



禾谷孢囊线虫 (*Heterodera avenae*)  
*Ha34609*基因功能预测  
**Prediction of Ha34609 Gene Function in  
*Heterodera avenae***

汇报人：坚晋卓

G08小组成员： 坚晋卓 彭新红

李云卿 闫大琦

汇报时间：2017-12-13

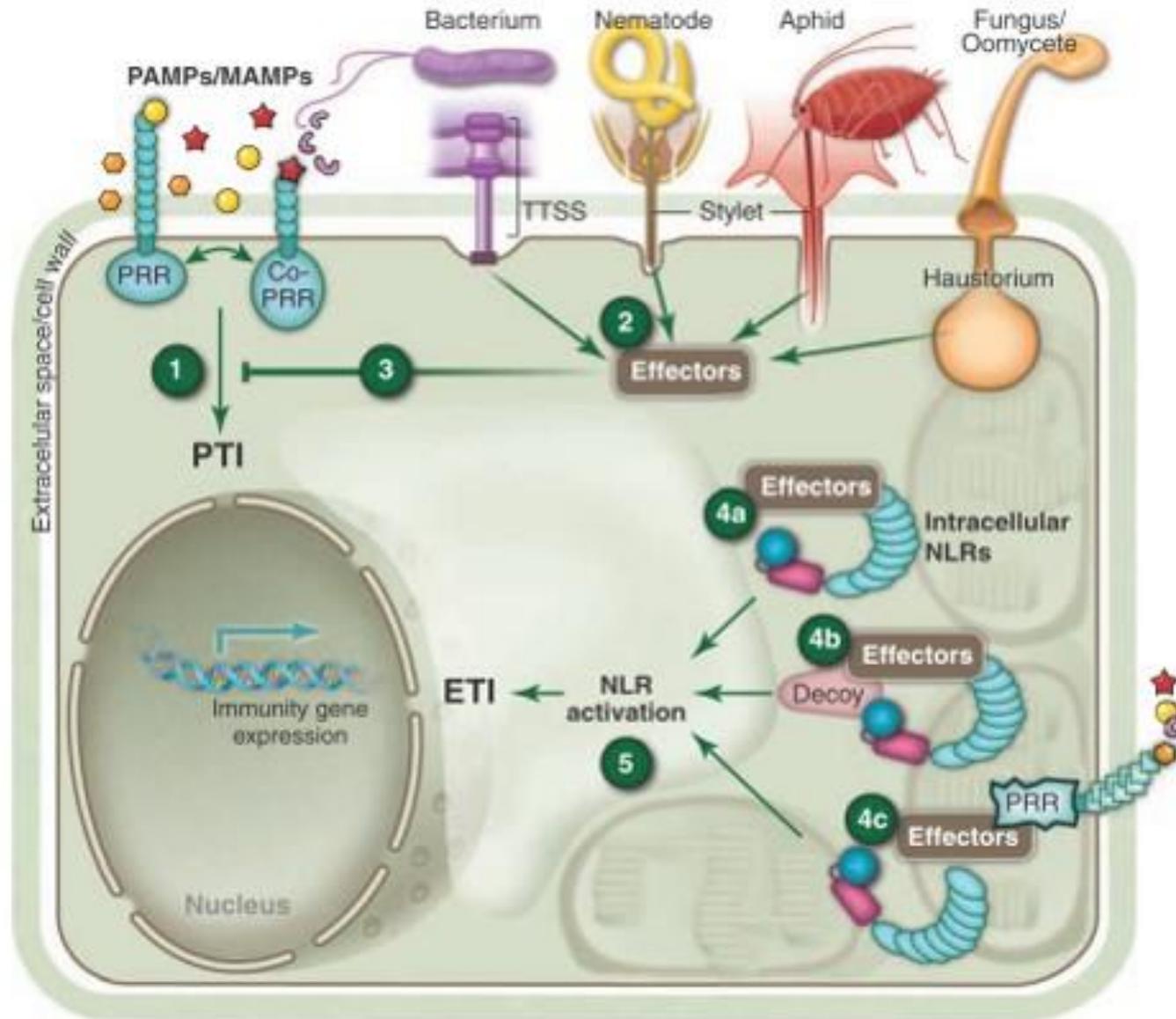


# 纲要

---

1. 研究背景
2. 组员课题相关
3. ORF片段预测
4. Signal peptide预测
5. 跨膜蛋白预测
6. 限制性内切酶位点分析
7. 植物表达载体质粒图谱分析
8. 用Primer5.0引物设计
9. 用DNAMAN检测引物特异性
10. 亚细胞定位的预测

# 1. 研究背景



# 1.研究背景

Beta-1,4 Endoglucanases  
pectate lyase  
Expansin  
cellulose binding protein

细胞壁修饰蛋白

Chorismate Mutase  
Annexin  
16A10  
PIN

细胞新城代谢与转录

细胞与靶标识别

Nuclear Localized  
Parasitism Proteins  
Ubiquitination

降低寄主防卫反应

14-3-3 Protein  
Venom Allergen Proteins  
Ran-Binding Protein  
Calreticulin

## 2.组员课题相关

---

G08D\_李云卿

棉花抗黄萎病相关基因的生物信息学分析及功能验证

对棉花抗病品种接种黄萎病菌VD080，在不同时间段取样、分析，筛选差异表达基因，并对候选基因进行后续的生物信息学分析和功能验证。

以候选基因Gh\_A06G1627为例进行结构和功能分析。

# Content

- ◎ 找相似序列——Blast
- ◎ 多序列比对——MEGA7
- ◎ 建进化树——MEGA7
- ◎ CDs序列转换成氨基酸序列——EMBOSS Transeq
- ◎ 分析蛋白结构——Protscale
- ◎ 预测蛋白模型——Phy2 ( Swiss -pdb Viewer )
- ◎ 预测功能模块——SMART
- ◎ 预测蛋白——Predict Protein
- ◎ 预测互做蛋白——STRING
- ◎ 引物设计——Primer 5

## 2.组员课题相关

G08A\_彭新红

哈茨木霉Th-33 *thga3* 基因的功能研究

### 研究目的及意义:

G蛋白信号系统在调控真菌的生长、产孢、拮抗和次生代谢等生物学过程中发挥着重要作用。GaIII亚基能调控真菌细胞壁相关酶活性，从而影响真菌的重寄生和拮抗能力，同时它还能影响真菌的产孢。因此，我们的研究拟采用基因敲除、回复突变技术，结合转录组测序数据分析，研究哈茨木霉GaIII亚基*thga3* 基因对木霉菌株生长、产孢、重寄生作用的影响，从而明确该基因的功能。

## 2. 组员课题相关

### 可能用到的相关软件

- ◆ 利用Primer5软件设计引物
- ◆ 利用DNAMAN软件和NCBI中ORF finder分析 *thga3* 的开放阅读框和 Blast 比对
- ◆ 利用ExPASy中的protScale分析编码蛋白质的疏水性
- ◆ 利用ExPASy中的Protparam进行氨基酸理化性质分析
- ◆ 利用signal 4.1分析翻译后信号肽
- ◆ 利用ExPASy中的TMPRED和 EMBOSS Explorer中的Tmap分析跨膜结构
- ◆ 利用MEGA 7.0构建系统发育树
- ◆ EMBOSS Explorer中的Rmap分析酶切位点。
- ◆ 利用Phyre2 和Swiss-PdbViewer预测和分析蛋白模型

## 2.组员课题相关

### G08A\_闫大琦

#### WRKYs基因功能研究

#### 研究背景

植物具有复杂的基因表达调控体系来抵御病原物的侵染，主要体现为转录因子在转录水平上对各种信号传导途径的调控。转录因子WRKY是一类与植物抗病防卫反应相关的转录因子，为植物中特有的超基因家族，该家族成员都含有高度保守的 WRKYGQK结构域和C2HH/C锌指结构基序，并通过与顺式作用元件W-box:(T)(T)TGAC(C/T)结合来调控靶基因的表达。已有大量研究证实，WRKY基因在植物的防卫反应中起着重要的调控作用。

## 2.组员课题相关

- 利用Blast找相似序列
- 利用DNAMAN软件比对序列
- 利用ExPASy中的protScale分析编码蛋白质的疏水性
- 利用signal 4.1分析翻译后信号肽
- EMBOSS Explorer中的Tmap分析跨膜结构
- 利用MEGA 7.0构建系统发育树
- 利用Phyre2 预测和分析蛋白模型

## G08B\_坚晋卓

### 禾谷孢囊线虫 (*Heterodera avenae*) *Ha34609*基因功能研究

植物寄生线虫效应子由线虫腺体细胞分泌产生，通过口针将这些蛋白质分泌到寄主细胞中从而促进自身的寄生作用。这些基因主要是在线虫分泌腺细胞中特异性表达，且多数具有信号肽，不含跨膜结构域。

本实验从实验室已构建的禾谷孢囊线虫转录组数据中，克隆出推定效应因子*Ha34609*，效应子对于线虫寄生作用至关重要，但其功能和致病机理我们现在却知之甚少。采用亚细胞定位研究该基因在植物体内的表达部位，以此完成禾谷孢囊线虫*Ha34609*基因的表达特性分析。

# 3. ORF 片段预测

开放性阅读框分析：利用NCBI的ORF Finder程序对cDNA序列做开放性阅读框分析

The image shows the NCBI homepage with a navigation menu on the left and a main content area. The 'Analyze' button is highlighted with a red box. The 'Analyze' button is located in the 'Develop' section, which is part of the 'Submit', 'Download', 'Learn', 'Develop', 'Analyze', and 'Research' grid.

**NCBI Home**  
Resource List (A-Z)  
All Resources  
Chemicals & Bioassays  
Data & Software  
DNA & RNA  
Domains & Structures  
Genes & Expression  
Genetics & Medicine  
Genomes & Maps  
Homology  
Literature  
Proteins  
Sequence Analysis  
Taxonomy  
Training & Tutorials  
Variation

**Welcome to NCBI**  
The National Center for Biotechnology Information advances science and health by providing access to biomedical and genomic information.  
[About the NCBI](#) | [Mission](#) | [Organization](#) | [NCBI News & Blog](#)

**Submit**  
Deposit data or manuscripts into NCBI databases

**Download**  
Transfer NCBI data to your computer

**Learn**  
Find help documents, attend a class or watch a tutorial

**Develop**  
Use NCBI APIs and code libraries to build applications

**Analyze**  
Identify an NCBI tool for your data analysis task

**Research**  
Explore NCBI research and collaborative projects

**Popular Resources**  
[PubMed](#)  
[Bookshelf](#)  
[PubMed Central](#)  
[PubMed Health](#)  
[BLAST](#)  
[Nucleotide](#)  
[Genome](#)  
[SNP](#)  
[Gene](#)  
[Protein](#)  
[PubChem](#)

**NCBI News & Blog**  
NCBI to assist in Southern California genomics hackathon in January  
30 Nov 2017  
From January 10-12, 2018, the NCBI will help with a bioinformatics hackathon in  
December 6th NCBI Minute: Keeping Current and Getting Help with NCBI Resources

# Analyze

NCBI provides a wide variety of data analysis tools that allow users to manipulate, align, visualize and evaluate biological data.

## Selected Analysis Tools

(UDAKI)

<a href="#">Conserved Domain Search Service (CD Search)</a>	Identifies the conserved domains present in a protein sequence
<a href="#">Digital Differential Display (DDD)</a>	Identifies genes with significantly different expression levels by comparing EST profiles
<a href="#">Electronic PCR (e-PCR)</a>	Identifies sequence tagged sites (STSs) within DNA sequences
<a href="#">Frequency-weighted Link (FLink)</a>	Links a group of records in a source database to a ranked list of associated records in a destination database based on frequency-weighted statistics
<a href="#">Gene Expression Omnibus (GEO) BLAST</a>	Finds regions of local similarity between query sequences and GenBank sequences included on microarray or SAGE platforms in the GEO database
<a href="#">Genetic Codes</a>	Displays the genetic codes for organisms in the Taxonomy database in tables and on a taxonomic tree
<a href="#">Genome BLAST</a>	Finds regions of local similarity between query sequences and genome sequences
<a href="#">Genome ProtMap</a>	Maps each protein from a COG or VOG back to its genome
<a href="#">Genome Remapping Service</a>	Projects genome annotation data from one genomic assembly to another
<a href="#">Genome Workbench</a>	Integrated application for viewing and analyzing sequence data
<a href="#">LinkOut</a>	Allows third parties to link directly from PubMed abstracts and other Entrez database records to relevant web-accessible resources beyond the Entrez system
<a href="#">Map Viewer</a>	Provides special browsing capabilities for genomic maps and assembled sequences for a subset of organisms
<a href="#">Open Reading Frame Finder (ORF Finder)</a>	Suggests potential open reading frames in a DNA sequence
<a href="#">OSIRIS</a>	Facilitates the assessment of multiplex short tandem repeat (STR) DNA profiles based on laboratory-specific protocols
<a href="#">Phenotype-Genotype Integrator (PheGenI)</a>	Finds human phenotype/genotype relationships with queries by phenotype, chromosome location, gene, and SNP identifiers.
<a href="#">Primer-BLAST</a>	Uses Primer3 to design PCR primers to a sequence template
<a href="#">ProSplign</a>	Computes alignments of proteins to genomic nucleotide sequences
<a href="#">PSSM Viewer</a>	Displays and manipulates PSSM matrices from CDD records and PSI-BLAST
<a href="#">PubChem Standardization Service</a>	Processes chemical structures into a standard representation used by PubChem
<a href="#">PubChem Structure Search</a>	Searches PubChem Compound by chemical structure, SMILES, InChI or Molecular Formula

ORFfinder

PubMed

Search

### Open Reading Frame Finder

ORF finder searches for open reading frames (ORFs) in the DNA sequence you enter. The program returns the range of each ORF, along with its protein translation. Use ORF finder to search newly sequenced DNA for potential protein encoding segments, verify predicted protein using newly developed SMART BLAST or regular BLASTP.

This web version of the ORF finder is limited to the subrange of the query sequence up to 50 kb long. Stand-alone version, which doesn't have query sequence length limitation, is available for [Linux x64](#).

Examples (click to set values, then click Submit button) :

- NC\_011604 Salmonella enterica plasmid pWES-1; genetic code: 11; 'ATG' and alternative initiation codons; minimal ORF length: 300 nt
- NM\_000059; genetic code: 1; start codon: 'ATG only'; minimal ORF length: 150 nt



#### Enter Query Sequence

Enter accession number, gi, or nucleotide sequence in FASTA format:

```
gaaaacaacacatttggttcattaattaggagtttaaatcaatttaaggcaattaatttgccttaattaaaa
ATGAGTTTATCAATATTTTTGTGGTTGTTACCCATCGTTTCCTTTACTACATCCGCCGTTTACATAAATCGTATCAACAAC
aataattggaaatgtgaaatgcttataaatactttttcccttttttgctttaatttttggaaatattcgtttatgatgaa
```

#### Choose Search Parameters

Minimal ORF length (nt): 75

Genetic code: 1. Standard

ORF start codon to use:

- "ATG" only
- "ATG" and alternative initiation codons
- Any sense codon

Ignore nested ORFs:

#### Start Search / Clear

Submit

Clear

# 3. ORF 片段预测

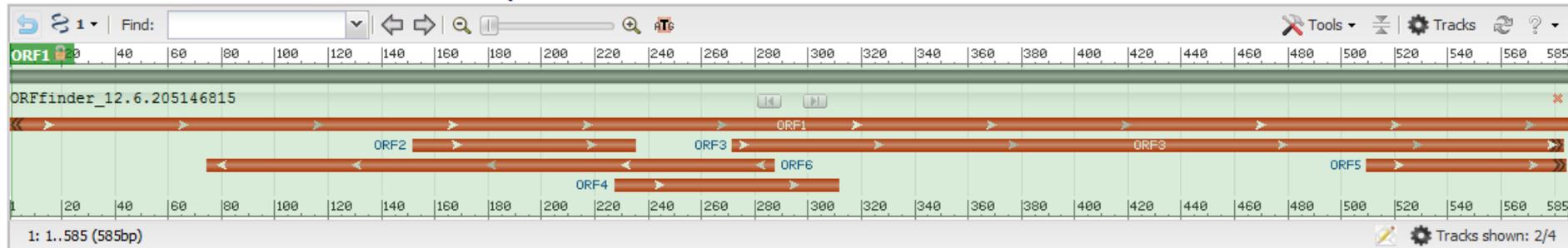
NCBI Resources  How To  Sign in to NCBI

ORFfinder PubMed  Search

## Open Reading Frame Viewer

### Sequence

ORFs found: 6 Genetic code: 1 Start codon: 'ATG' only



ORF1 (585 nt)

Display ORF as...

Mark

```
>1c1|ORF1 CDS
ATGAGTTTATCAATATTTTTGTGGTTGTTCCACCATCGTTTCCTTTACTAC
ATCCGCGCTTCACATAAAATCGTATCAACAACAGCAAAACAAACAGTGAGC
CGATGGTATTTGAAGGCGTTGAAAATGGGTGCGATTTAATCAAATGTTCA
AATGGCCAAACTTGGCGTATTCGAGTTGGTCTTGCAAATTTGGCGACAG
AGAAATTTGAGAAATTTGATTTTCCAAAATGTGTGACAAAGCGGAAT
TGAACCGAAACACGGACGACGATGGAATGGGCGCATTGTCCAGGATGGA
CCAGGATGTAATACTGTACACTGTAGCCACGGTTACAAAATGCCAAGTGCG
CATTTCAATTTCTAAGCTTGGCGATTTACCAGTATGCCCAATCAAATCGGAA
CCTTCCCTCAATGTGTGGSTCCAAAATGGCACATTTCAAACGCCAAGTTCA
ATTATCGAACAAAGCCCTGGGTGTGAAAAAATCCGACTAAATGTGAAGC
AGGCACAAAATGTGTACCCGCTGTTGGAATTCGGAATATGGAATTTTCG
AATGGTCTCAGTTCCTCTGGCCATTTTGCACCTTGA
```

Mark subset...

Marked: 0

Download marked set

as Protein FASTA

Label	Strand	Frame	Start	Stop	Length (nt   aa)
ORF1	+	1	<1	585	585   194
ORF3	+	2	272	>583	312   103
ORF6	-	2	287	75	213   70
ORF2	+	2	152	235	84   27
ORF4	+	3	228	311	84   27
ORF5	+	3	510	>584	75   24

Six-frame translation...

# 3. ORF 片段预测

NCBI Resources How To Sign in to NCBI

ORFfinder PubMed Search

### Open Reading Frame Viewer

Sequence

ORFs found: 8 Genetic code: 1 Start codon: 'ATG' only

1: 1..837 (837bp)

ORF3 (194 aa) Display ORF as... Mark

Mark subset... Marked: 0 Download marked set as Protein FASTA

Six-frame translation...

```
>1c1|ORF3
MSLSIFLWLFITIVSFTTSAASHKSYQQQQTINSEPMVFEGVENGCDLIKCS
NGQTCGIRVGLAKFGDREFEKDFPKCVTSKAELNRNIDDDGNGRIVIDG
PGCNTVHCSHGKQVRISSIKLGDLPYAQSIGTFPQCVPNGTFQTPSS
IIEQGPGEKLPFKCEAGTKCVTAVGIAGYGNLQWSQFFWPFCT
```

Label	Strand	Frame	Start	Stop	Length (nt   aa)
ORF3	+	2	74	658	585   194
ORF5	+	3	345	662	318   105
ORF7	-	1	360	148	213   70
ORF2	+	1	583	735	153   50
ORF1	+	1	301	384	84   27
ORF4	+	3	225	308	84   27

# 4.Signal peptide (信号肽) 预测

是引导新合成的蛋白质向分泌通路转移的短(长度5-30个氨基酸)肽链

打开abc



TOOLS



找到 [CBS tools](#)



SignalP 4.1 Server

CBS >> [CBS Prediction Servers](#) >> SignalP

### SignalP 4.1 Server

SignalP 4.1 server predicts the presence and location of signal peptide cleavage sites in amino acid sequences from different organisms: Gram-positive prokaryotes, Gram-negative prokaryotes, and eukaryotes. The method incorporates a prediction of cleavage sites and a signal peptide/non-signal peptide prediction based on a combination of several artificial neural networks.

View the [version history](#) of this server. All the previous versions are available on line, for comparison and reference.

**NEW (August 2017):** A book chapter on SignalP 4.1 has been published:

**Predicting Secretory Proteins with SignalP**  
Henrik Nielsen  
In Kihara, D (ed); *Protein Function Prediction* (Methods in Molecular Biology vol. 1611) pp. 59-73, Springer 2017.  
doi: [10.1007/978-1-4939-7015-5\\_6](#)  
PMID: [28451972](#)

FAQ	Article abstracts	Instructions	Output format	Performance	Data
-----	-------------------	--------------	---------------	-------------	------

#### SUBMISSION

Paste a single amino acid sequence or several sequences in [FASTA](#) format into the field below:

```
MSLSIFLWLFIVSFTISAASHRSYQQQINSEPMVFEVNGCDLIRCS  
IGQTCGIRVGLAKFGDREFEFKDFPKCVTSKAELNRNTDDGNGRIVTDG  
PGCNTVHC SHGYKQVRISISKLGDLPYAQSIGTFPQC VGPNGTFQTPSS  
IEQSPGCEKLPTRKCEAGTKCVIAVGI AKYGNLQWSQFFWPFCT
```

Submit a file in [FASTA](#) format directly from your local disk:  
浏览... 未选择文件.

**Organism group** ([explain](#))

- Eukaryotes
- Gram-negative bacteria
- Gram-positive bacteria

**Output format** ([explain](#))

- Standard
- Short (no graphics)
- Long
- All - SignalP-noTM and SignalP-TM output (no graphics)

**D-cutoff values** ([explain](#))

- Default (optimized for correlation)
- Sensitive (reproduce SignalP 3.0's sensitivity)
- User defined:  
 D-cutoff for SignalP-noTM networks  
 D-cutoff for SignalP-TM networks

**Method** ([explain](#))

- Input sequences may include TM regions
- Input sequences do not include TM regions

**Graphics output** ([explain](#))

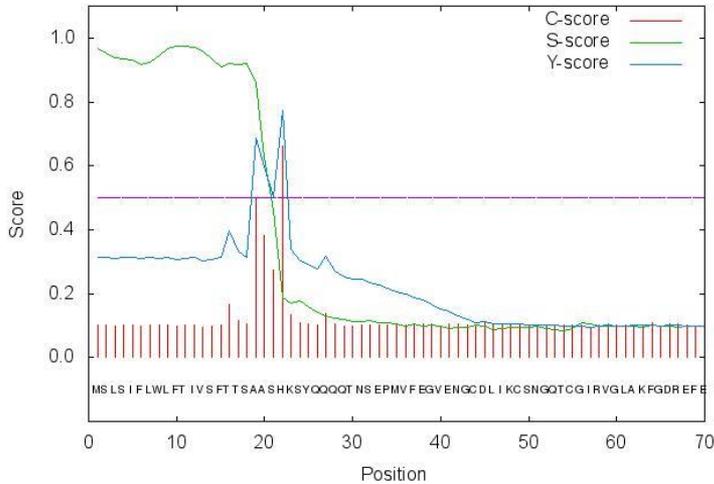
- No graphics
- PNG (inline)
- PNG (inline) and EPS (as links)

**Positional limits** ([explain](#))

- Minimal predicted signal peptide length. *Default: 10*
- N-terminal truncation of input sequence (0 means no truncation). *Default: Truncate sequence to a length of 70 aa*

# 4.Signal peptide预测

SignalP-4.1 prediction (euk networks): Sequence



Signal sequence prediction with scores over 3.50

```
>lcl|ORF3
MSLSIFLWLFTIVSFTTSAASHKSYQQQQQTNS EPMVFEV ENGCDLIKCS
NGQ from 1 to 70
```

Maximum score 6.6 at 19

Signal Sequence: MSLSIFLWLFTIVSFTTS  
Mature Sequence: AASHKSYQQQQQTNS EPMVFEV ENGCDLIK...

Other possible sites

Score 6.0 at 20

Score 5.1 at 21

Score 4.0 at 17

Score 3.7 at 15

# Measure	Position	Value	Cutoff	signal peptide?
max. C	22	0.661		
max. Y	22	0.773		
max. S	11	0.976		
mean S	1-21	0.900		
D	1-21	0.841	0.450	YES

Name=Sequence SP='YES' Cleavage site between pos. 21 and 22: AAS-HK D=0.841 D-cutoff=0.450 Networks=Signal

# [data](#)  
# [gnuplot script](#)

# 5.跨膜结构预测

打开abc



Tools



[TMHMM](#)

CENTER FOR BIOLOGICAL SEQUENCE ANALYSIS CBS

EVENTS NEWS RESEARCH GROUPS

STAFF CONTACT ABOUT CBS

[CBS](#) >> [CBS Prediction Servers](#) >> TMHMM

## TMHMM Server v. 2.0

### Prediction of transmembrane helices in proteins

---

#### SUBMISSION

Submission of a local file in [FASTA](#) format (HTML 3.0 or higher)

未选择文件。

OR by pasting sequence(s) in [FASTA](#) format:

```
CATTTC AATTTCTAAGCTTGGCGATTTACCGTATGCCCAATCAATCGGAA
CCTTCCCTCAATGTGTGGGTCCAAATGGCACATTTCAAACGCCAAGTTCA
ATTATCGAACAAAGGCCCTGGGTGTGAAAACTTCCGACTAAATGTGAAGC
AGGCACAAAATGTGTCACCGCTGTTGGAATTGCGAAATATGGAAATTTGC
AATGGTCTCAGTTCCTTCTGGCCATTTGCACTTGA
```

**Output format:**

Extensive, with graphics  
 Extensive, no graphics  
 One line per protein

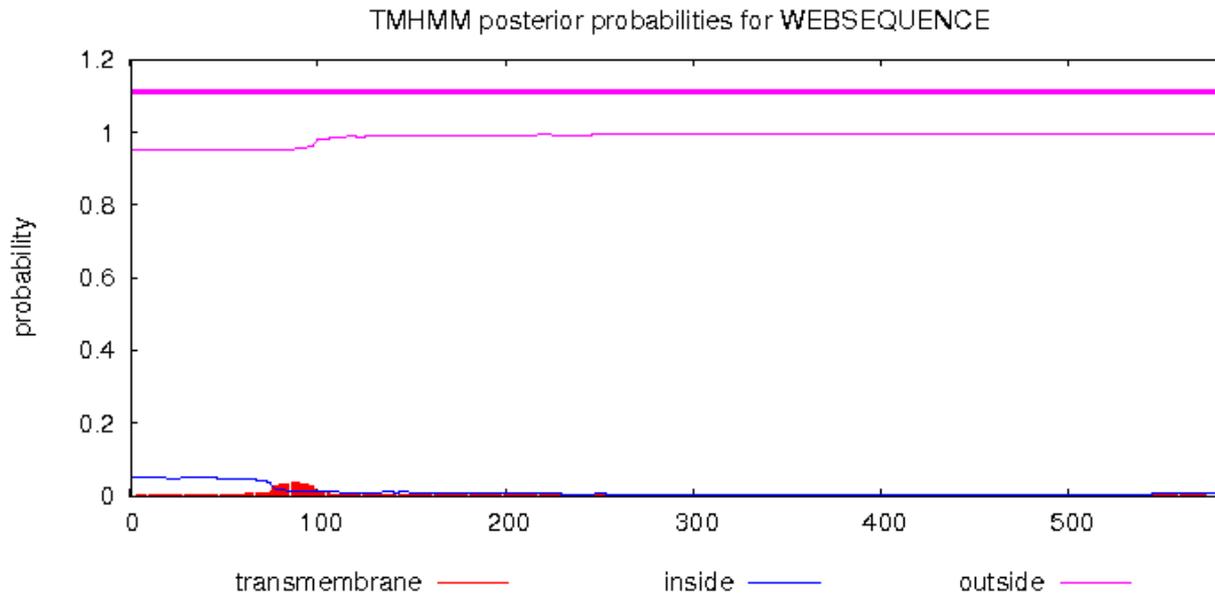
**Other options:**

# 5.跨膜结构预测

## TMHMM result

[HELP](#) with output formats

```
# WEBSEQUENCE Length: 585
# WEBSEQUENCE Number of predicted TMHs: 0
# WEBSEQUENCE Exp number of AAs in TMHs: 1.44465
# WEBSEQUENCE Exp number, first 60 AAs: 0.10711
# WEBSEQUENCE Total prob of N-in: 0.04904
WEBSEQUENCE TMHMM2.0 outside 1 585
```

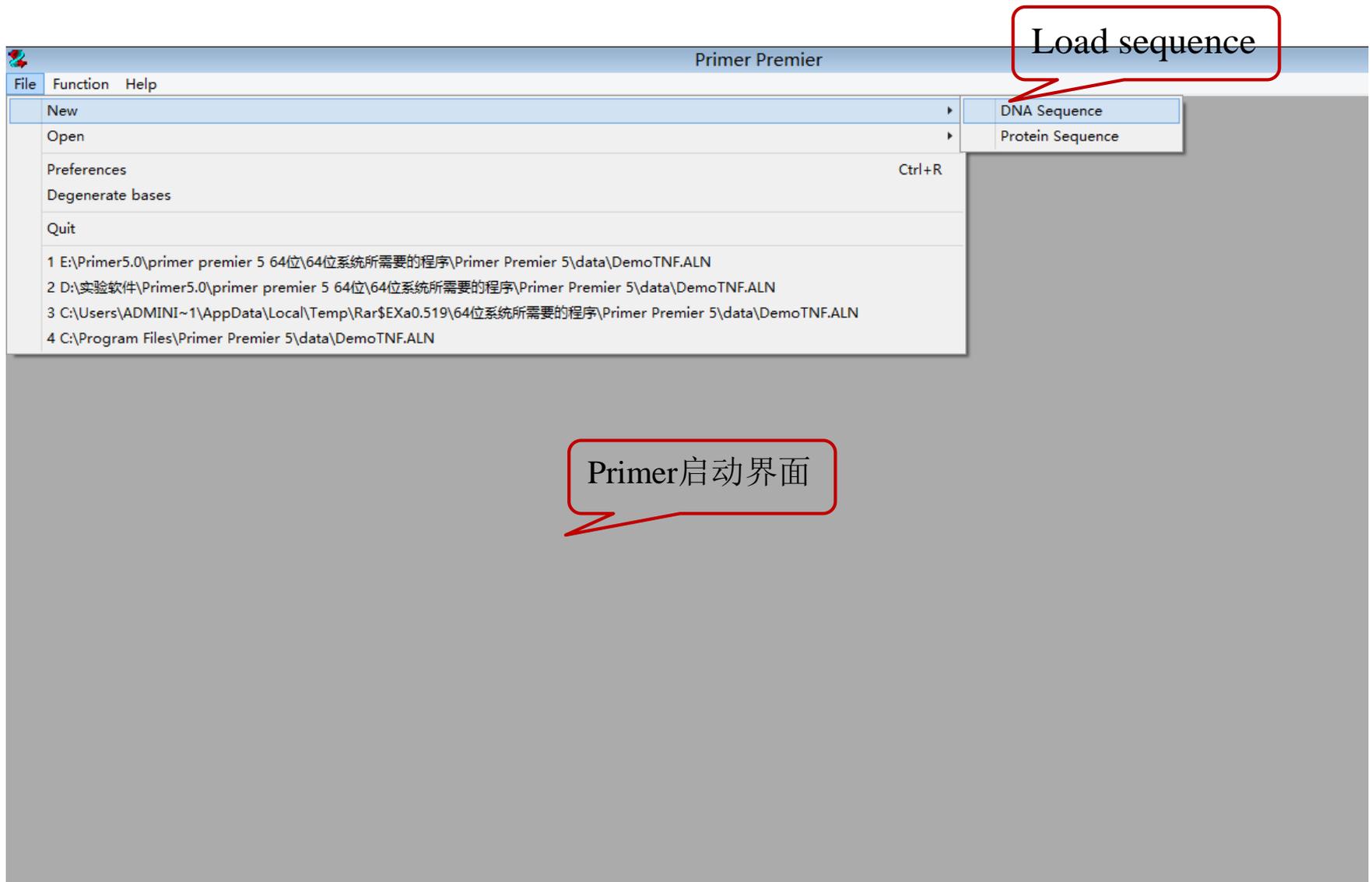


# [plot](#) in postscript, [script](#) for making the plot in gnuplot, [data](#) for plot

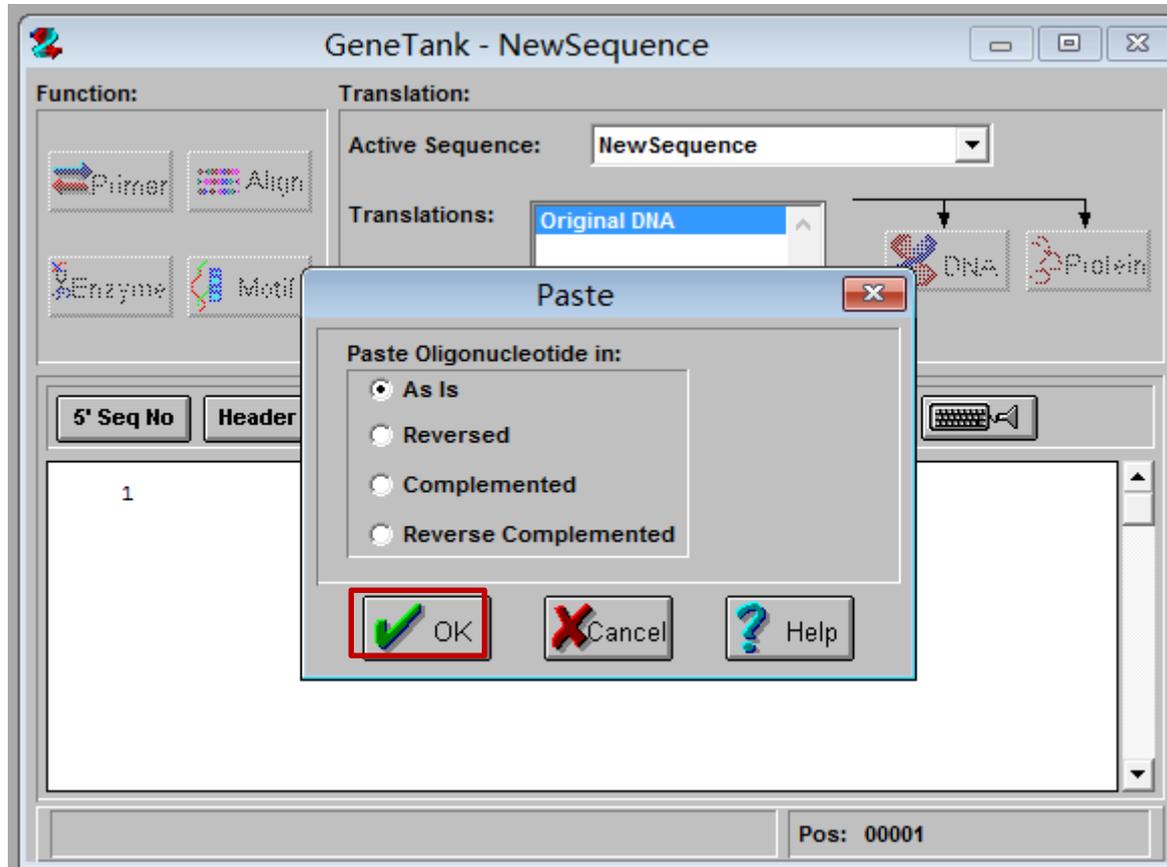
不具有跨膜结构域，说明该基因产物可能被线虫分泌到细胞外部

# 6.限制性内切酶位点分析

用primer5.0去分析目的片段上的酶切位点



## 6. 限制性内切酶位点分析



# 6.限制性内切酶位点分析

The screenshot shows the GeneTank - NewSequence software interface. The 'Function:' panel on the left contains buttons for 'Primer', 'Align', 'Enzyme' (highlighted with a red box), and 'Motif'. The 'Translation:' panel on the right shows 'Active Sequence: NewSequence' and a 'Translations:' list with 'Original DNA' selected. Below these panels is a control bar with buttons for '5' Seq No', 'Header', '3', '10', 'Find', 'Find Next', 'S', 'A', 'dsDNA', and a keyboard icon. The main text area displays a DNA sequence with line numbers 61 to 541. The sequence ends with 'CTTGA|'. The status bar at the bottom right shows 'Pos: 00586'.

Function: Primer, Align, Enzyme, Motif

Translation: Active Sequence: NewSequence

Translations: Original DNA

5' Seq No Header 3 10 Find Find Next S A dsDNA

```
61 TCACATAAAT CGTATCAACA ACAGCAAACA AACAGTGAGC CGATGGTATT TGAAGGCGTT
121 GAAAATGGGT GCGATTTAAT CAAATGTTCA AATGGGCAAA CTTGCGGTAT TCGAGTTGGT
181 CTTGCAAAAT TTGGCGACAG AGAATTTGAG AAATTTGATT TTCCAAAATG TGTGACAAGC
241 AAAGCGGAAT TGAACCGAAA CACGGACGAC GATGGAAATG GGCGCATTGT CACGGATGGA
301 CCAGGATGTA ATACTGTACA CTGTAGCCAC GGTACAAAT GCCAAGTGCG CATTTC AATT
361 TCTAAGCTTG GCGATTTACC GTATGCCCAA TCAATCGGAA CCTCCCTCA ATGTGTGGGT
421 CCAAATGGCA CATTTC A AAC GCCAAGTTCA ATTATCGAAC AAGGCCCTGG GTGTGAAAAA
481 CTTCCGACTA AATGTGAAGC AGGCACAAAA TGTGTCACCG CTGTTGGAAT TGCGAAATAT
541 GGAAATTTGC AATGGTCTCA GTTCTTCTGG CCAATTTGCA CTTGA|
```

Pos: 00586

## 6.限制性内切酶位点分析

GeneTank - NewSequence

Function: Translation:

Restriction Enzyme Analysis

Search Range: From: 1 to: 585

Show enzymes which occur <=

1  2  3  4  5  6

All Enzymes:

- Bal228I
- BalI
- BamHI**
- BamNxi
- BanI
- BanII
- BanIII
- BavAI
- BavAll
- BavBI
- BavBII
- BavI

Selected Enzymes:

- AvaI
- BstVI
- EcoRI
- HindIII
- NotI
- PstI
- SmaI
- TaqI
- XbaI
- XhoI
- XmaI

5' Seq No

61

121

181

241

301

361

421

481

541

Protein

STT

EGT

AGC

BGA

ATT

EGT

AAA

TAT

Add

Del

Edit

Filter

OK

Cancel

Help

## 6. 限制性内切酶位点分析

GeneTank - NewSequence

Function: Translation:

Active Sequence: NewSequence

Primer Align

### Restriction Sites

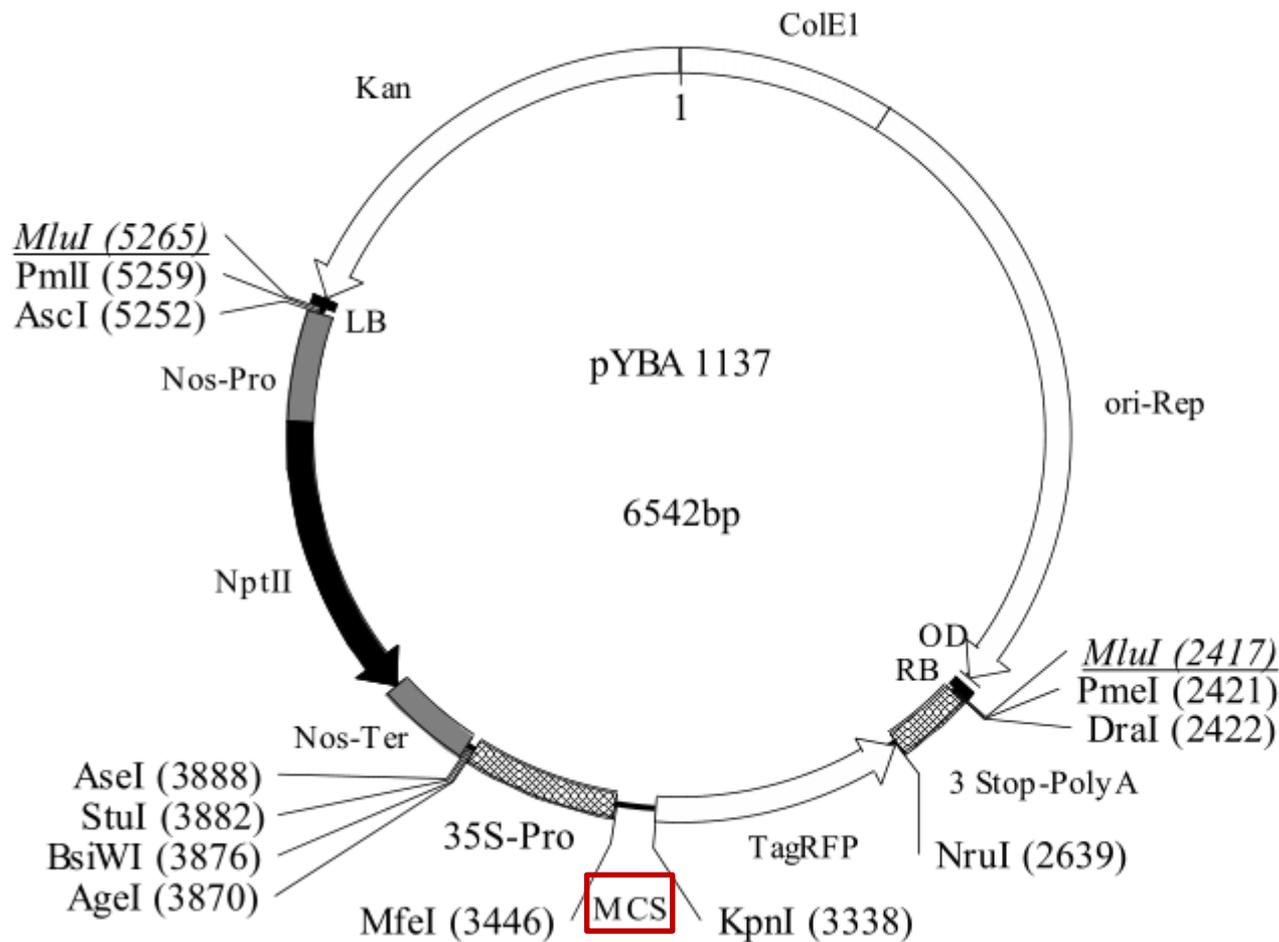
#	Name	Recognition Site	No	Pos
1	HindIII	A <sup>^</sup> AGCTT	1	365
2	TaqI	T <sup>^</sup> CGA	2	172,456

Table Seq Map Non-Cutters

541 GGAAATTG C AATGGTCTCA GTTCTTCTGG CCATTTGCA CTGA

Pos: 00586

# 7. 植物表达载体质粒图谱分析



- KpnI* (*Acc65I*)
- ApaI* (*PspOMI*)
- EcoO109I*
- XhoI*
- PspXI*
- SalI*
- ClaI*
- HindIII*
- EcoRV*
- EcoRI*
- PstI*
- SmaI* (*XmaI*)
- BamHI*
- SpeI*
- XbaI*
- NotI*
- AleI*
- SacI*
- MfeI*

# 7. 植物表达载体质粒图谱分析



## 7. 植物表达载体质粒图谱分析

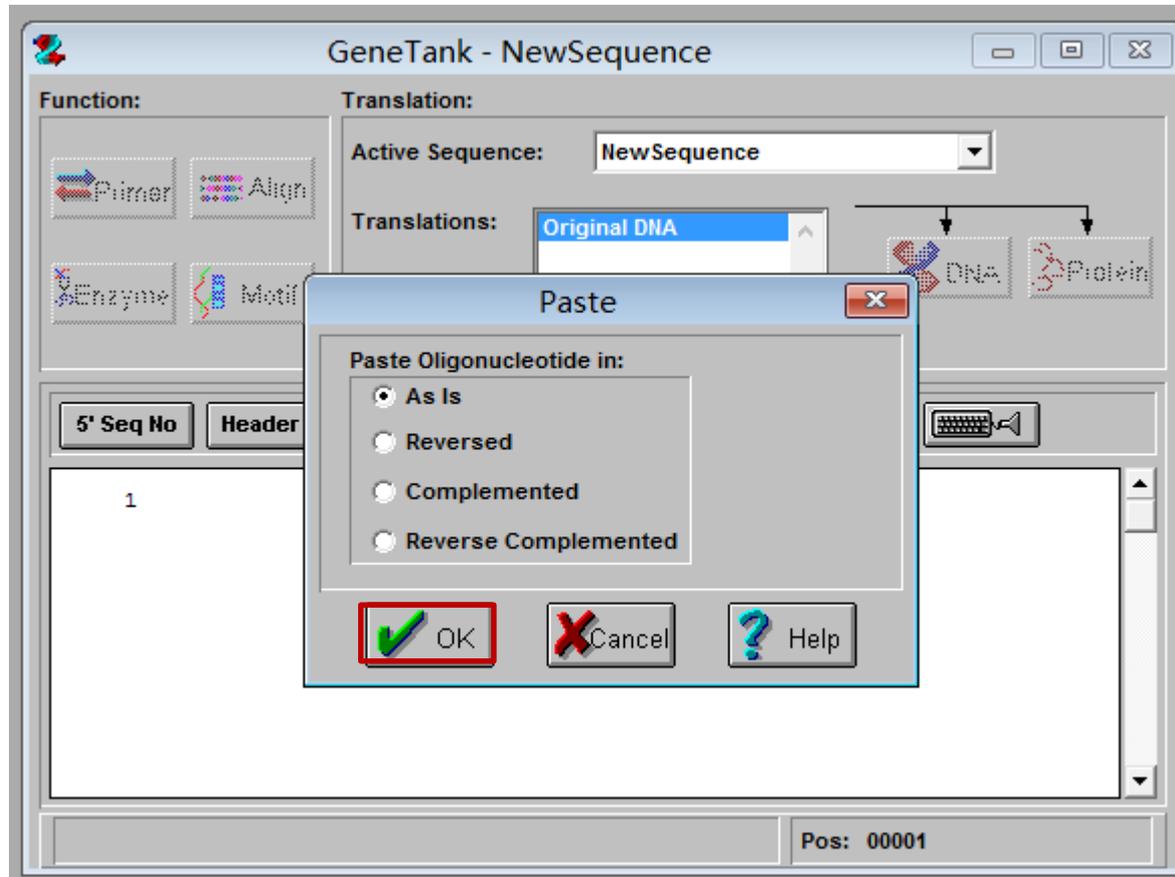
PCR 设计引物时酶切位点的保护碱基表 2

酶	寡核苷酸序列
BamH I	CGGATCCG CGGGATCCCG CGCGGATCCGCG
Xho I	CCTCGAGG CCCTCGAGGG CCGCTCGAGCGG

引物的结构就是（5→3）：

保护碱基+酶切位点+引物序列

## 8. 用Primer5.0引物设计



## 8. 用Primer5.0引物设计

The screenshot shows the GeneTank - NewSequence software interface. The window title is "GeneTank - NewSequence".

**Function:**

- Primer (highlighted with a red box)
- Align
- Enzyme
- Motif

**Translation:**

- Active Sequence: NewSequence
- Translations: Original DNA
- Buttons: DNA, Protein

**Navigation and Editing:**

- 5' Seq No
- Header
- 3 10
- Find
- Find Next
- S A
- dsDNA
- Speaker icon
- Keyboard icon

**DNA Sequence:**

```
61 TCACATAAAT CGTATCAACA ACAGCAAACA AACAGTGAGC CGATGGTATT TGAAGGCGTT
121 GAAAATGGGT GCGATTTAAT CAAATGTTCA AATGGGCAAA CTTGCGGTAT TCGAGTTGGT
181 CTTGCAAAAT TTGGCGACAG AGAATTTGAG AAATTTGATT TTCCAAAATG TGTGACAAGC
241 AAAGCGGAAT TGAACCGAAA CACGGACGAC GATGGAAATG GGCGCATTGT CACGGATGGA
301 CCAGGATGTA ATACTGTACA CTGTAGCCAC GGTTACAAAT GCCAAGTGCG CATTCAATT
361 TCTAAGCTTG GCGATTTACC GTATGCCCAA TCAATCGGAA CCTTCCCTCA ATGTGTGGGT
421 CCAAATGGCA CATTTCAAAC GCCAAGTTCA ATTATCGAAC AAGGCCCTGG GTGTGAAAAA
481 CTTCCGACTA AATGTGAAGC AGGCACAAAA TGTGTCACCG CTGTTGGAAT TGCGAAATAT
541 GGAAATTTGC AATGGTCTCA GTTCTTCTGG CCATTTTGCA CTTGA
```

**Pos: 00586**

# 8. 用Primer5.0引物设计

Primer Premier

Primer:

Direct Select:

3' TACTCAAATAGTTATAAAAACACCA 5'  
|||||  
5' ATGAGTTTATCAATATTTTTGTGGTTGTTCCACCATCGTTTCCTTTACTACATCCGCCGCTTCACATAAATCGTATCAACAAC 3'

M S L S I F L W L F T I V S F T T S A A S H K S Y Q Q Q

	Rating	Seq No	Length	Tm [°C]	GC%	Δ G [kcal/mol]	Activity [μg/OD]	Degeneracy	Ta Opt [°C]
Sense	87	1	25	54.8	24.0	-41.1	31.6	1	--
Anti-sense	87	25	25	54.8	24.0	-41.1	29.6	1	--
Product	32	--	25	58.9	24.0	--	--	--	32.7

	Hairpin	Dimer	False Priming	Cross Dimer	No Hairpins Found
Sense	None	Found	None	Found	<input type="button" value="All"/>
Anti-sense	Found	Found	None		

## 8. 用Primer5.0引物设计

Search Criteria

Search For:

PCR Primers     Sequencing Primers     Hybridization Probes

Search Type:

Sense Primer     Compatible with Sense Primer  
 Anti-sense Primer     Compatible with Anti-sense Primer  
 Both     Pairs

Search Ranges:

Sense Primer:  to

Anti-sense Primer:  to

PCR Product Size:  bp to  bp

Primer Length:  bp ?  bp

Search Mode:  Automatic     Manual

Search Parameters

# 8. 用Primer5.0引物设计

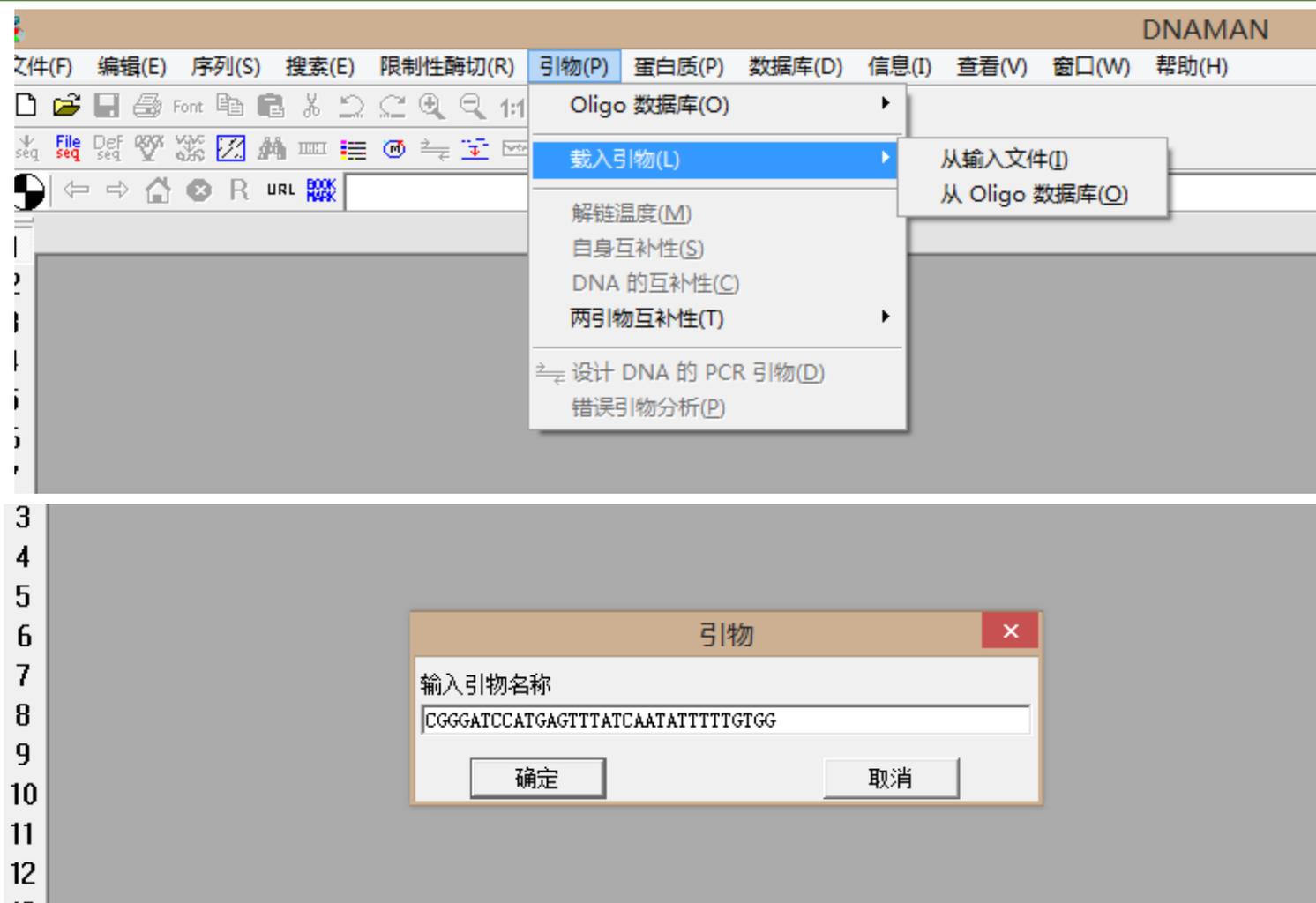
The screenshot displays the Primer Premier 5.0 software interface. The main window shows a DNA sequence with a primer pair highlighted. The sequence is: 3' **CGGATCC**ATGAGTTTATCAATATTTTGTGG 5' and 5' TTCCAAAATGTGTGACAAGCAAAGCGGAATTGAACCGAAACACGGACGACGATGGAAATGGCCGATTGTCACGGATGGACC 3'. The primer pair is (1) 95-113 and 250-267. The search results window shows 100 pairs found. The detailed primer table is as follows:

	Rating	Seq No	Length	Tm [°C]	GC%	Δ G [kcal/mol]	Activity [μg/OD]	Degeneracy	Ta Opt [°C]
Sense	100	95	18	51.2	50.0	-34.0	32.2	1	--
Anti-sense	100	250	17	51.0	47.1	-33.8	35.9	1	--
Product	100	--	156	82.4	39.7	--	--	--	48.0

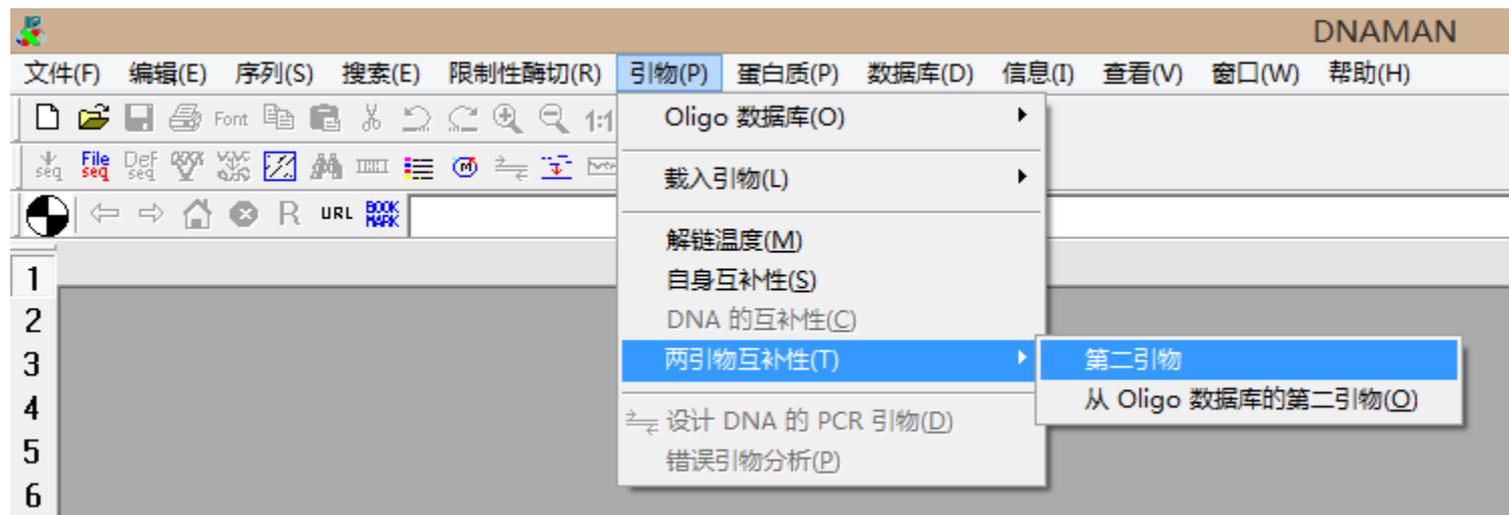
Below the table, there are buttons for Hairpin, Dimer, False Priming, and Cross Dimer, all set to None. A message box states "No Hairpins Found".

SP-Ha34609F(BamH1):CG**GGATCC**ATGAGTTTATCAATATTTTGTGG  
SP-Ha34609R(Xho1): CC**CTCGAG**AGTGCAAAATGGCCAGAAGA  
NSP-Ha34609F(BamH1):CG**GGATCC**ATGCATAAATCGTATCAACAACAG  
NSP-Ha34609R(Xho1):CC**CTCGAG**AGTGCAAAATGGCCAGAAGA

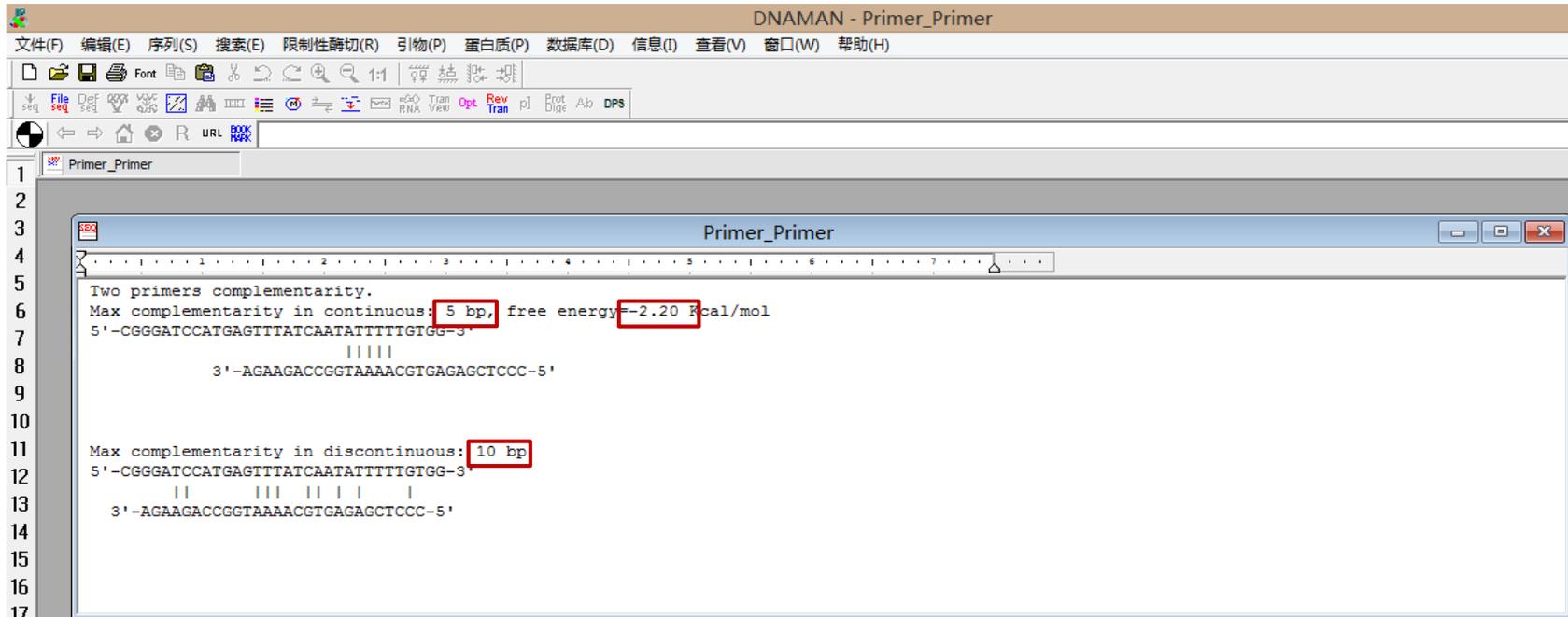
# 9.用DNAMAN检测引物特异性



# 9.用DNAMAN检测引物特异性



# 9.用DNAMAN检测引物特异性



The screenshot shows the DNAMAN software interface. The main window displays the following text:

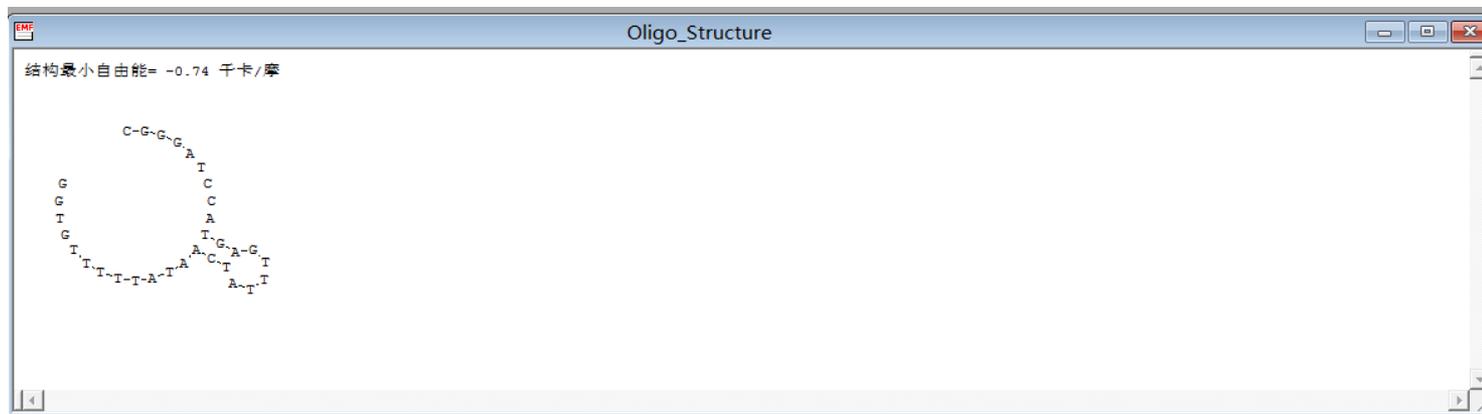
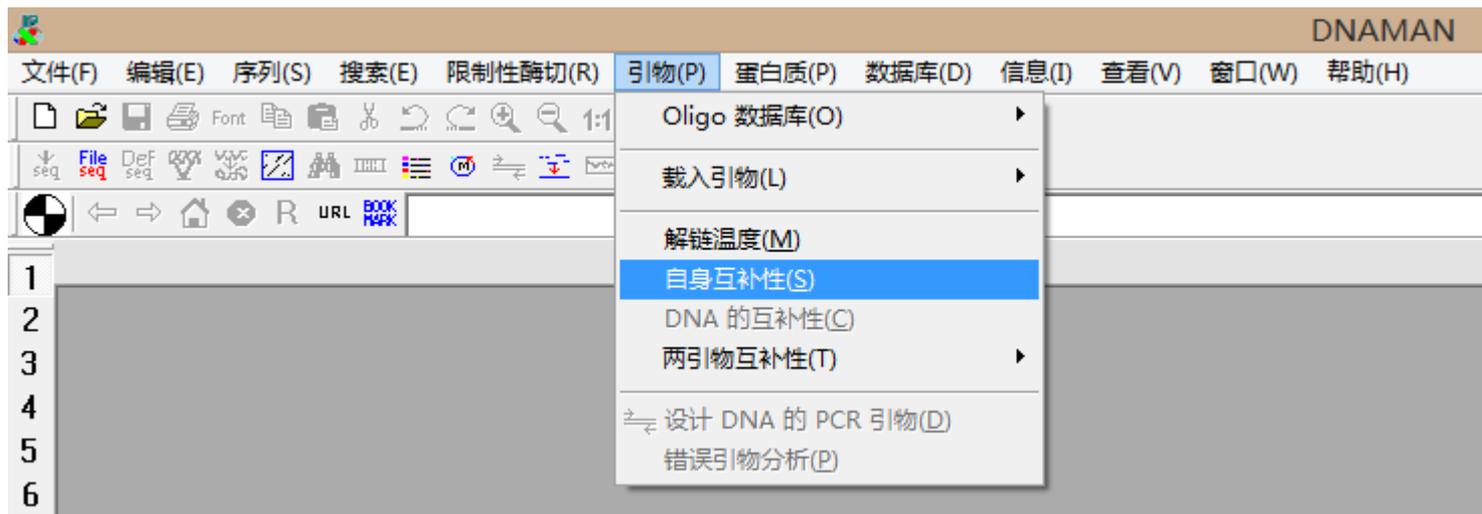
```
Two primers complementarity.  
Max complementarity in continuous: 5 bp, free energy = -2.20 Kcal/mol  
5'-CGGGATCCATGAGTTTATCAATATTTTGTGG-3'  
      |||||  
3'-AGAAGACCGGTAAAACGTGAGAGCTCCC-5'
```

Below this, it shows the discontinuous complementarity:

```
Max complementarity in discontinuous: 10 bp  
5'-CGGGATCCATGAGTTTATCAATATTTTGTGG-3'  
      ||      |||  ||  ||  |  
3'-AGAAGACCGGTAAAACGTGAGAGCTCCC-5'
```

The values "5 bp" and "10 bp" are highlighted with red boxes in the original image. The software interface includes a menu bar with options like "文件(F)", "编辑(E)", "序列(S)", "搜索(E)", "限制性酶切(R)", "引物(P)", "蛋白质(P)", "数据库(D)", "信息(I)", "查看(V)", "窗口(W)", and "帮助(H)". The toolbar contains various icons for file operations and sequence analysis. The status bar at the bottom shows "Primer\_Primer" and "URL".

# 9.用DNAMAN检测引物特异性



# 10.亚细胞定位的预测

 PREDICT PROTEIN OPEN

```
MLSIFLWLFITVSFTTSAASHKSYQQQTNSEPMVFEGVENGCGLIKCS
NGQTCGIRVGLAKFGDREFEKDFPKCVTSKAELNRNTDDDGNRIVTDG
PGCNTVHC SHGYKQVRISISKLGDLPYAQSIGTFPQCVGPNGTFTQTPSS
IIEQGPGEKLPKCEAGTKCVTAVGIAKYGNLQWSQFFWPFCT
```

 feedback

[Example Input 1 Example Input 2]

# 10.亚细胞定位的预测

- Flexibility >
- Disulphide Bridges >
- FUNCTION ANNOTATION
- Effect of Point Mutations >
- Gene Ontology Terms >
- Subcellular Localization >**
- Binding Sites >
- ADDITIONAL SERVICES
- Literature Search >
- HELP
- Site Tutorial >

feedback

**What am I seeing Here?** This viewer shows a cell schematic with the predicted subcellular localization compartment. The predicted compartment is also referenced by a GO ID for better clarity.

Domains: Archaea Bacteria **Eukarya**



Predicted localization for the Eukarya domain: Secreted (GO term ID: [GO:0005576](#)) Prediction confidence 45



THANKS

谢谢!

