

应用ExPASy相关工具对GCR2蛋白 进行分析

——兼述GCR2膜结合G蛋白偶联受体之争

Pku09s2c1 张 翼

Pku09s2c2 余 继 (组长)

Pku09s2d1 操昭煦

pku09s2d2 刘博超

2009-03-25

A G Protein–Coupled Receptor Is a Plasma Membrane Receptor for the Plant Hormone Abscisic Acid

Xigang Liu,^{1,2} Yanling Yue,¹ Bin Li,³ Yanli Nie,¹ Wei Li,² Wei-Hua Wu,³ Ligeng Ma^{1,2*}

23 MARCH 2007 VOL 315 SCIENCE www.sciencemag.org

Comment on “A G Protein–Coupled Receptor Is a Plasma Membrane Receptor for the Plant Hormone Abscisic Acid”

Christopher A. Johnston,³ Brenda R. Temple,² Jin-Gui Chen,⁴ Yajun Gao,⁴ Etsuko N. Moriyama,³ Alan M. Jones,^{1,3*} David P. Siderovski,^{1*} Francis S. Willard^{1*}

9 NOVEMBER 2007 VOL 318 SCIENCE www.sciencemag.org

Response to Comment on “A G Protein–Coupled Receptor Is a Plasma Membrane Receptor for the Plant Hormone Abscisic Acid”

Xigang Liu,¹ Yanling Yue,¹ Wei Li,^{1,2} Ligeng Ma^{1*}

9 NOVEMBER 2007 VOL 318 SCIENCE www.sciencemag.org

Background

1. G蛋白偶联受体是已知的最大的蛋白家族，人类基因组编码超过800个G蛋白偶联受体(1)，通过一系列信号转导，可以调控细胞增殖，分化，代谢，个体发育。它们也作为受体参与视觉，嗅觉，味觉等感受器官对环境的感受。拟南芥只编码25个候选的G蛋白偶联受体，其功能大都不清楚。拟南芥中仅有的一个典型的G α 突变拟南芥的研究表明，它同植物激素ABA的信号转导有关。
2. 脱落酸（ABA）是植物的一个很重要的激素，对于植物逆境生存有重要的作用。他通过关闭气孔，调节代谢来，抑制发芽和幼苗生长等生理调控来抵御脱水，寒冷以及高盐环境。研究证明，脱落酸通过细胞膜和细胞内受体调控生理活动。但一直都没有找到该膜受体，这样G蛋白上游的受体就成了大家的焦点。07年科学杂志上发表的一篇关于GCR2是ABA受体的文章（2），引起了广泛关注。但有人对于GCR2是G蛋白偶联受体提出了质疑，讨论的一个热点是用生物信息的方法预测跨膜蛋白的问题（3，4，5）。

我们的工作

- 第一部分 预测跨膜结构
- 第二部分 GCR2蛋白的Blast分析和motif预测
- 第三部分 序列比对
- 第四部分 搜索GCR2三级结构\推测GCR2四级结构
- 第五部分 讨论和结论

第一部分 预测跨膜结构

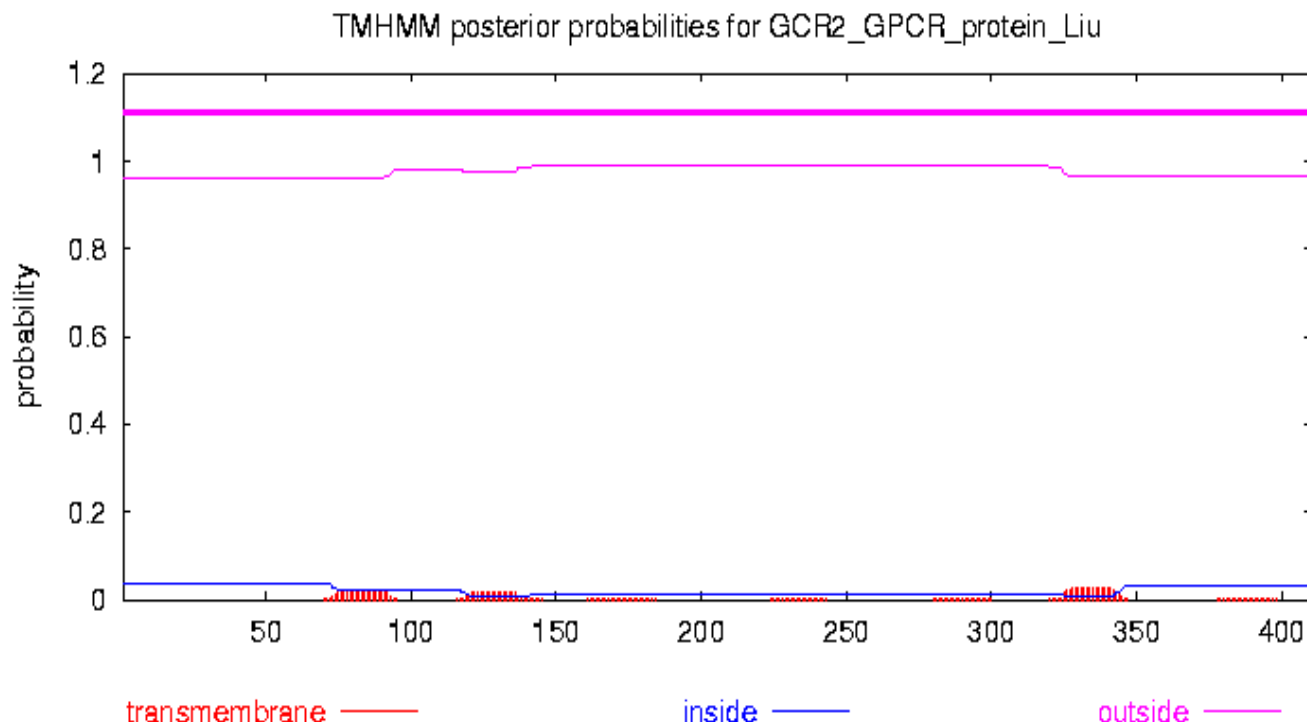
Topology prediction

- [NetNES](#) - Leucine-rich nuclear export signals (NES) in eukaryotic proteins
- [PSORT](#) - Prediction of protein subcellular localization
- [SecretomeP](#) - Non-classical and leaderless secretion of proteins
- [TargetP](#) - Prediction of subcellular location
- [TatP](#) - Twin-arginine signal peptides
- [DAS](#) - Prediction of transmembrane regions in prokaryotes using the Dense Alignment Surface method (Stockholm University)
- [HMMTOP](#) - Prediction of transmembrane helices and topology of proteins (Hungarian Academy of Sciences)
- [PredictProtein](#) - Prediction of transmembrane helix location and topology (Columbia University)
- [SOSUI](#) - Prediction of transmembrane regions (Nagoya University, Japan)
- [TMAP](#) - Transmembrane detection based on multiple sequence alignment (Karolinska Institut, Sweden)
- [TMHMM](#) - Prediction of transmembrane helices in proteins (CBS; Denmark)
- [TMPred](#) - Prediction of transmembrane regions and protein orientation (EMBNET-CH)
- [TopPred](#) - Topology prediction of membrane proteins (France)

ExPASy 的八个预测跨膜结构的软件

TMHMM (v2.0)

```
# GCR2_GPCR_protein_Liu Length: 410
# GCR2_GPCR_protein_Liu Number of predicted TMHs: 0
# GCR2_GPCR_protein_Liu Exp number of AAs in TMHs: 1.25279
# GCR2_GPCR_protein_Liu Exp number, first 60 AAs: 0
# GCR2_GPCR_protein_Liu Total prob of N-in: 0.03772
GCR2_GPCR_protein_Liu TMHMM2.0 outside 1 410
```



(2) 如果预测跨膜螺旋中的氨基酸残基数大于18, 则说明很有可能存在跨膜序列或者信号肽, 而GCR2分析结果为1.25279个, 从一个参数表明不具有明显的跨膜结构或信号肽;

(3) 如果预测蛋白前60个氨基酸中跨膜螺旋中的氨基酸残基数有数个, 则说明该蛋白N端可能存在信号肽, 而不是跨膜结构, 而GCR2分析结果为0, 证明至少其N端既不存在信号肽, 也不存在跨膜结构;

(4) 综合图及以上结果, 该软件预测GCR2不具有跨膜结构。

- 作者最开始使用TMpred 和DAS 软件程序推测 GCR2 是一个7 次跨膜的蛋白，并进一步使用12 种不同的软件程序（TMHMM, SOSUI 和DAS TMfilter 等）去推测GCR2 的拓扑结构，其中9 种显示GCR2 是具有不同跨膜结构哉的跨膜蛋白。TMHMM 已经错误推测了很多跨膜结构域，并且拟南芥中唯一一个被报告的GCR1 也是由SOSUI 预测出是三次跨膜的蛋白。另外，已知的跨膜蛋白中14%不能被可用的软件正确推测，所以，膜蛋白的计算机推测还不是一个成熟的技术方法，而主要是用来作为对实验测试的假设

SOSUI和SOSUISignal

- 【功能】 预测蛋白质的疏水性以及信号肽
- 【程序路径】 http://bp.nuap.nagoya-u.ac.jp/sosui/sosui_submit.html
- http://bp.nuap.nagoya-u.ac.jp/sosui/sosuisignal/sosuisignal_submit.html
- 【输入参数】 默认

SOSUI

Query title : GCR2/GPCR_protein_Liu Bochao_2 - 3-16

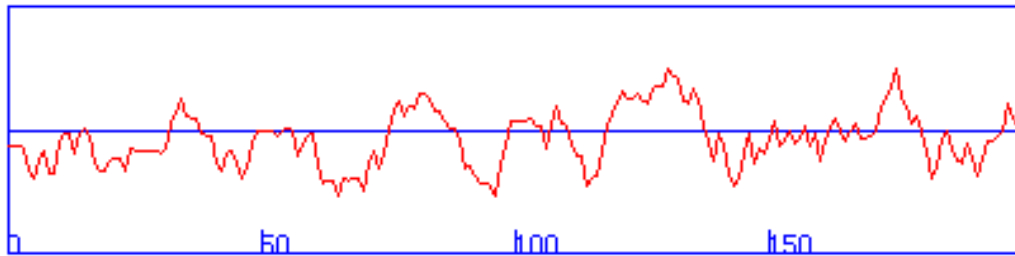
Total length : 410 A. A.

Average of hydrophobicity : -0.257073

This amino acid sequence is of a SOLUBLE PROTEIN.

Display Options

[Hydropathy profile]



(1) GCR2平均疏水性是-0.257，没有显著的疏水性；

(2) 从图中可见，GCR2也没有明显的疏水性集中的区域，而疏水性集中连续的区域可能是跨膜结构；

SOSUIsignal

SOSUIsignal Result

Query title : GCR2/GPCR_protein_Liu Bochao_2009-03-16

Total length : 410 A. A.

Average of hydrophobicity : -0.257073

This amino acid sequence has no signal peptide.

This amino acid sequence is of a SOLUBLE PROTEIN.

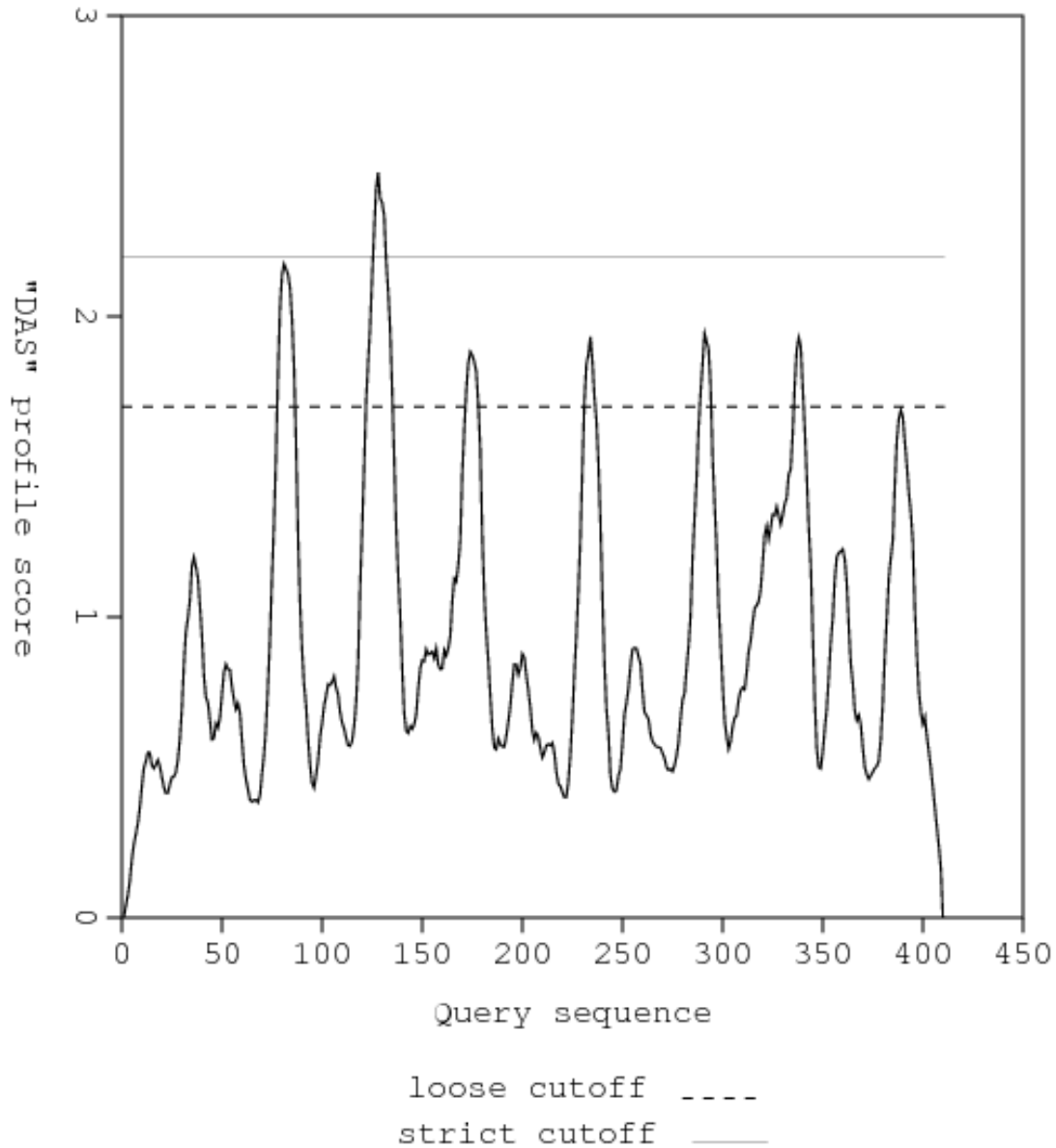
(3) GCR2没有信号肽;

(4) GCR2是一个可溶性蛋白;

(5) 综合以上, 结论与TMHMM软件分析结果相同, 该软件预测GCR2不具有跨膜结构以及信号肽

- 针对以上分析，马力耕等人在《Response to Comment on “A G Protein-Coupled Receptor Is a Plasma Membrane Receptor for the Plant Hormone Abscisic Acid”》作者最开始使用TMpred 和DAS 软件程序推测GCR2 是一个7 次跨膜的蛋白，并进一步使用12 种不同的软件程序（TMHMM，SOSUI 和DAS TMfilter 等）去推测GCR2 的拓扑结构，其中9 种显示GCR2 是具有不同跨膜结构域的跨膜蛋白。TMHMM 已经错误推测了很多跨膜结构域，并且拟南芥中唯一一个被报告的GCR1 也是由SOSUI 预测出是三次跨膜的蛋白。另外，已知的跨膜蛋白中14%不能被可用的软件正确推测，所以，膜蛋白的计算机推测还不是一个成熟的技术方法，而主要是用来作为对实验测试的假设。

DAS



- 在原核生物中利用**Dense Alignment Surface method**预测蛋白质的跨膜结构域
- 如果按照严格的划分线，只有一个跨膜结构域，按宽松的划分线，会有7个跨膜结构域。但是，这个是由于用于预测原核生物的软件，对于作为高等真核生物拟南芥的GCR2蛋白，可能并不准确，刘西岗等人以此结果为其依据之一，诚为不妥。

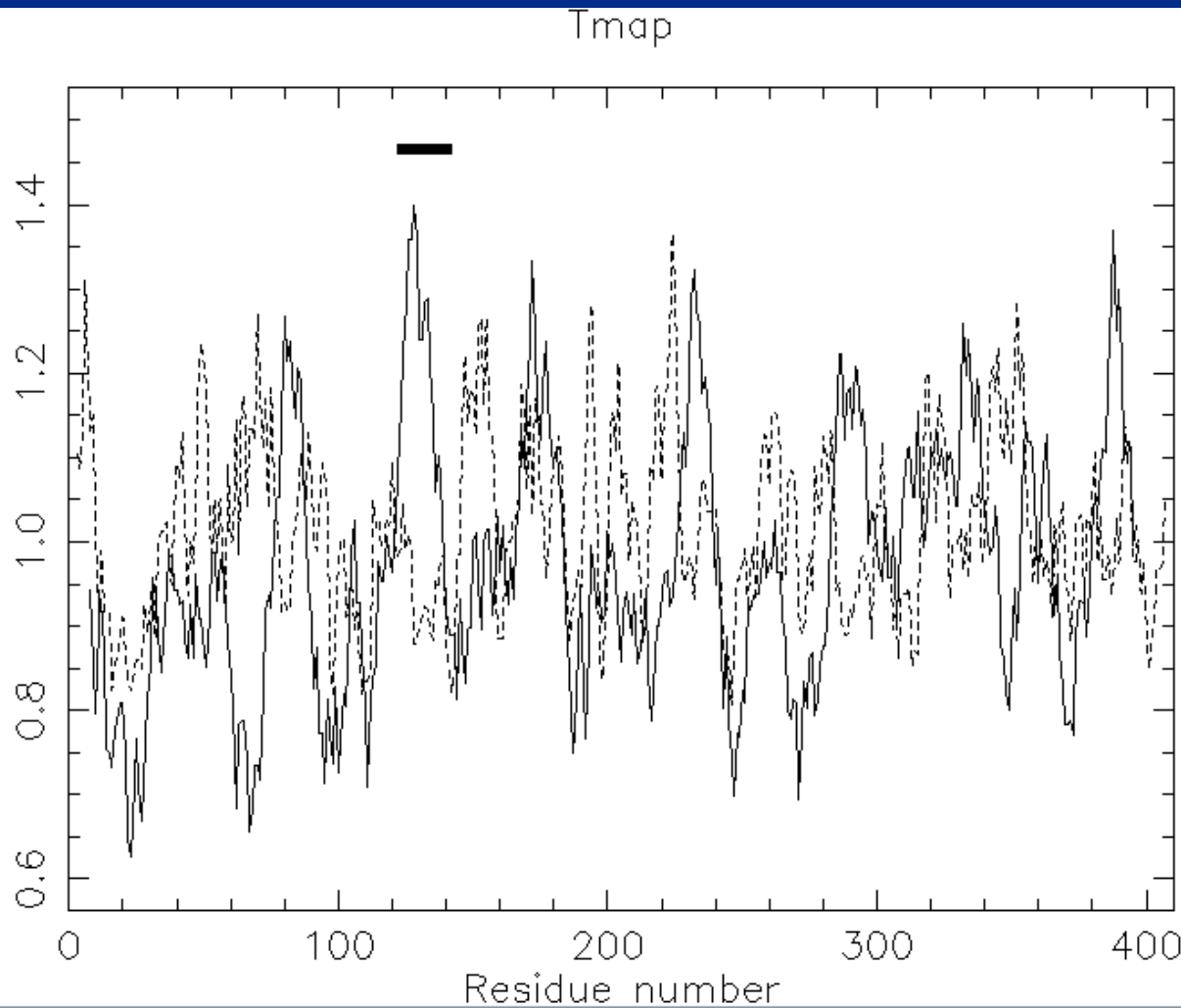
HMMTOP

- 【功能】 预测蛋白质的跨膜螺旋以及拓扑结构
- 【处理结果】
- Length: 410
- N-terminus: OUT
- Number of transmembrane helices: 0
- Transmembrane helices:
- Total entropy of the model: 17.0206
- Entropy of the best path: 17.0212

TMAP

【功能】基于多序列比对监测蛋白质的跨膜结构

【处理结果】



【结果分析】

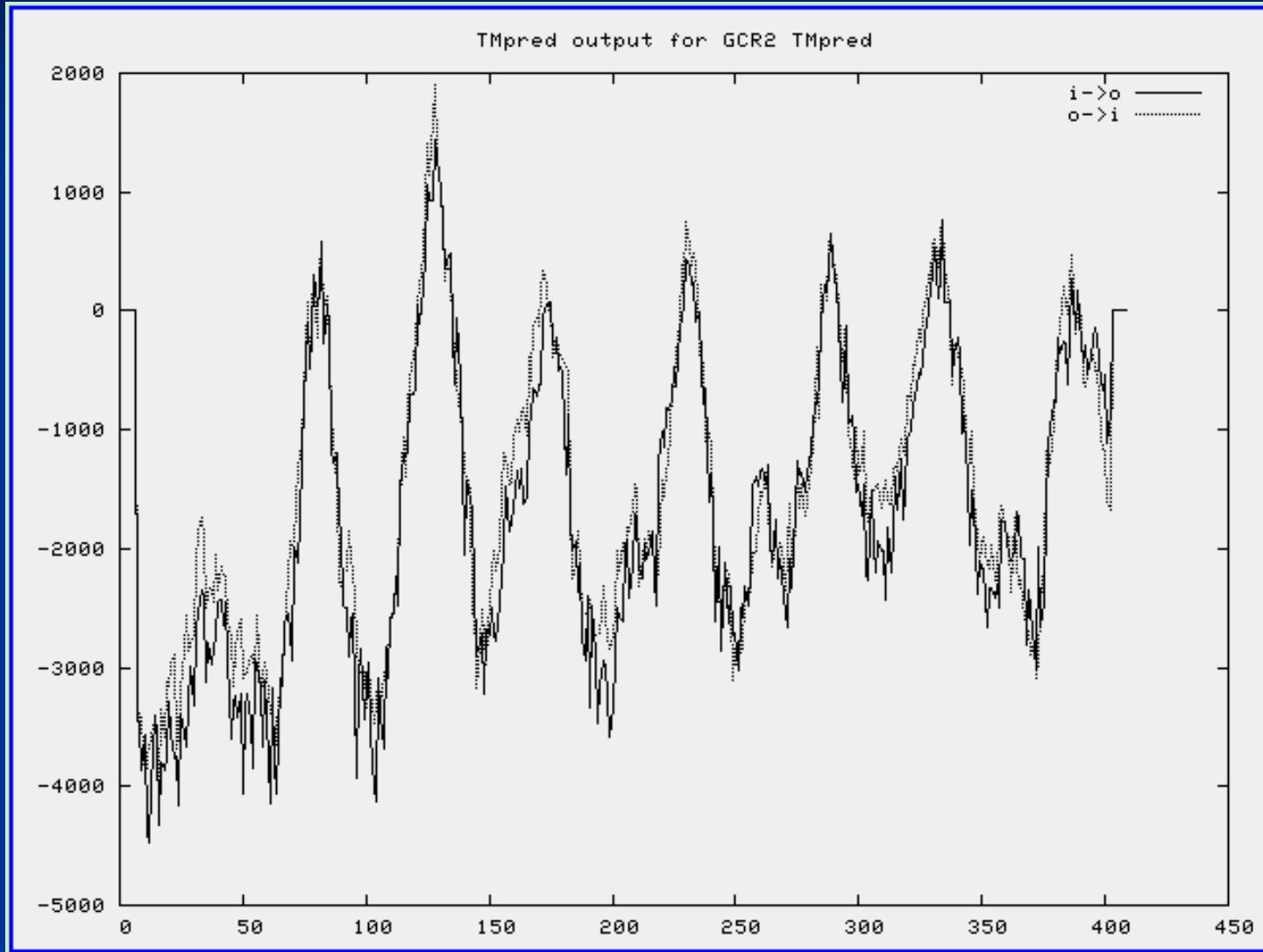
(1) 由TMAP预测出来，也只有一个跨膜结构域，与DAS预测的那个跨膜结构域位置相同，可以互相印证。

(2) TMAP程序是通过两个参数来计算蛋白是跨膜蛋白的可能性的：第一个是蛋白中段，疏水性残基的比例；另一个参数是末端区域跨膜序列的跨度。

(3) 所以，图中实线代表的是蛋白在中段形成跨膜蛋白的可能性，而虚线表示在末端形成跨膜蛋白的可能性。这些可能性，是针对蛋白的每个氨基酸残基计算的，所以横坐标表示残基的顺序，纵坐标表示成为跨膜蛋白的倾向性。

(4) 而图上方的黑条，则是程序自己预测的最可能成为跨膜区域的地方。

TMpred



有7个跨膜结构域，而第1、3、7是不显著的跨膜结构域。

TopPred

unknown

Structure no. 1

Segments included: 1 2 3

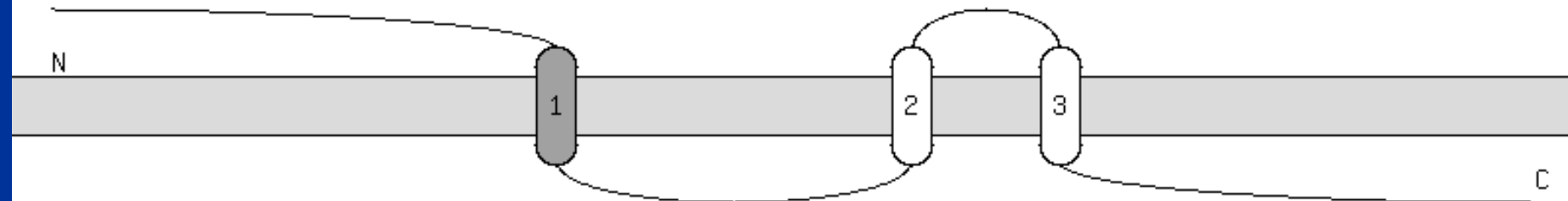
□ Segment Putative
■ Segment Certain

CYTOPLASM

L1 = 118
KR = 18

L1 = 35
KR = 5

N



L1 = 83
KR = 14

L1 = 111
KR = 13

C

EXTRACELLULAR

L1: Loop length

KR: Number of Lys and Arg

一个确定的跨膜结构域以及两个不确定的跨膜结构域

Summary

软件	结果	评注
TMHMM(v2.0)	无跨膜结构	
SOSUI	没有明显的疏水性集中的区域	
PredictProtein	五个跨膜螺旋	都很短
DAS	7	6个不确定
HMMTOP	0	
TMAP	1	
TMpred	7	6个不确定
TOPpre	3	2个不确定

讨论

- 四个预测软件认为它不是跨膜蛋白。
- DAS, TMpred, TOPpre和TMAP认为其有跨膜, 但跨膜数不定, 尤其是当对跨膜结构的阈值设定的较高时, DAS, TMpred, TOPpre认为它是一次跨膜。
- 预测软件预测的结果不能判定其是否为跨膜蛋白

关于跨膜结构预测软件

Moller等人曾经比较过各种跨膜预测软件（7），一个有趣的实验是他们用已经被SwissProt收录为GPCR的833个蛋白测试各种软件，结果如下。

软件名	准确预测为七次跨膜的个数	准确预测的比率
ALOM2	29	3%
DAS	24	3%
HMMTOP	712	85%
MEMSAT1.5	551	66%
TMHMM1.0	707	85%
TMHMM2.0	711	85%

第二部分 GCR2蛋白的Blast分析和motif预测

1 blast分析结果:

对GCR2蛋白序列进行序列相似性搜索，希望从中得到关于GCR2结构或功能的线索，不同搜索方法**blast, fasta, MPsrch**得到的结果大致相同，其中大多数相似蛋白功能尚未确定。就已知蛋白来推测GCR2可能为跨膜的受体（与GPCR相似），也可能为外周某种功能的酶（与LANC/ LANC-like）。而其他一些搜索软件**Propsearch**的结果则更多提示GCR2可能和LANC更接近的线索。

Blast

- **PSI-Blast---position-specific iterated blast,位点特异叠代比对**

PSI-BLAST的基本思路在于根据最初的搜索结果，依照预先定义的相似性阈值将序列分成不同的组，构建一个位点特异性的序列谱，并通过多次叠代不断改进这一序列谱以提高搜索的灵敏度。但缺点在于运算复杂，重复序列对结果有干扰等。有研究表明这种方法可以有效的找到很多序列差异较大而结构功能相似的相关蛋白，所以对GCR2尝试此法。

PSI-Blast结果

■ 默认参数-首次检索结果

Sequences with E-value BETTER than threshold

			Score (Bits)	E Value	
NEW	<input checked="" type="checkbox"/>	ref NP_175700.2 GCR2/GPCR (G PROTEIN COUPLED RECEPTOR); absc...	850	0.0	UG
NEW	<input checked="" type="checkbox"/>	qb AAG52264.1 AC019018_1 putative G protein-coupled receptor;...	830	0.0	
NEW	<input checked="" type="checkbox"/>	ref XP_002329106.1 predicted protein [Populus trichocarpa] >...	610	8e-173	G
NEW	<input checked="" type="checkbox"/>	ref XP_002284781.1 PREDICTED: hypothetical protein [Vitis vi...	604	5e-171	
NEW	<input checked="" type="checkbox"/>	qb EEF28136.1 catalytic, putative [Ricinus communis]	572	3e-161	
NEW	<input checked="" type="checkbox"/>	qb ABQ14360.1 abscisic acid ABA receptor [Populus trichocarpa]	549	2e-154	
NEW	<input checked="" type="checkbox"/>	ref XP_002263345.1 PREDICTED: hypothetical protein [Vitis vi...	541	3e-152	
NEW	<input checked="" type="checkbox"/>	ref NP_850003.1 GCL2 (GCR2-LIKE 2); catalytic [Arabidopsis t...	541	4e-152	UG
NEW	<input checked="" type="checkbox"/>	emb CAN64758.1 hypothetical protein [Vitis vinifera]	540	8e-152	
NEW	<input checked="" type="checkbox"/>	qb EEF44459.1 catalytic, putative [Ricinus communis]	539	2e-151	
NEW	<input checked="" type="checkbox"/>	ref XP_002319360.1 predicted protein [Populus trichocarpa] >...	521	5e-146	G
NEW	<input checked="" type="checkbox"/>	qb ACN31611.1 unknown [Zea mays]	515	3e-144	
NEW	<input checked="" type="checkbox"/>	qb AAL58268.1 AC068923_10 putative lanthionine synthetase [Or...	509	2e-142	
NEW	<input checked="" type="checkbox"/>	qb ABB47829.2 Lanthionine synthetase C-like protein, express...	509	2e-142	
NEW	<input checked="" type="checkbox"/>	ref NP_001064963.1 Os10g0497900 [Oryza sativa (japonica cult...	508	6e-142	UG
NEW	<input checked="" type="checkbox"/>	qb ABB47830.2 Lanthionine synthetase C-like protein, express...	507	7e-142	
NEW	<input checked="" type="checkbox"/>	qb ACG34942.1 lanC-like protein 2 [Zea mays]	503	1e-140	
NEW	<input checked="" type="checkbox"/>	qb AAD20918.1 hypothetical protein [Arabidopsis thaliana]	499	1e-139	
NEW	<input checked="" type="checkbox"/>	qb EAZ16566.1 hypothetical protein OsJ_32039 [Oryza sativa J...	493	1e-137	
NEW	<input checked="" type="checkbox"/>	emb CAD41202.1 OSJNBa0074L08.13 [Oryza sativa (japonica cult...	459	2e-127	
NEW	<input checked="" type="checkbox"/>	emb CAJ86377.1 OSIGBa0155K17.4 [Oryza sativa (indica cultiv...	457	9e-127	

Sequences with E-value BETTER than threshold

Sequences producing significant alignments:		Score (Bits)	E Value	
<input checked="" type="checkbox"/>	ref NP_175700.2 GCR2/GPCR (G PROTEIN COUPLED RECEPTOR); absc...	600	6e-170	UG
<input checked="" type="checkbox"/>	qb AAG52264.1 AC019018_1 putative G protein-coupled receptor;...	583	1e-164	
<input checked="" type="checkbox"/>	ref XP_002329106.1 predicted protein [Populus trichocarpa] >...	548	5e-154	G
<input checked="" type="checkbox"/>	qb EEF28136.1 catalytic, putative [Ricinus communis]	543	1e-152	
<input checked="" type="checkbox"/>	ref XP_002284781.1 PREDICTED: hypothetical protein [Vitis vi...	542	2e-152	
<input checked="" type="checkbox"/>	ref XP_002263345.1 PREDICTED: hypothetical protein [Vitis vi...	535	3e-150	
<input checked="" type="checkbox"/>	emb CAN64758.1 hypothetical protein [Vitis vinifera]	535	4e-150	
<input checked="" type="checkbox"/>	qb ACN31611.1 unknown [Zea mays]	530	1e-148	
<input checked="" type="checkbox"/>	qb EEF44459.1 catalytic, putative [Ricinus communis]	524	6e-147	
<input checked="" type="checkbox"/>	qb AAL58268.1 AC068923_10 putative lanthionine synthetase [Or...	519	2e-145	
<input checked="" type="checkbox"/>	ref NP_001064963.1 Os10g0497900 [Oryza sativa (japonica cult...	516	2e-144	UG
<input checked="" type="checkbox"/>	qb ABB47829.2 Lanthionine synthetase C-like protein, express...	516	2e-144	
<input checked="" type="checkbox"/>	ref NP_850003.1 GCL2 (GCR2-LIKE 2); catalytic [Arabidopsis t...	514	7e-144	UG
<input checked="" type="checkbox"/>	qb ABB47830.2 Lanthionine synthetase C-like protein, express...	513	1e-143	
<input checked="" type="checkbox"/>	qb EAZ16566.1 hypothetical protein OsJ_32039 [Oryza sativa J...	511	3e-143	
<input checked="" type="checkbox"/>	qb ACG34942.1 lanC-like protein 2 [Zea mays]	505	2e-141	
<input checked="" type="checkbox"/>	ref XP_002319360.1 predicted protein [Populus trichocarpa] >...	501	3e-140	G
<input checked="" type="checkbox"/>	qb ABQ14360.1 abscisic acid ABA receptor [Populus trichocarpa]	489	1e-136	
<input checked="" type="checkbox"/>	ref NP_598498.1 LanC (bacterial lantibiotic synthetase compo...	476	1e-132	UG
<input checked="" type="checkbox"/>	emb CAD41202.1 OSJNBa0074L08.13 [Oryza sativa (japonica cult...	475	3e-132	
<input checked="" type="checkbox"/>	ref NP_001014209.1 LanC lantibiotic synthetase component C-1...	475	3e-132	UG
<input checked="" type="checkbox"/>	qb AAD20918.1 hypothetical protein [Arabidopsis thaliana]	474	6e-132	
<input checked="" type="checkbox"/>	ref NP_001095726.1 LanC lantibiotic synthetase component C-1...	472	2e-131	UG

- 2次检索结果
- 但未得到新的相似蛋白
(若有新蛋白则前标显示new)

Fasta3

- Fasta的基本思路是识别与代查序列相匹配的很短的序列片段，称为k-tuple。蛋白质序列数据库搜索时，短片段的长度一般是1-2个残基长。Fasta程序采用渐进（heuristic approach）算法将位于同一对角线上相互接近的短片段连接起来。也就是说，通过不匹配的残基将这些匹配残基片段连接起来，以便得到较长的相似性片段，与blast类似。

Fasta3:

取代矩阵选择
Blosum62，
其他参数默
认

Alignment	DB:ID	Source	Length	Identity%	Similar%	Overlap	E()
1 <input type="checkbox"/>	UNIPROT:Q9C929_ARATH	Putative G protein-coupled	401	100.0	100.0	401	0
2 <input type="checkbox"/>	UNIPROT:A7P1J1_VITV	Chromosome chr19 scaffold_	407	69.0	89.0	410	1.2e-154
3 <input type="checkbox"/>	UNIPROT:A5HIY4_POPTR	Abecisic acid ABA receptor	443	73.8	90.0	359	9.9e-143
4 <input type="checkbox"/>	UNIPROT:Q8VZQ6_ARATH	Putative uncharacterized p	405	64.0	86.9	411	5.1e-139
5 <input type="checkbox"/>	UNIPROT:A7PWY6_VITV	Chromosome chr12 scaffold_	412	63.8	84.7	412	1.3e-137
6 <input type="checkbox"/>	UNIPROT:A5AP09_VITV	Putative uncharacterized p	412	63.8	84.7	412	1.6e-137
7 <input type="checkbox"/>	UNIPROT:Q337H0_ORYSJ	Lanthionine synthetase C-I	399	63.4	82.9	410	1e-132
8 <input type="checkbox"/>	UNIPROT:Q8W3F0_ORYSJ	Putative lanthionine synth	446	63.4	82.9	410	1.1e-132
9 <input type="checkbox"/>	UNIPROT:Q337G9_ORYSJ	cDNA clone:J033048L17, ful	401	63.3	82.8	412	3.6e-132
10 <input type="checkbox"/>	UNIPROT:Q0IWN8_ORYSJ	Os10g0497900 protein (Frag	447	63.3	82.8	412	3.9e-132
11 <input type="checkbox"/>	UNIPROT:Q9SKU0_ARATH	Putative uncharacterized p	371	66.5	89.1	358	3.3e-129
12 <input type="checkbox"/>	UNIPROT:B6TCV9_MAIZE	LanC like protein 2 OS-Zea	401	64.6	84.0	376	2.5e-128
13 <input type="checkbox"/>	UNIPROT:A3C653_ORYSJ	Putative uncharacterized p	461	58.8	80.0	420	3.6e-125
14 <input type="checkbox"/>	UNIPROT:A2XVS3_ORYSI	Putative uncharacterized p	422	55.3	78.0	423	1.2e-116
15 <input type="checkbox"/>	UNIPROT:Q7X7D6_ORYSJ	OSJNBa0074LD8.13 protein (422	55.3	78.0	423	1.2e-116
16 <input type="checkbox"/>	UNIPROT:Q00RJ6_ORYSA	OSIGBa0155K17.4 protein OS	422	55.1	77.8	423	1.2e-116
17 <input type="checkbox"/>	UNIPROT:A9TEG3_PHYPA	Predicted protein OS=PhySc	395	52.2	76.5	412	1.8e-107
18 <input type="checkbox"/>	UNIPROT:A9RX77_PHYPA	Predicted protein OS=PhySc	407	51.8	76.0	413	3.2e-105
19 <input type="checkbox"/>	UNIPROT:A9NV23_PICSI	Putative uncharacterized p	427	50.6	76.0	429	2.4e-78
20 <input type="checkbox"/>	UNIPROT:A5BN66_VITV	Putative uncharacterized p	421	47.9	72.1	420	1.3e-68
21 <input type="checkbox"/>	UNIPROT:A7R2I8_VITV	Chromosome undetermined sc	421	48.1	72.4	420	1.3e-68

MPsrch

- MPsrch采用穷尽算法来寻找相似序列，灵敏性比blast，fasta要高，但在对GCR2的检索中，三者结果类似

MPsrch结果:
BL62,其他
参数默
认

Alignment	DB:ID	Description	Length	Match%	Query Match%	Score	Pred.No.
1 <input type="checkbox"/>	UNIPROT:Q9C929_ARATH	SubName: Full=Putative G protein-coupled receptor; 80093-78432;	401	100.0	97.7	2125	0.00e+00
2 <input type="checkbox"/>	UNIPROT:A7P1J1_VITVI	SubName: Full=Chromosome chr19 scaffold_4, whole genome shotgun sequence;	407	69.0	69.1	1503	0.00e+00
3 <input type="checkbox"/>	UNIPROT:A7PWY6_VITVI	SubName: Full=Chromosome chr12 scaffold_36, whole genome shotgun sequence;	412	63.8	62.9	1367	0.00e+00
4 <input type="checkbox"/>	UNIPROT:A5AP09_VITVI	SubName: Full=Putative uncharacterized protein;	412	63.8	62.8	1366	0.00e+00
5 <input type="checkbox"/>	UNIPROT:A5HIY4_POPTR	SubName: Full=Abscisic acid ABA receptor;	443	73.8	62.7	1363	0.00e+00
6 <input type="checkbox"/>	UNIPROT:Q8VZQ6_ARATH	SubName: Full=Putative uncharacterized protein At2g20770;	405	64.3	62.3	1354	0.00e+00
7 <input type="checkbox"/>	UNIPROT:Q337H0_ORYSJ	SubName: Full=Lanthionine synthetase C-like protein, expressed;	399	65.8	60.8	1321	0.00e+00
8 <input type="checkbox"/>	UNIPROT:Q8W3F0_ORYSJ	SubName: Full=Putative lanthionine synthetase;	446	65.8	60.8	1321	0.00e+00
9 <input type="checkbox"/>	UNIPROT:Q0IWN8_ORYSJ	SubName: Full=Os10g0497900 protein;Flags: Fragment;	447	65.7	59.9	1303	0.00e+00
10 <input type="checkbox"/>	UNIPROT:Q337G9_ORYSJ	SubName: Full=cDNA clone:J033048L17, full insert sequence;SubName: Full=Lanthionine synthetase C-like protein, expressed;	401	65.7	59.9	1303	0.00e+00
11 <input type="checkbox"/>	UNIPROT:Q9SKU0_ARATH	SubName: Full=Putative uncharacterized protein At2g20770;	371	66.5	59.3	1290	0.00e+00
12 <input type="checkbox"/>	UNIPROT:B6TCV9_MAIZE	SubName: Full=LanC-like protein 2;	401	64.6	59.1	1284	0.00e+00
13 <input type="checkbox"/>	UNIPROT:A3C653_ORYSJ	SubName: Full=Putative uncharacterized protein;	461	66.1	57.0	1239	0.00e+00
14 <input type="checkbox"/>	UNIPROT:A2XVS3_ORYSJ	SubName: Full=Putative uncharacterized protein;	422	60.6	50.9	1107	0.00e+00
15 <input type="checkbox"/>	UNIPROT:Q7X7D6_ORYSJ	SubName: Full=OSJNBa0074L08.13 protein (Putative uncharacterized protein) (OSJNBa0081C01.1 protein);	422	60.6	50.9	1107	0.00e+00
16 <input type="checkbox"/>	UNIPROT:Q00RJ6_ORYSA	SubName: Full=OSIGBa0155K17.4 protein;	422	60.6	50.9	1107	0.00e+00
17 <input type="checkbox"/>	UNIPROT:A9TEG3_PHYPA	SubName: Full=Predicted protein;	395	56.2	49.2	1070	0.00e+00
18 <input type="checkbox"/>	UNIPROT:A9RX77_PHYPA	SubName: Full=Predicted protein;	407	51.6	46.8	1017	0.00e+00

19	UNIPROT:A9NV23_PICSI	Putative uncharacterized p	427	50.6	76.0	429	2.4e-78
20	UNIPROT:A5BN66_VITVI	Putative uncharacterized p	421	47.9	72.1	420	1.3e-68
21	UNIPROT:A7R2I8_VITVI	Chromosome undetermined sc	421	48.1	72.4	420	1.3e-68
22	UNIPROT:Q9LKK8_SOLOH	Putative 7-transmembrane G	424	48.6	74.8	385	1.1e-63
23	UNIPROT:Q9FJN7_ARATH	G protein-coupled receptor	433	45.6	73.6	397	7.7e-60
24	UNIPROT:Q940G7_ARATH	G protein-coupled receptor	433	45.6	73.6	397	7.7e-60
25	UNIPROT:Q1LZH8_BOVIN	LanC lantibiotic synthetas	399	38.5	67.8	382	7.7e-58
26	UNIPROT:LANC2_MOUSE	LanC-like protein 2 OS=Mus	450	39.4	64.9	436	5.6e-56
27	UNIPROT:B2R602_HUMAN	cDNA, FLJ92709, highly sim	399	40.1	67.6	352	6.5e-56
28	UNIPROT:Q6FHH6_HUMAN	LANCL1 protein (Fragment)	399	39.1	66.8	373	6.5e-56
29	UNIPROT:Q53TN2_HUMAN	Putative uncharacterized p	399	40.1	67.6	352	6.5e-56
30	UNIPROT:LANC1_HUMAN	LanC-like protein 1 OS=Homo	399	40.1	67.6	352	6.5e-56
31	UNIPROT:LANC1_MOUSE	LanC-like protein 1 OS=Mus	399	38.6	68.1	376	1.1e-55
32	UNIPROT:B2KGR1_MOUSE	LanC (Bacterial lantibioti	399	38.6	68.1	376	1.1e-55
33	UNIPROT:Q68FQ9_RAT	LanC (Bacterial lantibiotic	450	39.7	64.2	438	3.1e-55
34	UNIPROT:LANC1_RAT	LanC-like protein 1 OS=Rattus	399	38.8	67.3	376	3.6e-55
35	UNIPROT:B6NDC7_BRAFL	Putative uncharacterized p	403	43.8	68.5	356	4.7e-55
36	UNIPROT:Q5RD57_PONAB	Putative uncharacterized p	399	39.5	67.6	352	6e-55
37	UNIPROT:B6MN61_BRAFL	Putative uncharacterized p	403	43.8	68.3	356	6e-55
38	UNIPROT:B4DWZ7_HUMAN	cDNA FLJ56802, highly simi	432	39.1	64.5	437	1.3e-54
39	UNIPROT:Q6NSL4_HUMAN	LanC lantibiotic synthetas	450	38.9	64.3	437	1.4e-54
40	UNIPROT:LANC2_HUMAN	LanC-like protein 2 OS=Homo	450	39.1	64.5	437	1.4e-54

■ 接前页结果

两两比对结果示意

- 选取相似性高且功能已知的蛋白显示具体比对结果：

左图杨树ABA受体与GCR2比对, 右图水稻LANC-like protein与GCR2 比对, 图中*表示完全一致, ./..表示含有保守氨基酸

```

RESULT 5
ID ASHIY4_POPTR Unreviewed; 443 AA.
DE SubName: Full=Abscisic acid ABA receptor;

DB 11; Score 1363; Match 73.8%; QryMatch 62.7%; Pred. No. 0.00e+00;
Matches 265; Conservative 44; Mismatches 44; Indels 6; Gaps 2;

Db 86 EVVRETWGMGSKRVKDYSLYTGALGTAYLLFKAYQVTKNQNDLNLCSDIKACDSVSRDS 145
Qy 57 kvvwetwersgkrvrtdynlytgvlgtagyllfksyqvtrneddlklclenveacdvasrds 116

*****

Db 146 GRVTFICGLTGCGRAGVLCALGAVAAKHAGDGRLLDHYLAQFKEIKLPSDLPYELLYGRAG 205
Qy 117 ervtficgy----agvcalgavaakclgddqlydrylarfrgirlpsdlpyellygrag 171

****

Db 206 FLWACSFNKHIGKETISTSRKRAVDEIIKAGR-LARKGKCPLLYEWHGKRYWGAAHGL 264
Qy 172 ylwacflnkhiggesissermrsvveifragrqlgnkgctplmyewhgkrywgaahgl 231

****

Db 265 AGIMLVLMDELLETPDEVEDVKGTLHYMIKMHFPGNYPSSSEGKSDCLVHWCHGAPGIAL 324
Qy 232 agimvnlmhtelepdeikdvkgtlsymiqnrfpsgnylssegksdrlvhwchgapgval 291

**

Db 325 TLAKAAEVFGDKEFLQAAMEAGEVWVNRGLLKRVGICHGISGNTYVFLSLYRLTGWVEYL 384
Qy 292 tlwkaaqvyntkefveaameagevwsrgllkrvgichgisgntyvflslyrltrmpkyl 351

*****

Db 385 YRAKAFACFLHDRAQKLVSECKMHGGDRPYSLFEGFGGMAYLFLDILEPSEARFPAYEL 443
Qy 352 yrakafasflldksekliisegqmhgdrpfslfegfggmaymlldmndptqalfpgyel 410
    
```

```

RESULT 7
ID Q337HO_ORYSJ Unreviewed; 399 AA.
DE SubName: Full=Gamma-lanthionine synthetase C-like protein, expressed;

DB 12; Score 1321; Match 65.8%; QryMatch 60.8%; Pred. No. 0.00e+00;
Matches 254; Conservative 44; Mismatches 88; Indels 0; Gaps 0;

*****

Db 14 VAEAPDGGRGLLSLPYSSLSERLLRAALRIKDKVMEETWTRARRQVTDYTLTYTGALGTAL 73
Qy 25 vteckdsltllslpyksfsekhrhryalsikdkvwetwersgkrvrtdynlytgvlgtagy 84

*****

Db 74 LLFKSFQVTGNRADLALAGDIVKECDAASRGLPFLTFICRAGVLCALGAVIAKHQNDQLL 133
Qy 85 llfksyqvtrneddlklclenveacdvasrdservtficgyagvcalgavaakclgddq 144

**

Db 134 LTHYLSSFDELIIVTEKVPNELLYGRAGYLWACLFLNTHLGEKTIPHEHITSVAKDIIDEG 193
Qy 145 ydrylarfrgirlpsdlpyellygragylwacflnkhiggesissermrsvveifrag 204

**

Db 194 RKLAKGKNCPLMYEWHGKRYWGAAHGLAGIMVLMHTELEPDEIKDVKGTLSYMIQNRFP 253
Qy 205 rqlgnkgctplmyewhgkrywgaahglagimvnlmhtelepdeikdvkgtlsymiqnrfp 264

****

Db 254 TGNYPSSGESDRLVHWCHGAPGVALTLAKAYQVFHDEHFQTAEEAAEVWVNRGLLKR 313
Qy 265 sgnylssegksdrlvhwchgapgvaltlvkaaqvyntkefveaameagevwsrgllkr 324

*****

Db 314 VGICHGVSNGNAYVFLSLYRLTGWVEYLRAKAFACFLLEKADQLIADGAMHGGDHPFSLF 373
Qy 325 vgichgisgntyvflslyrltrmpkylrakafasflldksekliisegqmhgdrpfslf 384

**

Db 374 EGRAGMAYLLDMVSPSESKFPAYEL 399
Qy 385 egiggmaymlldmndptqalfpgyel 410
    
```

2 motif检索

应用ExPaSy中相关软件对GCR2进行motif相似性的检索(Pfam, HITS, FingerPRINTSCAN, InterProscan, SMART, Propsearch)

Pfam

大规模蛋白家族数据库，由Sanger研究中心维护，输入序列即可，E value设为1.0，结果得到GCR2中含有LANC-like domain（69-410）

Sequence search results

[Show](#) the detailed description of this results page.

We found **1** Pfam-A match to your search sequence (**all** significant). You did not choose to search for Pfam-B matches. The graphic below shows the arrangement of matches on your sequence:



[Show](#) the search options and sequence that you submitted.

[Return](#) to the search form to look for Pfam domains on a new sequence.

Significant Pfam-A Matches

[Show](#) or [hide](#) all alignments.

Pfam-A	Description	Entry type	Sequence		HMM		Bits score	E-value	Alignment mode	Predicted active sites	Show/hide alignment
			Start	End	From	To					
LANC like	Lanthionine synthetase C-like protein	Family	69	410	1	404	459.6	4.4e-135	ls	n/a	Show

Pfam中关于LANC-like的特征注释:

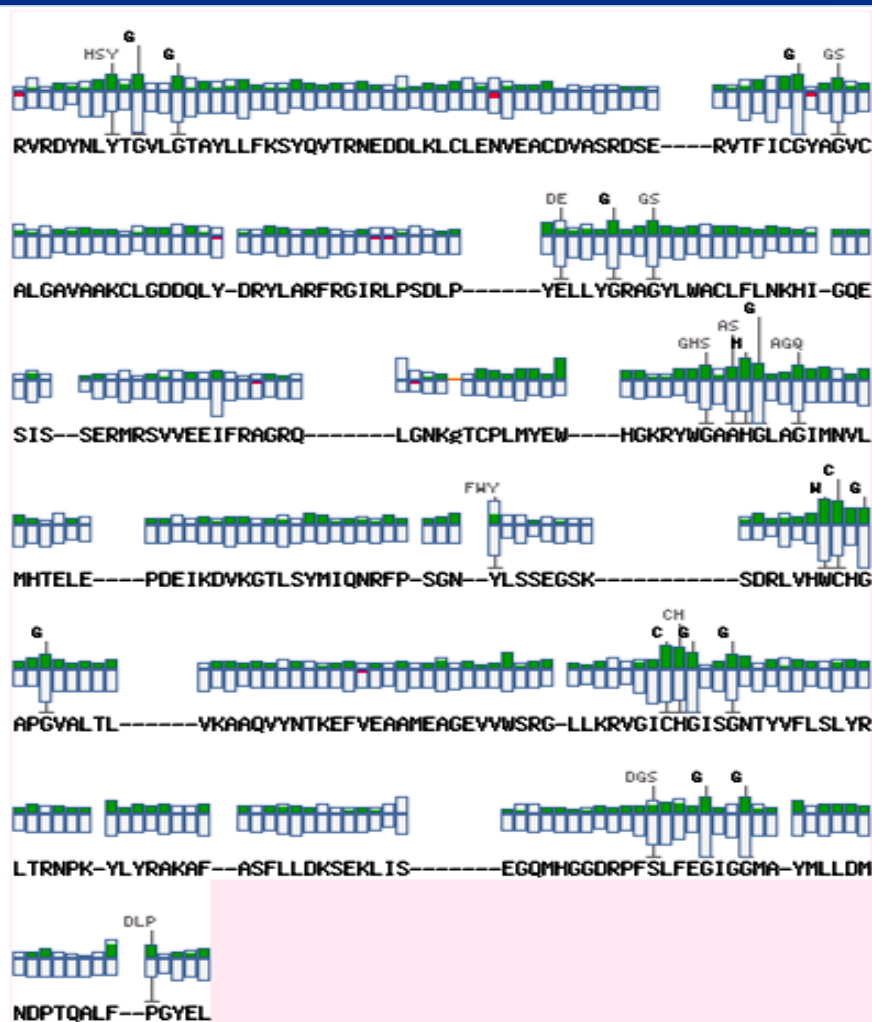
- Lanthionines are thioether bridges that are putatively generated by dehydration of Ser and Thr residues followed by addition of cysteine residues within the peptide. This family contains the lanthionine synthetase C-like proteins 1 and 2 which are related to the bacterial lanthionine synthetase components C (LanC). LANCL1 (P40 seven-transmembrane-domain protein) and LANCL2 (testes-specific adriamycin sensitivity protein) are thought to be peptide-modifying enzyme components in eukaryotic cells. Both proteins are produced in large quantities in the brain and testes and may have role in the immune surveillance of these organs [1]. Lanthionines are found in lantibiotics, which are peptide-derived, post-translationally modified antimicrobials produced by several bacterial strains [2]. This region contains seven internal repeats
- 其中提到的p40也具有跨膜结构（虽然有研究提出异议），所以单凭LANC-like结构域还不能说明GCR2不是GPCR

HITs

motif搜索，显示蛋白序列与motif间的关系

参数：默认，在全部数据库中搜索

结果中有比较完整的长序列一致性的显示GCR2的部分区域（69-410）与LANC和LANC-like domain比较相似：



Status: !
pos.: 69-410
raw-score =
457.7
N-score =
142.746
E-value = 3.8e-
136

pfam_fs:LANC_like
Lanthionine
synthetase C-like
protein
[[entry](#)]

FingerPRINTSCAN

根据motifs的指纹特征来确认某个蛋白家族

参数: FPScan/FPScan-fam:e value 0.0001(默认); database:

PRINTS39_0; Matrix: blos62; Distance variance 10%;

结果如下, 仍然显示可能存在LANC样区域

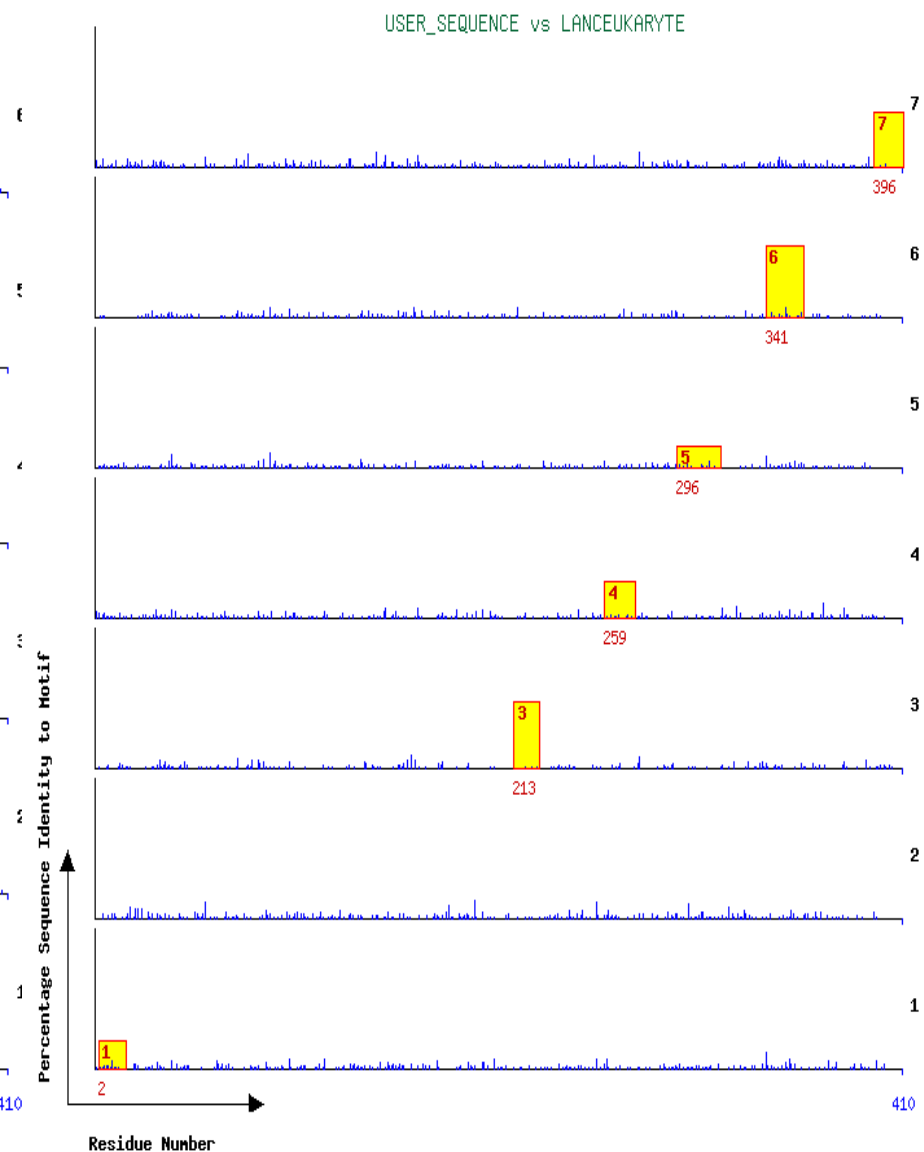
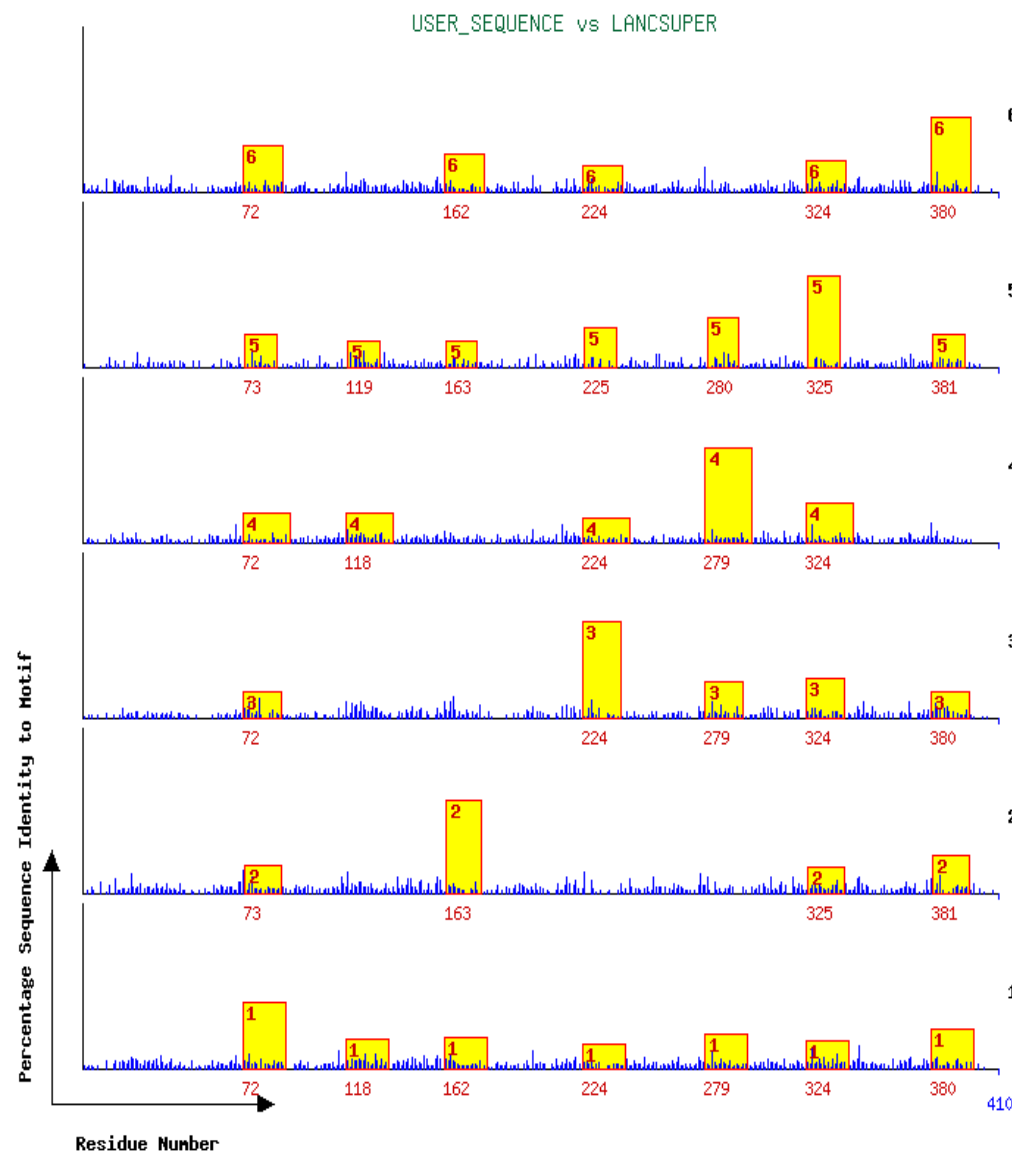
PRINTS39_0 and matrix blos62

Scan of sequence: USER_SEQUENCE

Highest scoring fingerprints for your query

Fingerprint	E-value	GRAPHScan	Motif3D
<u>LANCSUPER (relations)</u>	5.379991e-50	<u>Graphic</u>	
<u>LANCEUKARYTE (relations)</u>	3.816264e-18	<u>Graphic</u>	

GCR2与原核（左）真核（右）LANC家族蛋白的特征序列一致性



InterProScan









整合PROSITE, Pfam, PRINTS and other family and domain databases的蛋白特征查询

输入序列, 参数默认: 仍得到69-410LANC-like family相似序列, 同时245-363为six-hairpin glycosidase domain相似区

InterProScan Results

Table View Raw Output XML Output Original Sequences SUBMIT ANOTHER JOB

SEQUENCE: [Sequence_1](#) CRC64: DB8240DC0A60D6F9 LENGTH: 410 aa

InterPro IPR007822 Family  	Lanthionine synthetase C-like PF05147  LANC_like
InterPro IPR008928 Domain  	Six-hairpin glycosidase-like SSF48208  Six-hairpin glycosidases
noIPR unintegrated	unintegrated PTHR12736  LANC-LIKE PROTEIN PTHR12736:SF4  LANC-LIKE

SMART

simple modular architecture research tool, 一个综合的基因组和蛋白结构功能预测工具, 包括了酶催化位点, 多结构域预测, 蛋白相互作用预测, 蛋白无规则结构预测等。结果中1-22氨基酸序列区为无规则结构区 (from DisEMBL); 23-68未知; 69-410为LANC-like domain (from Pfam) 图略; 下图则显示在阈值一下还有一些低可能性的motifs

Name	Begin	End	E-value	Reason
intrinsic disorder	51	76	-	overlap
PA2c	58	153	2.52e+02	threshold
H4	68	130	3.33e+03	threshold
PSA	75	139	3.14e+03	threshold
DSL	87	138	1.93e+03	threshold
DISIN	104	149	2.12e+03	threshold
EGF_CA	105	139	2.73e+03	threshold
intrinsic disorder	108	116	-	overlap
FOLN	120	139	3.33e+02	threshold
TDU	194	211	1.39e+03	threshold
intrinsic disorder	202	211	-	overlap
SMR	214	295	6.24e+02	threshold
acidPPc	232	338	1.94e+03	threshold
cNMP	239	369	1.23e+03	threshold
A1pp	247	369	1.29e+03	threshold
ZnF_CHCC	275	313	2.55e+03	threshold
HNS	340	369	5.04e+03	threshold
intrinsic disorder	371	379	-	overlap

Propsearch

- Sequence homology using amino acid composition. (by Uwe Hobohm and Chris Sander), 相对于利用数据库中的蛋白质序列的比对, 该程序利用对蛋白质氨基酸组成, 残基性质, 等电点及分子量等几十组性质来查找可能的同源蛋白家族, 该方法对于找到一些低一致性 (如25%以下, 用blast无法分析的) 蛋白间的同源性有一定的帮助. 数据库中蛋白和查询蛋白均用向量表示 (根据各种比对所需性质), 结果用查询蛋白与已知蛋白间的距离 (DIST) 分组, 距离越小表示同源性越高。如0-1.3表示相似可能性为99.9%, 1.3-7.5可能性99.6%, 而8.7-10表示相似可能性为87%。

Propsearch 结果，下列结果并无直接显示GCR2与LANC有相似结果(所用数据库和算法限制?)，但从结果中蛋白的性质看，GCR2的功能或者结构也可能更接近酶的性质

Rank	ID	DIST	LEN2	POS1	POS2	pI	DE
1	cisy_sulso	5.94	377	1	377	7.86	Citrate synthase (EC 4.1.3.7).
2	actz_drome	5.97	376	1	376	7.32	Actin-like protein 87C.
3	actz_human	6.00	376	1	376	6.63	Alpha-centractin (Centractin) (Centrosome-associated actin homolog)
4	syh_trepa	6.14	442	1	442	7.81	Histidyl-tRNA synthetase (EC 6.1.1.21) (Histidine--tRNA ligase)
5	ygde_ecoli	6.23	366	1	366	7.33	Hypothetical protein ygdE.
6	cisy_syny3	6.28	397	1	397	5.98	Citrate synthase (EC 4.1.3.7).
7	vnfd_azovi	6.30	473	1	473	6.38	Nitrogenase vanadium-iron protein alpha chain (EC 1.18.6.1)
8	trb2_pyrab	6.30	451	1	451	6.67	Tryptophan synthase beta chain 2 (EC 4.2.1.20).
9	nucm_pylli	6.32	398	1	398	6.99	NADH-ubiquinone oxidoreductase 49 kDa subunit (EC 1.6.5.3)
10	tpl_escin	6.33	456	1	456	6.41	Tyrosine phenol-lyase (EC 4.1.99.2) (Beta-tyrosinase).
11	acty_human	6.38	376	1	376	6.38	Beta-centractin (Actin-related protein 1B) (ARP1B).
12	tpl_symth	6.42	458	1	458	6.77	Tyrosine phenol-lyase (EC 4.1.99.2) (Beta-tyrosinase).
13	tpl_citfr	6.45	456	1	456	6.30	Tyrosine phenol-lyase (EC 4.1.99.2) (Beta-tyrosinase).
14	nucm_human	6.46	463	1	463	7.54	NADH-ubiquinone oxidoreductase 49 kDa subunit, mitochondrial precursor
15	tpl_erwhe	6.54	456	1	456	6.34	Tyrosine phenol-lyase (EC 4.1.99.2) (Beta-tyrosinase).
16	nud1_rhime	6.56	396	1	396	6.46	NADH dehydrogenase I chain D 1 (EC 1.6.5.3) (NADH-ubiquinone
17	nucm_caeel	6.56	482	1	482	6.47	Probable NADH-ubiquinone oxidoreductase 49 kDa subunit, mitochondrial
18	bioa_metja	6.57	464	1	464	6.23	Adenosylmethionine-8-amino-7-oxononanoate aminotransferase
19	pafa_chick	6.58	422	1	422	6.44	Platelet-activating factor acetylhydrolase precursor (EC 3.1.1.47)
20	dcd_drosi	6.60	437	1	437	6.65	Aromatic-L-amino-acid decarboxylase (EC 4.1.1.28) (AADC) (DOPA)

motif检索结果

- (Pfam, HITS, FingerPRINTSCAN, InterProscan, SMART, Propsearch)
- 均得到GCR2的69-410区含LANC-like相似序列

第三部分 多序列比对

GCR2和其他GPCR比对（结果未显示）

最近报道了拟南芥中的两个新的ABA GPCR受体GTG1和GTG2（8），他们同人的GPR89的序列一致性达45%和68%。用needle分析GTG1和GPR89同GCR2的序列一致性为1.8%和7.6%。Pandey等用PROSITE分析的GTG1/2的一个序列特征GTG1的382–411序列处是比较保守的ATP-/GTP-binding region，GCR2在该处同GTG1有较高的保守性，但是两端有缺失的序列。

第四部分 搜索GCR2三维结构 (三级结构)

■ 1 iMolTalk

■ 【功能】 搜索同源蛋白的结构

■ 【程序路径】 <http://www.imoltalk.com/>

■ 【输入参数】 默认

■ 【处理结果】

```
Database PDBChainSaw was created Mar 11, 2009 2:15 PM
Number of sequences with aa count ≥ 20: 125,308
Database iMolTalkMTDBpublished was created Mar 11, 2009 2:15 PM
Number of sequences with aa count ≥ 20: 0

Select a homologous structure:

■ ■ 40% 6.4e-05 PDB: 2G0D 2G0D65  [2.21, BIOSYNTHETIC PROTEIN,
226 GAAHGLAGIMNVLMHTELEPDEIKDVKGTLSYMI-----QNRFPSGNYLSSEGSK 275
203 GLAHGLAGVGCILAYAHIRGYSNEASLSALQKIIFIYEKFELERRKQFLWRDGLVADELK 262

276 SDRLVH-----WCHGAPGVALTLVKAAQVYNTKEFVEAAMEAGEVVWSRGL-LKRVG 326
263 KEKVIREASFIRDAWCYGPGISLLYLYGGLALDNDYFVDKAEKILESAMQRKLGIDSYM 322

327 ICHGISGNTYVFLSLYRLTRNPKYLYRAKAFASFLLD---KSEKLISEGQMHGDRPFSL 383
323 ICHGYSG-----LIBICSLFKRLLNTKFDSYMEEFNVNSEQILEE---YGDESGTGF 372

384 FEGIGMAYML 394
373 LEGISCILVL 383
```

■ 【结果分析】

■ (1) 找到同源性

■ 最高的结构信息为NisC蛋白 (PDB ID: 2G0D);

■ (2) 为了更加形象观察其相似性, 使用 [Phyre \(Successor of 3D-PSSM\)](#) 软件进一步预测相三维结构的同源物。

2 Phyre (Successor of 3D-PSSM)

- 【功能】通过轮廓匹配以及二级结构寻找三维结构类似的蛋白
- 【程序路径】 <http://www.sbg.bio.ic.ac.uk/~phyre/index.cgi>
- 【输入参数】默认
- 【处理结果】

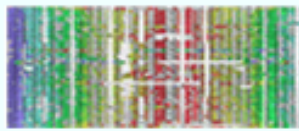
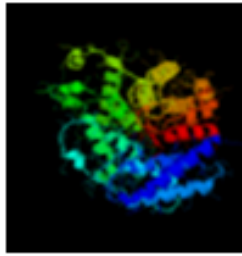
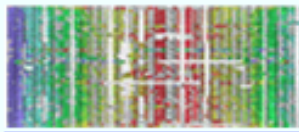
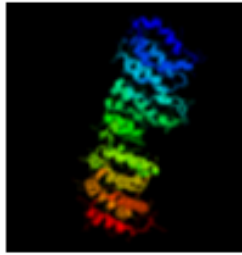

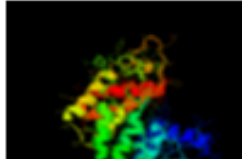
QuickPhyre Results for Job GCR2_____

Please cite:

Protein structure prediction on the web: a case study using the Phyre server.
Kelley LA & Stenberg MJE *Nature Protocols*. 4, 363 - 371 (2009)

Email:	mammalia@163.com
Job Code:	55e7b5dfcd4a4368
Description:	GCR2_____
Date:	Mon Mar 16 09:39:40 GMT 2009

Fold Recognition

View Alignments	SCOP Code	View Model	E-value	Estimated Precision	Bio Text	Fold/PDB descriptor	Superfamily	Family
	c2g0dA ↓ (length: 409) 15% i.d.		2.5e-27	100 %	n/a	PDB header: biosynthetic protein	Chain: A: PDB Molecule: nisin biosynthesis protein nisc;	PDBTitle: nisin cyclase
	d1b3ua ↓ (length: 588) 7% i.d.		0.00018	95 %	n/a	alpha-alpha superhelix	ARM repeat	HEAT repeat
	d1nc5a (length:		0.00023e	95 %	n/a	alpha/alph	Six-hairpin	Hypothetical protein

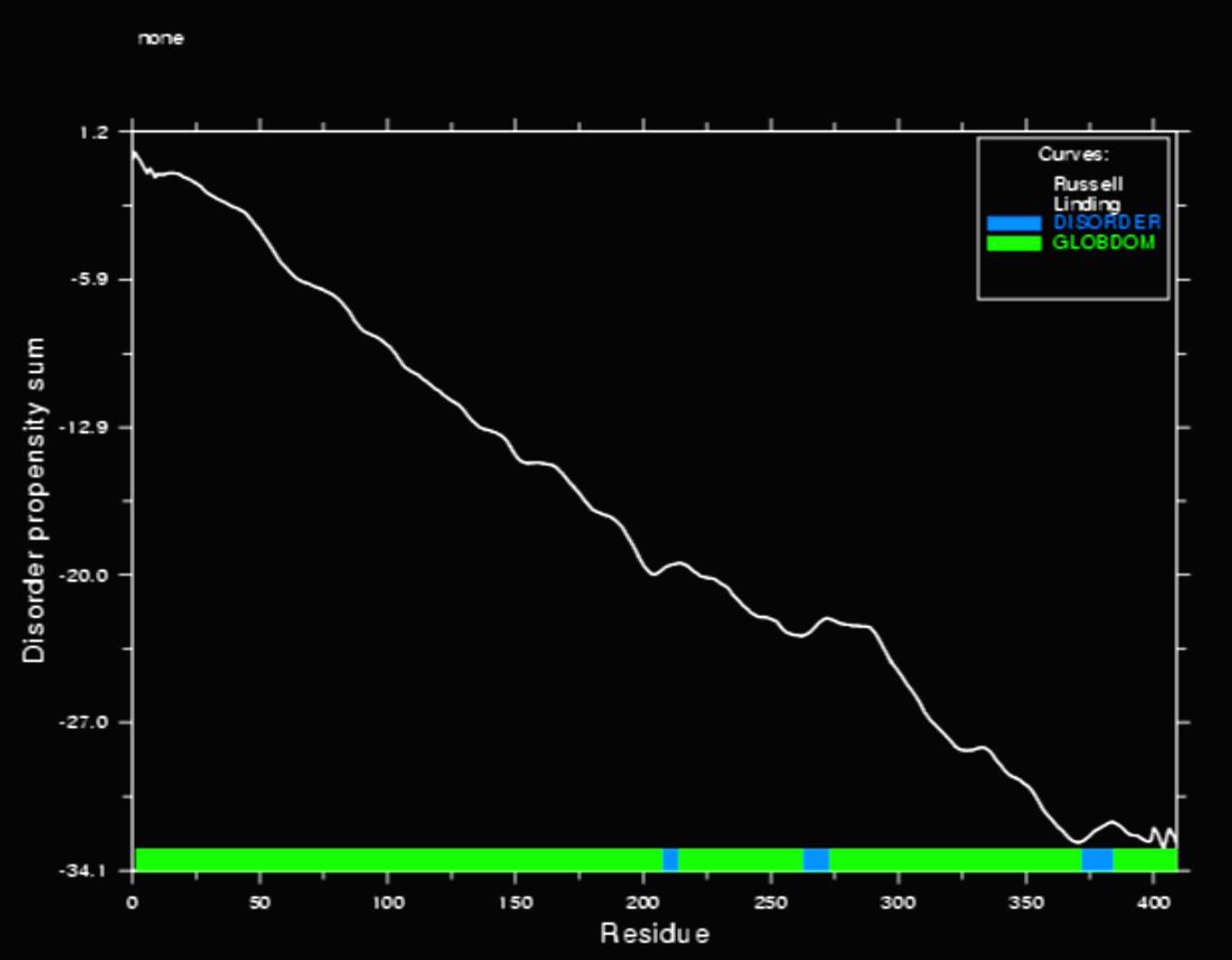
第四部分 搜索GCR2三维结构 (四级结构)

1 GlobPlot

- 【功能】 预测蛋白质序列中的紊乱区、球蛋白区等
- 【程序路径】 <http://globplot.embl.de/>
- 【输入参数】 默认

- Disordered by Russell/Linding definition
- >none_Disorder 208-214, 263-273, 372-384
- mgerffrnem pefvpedlsg eeetvteckd sltkllslpy ksfseklhry alsikdkvww etwersgkrv rdynlytgvl gtayllfksy qvtrneddlk lclenveacd vasrdservt ficgyagvca lgavaakclg ddqlydryla rfrgirlpsd lpyellygra gylwacfln khigqesiss ermrvveei fragrql**GNK GTCPLmyewh** gkrywgaahg lagimnvlmh telepdeikd vkgtlsymiq nr**FPSGNYLS SEG**sksdrlv hwchgagva ltlvkaaavy ntkefveaam eagevwsrg llkrvgichg isgntyvfls lyrtrnpsy lyrakafasf lldkseklis e**GQMHGGDRP FSLF**egiggm aymlldmndp tqalfpgyel

- Potential globular domains (GlobDoms) by Russell/Linding definition
 - >none_GlobDoms 2-410
- mGERFFRNEM PEFVPEDLSG EEETVTECKD SLTKLLSLPY KSFSEKLHRY ALSIKDKVWW ETWERSGKRV RDYNLYTGVL GTAYLLFKSY QVTRNEDDLK LCLENVEACD VASRDSERT FICGYAGVCA LGAVAAKCLG DDQLYDRYLA RFRGIRLPSD LPYELLYGRA GYLWACLFLN KHIGQESISS ERMRSVVEEI FRAGRQLGNK GTCPLMYEWH GKRYWGAAHG LAGIMNVLHM TELEPDEIKD VKGTLSYMIQ NRFPSGNYLS SEGSKSDRLV HWCHGAPGVA LTLVKAAQVY NTKEFVEAAM EAGEVWSRG LLKRVGICHG ISGNTYVFLS LYRLTRNPKY LYRAKAFASF LLDKSEKLIS EGQMHGGDRP FSLFEGIGGM AYMLLDMNDP TQALFPGYEL



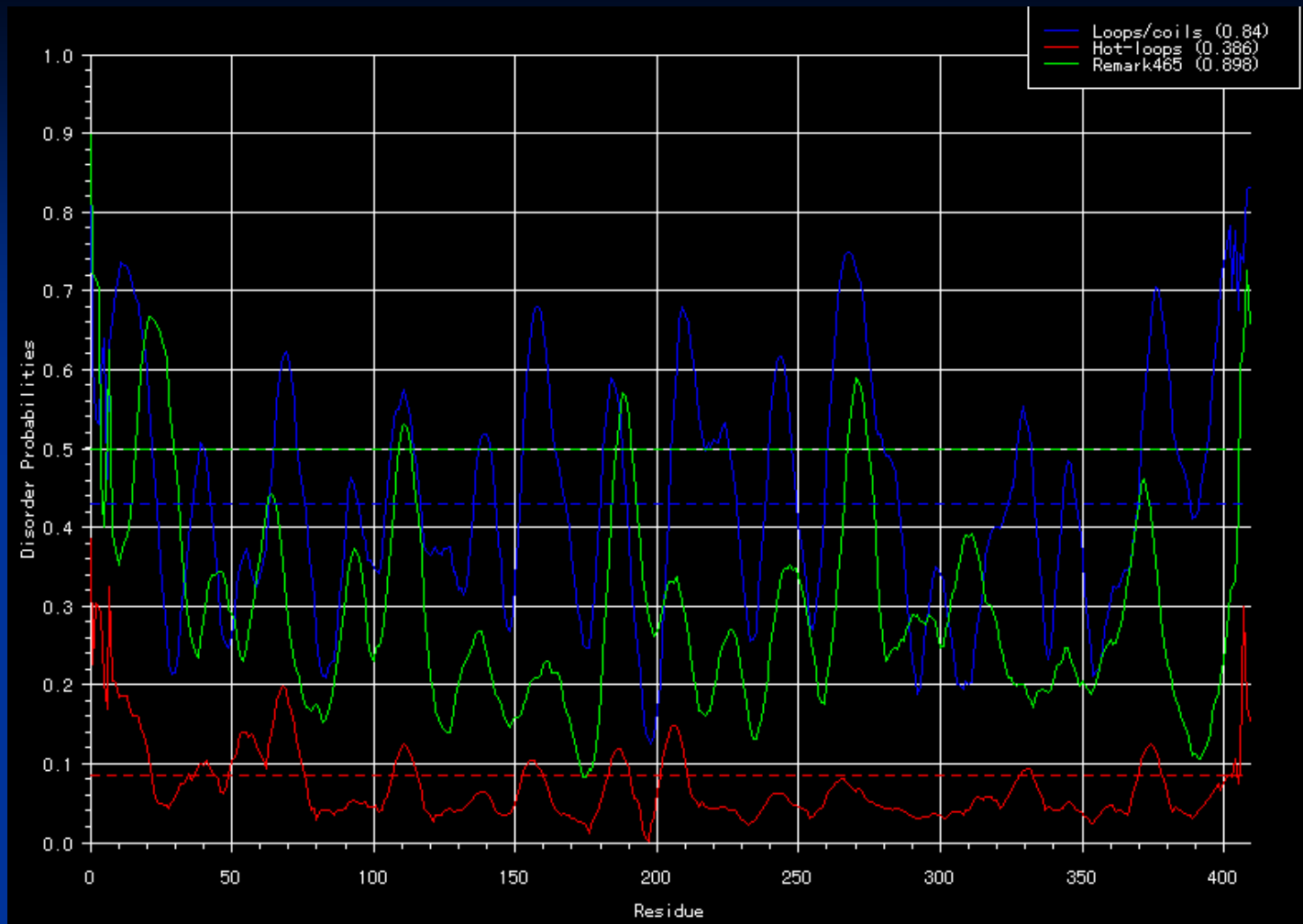
【结果分析】

(1) 找到3个较为不稳定的区域，即208-214, 263-273, 372-384，这些位置，由于可变性，很可能参与该蛋白的酶活性或者蛋白相互作用，在点突变法探寻蛋白功能机制时应当重点关注；

(2) 2-410为潜在的球蛋白区，而该蛋白全场也是410个氨基酸残基而已，佐证了关于该蛋白不具有跨膜结构的分析。

2 DisEMBL 1.5

- 【功能】 预测蛋白质序列中的紊乱区、球蛋白区等，
- 【程序路径】 <http://i.moltalk.org/>
- 【输入参数】 默认



【结果分析】

(1) GCR2含有十多个loop或者coil结构，说明其蛋白多螺旋，印证了之前对其球蛋白的判断；

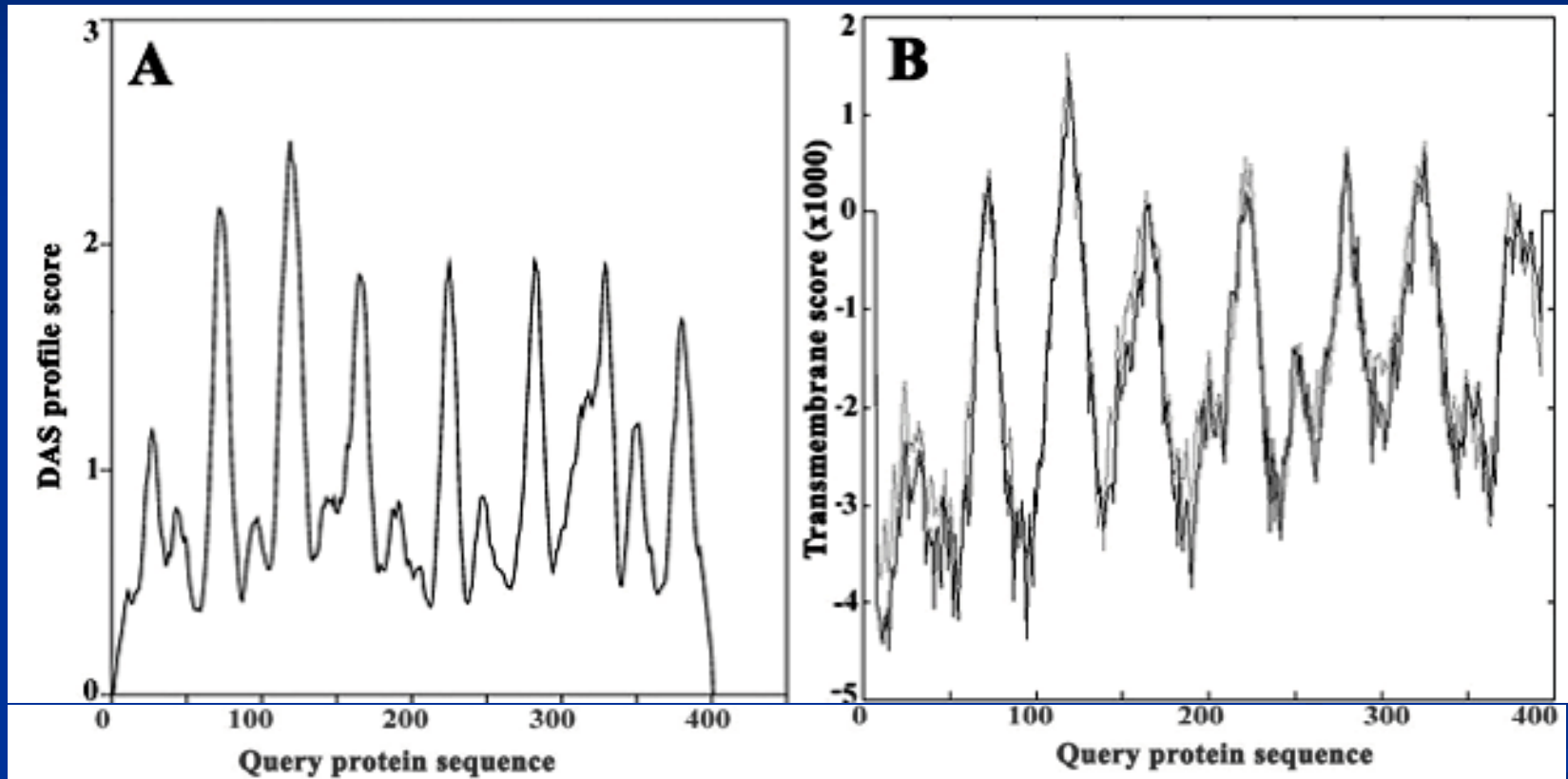
(2) 在1、40、55、65、110、150、180、200、350等处存在较显著的Hot-loops结构，很可能参与该蛋白的酶活性或者蛋白相互作用，在探寻蛋白功能机制时应当重点关注。

第五部分 讨论和体会

GCR2膜结合G蛋白偶联受体之争

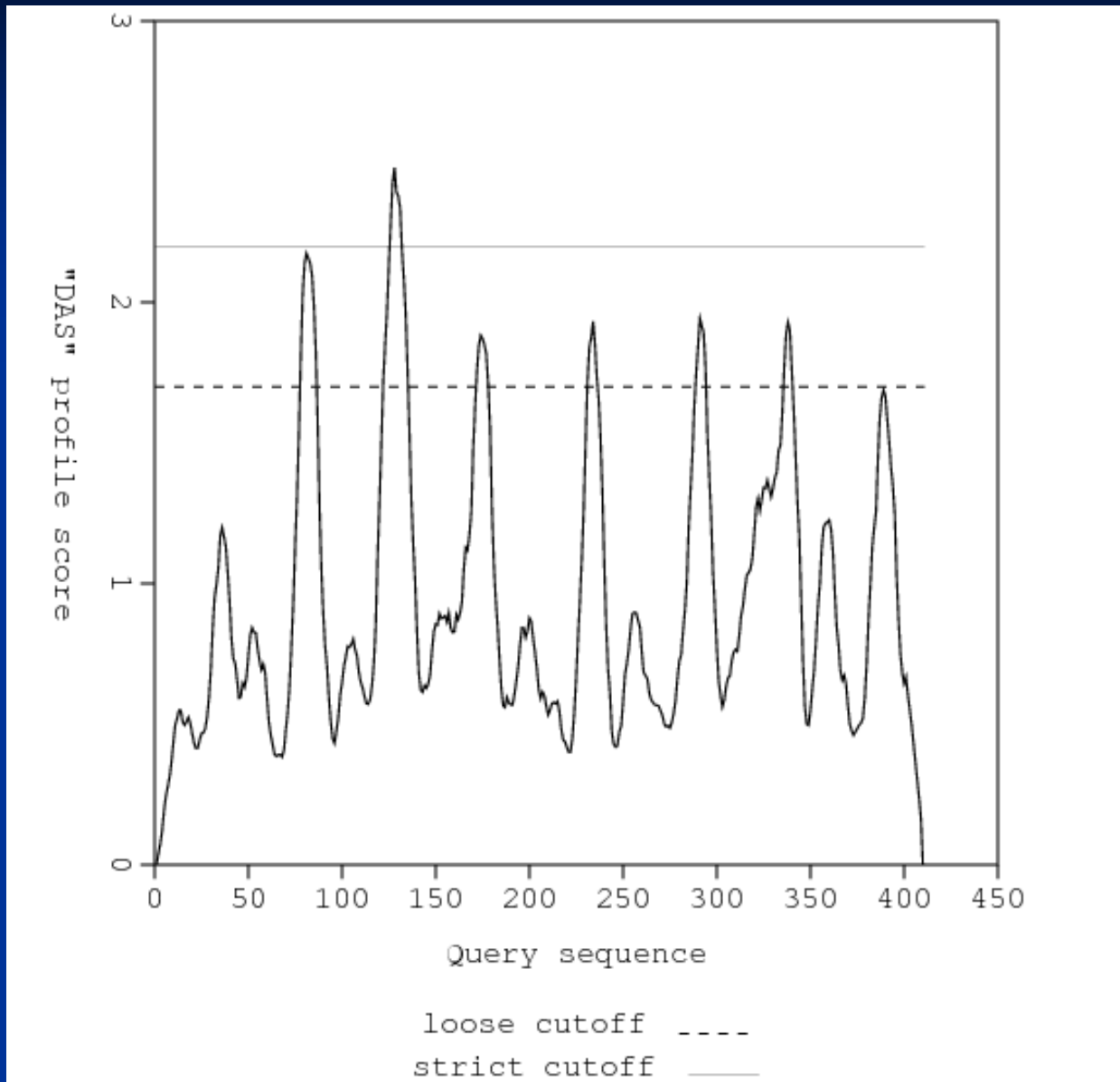
- 在拟南芥中的一个标准的G蛋白的通路是由一个G α 亚基、一个G β 亚基和两个G γ 亚基组成的。这些亚基在植物系统中研究得比较少，目前只有一个确定的G蛋白受体GCR1。
- 而刘西岗等人为了证明GCR2也是一个参与植物激素ABA代谢的G蛋白受体，是一个结合G α 亚基（GPA1）的蛋白。
- 他们07年在《A G protein-coupled receptor is a plasma membrane receptor for the plant hormone abscisic acid》中提出了一系列证据。

生物信息学方法预测（DAS和TMpred）



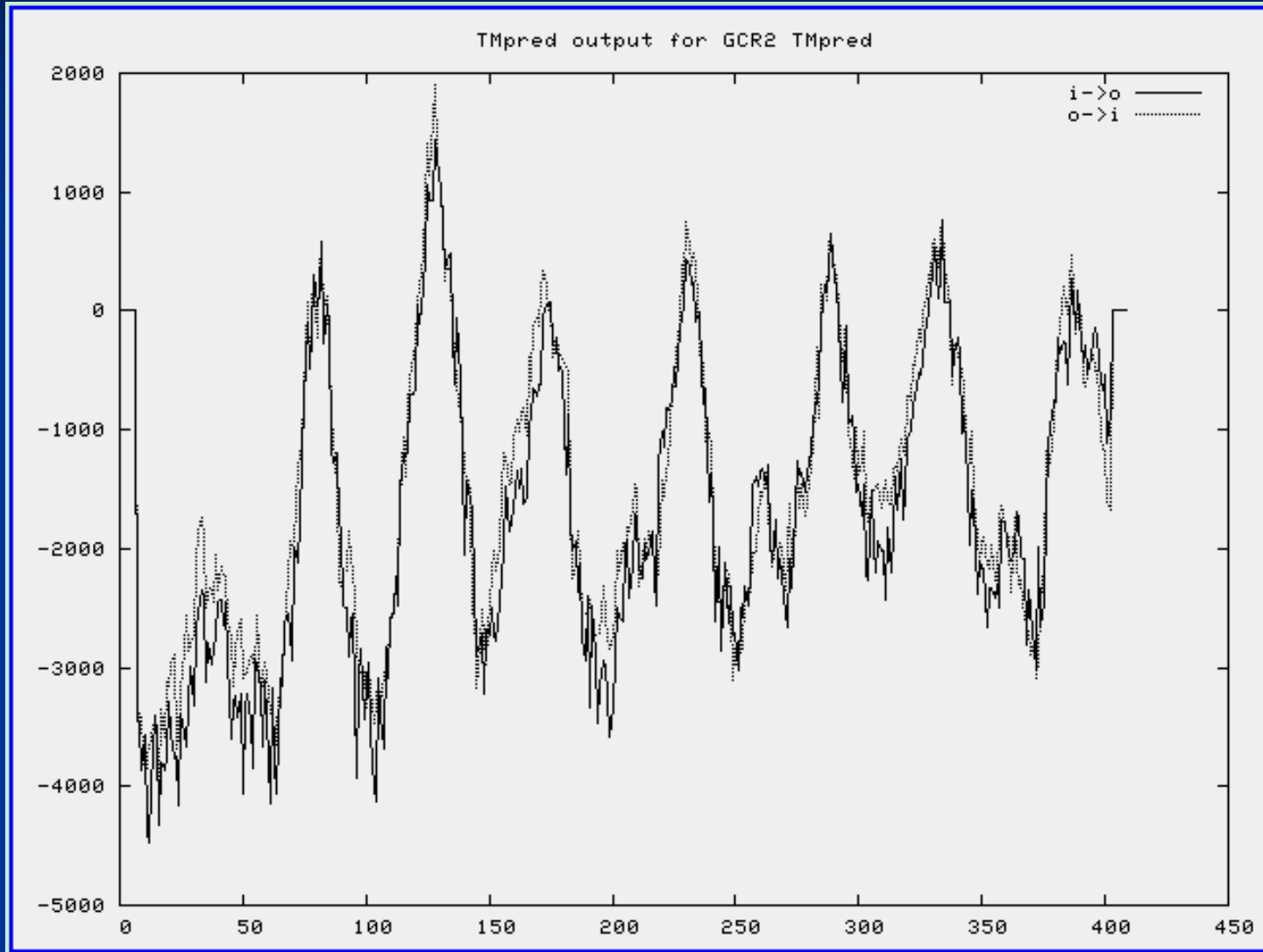
Liu X, Yue Y, Li B, Nie Y, Li W, Wu WH, Ma L. A G protein-coupled receptor is a plasma membrane receptor for the plant hormone abscisic acid. *Science*. 2007 315(5819): 1712-6.

DAS



在原核生物中利用**Dense Alignment Surface method**预测蛋白质的跨膜结构域

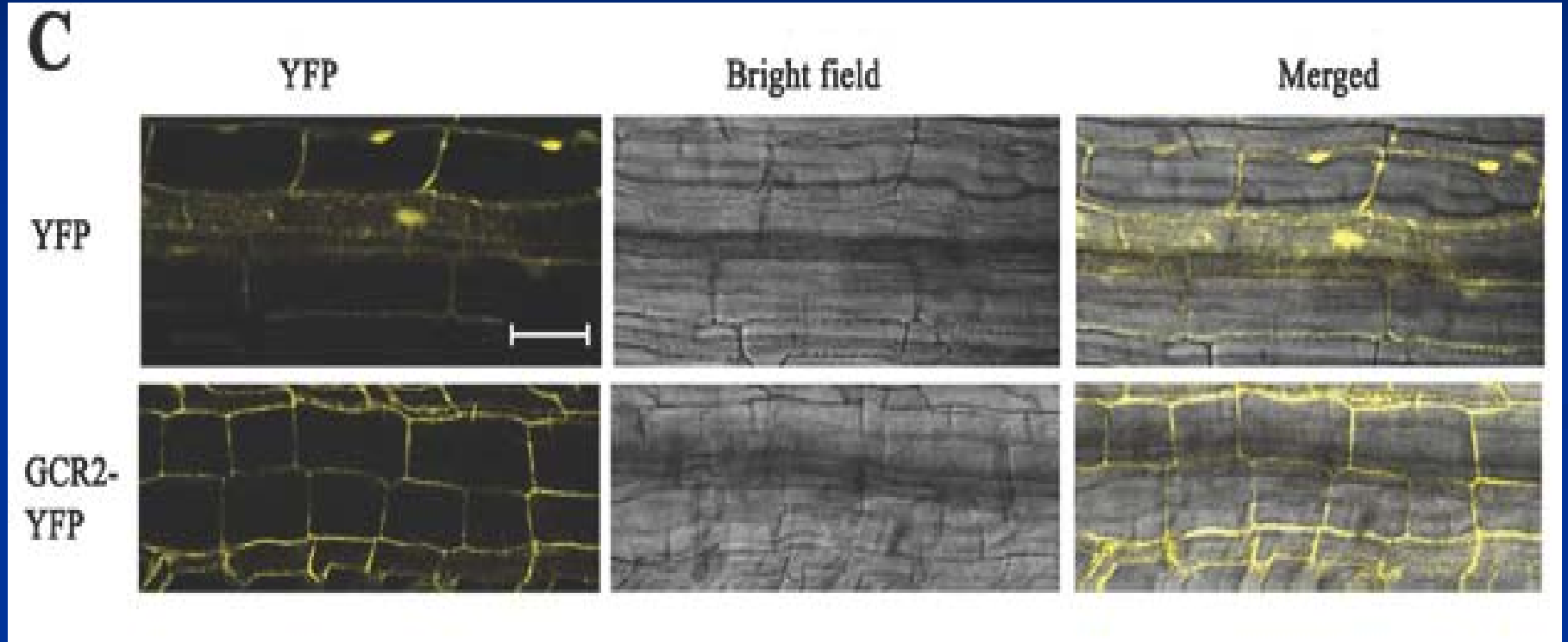
TMpred



有7个跨膜结构域，而第1、3、7是不显著的跨膜结构域。

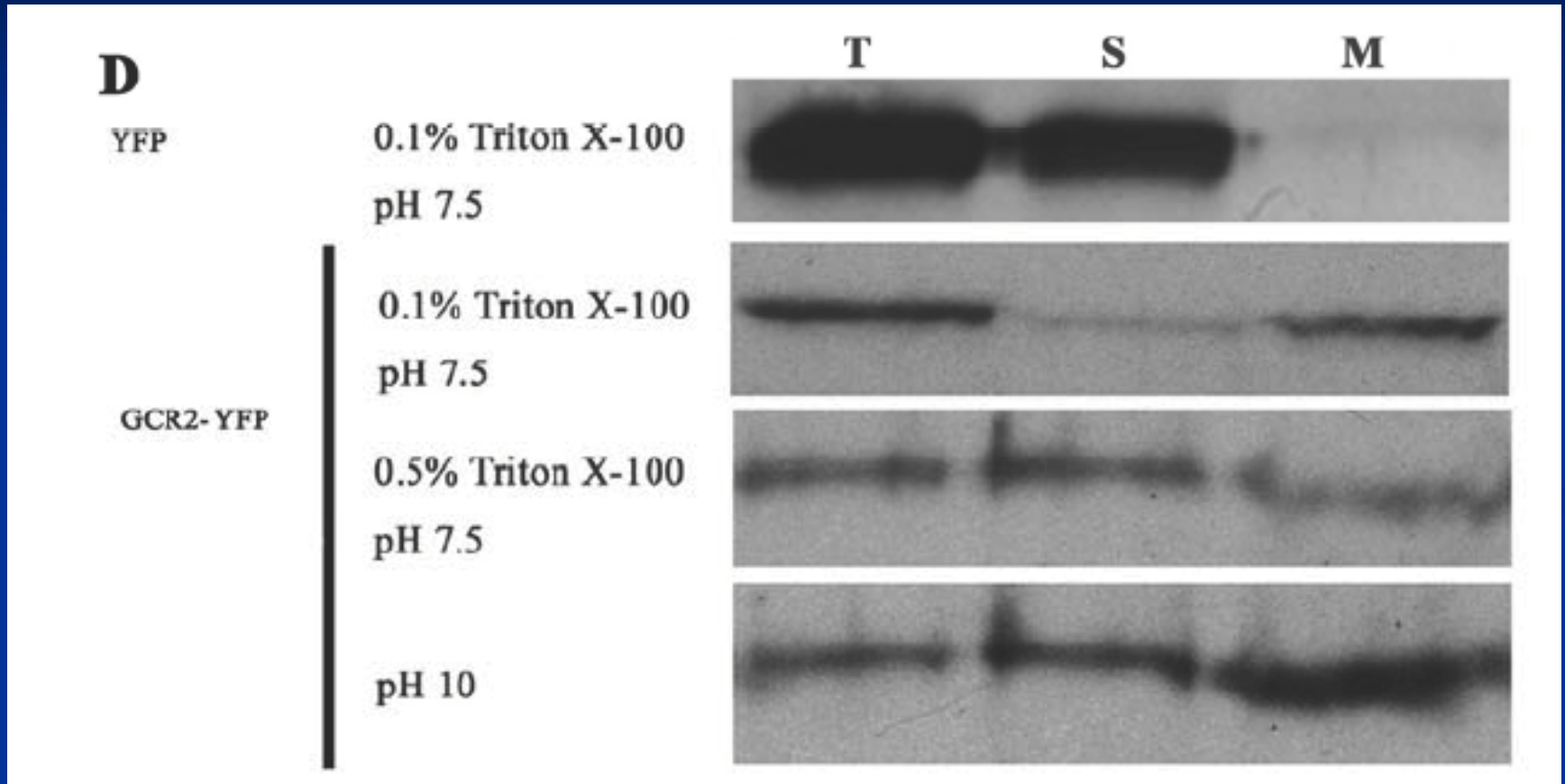
直接观察GCR2在植物根中的亚细胞定位

YFP标记GCR2蛋白（GCR2-YFP）



Liu X, Yue Y, Li B, Nie Y, Li W, Wu WH, Ma L. A G protein-coupled receptor is a plasma membrane receptor for the plant hormone abscisic acid. *Science*. 2007 315(5819): 1712-6.

western blotting检测在全细胞、胞质、细胞膜上， GCR2-YFP的分布



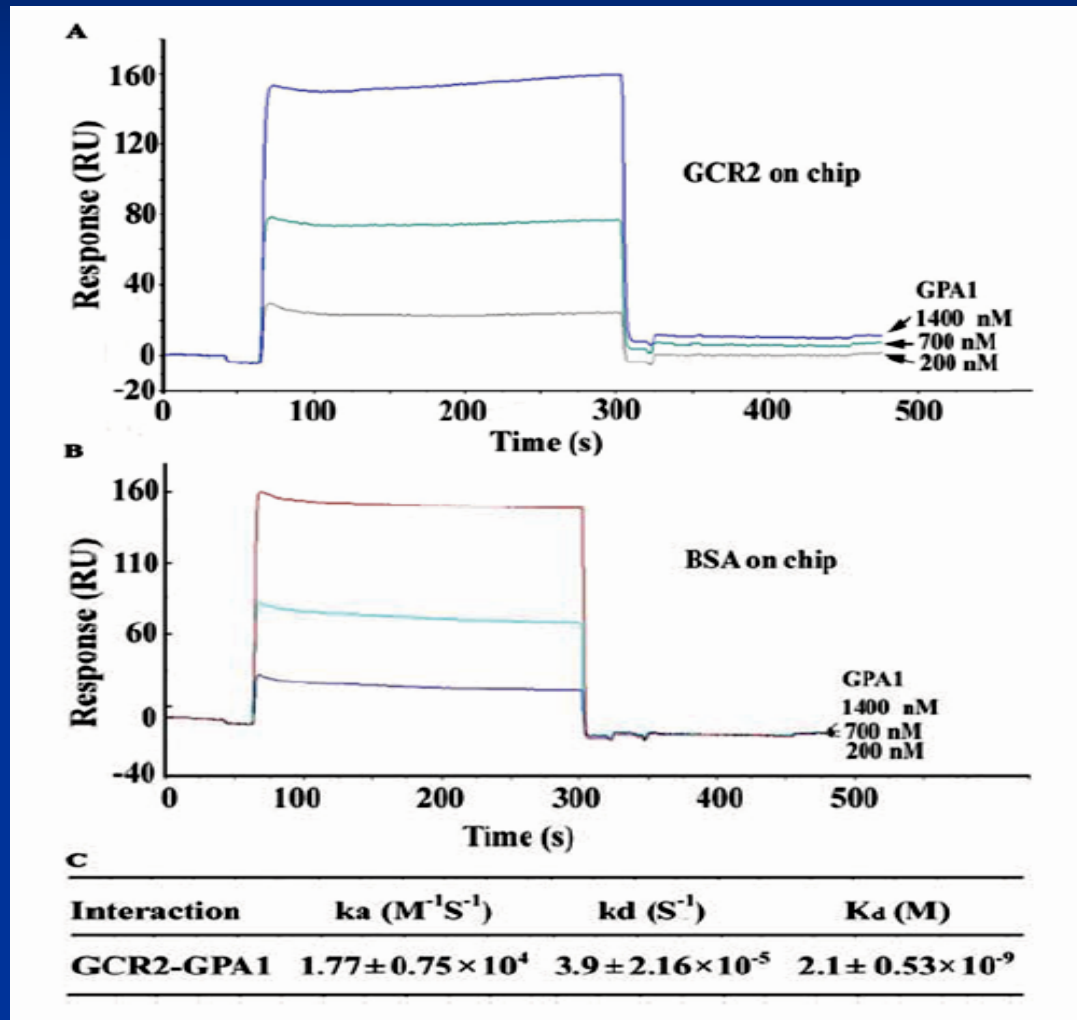
用变性剂和高pH溶液洗细胞，发现GCR2-YFP仍然主要定位在膜组分上，证明GCR2是一个结合在膜上的蛋白。

Liu X, Yue Y, Li B, Nie Y, Li W, Wu WH, Ma L. A G protein-coupled receptor is a plasma membrane receptor for the plant hormone abscisic acid. *Science*. 2007 315(5819): 1712-6.

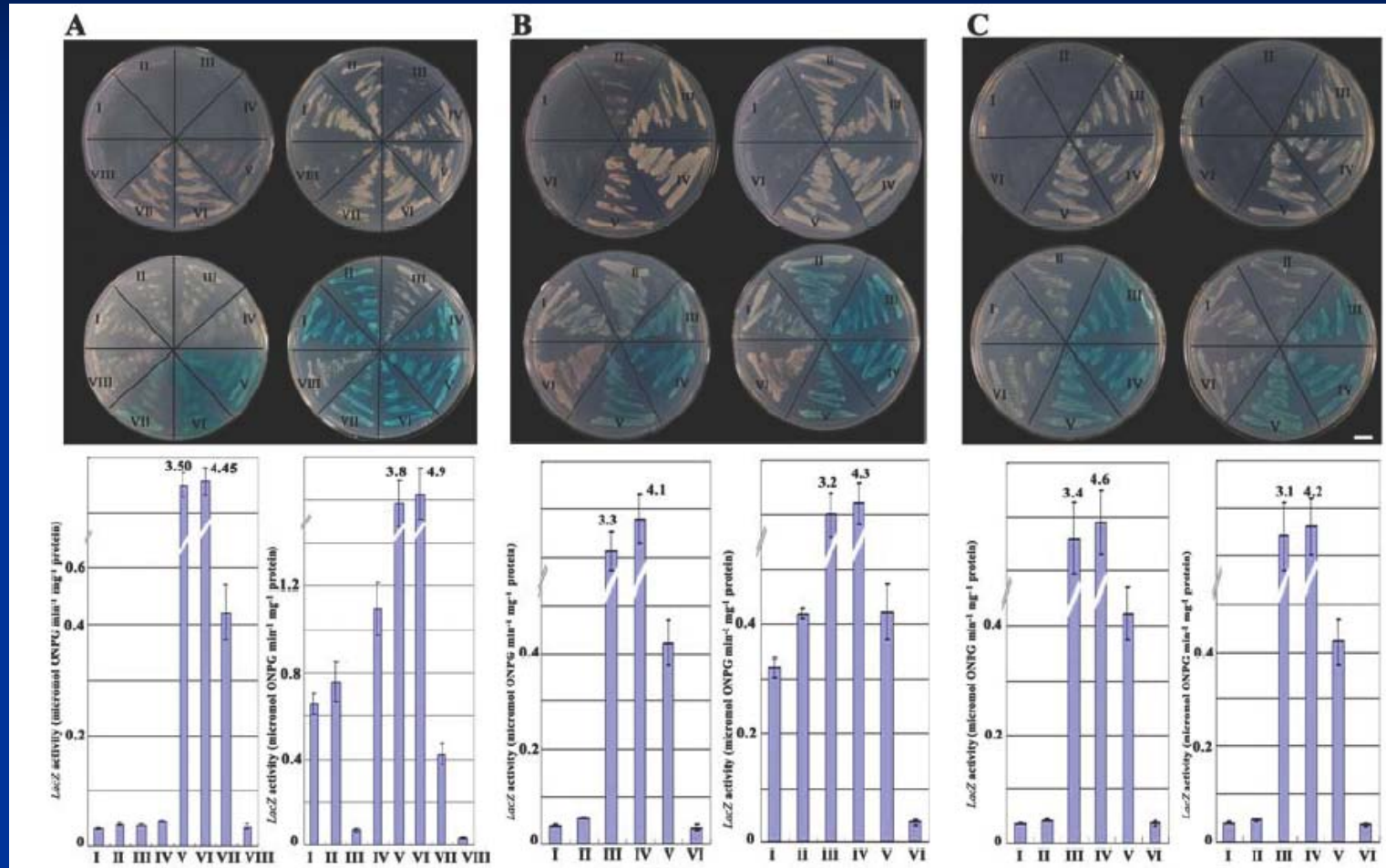
GCR2与G α 亚基（GPA1）的相互作用

- 为证明GCR2是一个参与ABA代谢通路的G蛋白受体，并推测其与G α 亚基存在相互作用，作者采用了4种方法：
 - SPR assay
 - split-ubiquitin Y2H
 - 双分子荧光共振
 - Co-IP

SPR assay: GCR2能与GPA1结合, BSA作为负对照不能与GPA1结合。

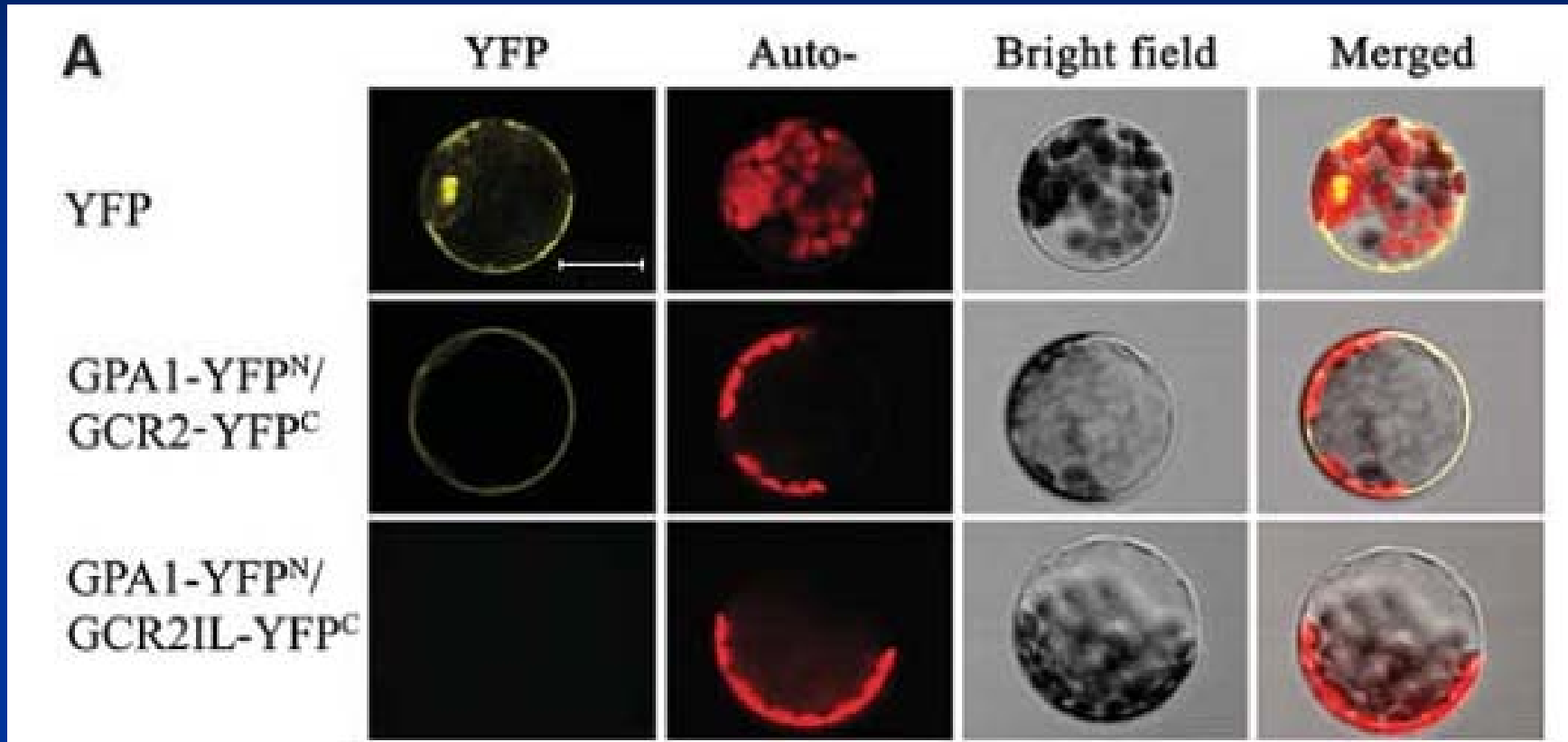


Split-ubiquitin Y2H



证明了全长的GCR2与GPA1的相互作用、GCR2的free C端（290-401）与GPA1的相互作用，但是N端与GPA1没有相互作用。

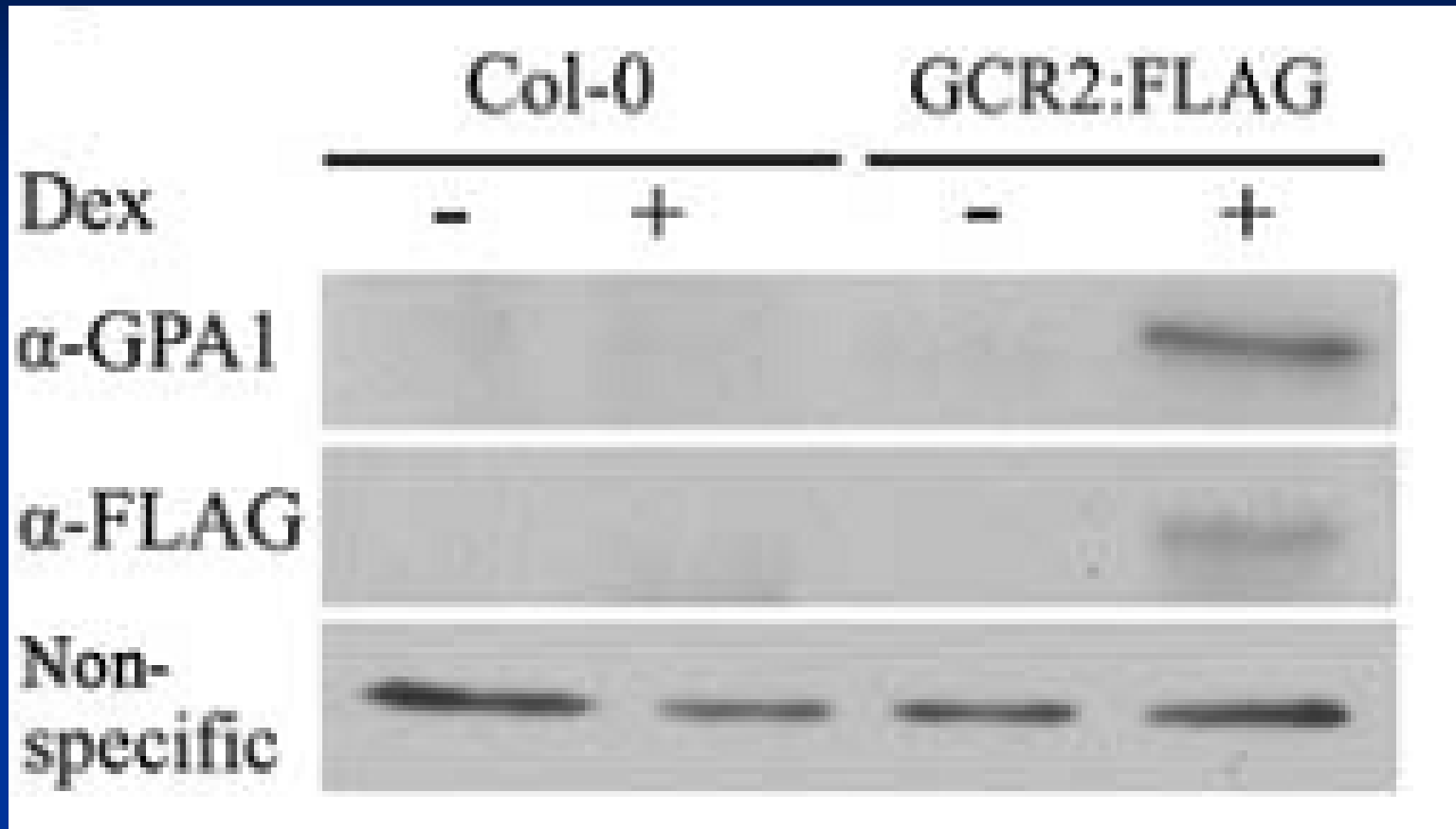
双分子荧光共振



Liu X, Yue Y, Li B, Nie Y, Li W, Wu WH, Ma L. A G protein-coupled receptor is a plasma membrane receptor for the plant hormone abscisic acid. *Science*. 2007 315(5819): 1712-6.

GCR2IL-YFPC是将GCR2的C端290-401氨基酸删除后的蛋白。可见，GCR2与GPA1存在相互作用，但删除C端后就没有了相互作用，说明GCR2的C端与GPA1发生了相互作用。

Co-IP assay



Liu X, Yue Y, Li B, Nie Y, Li W, Wu WH, Ma L.A G protein-coupled receptor is a plasma membrane receptor for the plant hormone abscisic acid. *Science*. 2007 315(5819): 1712-6.

FLAG是融合于GCR2的一个抗体标签。可见GCR2和GPA1可以共沉淀下来。

对“GCR2是G蛋白受体”的挑战和辩护

- 后来Johnston等人在《Comment on “A G Protein–Coupled Receptor Is a Plasma Membrane Receptor for the Plant Hormone Abscisic Acid”》中对其提出了质疑：

凭借硅模型（silica modeling）和生物信息学来推测GCR2既不是跨膜蛋白也不是G蛋白偶联受体（GPCR），而是细菌羊毛硫氨酸合成酶（LanC）的一个植物同源物

刘西岗等人在《Response to Comment on “A G Protein–Coupled Receptor Is a Plasma Membrane Receptor for the Plant Hormone Abscisic Acid”》对Johnston等人完全用生物信息学的方法得出的结论指出是不合理的，并进一步提供出实验的证据支持自己的观点，即GCR2可能是一个新的G蛋白偶联受体。

GCR2-YFP 融合蛋白的定位

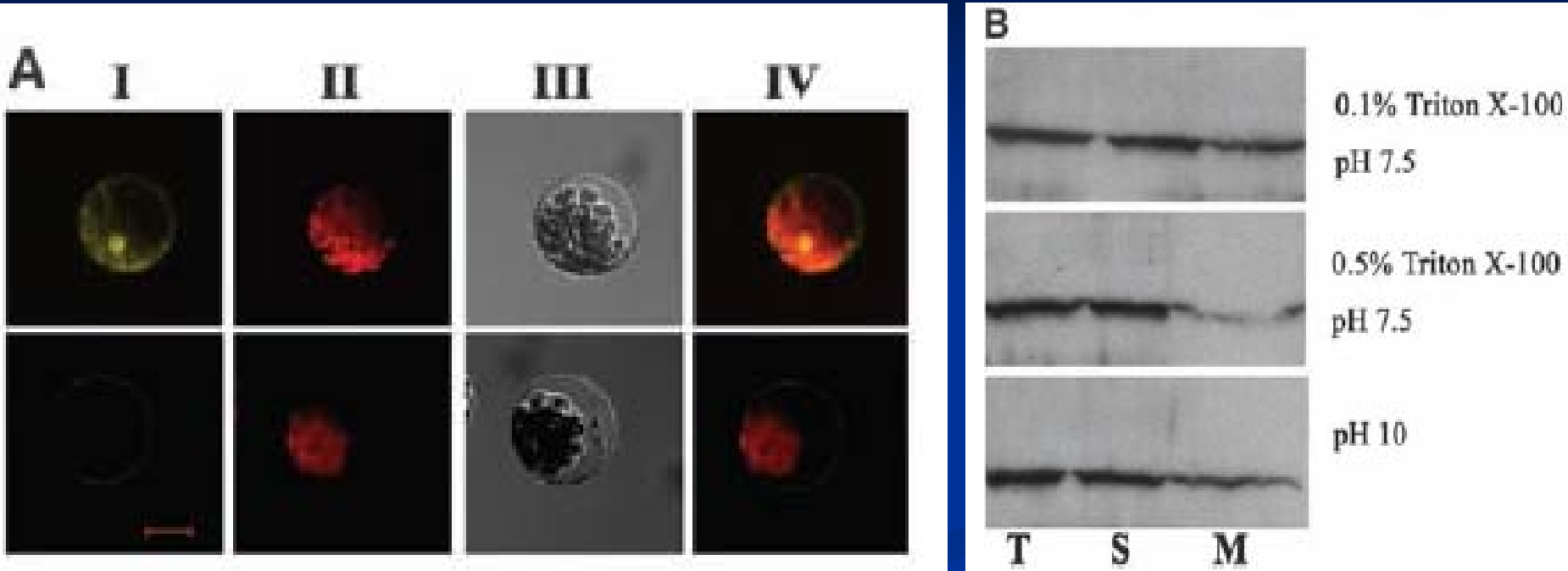


Fig.1 (A) GCR2 is localized in the plasma membrane in the protoplast expressing 35S::GFP (top) and dex-inducible promoter::GCR2-YFP (bottom). I, YFP fluorescence; II, chloroplast fluorescence; III, bright field; IV, merge of I and II. (B) GCR2 is predominantly associated with the membrane fraction. Total protein (T) was isolated from transgenic plants expressing dex-inducible promoter::GCR2-YFP. Proteins were fractionated into soluble (S) or membrane (M) fractions. Equal amounts of protein were separated on SDS-PAGE and subjected to immunoblotting using antibodies to GFP. The samples were isolated from the transgenic plant under dexinductive conditions.

- 作者集中在GCR2 的生物化学特性上进行研究。发现GCR2-YFP 融合蛋白只定位在质膜上，即使被诱导高表达也是如此(1A)，并且在细胞分选和免疫印迹分析中也和膜结构联系在一起。尽管使用清洁剂（0.1% or 0.5% Triton at Ph7.5 ）或者更高的缓冲液（pH10）清洗，GCR2 依然会留在质膜上，而人体LanC 同源物LANCL1 已被除去。

GCR2 与GPA1以及GCR2与Gb subunit 的相互作用

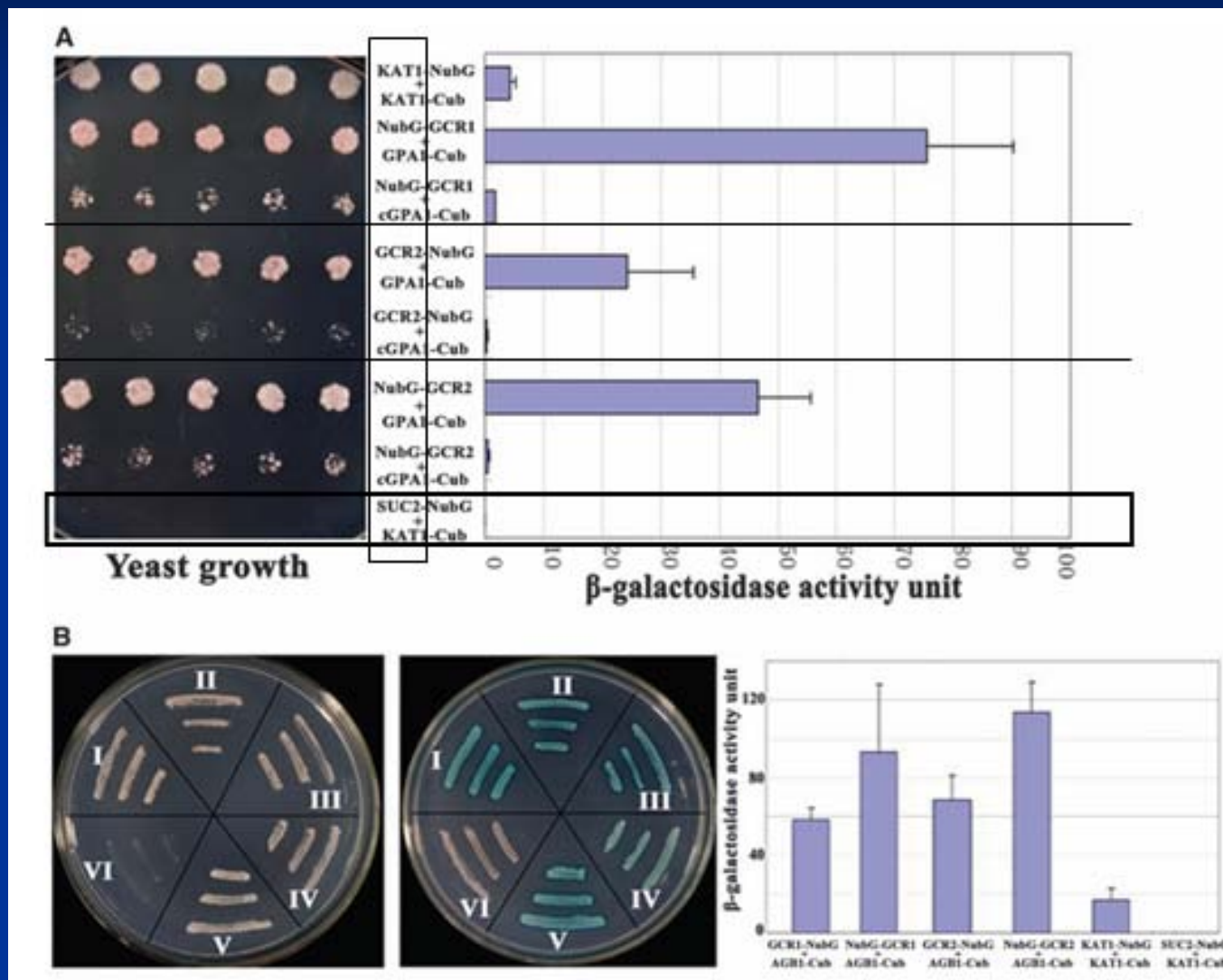


Fig.2

- (A) The dependence of intrinsic GTPase activity for physical interaction between GPA1 and GCR2 or GCR1 shown by Yeast growth assay (left) and LacZ activity assay (right). KAT1-NubG + KAT1-Cub, positive control; SUC2-NubG + KAT1-Cub, negative control; cGPA1, Q222L-mutated GPA1 defect in GTPase activity. (B) GCR2 and GCR1 interact with the Gb subunit(AGB1) by split-ubiquitin assay in yeast. (Left) Yeast growth assay. (Middle) X-gal overlay assay. (Right) Corresponding LacZ activity for each yeast strain. I, GCR1-NubG + AGB1-Cub; II, NubG-GCR1 + AGB1-Cub; III, GCR2-NubG + AGB1-Cub; IV, NubG-GCR2 + AGB1-Cub; V, NubG + KAT1-Cub, positive control; VI, SUC2-NubG + KAT1-Cub, negative control. b-galactosidase activity unit: Miller unit.

- 已经被证实G蛋白的 α 亚基的内在GTP酶活性会调节它与G蛋白偶联受体的结合。确实，拟南芥G蛋白 α 亚基GPA1的GTP酶活性就会调节其与GCR2和GCR1的相互作用。另外，与一些G蛋白偶联受体相似的是，GCR2能分别与Ga (Fig.2A)和G β (Fig.2B)相互作用。(负对照仍有不足！)

- 在知道GCR2 与细菌的LanC 蛋白只有较低的序列相似性的基础上，Johnston等人认为GCR2 是细菌的羊毛硫氨酸合成酶的植物同源物而不是G 蛋白偶联受体。然而作者有7 种证据支持GCR2 在生物化学上显著不同于LANCL 蛋白，包括GCR2 和LANCL1 之间膜结合特性上和亚细胞定位上的不同。LANCL1-GFP 只能定位在细胞溶质和核仁内而不能在质膜上，然而LANCL1-YFP 只能定位在质膜上。最后，在与Ga和G β 亚基的相互关系上也有不同。还没有证据证明LANCL1 和G 蛋白之间的相互作用，而GCR2 可以与Ga和G β 亚基有相互作用。

- Johnston等人认为GCR2 与乳酸乳球菌LanC 蛋白即乳酸链球菌素环化酶(NisC)结构的相似性可以为GCR2 是LanC 蛋白超家族的一员提供证据。GCR2 可能是一种LANC 的同源物，但也有好几种有类似相似性的同源性蛋白显示出植物和细菌功能和生理生化特性上的不同。另外，ERF/AP2 转录因子基因家族从细菌的HNH 核酸内切酶进化而来，而B3 结构域转录因子基因家族（拟南芥的37 个成员）显示出与大肠杆菌ECORII 的序列相似性。因此，很难只依据计算机模型去估计生理生化功能。而作者提供了好几个生理生化方面的证据证明GCR2 是ABA 的受体和与拟南芥中Ga和G β 亚基的相互关系。GCR2 其它的特性可能与经典的G 蛋白偶联受体不相一致，因此认为GCR2 这种蛋白可能是一种新的G 蛋白偶联受体。

我们的结论

- 综上，刘西岗等人的文章中提出了很多支持自己观点的实验证据（2，4），但也有人有反对的实验数据（5，6）。而生物信息预测的结果不能确证GCR2是七次跨膜的蛋白，且生物信息学方法自身又有很多的局限，所以GCR2是否为跨膜蛋白或GPCR关键在于进一步的实验数据。

我们的体会

- 正确并适当的使用生物信息学工具有助于科研工作；
- 盲目使用生物信息学工具而忽视前提假定导致错误；
- 生物信息学工具预测的结果需要“湿实验”加以验证；
- 科研工作一下结论需慎重；
- 学术之争有助于加深认识

分工表

姓名	座位号	工作
张翼	C1	<ol style="list-style-type: none">1.第二部分， GCR2蛋白的blast分析和motif预测。2.第三部分（参与）。
余继（组长）	C2	<ol style="list-style-type: none">1.背景讨论（全部）。2.第三部分，多序列比对。3.第五部分，跨膜软件讨论。
操昭煦	D1	<ol style="list-style-type: none">1.第一部分，跨膜结构软件预测中的TMAP， TMpred ， TopPred ， DAS ， HMMTOP,总结各软件的分析结果。2.第五部分，结合实验数据讨论。
刘博超	D2	<ol style="list-style-type: none">1.第一部分，跨膜结构软件预测中的TMHMM (v2.0) 和 SOSUI和SOSUIsignal;2.第四部分 搜索GCR2三维结构。3.第五部分，结合实验数据讨论。

参考文献

- 1. Erwin Grill and Alexander Christmann. BOTANY: A Plant Receptor with a Big Family. *Science* . 2007: 1676-1677.
- 2. Liu X, Yue Y, Li B, Nie Y, Li W, Wu WH, Ma L. A G protein-coupled receptor is a plasma membrane receptor for the plant hormone abscisic acid. *Science*. 2007 315(5819): 1712-6.
- 3. Christopher A. Johnston, Brenda R. Temple, Jin-Gui Chen, Yajun Gao, Etsuko N. Moriyama, Alan M. Jones, David P. Siderovski, and Francis S. Willard. Comment on "A G Protein-Coupled Receptor Is a Plasma Membrane Receptor for the Plant Hormone Abscisic Acid". *Science* 2007 318 (5852): 914c.
- 4. Xigang Liu, Yanling Yue, Wei Li, and Ligeng Ma. Response to Comment on "A G Protein-Coupled Receptor Is a Plasma Membrane Receptor for the Plant Hormone Abscisic Acid" . *Science*. 2007 318 (5852): 914d.
- 5. Gao Y, Zeng Q, Guo J, Cheng J, Ellis BE, Chen JG. Genetic characterization reveals no role for the reported ABA receptor, GCR2, in ABA control of seed germination and early seedling development in *Arabidopsis*. *Plant J*. 2007 52(6):1001-13.
- 6. Guo J, Zeng Q, Emami M, Ellis BE, Chen JG. The GCR2 gene family is not required for ABA control of seed germination and early seedling development in *Arabidopsis*. *PLoS ONE*. 2008 3(8): e2982
- 7. Möller S, Croning MD, Apweiler R. Evaluation of methods for the prediction of membrane spanning regions. *Bioinformatics*. 2001 17(7): 646-53.
- 8. Pandey S, Nelson DC, Assmann SM. Two novel GPCR-type G proteins are abscisic acid receptors in *Arabidopsis*. *Cell*. 2009 136(1):136-48.
- 9. Li B, Yu JP, Brunzelle JS, Moll GN, van der Donk WA, Nair SK. Structure and mechanism of the lantibiotic cyclase involved in nisin biosynthesis. *Science*. 2006 311(5766):1464-7.

Thank You!