

TLR8序列分析和结构预测

Sequence Analysis and Structure Prediction of TLR8

第四组： 银行 马妍妍 管哲

2012-6-8

报告内容

- TLRs简介
- TLR8的分析

TLRs简介

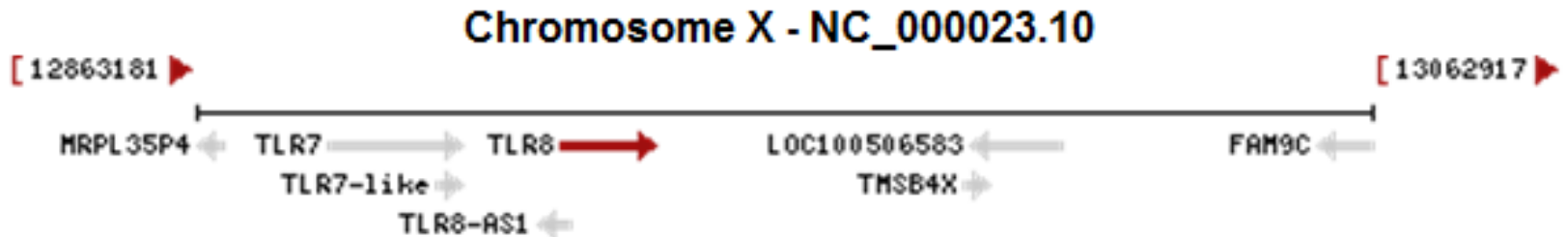
- **Toll like receptor** (toll样受体) 在固有免疫中起重要作用的一类膜蛋白家族。在人体细胞中, 这个家族包括了10个蛋白。
- 这个家族的蛋白包括了胞外的受体结构域, 跨膜结构域和胞内的**TIR**结构域。

TLR8的分析

- 位于X染色体上
- 有2个外显子
- TLR8全长具有1041个氨基酸

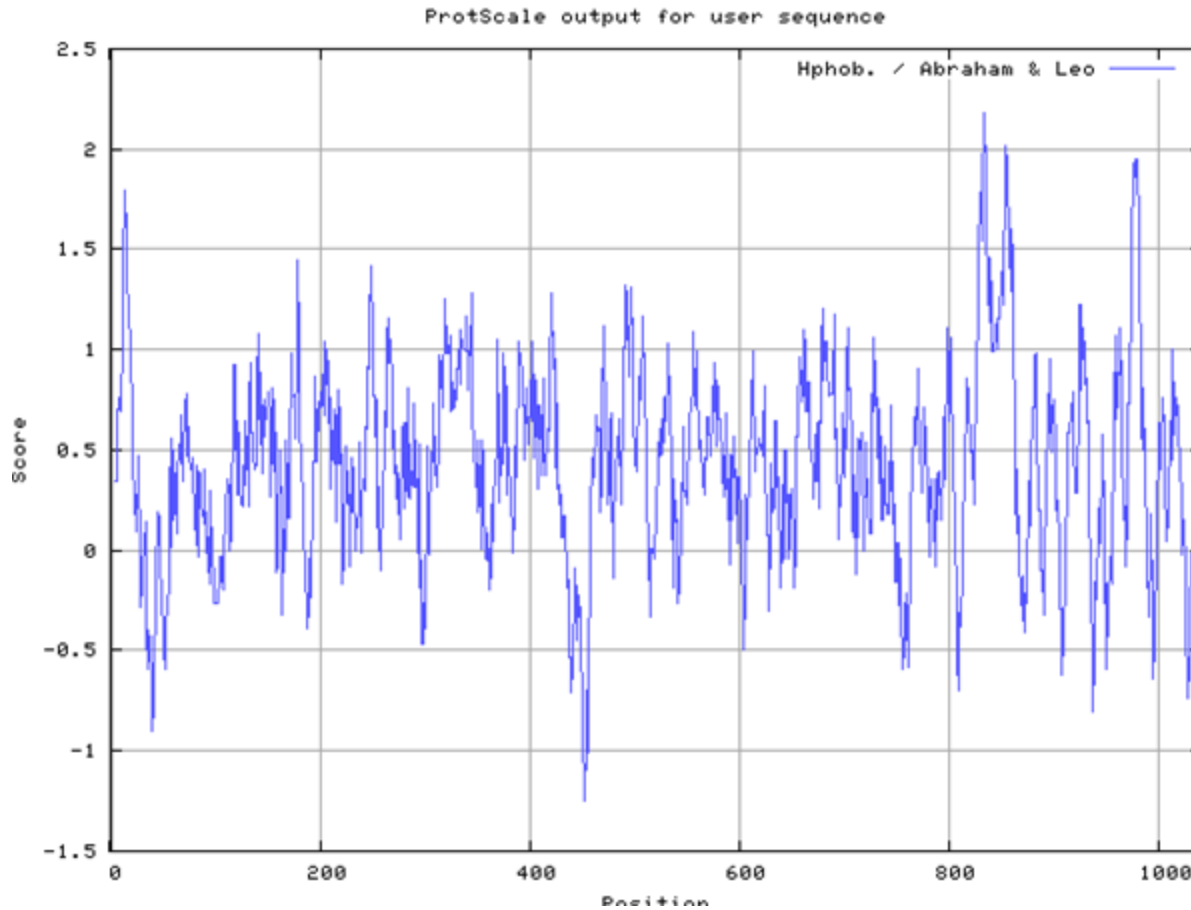
Location: Xp22

Sequence: Chromosome: X; NC_000023.10 (12924739..12941288)



- 蛋白性质分析
 - 疏水性分析
 - 信号肽的预测
 - 亚细胞定位
 - 跨膜分析
 - 糖基化修饰的预测
- 二级结构预测
- 3D结构预测
- 多序列比对及系统发育树的构建

蛋白性质分析

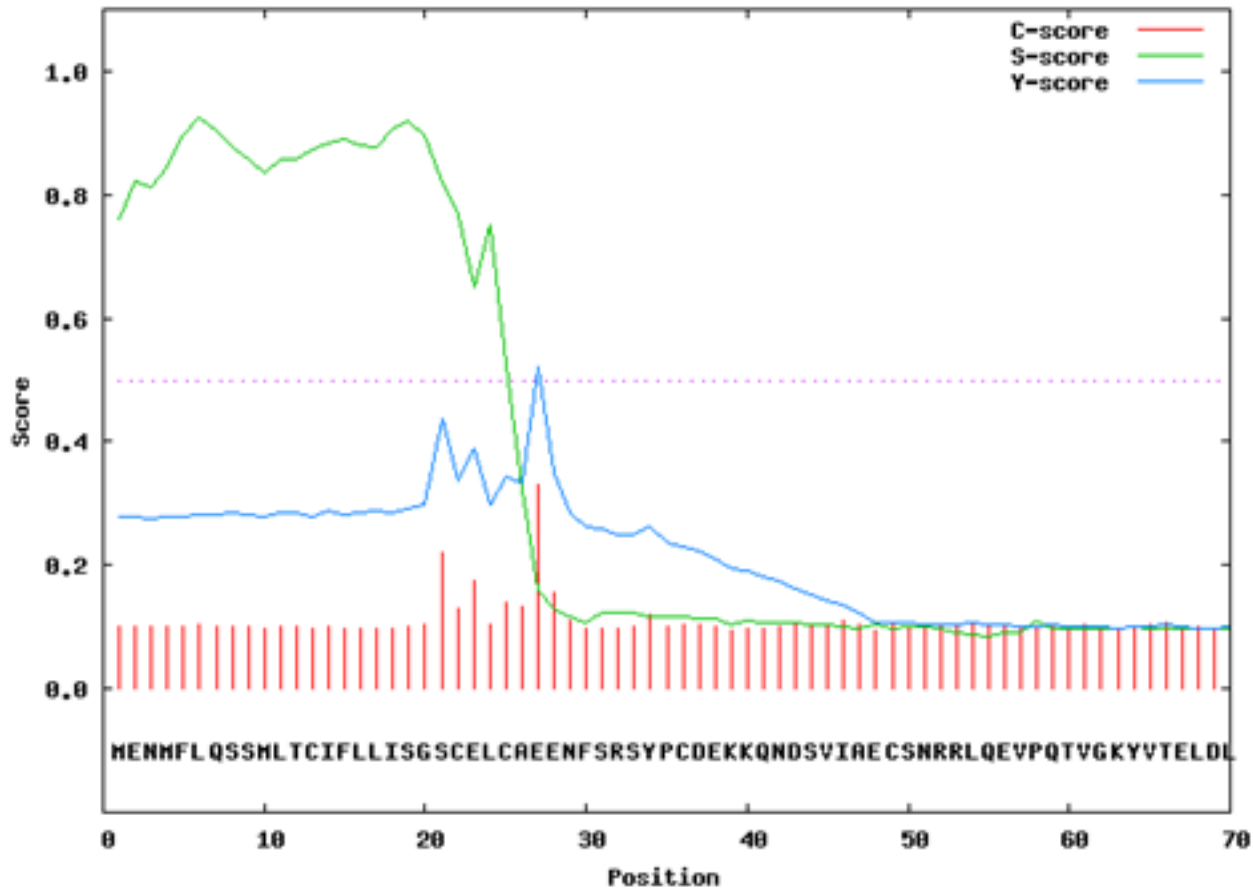


ProtScale疏水性分析：比较疏水

```
# SignalP-4.0 euk predictions
```

```
>Sequence
```

SignalP-4.0 prediction (euk networks): Sequence



```
# Measure Position Value Cutoff signal peptide?  
max. C 27 0.330  
max. Y 27 0.520  
max. S 6 0.924  
mean S 1-26 0.816  
D 1-26 0.680 0.450 YES
```

```
Name=Sequence SP='YES' Cleavage site between pos. 26 and 27: LCA-EE D=0.680 D-cutoff=0.450 Networks=SignalP-noTM
```

signalP 4.0 Server信号肽预测：1—26位氨基酸为信号肽

```
### targetp v1.1 prediction results #####  
Number of query sequences: 1  
Cleavage site predictions not included.  
Using NON-PLANT networks.
```

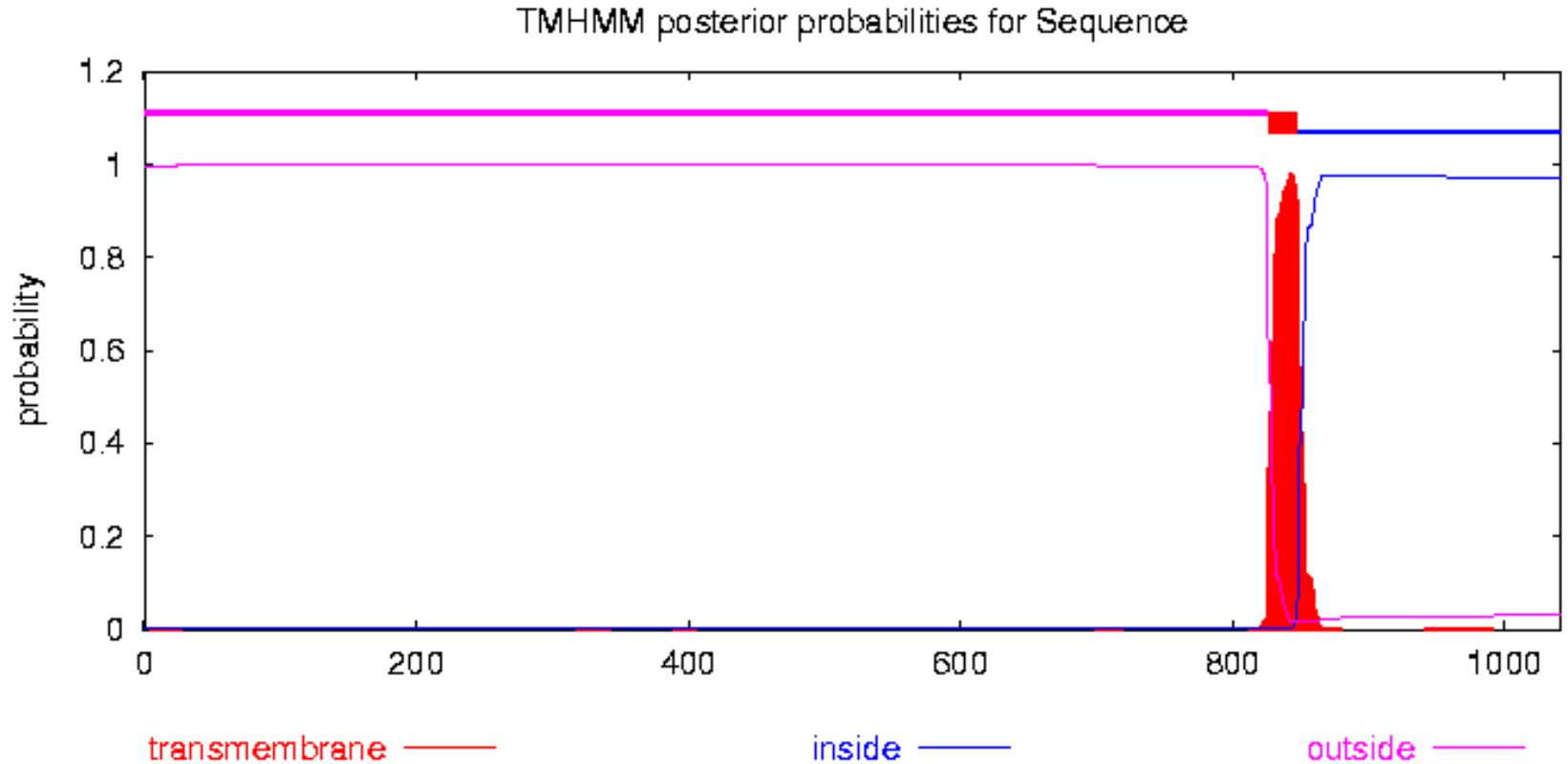
Name	Len	mTP	SP	other	Loc	RC
Sequence	1041	0.017	0.926	0.113	S	1
cutoff		0.000	0.000	0.000		

TargetP亚细胞定位预测：SP值最高，定位于分泌系统


```

# Sequence Length: 1041
# Sequence Number of predicted TMHs: 1
# Sequence Exp number of AAs in TMHs: 23.13784
# Sequence Exp number, first 60 AAs: 0.04149
# Sequence Total prob of N-in: 0.00310
Sequence      TMHMM2.0      outside      1      825
Sequence      TMHMM2.0      TMhelix      826      848
Sequence      TMHMM2.0      inside      849      1041

```



TMHMM预测跨膜螺旋：有一个跨膜螺旋

Asn-Xaa-Ser/Thr sequons in the sequence output below are highlighted in **blue**.
 Asparagines predicted to be N-glycosylated are highlighted in **red**.

```

Name: gi_20302168_ref_NP_619542.1 Length: 1041
MENMFLQSSMLTCIFLLISGSCELCAEENFSRSYPCDEKKQNSVIAECSNRRLQEVPPQTVGKYVTELDLSDNFITHITN 80
ESFQGLQNLTKINLNHNPNVQHONGNPGIQSNGLNITDGAFLNLKNLRELLLEDNQLPQIPSGLPESLTELSLIQNNIYN 160
ITKEGISRLINLNLYLAWNCYFNKVCEKTNIEDGVFETLTNLELLSLSFNSLSHVPPKLPSSLRKLFLSNTQIKYISEE 240
DFKGLINLTLDDLSGNCPRCFNAPFPCVPCDGGASINIDRFAFQNLTLQLRYLNLSSTSLRKINAAWFKNMPHLKVLDFEF 320
NYLVGEIASGAFLTMLPRLEILDLSFNFIKGSYPQHINISRNFSKLLSLRALHLRGYVFQELREDDFQPLMQLPNLS TIN 400
LGINFIKQIDFKLFQNFNSLEI IYLSENRISPLVKDTRQSYANSSSFQRHIRKRRSTDFEFDPHSNFYHFTRPLIKPQCA 480
AYGKALDLSLNSIFFIGPNQFENLPDIACLNLSANSNAQVLSGTEFSAI PHVKYLDLTNNRLDFDNASALTELS DLEVLD 560
LSYNSHYFRIAGVTHHLEFIQNFTNLKVLNLSHNNIYTLTDKYNLESKSLVELVFSGNRLDILWNDDDNRYISIFKGLKN 640
LTRLDLSLNRKHIPNEAFLNLPASLTELHINDNMLKFFNWTLLQQFPRELELLDRGNKLLFLTDSLSDFTSSLRTLLLS 720
HNRI SHLPSGFLSEVSSLKHLDLSSNLLKTINKSALETKTTTKLSMLELHGPNPECTCDIGDFRRWMDEHLNVKI PRLVD 800
VICASPGDQRGKSIVSLELTTCVSDVTAVILFFFITTMVMLAALAHHLFYWDVWFIYNVCLAKVKGYRSLSTSQT FY 880
DAYISYDTKDASVTDWVINELRYHLEESRDKNVLLCLEERDWD PGLAIDNMQSINQSKKTVFVLTKKYAKSWNFKTA F 960
YLALQRLMDENMDVIIIFILLEPVLQHSQYLRLRQRICKSSILQWPDNPKAEGLFWQTLRNVVLTENDSRYNNMYVDSIKQ 1040
Y
. . n . . . . . N . . . . . N . . . . . n . . . . . n . . . . . N 80
. . . . . N . . . . . n . . . . . n . . . . . n . . . . . n . . . . . N . . . . . n . . . . . nn . . . . . N 160
. . . . . n . . . . . n . . . . . n . . . . . n . . . . . n . . . . . n . . . . . n . . . . . n . . . . . 240
. . . . . N . . . . . n . . . . . n . . . . . n . . . . . n . . . . . N . . . . . N . . . . . 320
n . . . . . n . . . . . n . . . . . n . . . . . N . . . . . N . . . . . n . . . . . N . . . . . n 400
. . . . . n . . . . . N . . . . . n . . . . . n . . . . . n . . . . . N . . . . . n . . . . . n . . . . . 480
. . . . . n . . . . . n . . . . . n . . . . . N . . . . . n . . . . . nn . . . . . 560
. . . . . n . . . . . N . . . . . n . . . . . N . . . . . n . . . . . n . . . . . n . . . . . n . . . . . N 640
. . . . . n . . . . . n . . . . . n . . . . . n . . . . . N . . . . . n . . . . . n . . . . . 720
. n . . . . . n . . . . . n . . . . . n . . . . . n . . . . . n . . . . . n . . . . . n . . . . . 800
. . . . . n . . . . . n . . . . . n . . . . . n . . . . . n . . . . . n . . . . . n . . . . . 880
. . . . . n . . . . . n . . . . . n . . . . . n . . . . . n . . . . . n . . . . . n . . . . . 960
. . . . . n . . . . . n . . . . . n . . . . . n . . . . . n . . . . . n . . . . . n . . . . . 1040
. . . . . n . . . . . n . . . . . n . . . . . n . . . . . n . . . . . n . . . . . n . . . . . 1120

```

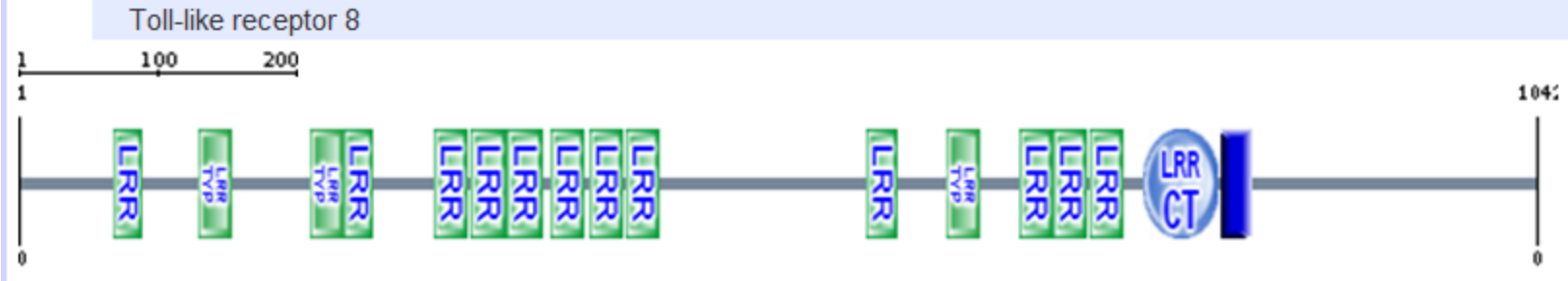
NetNGlyc 1.0server预测糖基化位点

garnier二级结构预测

```
. 10 . 20 . 30 . 40 . 50
MEMMFLQSSMLTCLIFLLISGSCELCAEENFSRSPYCDKQNDSVIAECS
helix HHHHHHHHHHH HHHHHH H H
sheet EEEEE EEEE
turns TTTT TTTT TTTTTT TT
coil C
. 60 . 70 . 80 . 90 . 100
NRRLQEVPTVGVKVTLDLSDNFITHITNESFQGLQNLTKINLNNPNV
helix HHHHHH H
sheet EEE EEEEE EE EEE E
turns TTTT TT T TTTT T T T
coil C CCCC C CCC C CCC
. 110 . 120 . 130 . 140 . 150
QHQCNGPCIQSNGLNITDGAFLNLKMLRELLLEDNQLPQIPSGLPESLTE
helix HHHHHHHHHHH HHH
sheet EE EEE EEEE E
turns TT T TT T
coil CCC CCCC CC CCCCCC
. 160 . 170 . 180 . 190 . 200
LSLIQNNIYNITKEGISRLINLKNLYLAWNCYFNKVCETNIEDGVFETL
helix HHHH HH HHH HHHH
sheet E EE EEEEE EEEEE
turns TTT TT T TTTTTTT T
coil CC CC CC CC
. 210 . 220 . 230 . 240 . 250
TNLELLSLSFNLSLHVPPKLPSSRLKFLSNTQIKYISEEDFKGLINLTL
helix HHHHHHHHH HHHHHHHHH
sheet EEEEE EEE EEE
turns TT TT TT T T
coil C CCC CC CC CCC
. 260 . 270 . 280 . 290 . 300
LDLSCNCPRCFNAPFCVPCDGGASINIDRFQNLTLQLRYLNLSSSTSLR
helix E EE EEEEE EEEE E
turns TTTT TTTTTTTT TTT TT T
coil C C C CC C CCCCCC
. 310 . 320 . 330 . 340 . 350
KINAAWFKNMPHLKVLDFNLYVGBIASGAFITMLPRLEILDLSFNVIK
helix HHHHHH HHHHHHHHHHHHHHHHHHH HHHHHHHH
sheet EE EE E E
turns TTTT
coil CCC C
. 360 . 370 . 380 . 390 . 400
GSYPQHINISRNFSKLLSLRALHLRGYVVFQELREDDFQPLMQLPNLSTIN
helix HHHHHH H HHH
sheet E EEEE EEEE EE EEEE
turns T T TTTT TT TTTT T T
coil CC C C C C CC C
. 410 . 420 . 430 . 440 . 450
LGINFIKIDFKLQNFNLEIIVLSENRI SFLVKDTRQSYANSSFPQRH
helix H HH
sheet EE EEEEE EEE EEEEE EE EEE E
turns T TTTT T T T T
coil C CCCCC C CC C CCC
. 460 . 470 . 480 . 490 . 500
IRKRRSTDFEFDPHSNFYHFRPLIKPQCAAYGKALDLSNSIFFIGPNQ
helix HHHHHHHH
sheet E EEE EEEEE
turns TTTT TT T TT TTTT TTT T
coil CC C C C C C CCCC
```

```
. 510 . 520 . 530 . 540 . 550
FENLPDIACLNLNSANSNAQVLSCTEFSIAPHVKYLDLTNNRLDFDNASAL
helix HH HHHH HHHH
sheet EEE EEEEE EEEEE
turns TT TT T TT TTT
coil CC CCCCC CCC C C C
. 560 . 570 . 580 . 590 . 600
TELSDLVELVLDLSYNSHYFRIACVTHHLEFIQNFTNLKVLMLSHNNIYTLT
helix HHHHHHHHHH HHHHHH
sheet EEEEE EEEEE EEEEE
turns TTTT T T
coil CCCCC CCC CCC
. 610 . 620 . 630 . 640 . 650
DKYNLESKSLVELVFGNRLDILWDDDRYISIFKGLKMLRLDLSLNR
helix HHHHHH H HHHHHH
sheet EEEE EEEE EEEEE EEEE
turns T TT TTTT
coil CC C CC C CCCC
. 660 . 670 . 680 . 690 . 700
LKHIPNEAFNLNLPASLTELHINDNMLKFFNWTLQQFPRLELLDLRGNKL
helix HHHH HHHHHHHHHHHH HH H HHHHHHHH
sheet E E
turns TT T T T T
coil CCC CC CCC CCC
. 710 . 720 . 730 . 740 . 750
LFLTDSLSDFTSSLRTLLSHNRISHLPSCFLSEVSSLKHLDLSSNLLKT
helix HHHH HHHHHHHHHHHHHHHHH
sheet EEE E
turns T TT TTTT T
coil CCC CCC CC CCC CC
. 760 . 770 . 780 . 790 . 800
INKSALETKTITTKLSMLELHGNPFECTCDIGDFRWMDBHLNVKIPRLVD
helix HHHHHHHHHHHHHHHH HH HHHHH H
sheet E E EEEEE
turns TTT TTT TTTT T TT
coil C C C C
. 810 . 820 . 830 . 840 . 850
VICASPGDQRGKSIVSLELTTCVSDVTAVILFFFTITTMVMLAALAH
helix HH HHHHHHHHHHHHHHHHH
sheet EEEE EEEEE EE EEEEE
turns TTTTTT TT
coil C
. 860 . 870 . 880 . 890 . 900
LFYWDVWFYINWVCLAKVKYRSLSTSQTFFDAYISYDTKDASVTDVWINE
helix H HHHH
sheet E EEEEE EEEEEEEEE EEE EE
turns TTTTT TTT TTTT T
coil CCCCC C
. 910 . 920 . 930 . 940 . 950
LRVHLEESRDKNVLLCLEERDWDPLAIDNLMQSQINQSKTVFVLTQKY
helix H HHHHHHHHHHHH HHH H HHHH HHHH
sheet EE EEEEE EEEE
turns TTT T T TT
coil C CC C C
. 960 . 970 . 980 . 990 . 1000
AKSWNFKTAFYLALQRLMDENMDVIIFILPEVLQHSQYLRRLQRICKSS
helix HH HHHHHHHHHHHHHHHH HH HHHH
sheet EEE EEE
turns TT TT TTT TTTTT
coil C CC C C
. 1010 . 1020 . 1030 . 1040
ILQWPDNPKAEGFLFWQTLRNVLTENDSRYNMYVDSIKQY
helix HHHH
sheet EE EEE EEEEE
turns T T TTTTT
coil CCCC CC CCCC
```

Domains within *Homo sapiens* protein TLR8_HUMAN (Q9NR97)



Name	Begin	End	E-value
LRR	66	85	3.55e+01
LRR_TYP	124	147	3.83e-02
LRR_TYP	200	223	2.05e-02
LRR	224	244	1.76e+02
LRR	286	309	7.79e+00
LRR	310	334	2.68e+01
LRR	336	360	1.58e+02
LRR	366	389	1.76e+02
LRR	393	416	1.19e+01
LRR	417	440	4.58e+01
LRR	583	605	1.45e+01
LRR_TYP	638	661	6.42e-04
LRR	687	710	3.65e+01
LRR	711	734	5.26e+00
LRR	735	758	4.97e+00
LRRCT	772	823	4.15e-08
transmembrane	826	848	-

SMART结构功能 域预测

3D结构预测

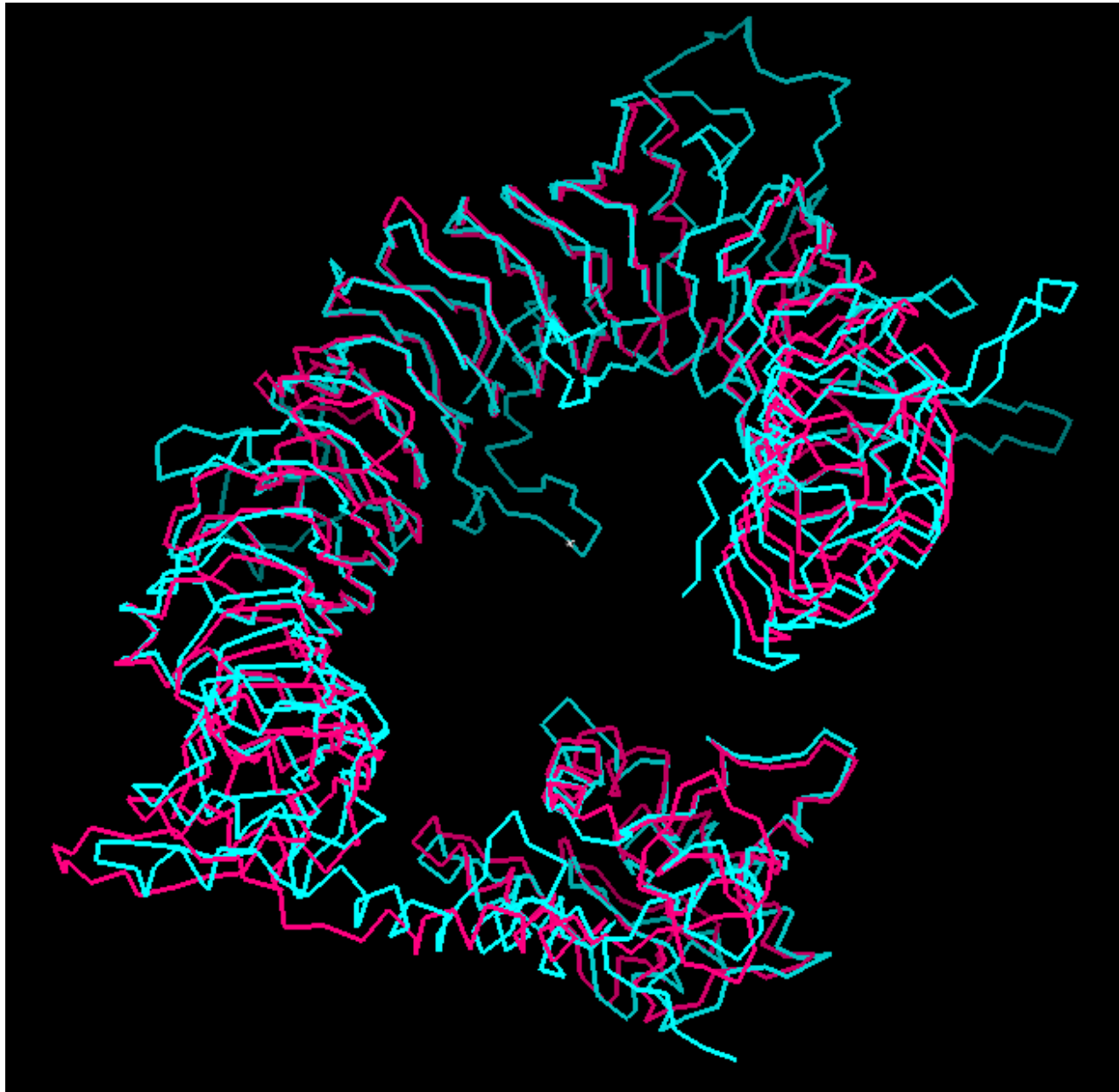


Swiss Model网站预测



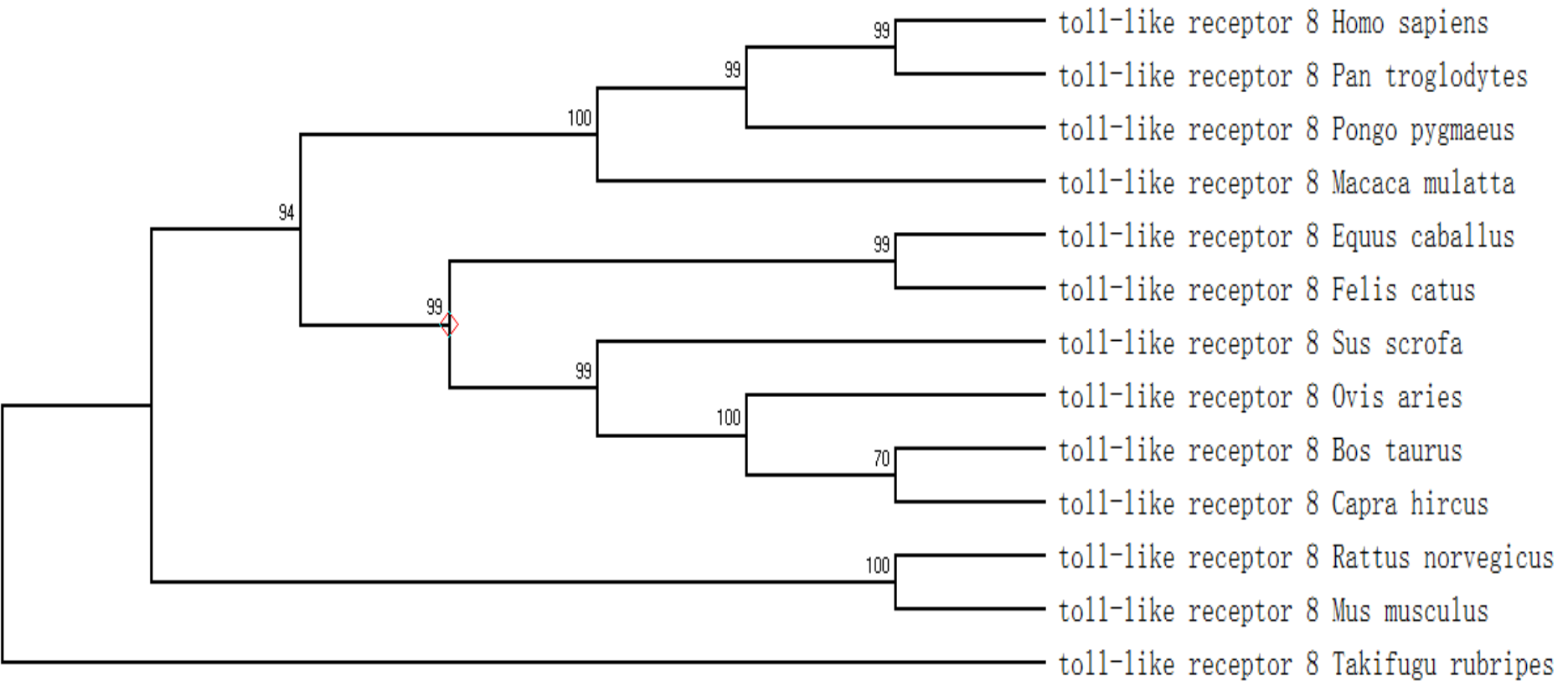
ESyPred3D网站预测

预测的结构对比



粉红色的是ESyPred3D网站预测，蓝色的的是Swiss Model 网站预测

系统发育树的构建



谢谢!

