

# 一个新发现的禾谷镰刀菌低毒 病毒的生物信息学分析

指导教师：罗静初

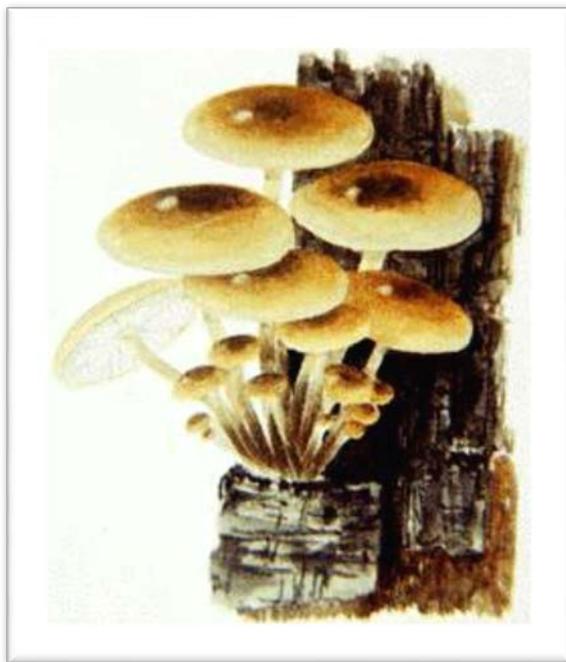
组员：王小奇 韩强 王颢潜 张易

2013.6.22

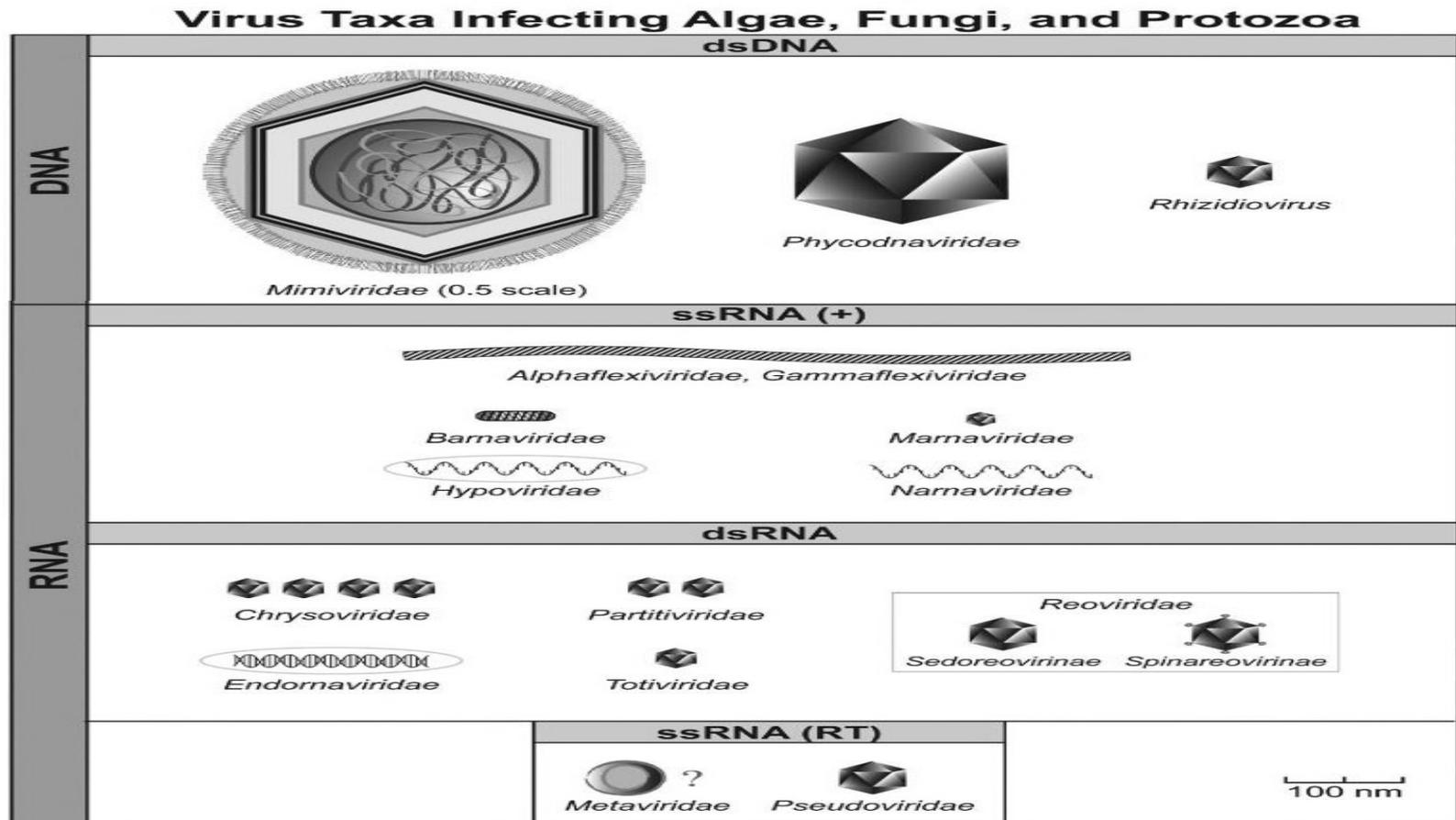
# 研究背景

- 真菌病毒的研究进展

真菌病毒（viruses of fungi）以真菌为宿主的病毒。1962年英国的霍林斯等在电子显微镜下从栽培蘑菇Agaricus中发现了与病害有关的3种病毒。



截止2009年，Ghabrial和Suzuki总结了已报道的182种真菌病毒，同时也总结了侵染植物病原真菌的46种病毒。



# 真菌病毒的研究并未受到重视

- 一、真菌病毒侵染不表现出明显的症状，因此不容易被人所发掘和重视，而且大部分的真菌学家在接触到菌落异常的菌株时不会想到可能是由于真菌病毒侵染导致的。
- 二、研究真菌病毒的病毒学家比较少。

# 传统防治的缺点及真菌病毒防治的优点

## 传统防治缺点：

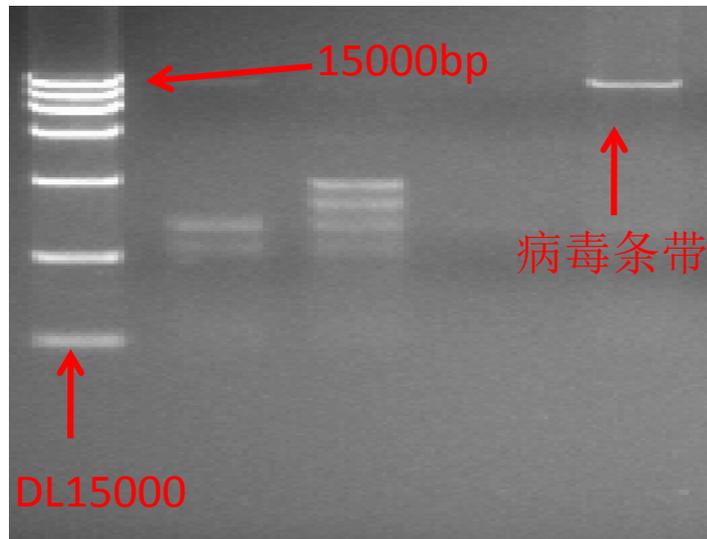
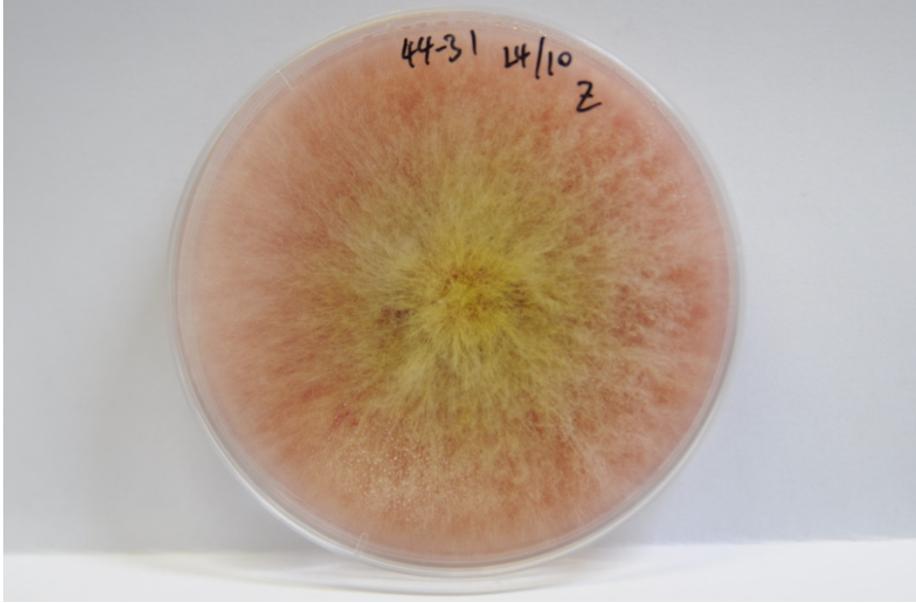
- 1、防效有限
- 2、易产生抗药性
- 3、环境不友好

## 真菌病毒防治优点：

除不具有以上缺点外还具有一下两点优点

- 1、具有广泛的生防应用潜力
- 2、可作为探究真菌和影响真菌与寄主植物的相互作用的工具

# 病毒的筛选



# 拟通过生物信息学手段解决的问题

- 鉴定新病毒
- 分析新病毒的进化关系
- 新病毒编码蛋白预测

# BLASTn结果

Search a nucleotide database using a nucleotide query

▶ NCBI/ BLAST/ blastn suite/ Formatting Results - WBASBN4N014

[Edit and Resubmit](#) [Save Search Strategies](#) ▶ [Formatting options](#) ▶ [Download](#)

 [How to read this page](#) [Blast report description](#)

## Nucleotide Sequence (11018 letters)

**Query ID** |d|32635  
**Description** None  
**Molecule type** nucleic acid  
**Query Length** 11018

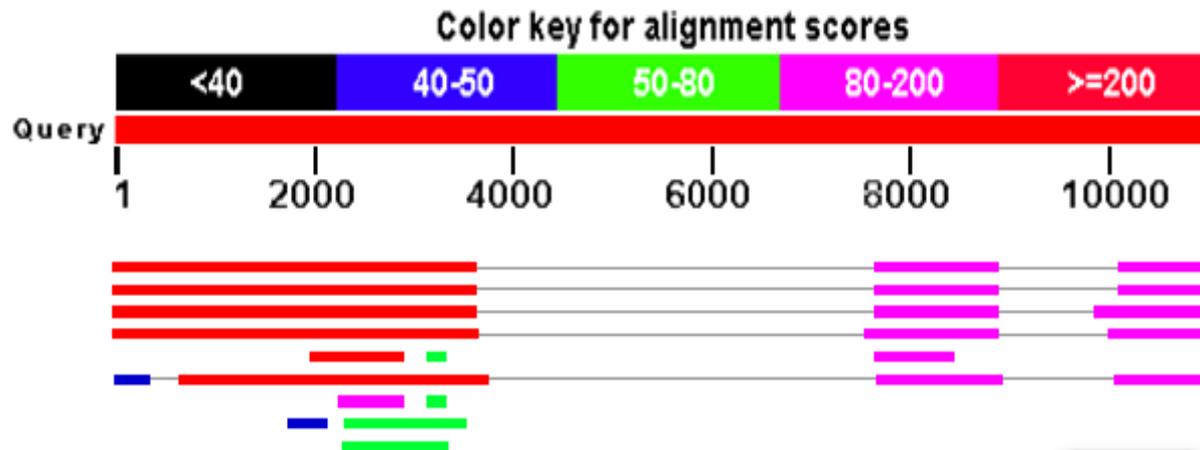
**Database Name** nr  
**Description** Nucleotide collection (nt)  
**Program** BLASTN 2.2.28+ ▶ [Citation](#)

 No significant similarity found. For reasons why, [click here](#)

Other reports: ▶ [Search Summary](#)

# BLASTX结果

Search protein database using a translated nucleotide query



	Description	Max score	Total score	Query cover	E value	Max ident	Accession
<input type="checkbox"/>	<a href="#">RecName: Full=ORFB polyprotein; Contains: RecName: Full=Papain-like protease p48; Contains: RecName: Full=Putative RNA-directed RNA</a>	320	627	53%	3e-84	31%	<a href="#">Q9YTU2.1</a>
<input type="checkbox"/>	<a href="#">hypothetical protein [Cryphonectria hypovirus 1] &gt;sp Q04350.1 POLB_CHPVE RecName: Full=ORFB polyprotein; Contains: RecName: Full=Pa</a>	318	624	53%	1e-83	31%	<a href="#">NP_041091.1</a>
<input type="checkbox"/>	<a href="#">ORFB [Cryphonectria hypovirus 1]</a>	316	622	55%	8e-83	29%	<a href="#">ABI64296.1</a>
<input type="checkbox"/>	<a href="#">polyprotein [Cryphonectria hypovirus 2] &gt;gb AAA20137.1 ORFB [Cryphonectria hypovirus 2-NB58]</a>	288	588	54%	2e-74	29%	<a href="#">NP_613266.1</a>
<input type="checkbox"/>	<a href="#">putative RNA helicase [Aqaricus bisporus virus X]</a>	201	201	8%	6e-53	35%	<a href="#">CAD19173.1</a>
<input type="checkbox"/>	<a href="#">hypothetical protein FgHV1gp2 [Fusarium graminearum hypovirus 1]</a>	204	497	52%	5e-49	29%	<a href="#">AGC75065.1</a>

# CHV1核酸序列翻译的蛋白质



# 与Chv1相关的蛋白结构域仅比对上了ORFB

**Conserved domains on** [gi|68619864|sp|Q9YTU2] [View full result](#) ?

RecName: Full=ORFB polyprotein; Contains: RecName: Full=Papain-like protease p48; Contains: RecName: Full=Putative RNA-directed RNA polymerase/helicase

**Graphical summary** [show options](#) ?

Query seq. 1 500 1000 1500 2000 2500 3000 3164

ATP binding site ▲  
putative Mg<sup>++</sup> binding site ▲

Specific hits Peptidase DUF3525 DEXDc

Superfamilies Peptidase DUF3525 DEXDc

Multi-domains DEXDc

[Search for similar domain architectures](#) ? [Refine search](#) ?

**List of domain hits** ?

Description	PssmId	Multi-dom	E-value
[+]DEXDc[cd00046], DEAD-like helicases superfamily. A diverse family of proteins involved in ATP-dependent ...	238005	yes	3.49e-08
[+]DUF3525 super family[cl13472], Protein of unknown function (DUF3525); This family of proteins is functionally uncharacterized. This protein is found	152474	no	0e+00
[+]Peptidase_C8[pfam03569], Peptidase family C8;	112389	no	7.03e-116
[+]DEXDc[smart00487], DEAD-like helicases superfamily;	214692	yes	3.20e-07

# 蛋白质翻译纵览

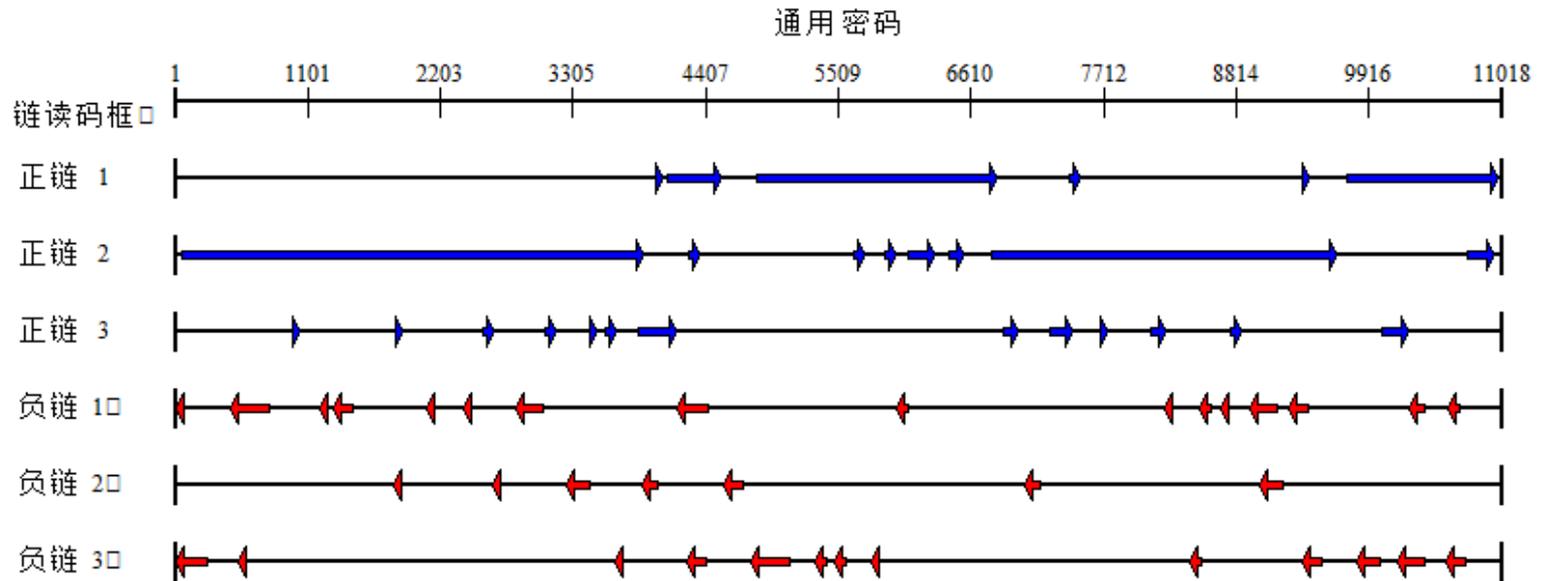


Fig. 由于目前测序工作还在进行中，因此整个测序结果并不准确，造成其具有多个蛋白结构域。

# 理化性质分析

根据蛋白质的序列我们利用expasy的ProtParam tool分析CHV1的理化性质  
(<http://web.expasy.org/protparam/>)

Ala (A) 228	7.2%	Arg (R) 209	6.6%
Asn (N) 116	3.7%	Asp (D) 181	5.7%
Cys (C) 38	1.2%	Gln (Q) 108	3.4%
Glu (E) 190	6.0%	Gly (G) 195	6.2%
His (H) 87	2.7%	Ile (I) 164	5.2%
Leu (L) 286	9.0%	Lys (K) 161	5.1%
Met (M) 79	2.5%	Phe (F) 153	4.8%
Pro (P) 200	6.3%	Ser (S) 187	5.9%
Thr (T) 157	5.0%	Trp (W) 81	2.6%
Tyr (Y) 127	4.0%	Val (V) 218	6.9%
Pyl (O) 0	0.0%	Sec (U) 0	0.0%

Estimated half-life: 半寿期

The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).

>20 hours (yeast, in vivo).

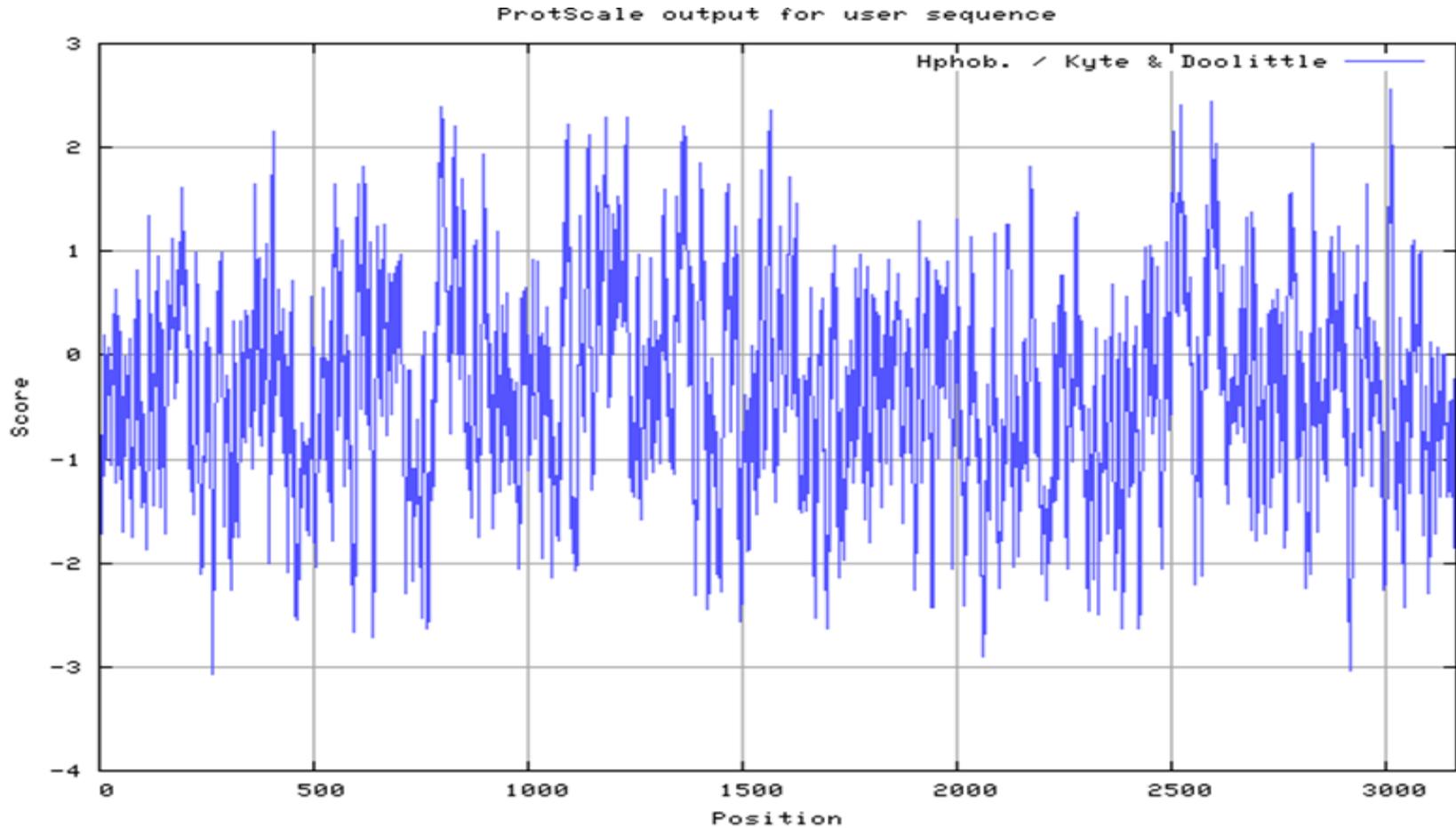
>10 hours (Escherichia coli, in vivo).

Instability index: 不稳定系数

The instability index (II) is computed to be 41.74

This classifies the protein as unstable.

# 蛋白质亲水性疏水性预测

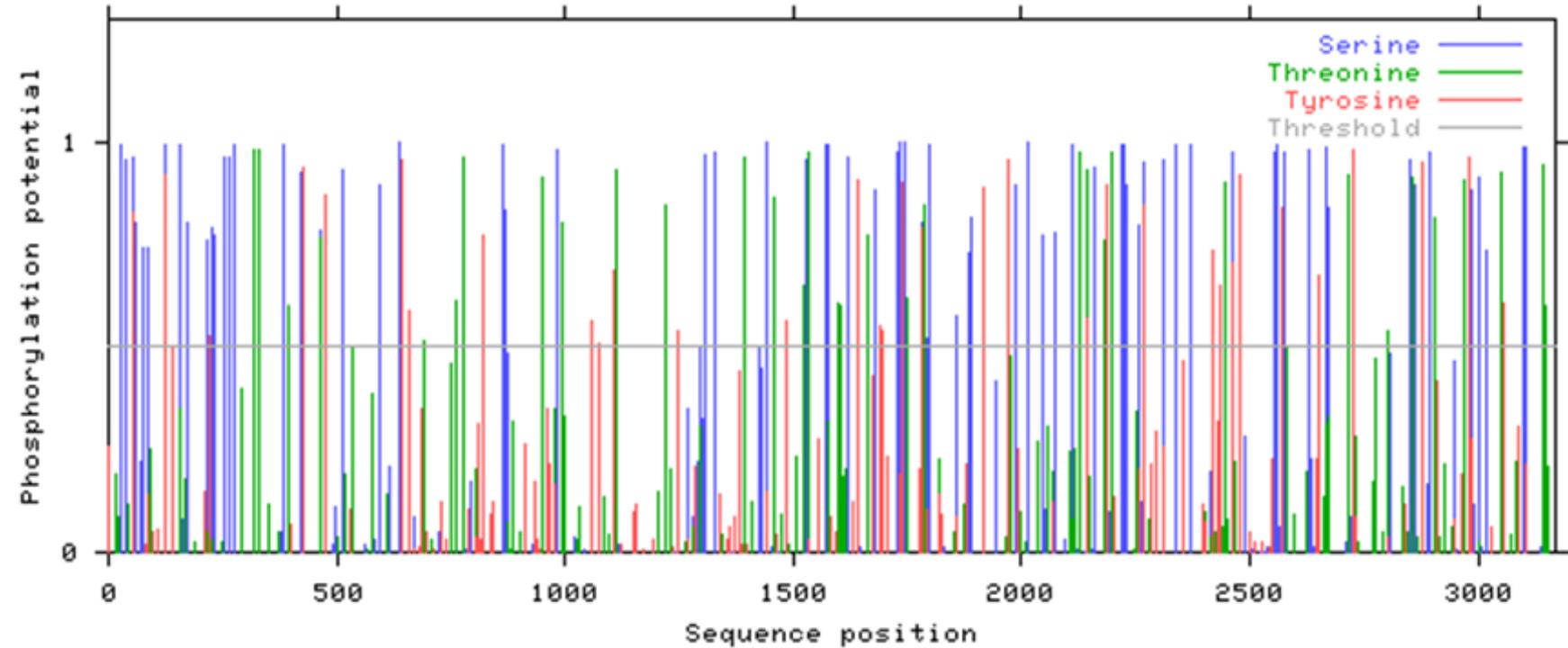


(<http://www.expasy.ch/tools/protscale.html>)

- 经ProtScale 分析， 蛋白疏水性最高分值为2.556， 亲水性最高分值为-3.087， 总的来说整条多肽链绝大多数氨基酸的分值较低， 表明亲水性氨基酸较均匀分布， 且多于疏水性氨基酸， 整条多肽链表现为亲水性氨基酸， 由此推测属亲水性蛋白

# 蛋白质磷酸位点的分析

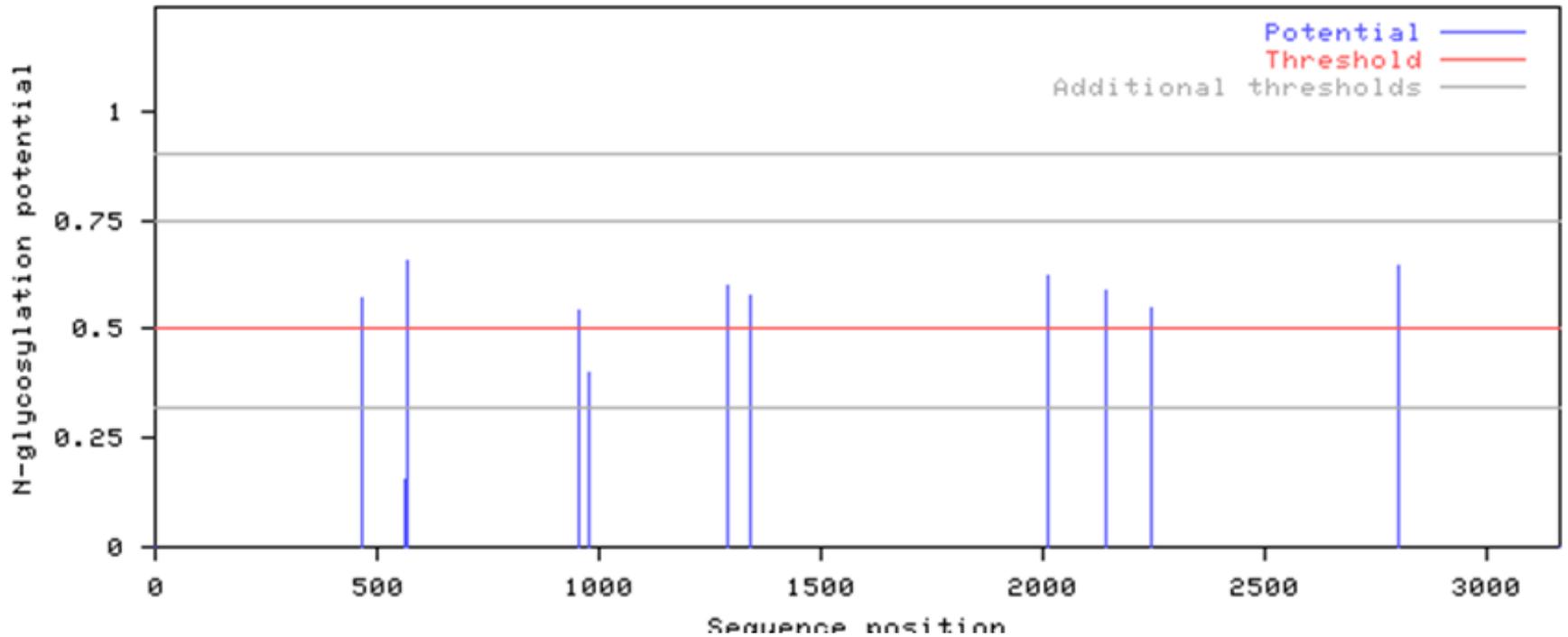
NetPhos 2.0: predicted phosphorylation sites in Sequence



蛋白质的磷酸位点 说明存在较多的磷酸位点  
(<http://www.cbs.dtu.dk/services/Net-Phos/>)

# 糖基化位点的分析

NetNGlyc 1.0: predicted N-glycosylation sites in Sequence



(Threshold=0.5)

SeqName	Position	Potential	Jury agreement	N-Glyc result
Sequence	464 NPST	0.5694	(6/9)	+ WARNING: PRO-X1.
Sequence	560 NCSP	0.1567	(9/9)	---
Sequence	566 NVTA	0.6567	(9/9)	++
Sequence	957 NATS	0.5447	(6/9)	+
Sequence	980 NRSG	0.4026	(7/9)	-
Sequence	1291 NPSE	0.6023	(8/9)	+ WARNING: PRO-X1.
Sequence	1342 NCTT	0.5744	(7/9)	+
Sequence	2009 NVTE	0.6227	(8/9)	+
Sequence	2144 NTSD	0.5884	(7/9)	+
Sequence	2243 NPTA	0.5460	(4/9)	+ WARNING: PRO-X1.
Sequence	2799 NLTV	0.6472	(9/9)	++

其中566 与 2799的氨基酸位点糖基化可能性最大

# 后续工作

- 得到全序列后利用生物信息学手段对其蛋白结构域进行准确分析
- 利用Mega等相关软件对该新病毒的进化关系进行分析
- 利用Mfold version等软件进行RNA的二级结构预测
- 双向电泳检测蛋白表达差异

谢谢指导