



中国农业科学院蔬菜花卉研究所

The institute of vegetables and flowers chinese academy of agricultural sciences



青花菜*PGIP*基因的生物信息学分析

- 汇报人：樊艳燕
- G02组：樊艳燕 刘炎霖
刘丹丹 蒲研君

2013.6.23



中国农业科学院蔬菜花卉研究所

The institute of vegetables and flowers chinese academy of agricultural sciences



- ❁ 研究背景
- ❁ PGIP的基本信息
- ❁ 蛋白序列分析
- ❁ 进化树构建
- ❁ 小结



一、研究背景



青花菜(*Brassica oleracea L. Var. italica Planch*)又名西兰花，意大利芥蓝、绿花菜，是十字花科蔬菜甘蓝的一个变种，原产地中海沿岸，一二年生草本植物，以绿色的花球供食用。

青花菜作为绿色鲜食蔬菜，具有极高的营养价值，其维生素C含量是柠檬的2倍，而且还含有大量维生素A、维生素B、维生素K、维生素U、尼古丁酸、钙、磷、铁、蛋白质、糖等。此外青花菜中所含的萝卜硫素是迄今为止在蔬菜中发现的最强的抗癌成分，特别是对肺癌、胃癌、结肠癌和直肠癌的抵抗。健康的人经常食用青花菜可以起到预防癌症的作用。



然而，由于其食用部分是幼嫩的花梗与小花蕾，呼吸代谢十分旺盛，极易衰老，表现为叶绿素降解而使花球黄化，茎和花蕾失水变松软萎蔫，严重影响其商品价值。因此，选育出耐贮藏的青花菜新品种具有重要的科学意义。



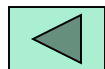
PGs

一种在细胞壁结构的改变中起重要作用的酶，参与果胶的降解，它与果实的软化密切相关，是果蔬贮藏研究中的热点。

PGIP

多聚半乳糖醛酸酶抑制蛋白，是一种富含亮氨酸的蛋白质，能够抑制多聚半乳糖醛酸酶的活性。

因而，*PGIP*基因成为目前果实耐贮藏基因工程研究的焦点。本研究以青花菜为材料，对它的*PGIP*基因进行生物信息学分析，为培育耐贮青花菜新品种奠定基础。



二、 PGIP 基本信息



PGIP是一种细胞表面和细胞间的受体蛋白，具有专一性；



属于富含亮氨酸重复（Leucine-rich Repeat，LRR）的蛋白质超家族；

- ❖ 一般有2-45个motifs组成

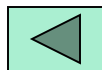
- ❖ 每个motif的长度一般为20-30AA

- ❖ 形成 β -折叠/ β -转角/ β -折叠/ α -螺旋结构，即LRR基序

- ❖ 高级结构通常折叠成弧形或马蹄形



通过LRR基序中暴露于外表面的氨基酸残基与PG活性位点处的氨基酸残基相互作用，从而抑制PG的活性。



三、蛋白质序列分析



1. 理化性质分析

2. 亲疏水分析

3. 跨膜区预测

4. 信号肽预测

5. 二级结构分析

6. 三级结构预测



使用ProtParam (<http://web.expasy.org/protparam/>) 对该蛋白进行理化性质分析

| | | | |
|-----------------|-----------------|------------------|-----------------|
| Ala (A) 12 3.6% | Gln (Q) 13 3.9% | Leu (L) 51 15.5% | Ser (S) 26 7.9% |
| Arg (R) 8 2.4% | Glu (E) 10 3.0% | Lys (K) 22 6.7% | Thr (T) 19 5.8% |
| Asn (N) 25 7.6% | Gly (G) 19 5.8% | Met (M) 6 1.8% | Trp (W) 5 1.5% |
| Asp (D) 20 6.1% | His (H) 9 2.7% | Phe (F) 14 4.2% | Tyr (Y) 9 2.7% |
| Cys (C) 8 2.4% | Ile (I) 23 7.0% | Pro (P) 19 5.8% | Val (V) 12 3.6% |

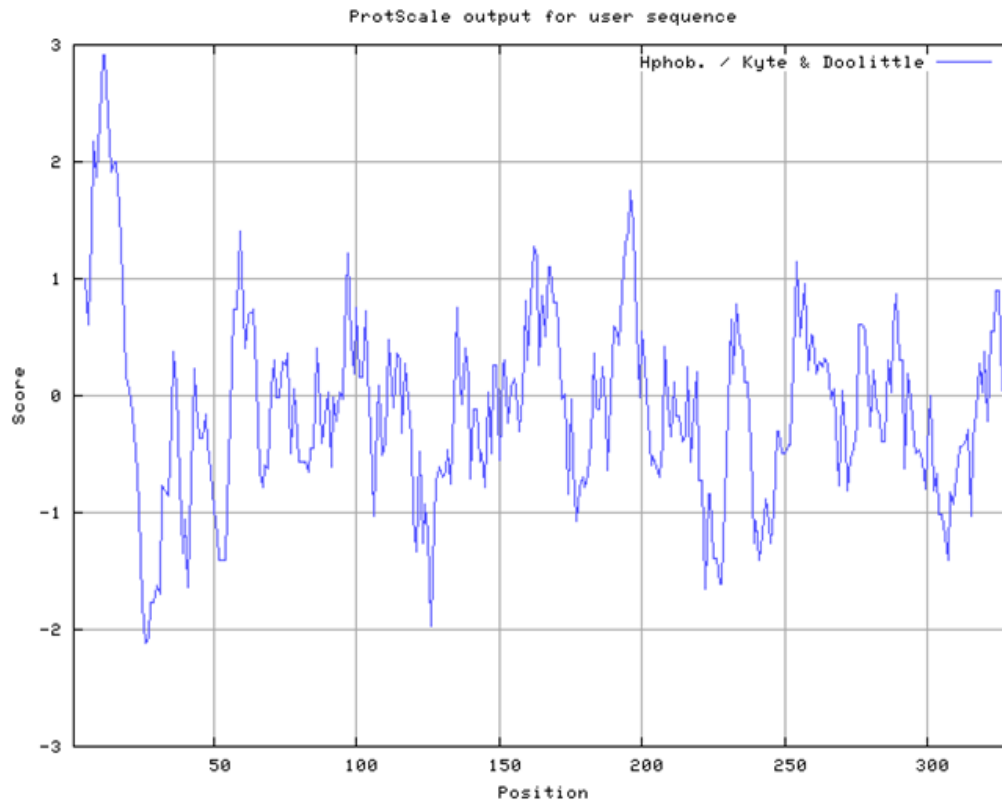
蛋白质的分子量为: 37035.7D
等电点: 7.09

带正电的氨基酸数目: 30
带负点的氨基酸数目: 30



使用ProtScale在线分析

(<http://web.expasy.org/protscale/>) 对该蛋白质进行亲疏水性分析，其结果如下：



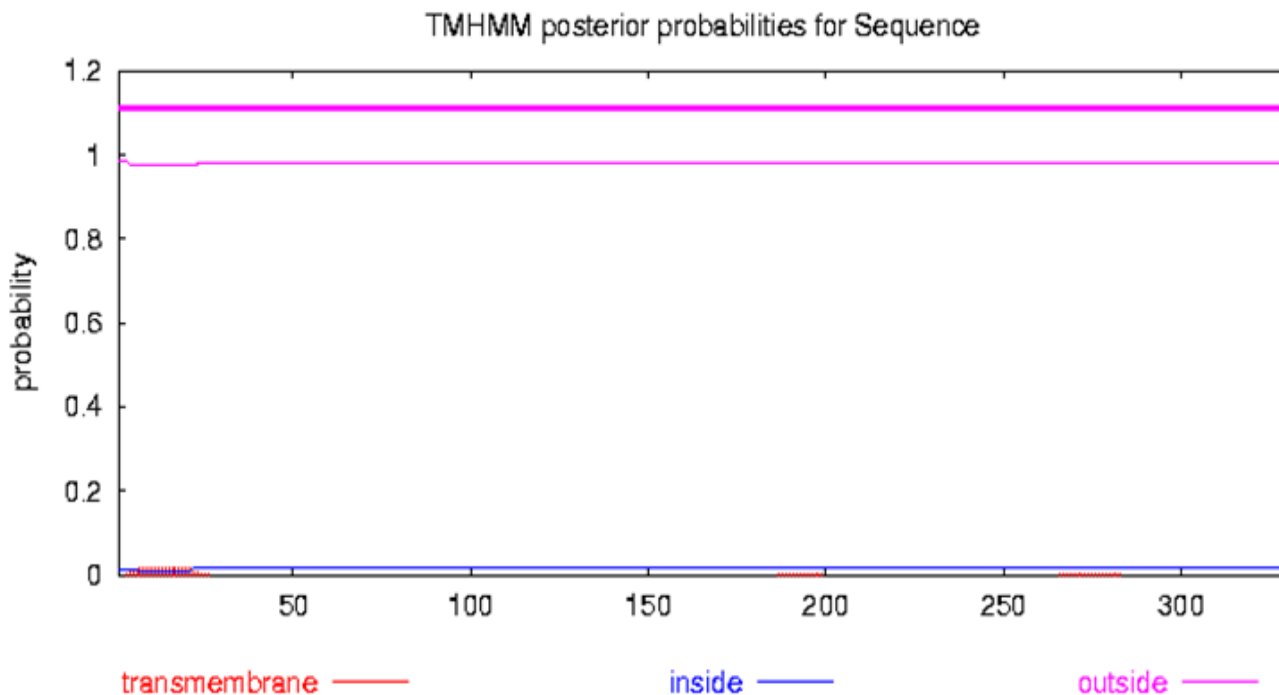
PGIP1 基因编码的蛋白的疏水性分析结构

由图可以看出该蛋白质在N末端有一个明显的疏水区域。

使用TMHMM对其进行跨膜区预测，结果如下图：



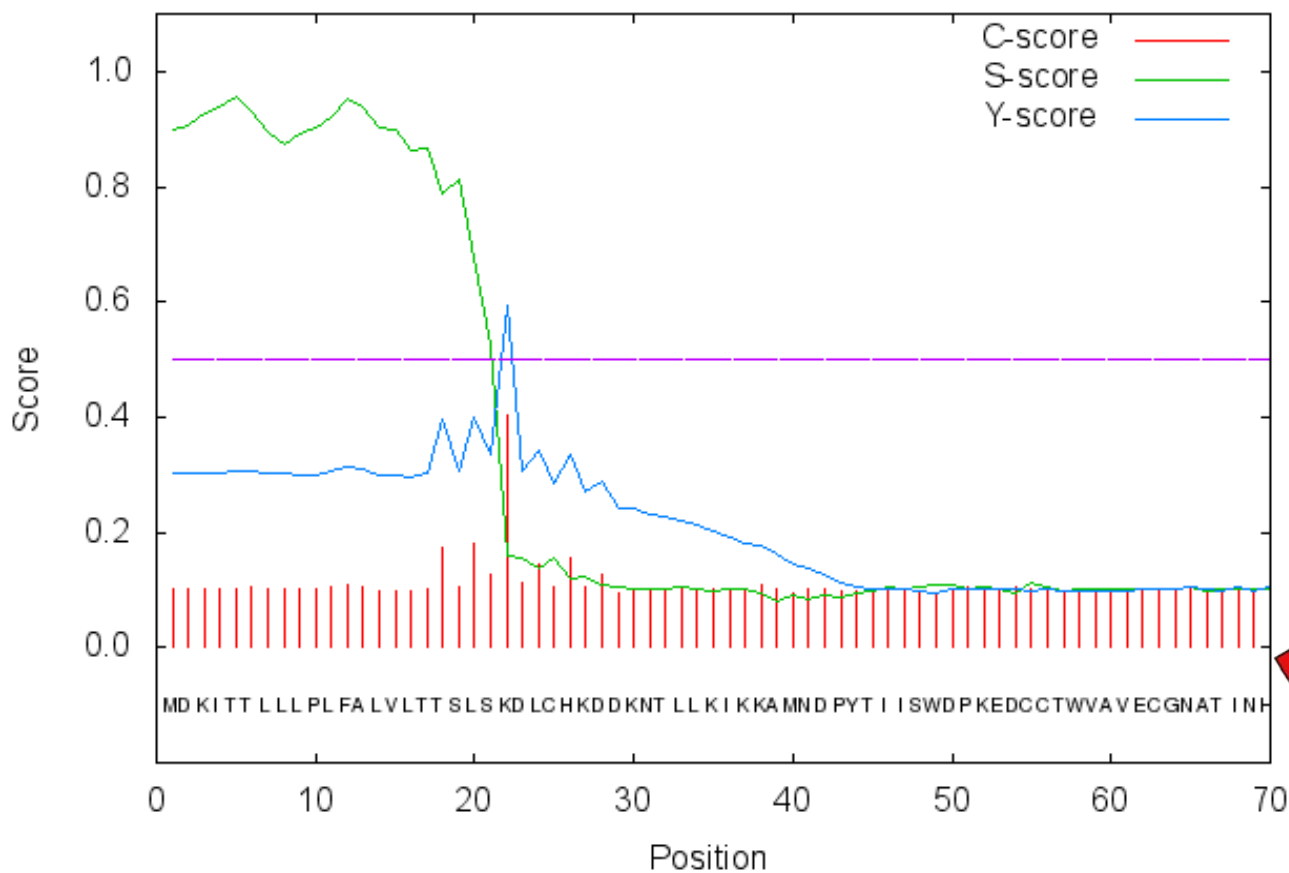
```
# Sequence Length: 330
# Sequence Number of predicted TMHs: 0
# Sequence Exp number of AAs in TMHs: 0.34780000000000000000000000000001
# Sequence Exp number, first 60 AAs: 0.34063
# Sequence Total prob of N-in: 0.01319
Sequence TMHMM2.0 outside      1   330
```



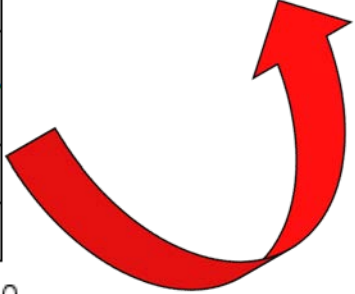
使用SignalP 4.1对其进行信号肽预测，结果如下图：



SignalP-4.1 prediction (euk networks): tr_G0ZRI6_G0ZRI6_BRAOT



剪切位点在
21-22bp之间

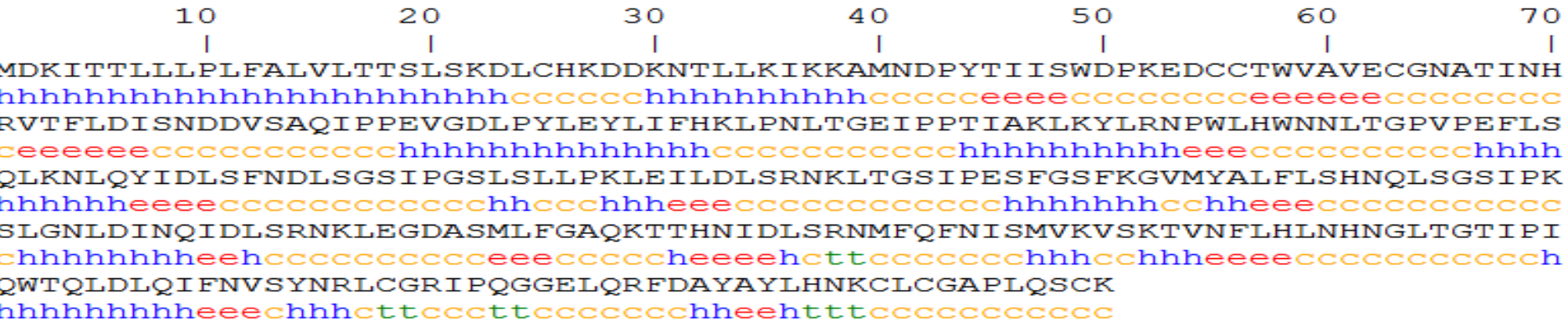


使用SPOMA在线分析

(http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html)

进行二级结构预测，其结果如下图：

View SOPMA in: [[AnTheProt \(PC\)](#)], [[Download...](#)] [[HELP](#)]



Sequence length : 330

SOPMA :

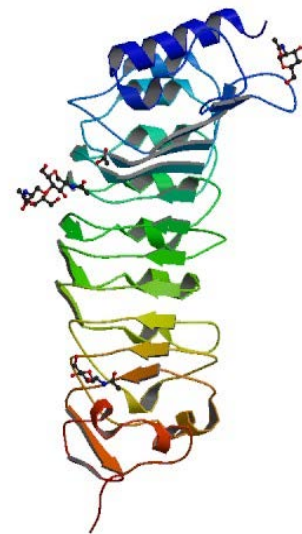
| | | | | | |
|-----------------------|------|---|-----|----|--------|
| Alpha helix | (Hh) | : | 114 | is | 34.55% |
| 3 ₁₀ helix | (Gg) | : | 0 | is | 0.00% |
| Pi helix | (Ii) | : | 0 | is | 0.00% |
| Beta bridge | (Bb) | : | 0 | is | 0.00% |
| Extended strand | (Ee) | : | 47 | is | 14.24% |
| Beta turn | (Tt) | : | 9 | is | 2.73% |
| Bend region | (Ss) | : | 0 | is | 0.00% |
| Random coil | (Cc) | : | 160 | is | 48.48% |
| Ambiguous states (?) | | : | 0 | is | 0.00% |
| Other states | | : | 0 | is | 0.00% |

PGIP1基因编码的蛋白的二级结构预测结果

SWISS-MODEL对该蛋白进行 三级结构预测：



1 青花菜*PGIP1*基因编码
蛋白 三级结构预测模型



2 菜豆*PGIP1*基因编码
蛋白的三级结构

四、进化树构建

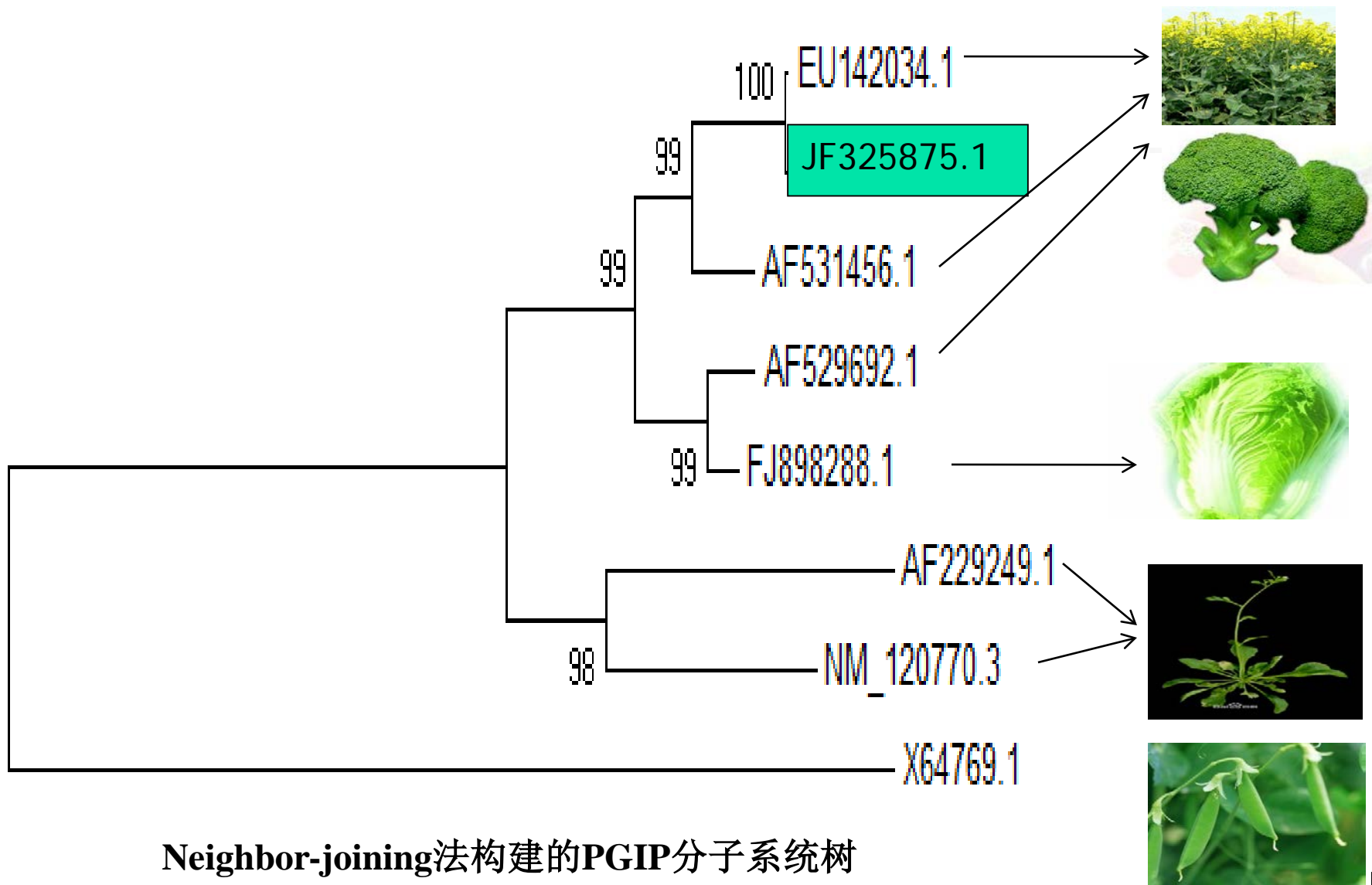


在青花菜中该基因的名字为*BoPGIP*，目前已克隆的是*BoPGIP1*，该基因全长1088bp，编码的蛋白质由330个氨基酸组成。

通过Blast搜索与*BoPGIP*基因同源的序列，然后分别选取甘蓝型油菜 (*Brassica.napus*)、白菜 (*Brassica.rapa*)、拟南芥 (*Arabidopsis.thaliana*) 以及菜豆 (*Phaseolus.vulgaris*) 的*PGIP*基因的序列做进化树分析。



| 物种 Species | 序列登录号 Assession No. | 物种 Species | 序列登录号 Assession No. |
|--------------------------------------|------------------------|-----------------------------|------------------------|
| <i>Brassica oleracea var.italica</i> | JF325875.1 | <i>Brassica .rapa</i> | FJ898288.1 |
| <i>Brassica .napus</i> | EU142034.1 | <i>Arabidopsis.thaliana</i> | NM_120770.3 |
| <i>Brassica .napus</i> | AF531456.1 | <i>Arabidopsis.thaliana</i> | AF229249.1 |
| <i>Brassica .napus</i> | AF529691.1 | <i>Phasedus vulgaris</i> | X64769 |



Neighbor-joining法构建的PGIP分子系统树



五、小结



蛋白序列分析发现：

- ★ 分子量为37035.7D；
- ★ N末端有一个疏水区；
- ★ 无跨膜区；
- ★ 有信号肽；
- ★ 该蛋白质二级结构预测符合LRR基序特征；
- ★ 三级结构为弧形。



通过进化树分析，我们发现该基因与甘蓝型油菜的*PGIP*基因亲缘关系最近；



由上面的分析我们可以大概了解该基因的一些基本信息，但是据文献报道，不同的*PGIP*基因在同一植物中的序列、结构、功能和表达上有差异，因此想真正了解果蔬的贮藏机理还需要许多具体的实验验证。

谢谢大家

