2.1	-2.0	-2.3	-2.1	-1.2	-1.1	1.4	1.4	-3.0	13.0	0.3	3.5	0.0	1.8	1.0	4.7	4.2	8.0	0.7	0.33	0.20		
10P	-4.3	-2.0	-5.2	-3.5	-5.9	-5.2	-3.4	-4.5	-2.4	-5.2	0.0	0.6	0.0	0.61	0.4	0.0	0.27	2.0	0.0	0.0	1.85	0.51		
11D	-4.5	-5.4	-6.4	-4.5	-0.6	-6.4	-5.7	9.3	-4.7	-6.5	0.0	0.0	0.0	0.0	3.0	0.0	0.96	1.0	0.0	0.0	2.89	0.65		
12D	-2.1	1.3	2.1	-1.1	0.3	-1.1	0.2	-3.1	0.4	-3.1	1.3	5.23	5.2	2.4	3.26	1.3	2.1	0.5	4.0	0.0	0.21	0.16		
13H	-3.5	-3.4	7.1	-3.3	4.4	-3.2	-3.1	4.2	-3.3	3.4	0.36	0.0	0.29	2.0	1.12	0.2	0.0	6.0	2.0	0.10	0.05	0.37		
14L	-4.2	-6.4	-4.5	-5.6	-5.1	6.5	0.2	-5.5	-3.4	-3.1	0.3	0.0	0.0	0.0	0.96	0.0	0.0	0.0	0.0	0.0	1.47	0.46		
15W	-4.4	-4.3	0.3	-3.4	-4.4	-4.2	-1.3	-5.1	-3.11	-1.1	0.26	0.2	2.0	1.1	0.0	0.0	1.2	0.0	6.1	5.4	0.5	1.99	0.54	
16V	-1.2	-3.4	1.3	-4.4	-4.5	0.3	1.1	-4.1	-2.3	0.2	6.2	1.0	3.0	0.0	0.55	3.0	3.6	0.4	0.0	3.3	13	0.59	0.26	
17H	-3.5	-3.4	3.2	-3.4	-4.4	-1.1	-1.3	-3.4	-2.1	8.2	-2.2	0.35	0.0	6.9	0.0	0.12	0.5	0.0	0.0	1.3	26	0.4	1.10	0.39
18Q	-2.4	-1.1	1.2	2.2	3.2	-4.3	0.3	-4.3	-1.2	-1.2	-2.3	1.28	2.8	1.10	13.27	0.0	0.2	1.0	0.0	3.0	1.1	2.0	0.40	0.23
19E	-2.1	-1.2	-1.2	-0.4	-4.3	-1.1	2.0	-4.6	-2.1	-5.3	-2.1	1.7	1.2	0.0	5.0	0.7	1.6	16	3.0	0.45	0.3	0.0	3.91	0.30
20G	1.3	-0.1	-3.2	-2.4	1.3	-1.2	-2.3	-1.3	3.0	2.3	-2.8	0.4	3.0	0.0	0.33	4.0	6.0	3.0	28	6.3	0.2	0.43	0.21	
21I	-2.3	-4.5	-3.4	-4.4	-4.3	0.4	2.0	-4.3	0.0	-1.5	0.1	0.0	0.0	0.0	1.0	21	5.0	5.0	3.0	0.5	1.2	56	0.76	0.30
22Y	-5.5	-5.1	-5.4	-5.6	-1.4	-4.4	-5.4	-2.6	-4.4	-1.9	4.0	0.0	0.4	0.0	0.0	0.0	0.0	0.0	0.91	0.0	0.0	2.49	0.63	

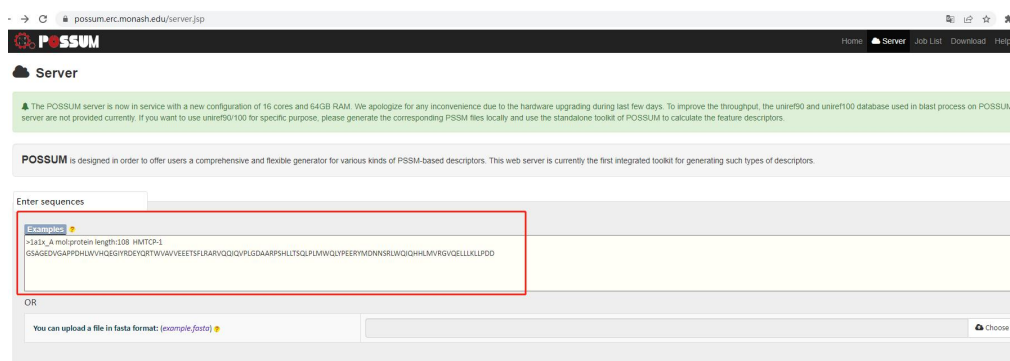
### Step 4. Parse it to a numpy array as shown in step5 in the following 1.2 online version

### 1.2 Online version

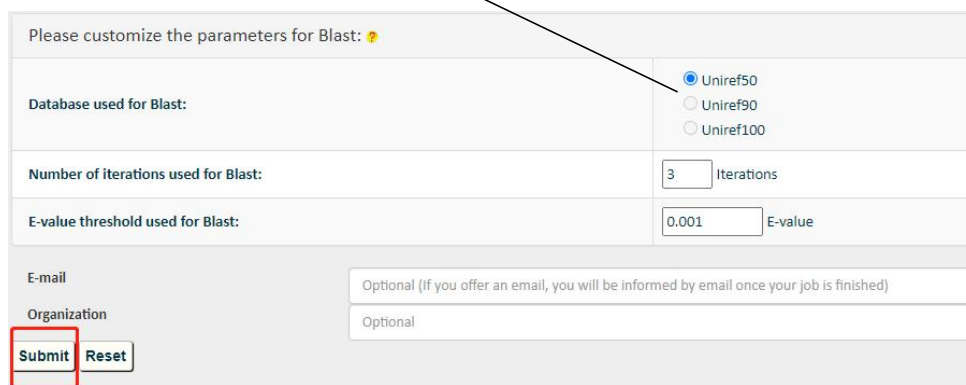
After several years of updates, source NR90 database we used is hard to find and download. If you cannot find a link to the original NR90 database in the NCBI like us, or if you want to avoid downloading the huge database for Psi-blast. you can use online web server [1]: <https://possum.erc.monash.edu/server.jsp>

[1] Wang J, Yang B et al. POSSUM: a bioinformatics toolkit for generating numerical sequence feature descriptors based on PSSM profiles. *Bioinformatics* 2017;33(17):2756-2758. DOI: 10.1093/bioinformatics/btx302.

### Step 1. Enter your sequence into the red box.



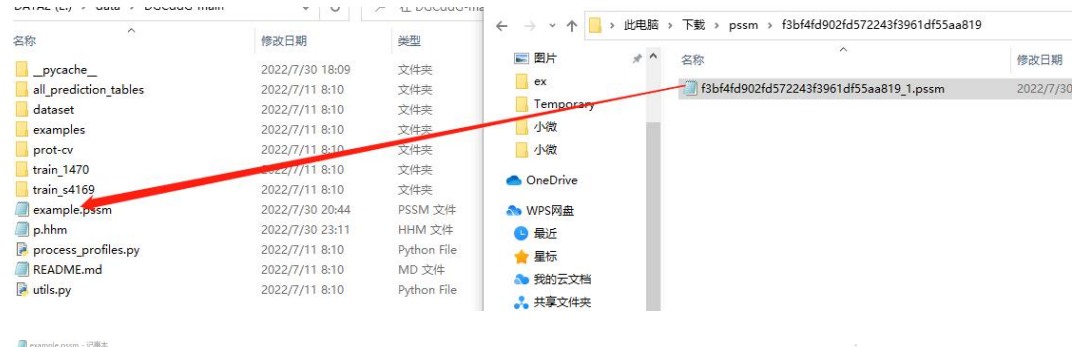
### Step 2. You can choose the database for blast. Submit and wait 1~10minutes, it depends on the number of online users.



### Step 3. When job finished, you can download the PSSM files.



**Step 4. You can decompress it and move it to the DGCddG program. In this example, we modify the name of the file. You can see the standard pssm files as follow.**



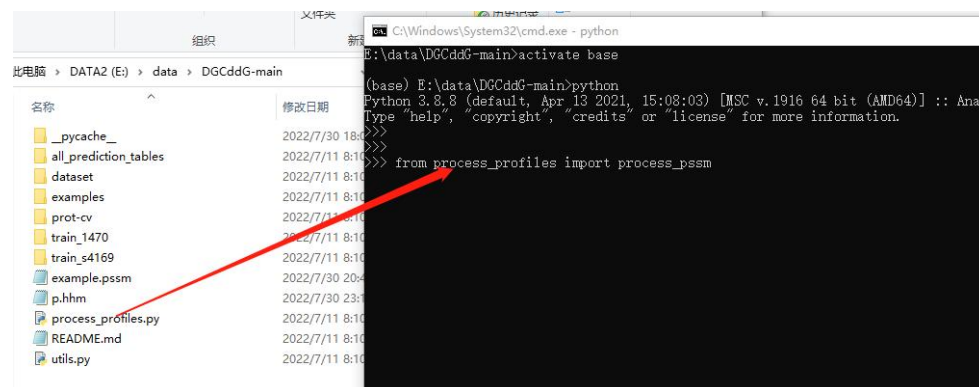
example.pssm - 记事本

文件(F) 编辑(E) 格式(O) 查看(V) 帮助(H)

Last position-specific scoring matrix computed, weighted observed percentages rounded down, information per position, and relative weight of gapless real matches to pseudocounts

	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	S	T	W	Y	V
1G	4	2	1	2	1	1	1	3	2	2	1	2	3	1	1	1	3	2	1
2S	2	1	0	1	1	1	1	1	2	2	1	1	3	3	2	1	20	0	0
3A	3	2	3	3	2	2	2	1	1	2	0	3	1	1	2	0	1	50	0
4G	0	2	1	2	2	1	1	2	2	3	3	0	2	1	2	3	1	5	21
5E	2	2	2	1	3	1	2	2	2	3	1	2	4	5	1	1	4	3	30
6D	1	1	1	2	3	1	0	3	6	3	2	2	3	4	1	2	4	1	6
7V	1	2	3	3	2	1	3	4	3	2	2	1	3	2	2	3	0	3	2
8G	0	1	0	2	3	2	2	4	2	3	4	2	3	3	1	2	3	1	4
9A	2	2	1	0	2	0	2	2	1	1	1	1	0	2	2	1	3	2	0
10P	3	2	2	1	4	2	2	4	7	4	1	2	3	5	2	3	4	1	4
11P	3	4	4	3	5	3	3	4	4	5	5	3	4	6	3	3	6	5	4
12D	2	1	1	4	5	1	1	0	3	0	2	2	2	3	3	0	1	4	2
13H	2	5	2	3	6	1	2	2	5	4	3	1	3	0	4	0	2	3	1
14L	3	4	5	5	3	4	4	5	4	0	6	4	1	1	4	4	3	3	1
15W	4	2	5	5	1	3	4	4	4	3	3	3	1	5	2	4	1	0	0
16V	0	1	2	3	1	3	3	3	2	4	0	2	1	2	3	1	1	2	1
17H	3	3	3	4	3	2	4	3	3	2	1	2	2	4	3	1	9	1	0
18Q	2	4	1	0	2	2	0	1	4	0	3	4	3	0	2	4	3	3	0
19E	2	1	2	1	4	1	0	3	1	3	0	3	1	4	6	2	1	4	3
20G	1	2	1	1	3	2	2	2	4	2	3	2	2	0	3	1	1	1	2
21I	2	3	4	4	2	3	4	4	4	3	1	3	3	1	4	3	0	3	0
22Y	4	4	4	0	4	3	4	5	0	3	3	4	3	5	5	4	0	9	3
23R	2	1	2	1	3	1	3	4	2	0	1	0	1	2	3	2	2	4	2
24D	4	4	0	8	5	2	0	3	3	5	6	3	5	6	3	2	3	6	5
25E	3	2	2	1	5	1	7	4	2	5	5	1	4	5	3	2	5	4	4
26Y	2	2	3	1	4	1	1	2	4	1	3	3	2	0	3	1	0	3	1
27Q	2	3	1	1	4	5	1	3	1	4	4	0	3	4	3	1	2	3	0
28R	3	7	2	3	5	1	2	4	2	5	4	1	3	5	4	2	3	5	3

**Step 5. Try to parse it to a numpy array. You can see in this example, the array is (108, 20), 108 is the residue number and 20 is features for every residue.**



此电脑 > DATA2 (E:) > data > DGCddG-main

名称 修改日期 类型

\_\_pycache\_\_ 2022/7/30 18:09 文件夹

all\_prediction\_tables 2022/7/11 8:10 文件夹

dataset 2022/7/11 8:10 文件夹

examples 2022/7/11 8:10 文件夹

prot-cv 2022/7/11 8:10 文件夹

train\_1470 2022/7/11 8:10 文件夹

train\_s4169 2022/7/11 8:10 文件夹

example.pssm 2022/7/30 20:44 PSSM 文件

p.hhm 2022/7/30 23:11 HHM 文件

process\_profiles.py 2022/7/11 8:10 Python 文件

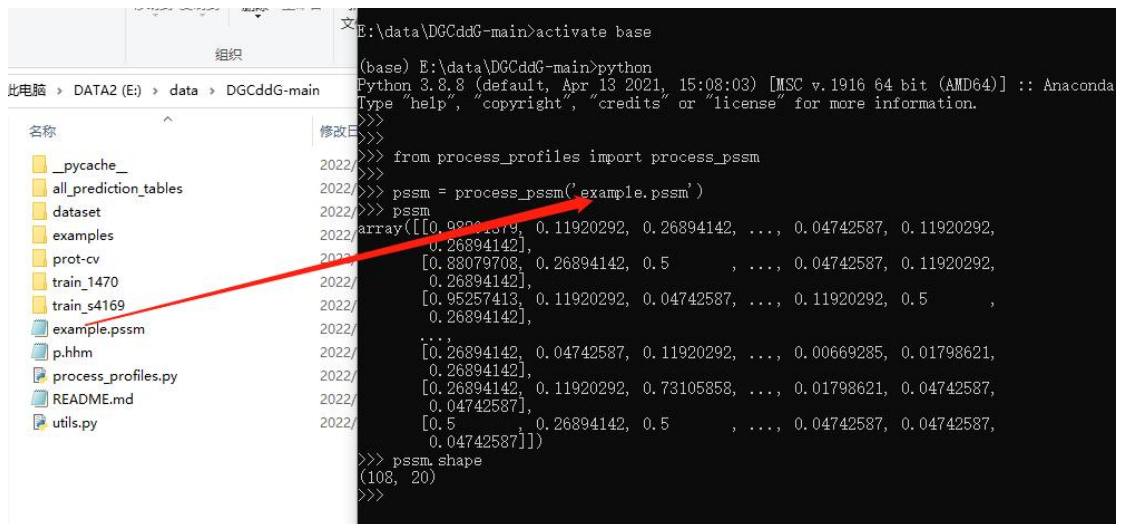
README.md 2022/7/11 8:10 MD 文件

utils.py 2022/7/11 8:10 Python 文件

```

C:\Windows\System32\cmd.exe - python
E:\data\DGCddG-main>activate base
(base) E:\data\DGCddG-main>python
Python 3.8.8 (default, Apr 13 2021, 15:08:03) [MSC v.1916 64 bit (AMD64)] :: Anaconda3
Type 'help', 'copyright', 'credits' or 'license' for more information.
>>> from process_profiles import process_pssm
  
```





```
E:\data\DGcddG-main>activate base
(base) E:\data\DGcddG-main>python
Python 3.8.8 (default, Apr 13 2021, 15:08:03) [MSC v.1916 64 bit (AMD64)] :: Anaconda
Type "help", "copyright", "credits" or "license" for more information.
>>>
>>> from process_profiles import process_pssm
>>> pssm = process_pssm('example.pssm')
>>> pssm
array([[0.88091379, 0.11920292, 0.26894142, ..., 0.04742587, 0.11920292,
        0.26894142],
       [0.88097708, 0.26894142, 0.5, ..., 0.04742587, 0.11920292,
        0.26894142],
       [0.95257413, 0.11920292, 0.04742587, ..., 0.11920292, 0.5,
        0.26894142],
       ...,
       [0.26894142, 0.04742587, 0.11920292, ..., 0.00669285, 0.01798621,
        0.26894142],
       [0.26894142, 0.11920292, 0.73105858, ..., 0.01798621, 0.04742587,
        0.04742587],
       [0.5, 0.26894142, 0.5, ..., 0.04742587, 0.04742587,
        0.04742587]])
>>> pssm.shape
(108, 20)
>>>
```

## 2. Generation of hhm profiles

The HH-suite is an open-source software package for sensitive protein sequence searching based on the pairwise alignment of hidden Markov models. It also incorporates hhblits.


**Step 1.** install hh-suite from Github: <https://github.com/soedinglab/hh-suite.git>

You should install this software with the following instructions:

To download the source code and compile the HH-suite execute the following commands:

```
git clone https://github.com/soedinglab/hh-suite.git
mkdir -p hh-suite/build && cd hh-suite/build
cmake -DCMAKE_INSTALL_PREFIX=. ..
make -j 4 && make install
export PATH="$(pwd)/bin:${(pwd)}/scripts:$PATH"
```

After your installation, you can see these files like us.



名称	大小	类型	修改时间	属性	所有者
..					
build		文件夹	2022/7/30, 16:28	drwxrwxr-x	teacher03
cmake		文件夹	2022/7/30, 16:26	drwxrwxr-x	teacher03
data		文件夹	2022/7/30, 16:26	drwxrwxr-x	teacher03
lib		文件夹	2022/7/30, 16:26	drwxrwxr-x	teacher03
scripts		文件夹	2022/7/30, 16:26	drwxrwxr-x	teacher03
src		文件夹	2022/7/30, 16:26	drwxrwxr-x	teacher03
README.md	5KB	MD 文件	2022/7/30, 16:26	-rw-rw-r--	teacher03
azure-pipelines.yml	8KB	YML 文件	2022/7/30, 16:26	-rw-rw-r--	teacher03
CMakeLists.txt	2KB	文本文档	2022/7/30, 16:26	-rw-rw-r--	teacher03
Dockerfile	2KB	文件	2022/7/30, 16:26	-rw-rw-r--	teacher03
LICENSE	34KB	文件	2022/7/30, 16:26	-rw-rw-r--	teacher03

You can also check hhblits in following path:

**1** is the path of software: **/home/teacher03/jiangvelu/hh-suite/build/bin/hhblits**

/mnt/sdc/home/heruji/database/uniprot20\_2016\_02/uniprot20\_2016\_02

The screenshot shows a Jupyter Notebook interface. On the left, a file explorer displays the contents of the directory 'E:\data\DGcddG-main'. The files listed are: '..', '\_pycache\_', 'all\_prediction\_tables', 'dataset', 'examples', 'prot-cv', 'train\_1470', 'train\_s4169', 'README.md' (5KB), 'process\_profiles.py' (2KB), 'utils.py' (23KB), 'p.hhm' (12KB), and 'example.pssm' (18KB). On the right, the code editor shows the following Python code:

```
>>>
>>>
>>>
>>> from process_profiles import process_hhm
>>> hhm = process_hhm('p.hhm')
>>> hhm
array([[1.          , 1.          , ..., 0.10001   , 0.          ,
        [0.29792979, 1.          , 1.          , ..., 0.2940294 , 0.          ,
        [0.          , 0.          , ..., 0.          , 0.          ,
        [0.1760176 , 1.          , 1.          , ..., 0.33593359, 0.          ,
        [0.          , 0.          , ..., 0.          , 0.          ,
        ...,
        [1.          , 1.          , 0.23182318, ..., 0.34933493, 0.          ,
        [0.          , 0.          , ..., 0.          , 0.          ,
        [0.45114511, 1.          , 0.10711071, ..., 0.33963396, 0.          ,
        [0.          , 0.          , ..., 0.          , 0.          ,
        [0.54145415, 1.          , 0.26722672, ..., 0.3440344 , 0.          ,
        [0.          , 0.          , ..., 0.          , 0.          ,
        ]])
>>> hhm.shape
(108, 30)
>>>
```

Red arrows point from the file explorer to the code editor. One arrow points from 'p.hhm' to the file path argument in the code. Another arrow points from the 'check the shape' text to the 'hhm.shape' output. A third arrow points from the 'dataset' file to the 'dataset' variable in the code.

Click the link in red box shown in the following picture:

## Protein

### Protein 3D Structure Prediction

[SPARKS-x: Fold Recognition](#)

### Protein Local Structural Prediction

[SPOT-Contact: Sequence-based contact map prediction](#)

[SPOT-1D: Latest method On sequence-based prediction of structural features for proteins](#)

[SPIDER3-Single \(Numpy\): Single-sequence-based prediction of structural features for proteins](#)

[SPOT-Disorder2: Sequence-based prediction of intrinsically disordered proteins: V2.0](#)

[SPOT-Disorder-Single: Single sequence-based prediction of intrinsically disordered proteins](#)

[SPOT-Contact-LM: Improving single-sequence-based prediction of protein contact map using a transformer language model](#)

Older versions, please use **SPOT-1D** instead. Links are provided for reference:

[SPIDER3: Sequence-based prediction of structural features for proteins](#)

[SPOT-Disorder: Sequence-based prediction of intrinsically disordered proteins](#)

[SPIDER2: Predicting secondary structure, local backbone angles, and ASA of proteins](#)

[SPINE-D: Protein disorder prediction](#)

[SPIDER-HSE: HSE \(alpha/beta\) prediction from protein sequence](#)

[SPIDER2 \(including HSE\): Predicting secondary structure, local backbone angles, ASA, and \[HSE\] of proteins](#)

[SPINE-X: Protein secondary Structure Prediction](#)

After your download you can see the right format ‘.hmm’ and ‘pssm’ files in the example directory “ex” of SPIDER3.

..		文件夹	
ex	40,050	? 文件夹	2020/1/18 7:01
network_files	105,572,136	? 文件夹	2020/1/18 7:01
script	252,216	? 文件夹	2020/1/18 7:01
README	78	? 文件	2020/1/18 7:01
run_list.sh	1,293	? SH 文件	2020/1/18 7:01

**Step 2.** You should **modify the paths** of ‘run\_list.sh’ in the red box so that the SPIDER3 can find them. NR90 and **uniprot20\_2016\_02** are two databases prepared for blast and hhblits.

Here, we provide our examples:

```
#NR90=/scratch1/NR90/nr90
NR90=/scratch1/uniref90/uniref90
HHDB=/scratch1/uniprot20_2016_02/uniprot20_2016_02
HHBLITS=/export/home/s2878100/local/hhsuite/bin/hhblits
psiblast=/export/home/s2878100/aspen/software/blast-2.2.30+/bin/psiblast

NR90_pub=/export/home/s2878100/aspen/pub/uniref90/uniref90
HHDB_pub=/export/home/s2878100/aspen/pub/uniprot20_2016_02/uniprot20_2016_02

#!/bin/bash

NR90=/mnt/sdc/user/heruji/NR90/nr90
HHDB=/mnt/sdc/user/heruji/uniprot20_2016_02/uniprot20_2016_02
HHBLITS=/mnt/sdc/user/heruji/hh-suite/build/bin/hhblits
psiblast=/home/lzhpc/user/hlq/nblast2.7/bin/psiblast

NR90_pub=$NR90
HHDB_pub=$HHDB
```

**Step 3.** After several update in these years, NR90 is hard to download from the NCBI. Maybe you can download uniref90 from this website as a substitution.

<https://ftp.uniprot.org/pub/databases/uniprot/uniref/uniref90/>



← → ↻ <https://ftp.uniprot.org/pub/databases/uniprot/uniref/uniref90/>

## Index of /pub/databases/uniprot/uniref/uniref90

<a href="#">Name</a>	<a href="#">Last modified</a>	<a href="#">Size</a>	<a href="#">Description</a>
<a href="#">Parent Directory</a>		-	
<a href="#">README</a>	2022-05-25 10:00	6.8K	
<a href="#">RELEASE.metalink</a>	2022-05-25 10:00	4.3K	
<a href="#">uniref.xsd</a>	2022-05-25 10:00	4.3K	
<a href="#">uniref90.dtd</a>	2022-05-25 10:00	1.7K	
<a href="#">uniref90.fasta.gz</a>	2022-05-25 10:00	33G	
<a href="#">uniref90.release.note</a>	2022-05-25 10:00	311	
<a href="#">uniref90.xml.gz</a>	2022-05-25 10:00	51G	

### Step 4. Run SPIDER3

```
#RUN  
./run_list.sh *.seq
```

You will see three files.

.pssm and .hhm and .spd3