

东方山羊豆赤霉素受体生物信息学分析

李俊 董坤哲

阿地力江 陆涛峰 晁跃辉

2011.12.24



植物利用激素对其生长发育和环境适应进行调节。在这些激素中，赤霉素（GAs）是植物激素四环、二萜类一大家族^[1]。

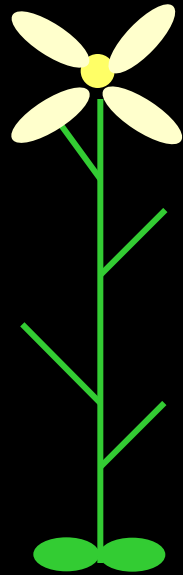
赤霉素途径的研究极大促进了农业发展，例如：上世纪，在水稻、小麦“绿色革命”中分别发挥作用的sd1、Rht1基因，它们分别与编码赤霉素生物合成途径、信号转导途径有关^[2, 3]。

目前在生物体内已发现的赤霉素共有127种，但只有部分具有生物学活性^[4]。

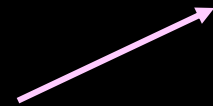
赤霉素(Gibberellins)

是植物生长调节物质, 在整个生命周期中扮演极其重要的作用

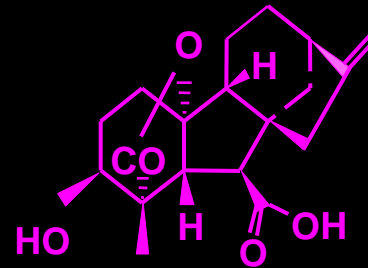
开花时间



花发育



果实成熟



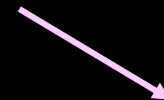
下胚轴伸长

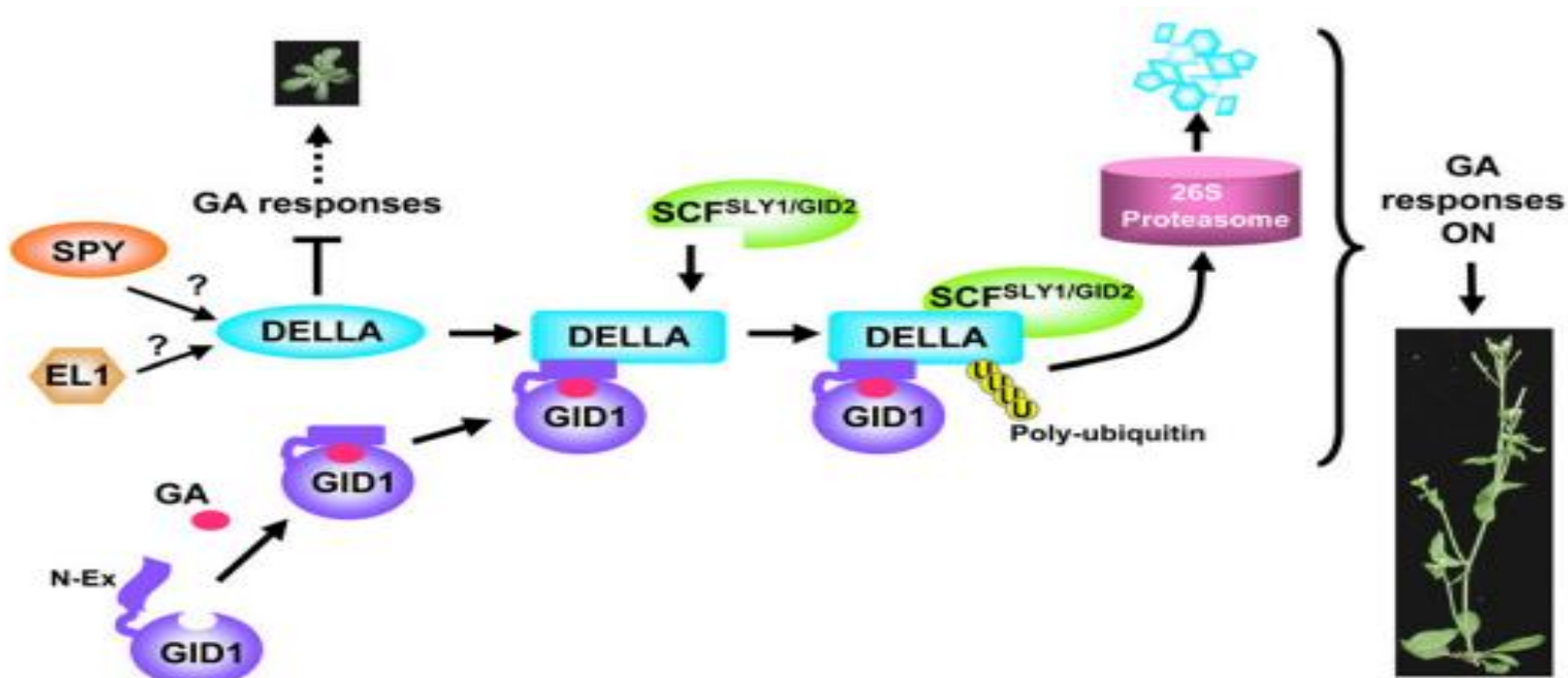


种子萌发



叶片发育





植物中GA信号传导模型 (Tai-ping Sun, 2010.)

东方山羊豆 (*Galega. Orientalis L.*) 为多年生豆科草本植物，茎中空，株高80~110cm；为根蘖型牧草，主根深达1 m，侧根发达。



1. 叶量丰富。花期（干重）茎叶比为1： 1.9，而同期新疆大叶苜蓿为1：0.71^[1]。
2. 粗蛋白含量高。干物质中粗蛋白质现蕾期为27.8% ， 始花期16.3 % ^[2]；与初花期的新疆大叶苜蓿相比，粗蛋白质含量高12.9% ， 粗脂肪高117% ， Ca和P分别高139% 和4.7% ^[1] 。
- 3.抗寒和抗旱性比苜蓿和三叶草强。在土壤水分不足的情况下仍然获得高于苜蓿和三叶草的产量^[2] 。

[1]张清斌、刘万胜, 东方山羊豆引种研究初报. 中国草地, 2001. **23**(4): p. 17-20.

[2]张自和, 东方山羊豆的生物学特性与栽培技术. 草原与草坪, 2002. **1**: p. 19-20.

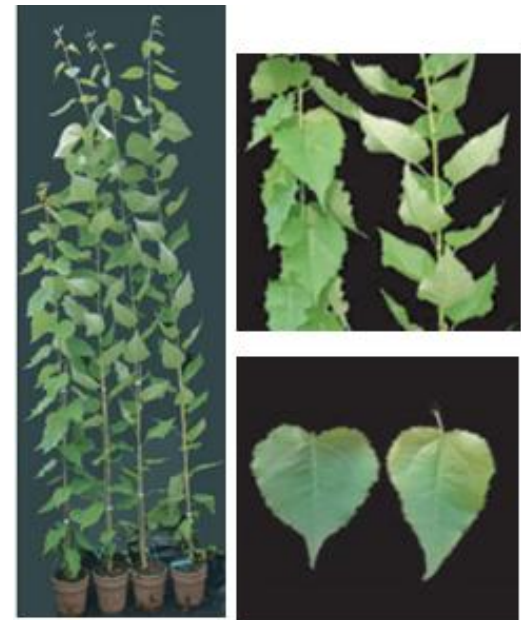
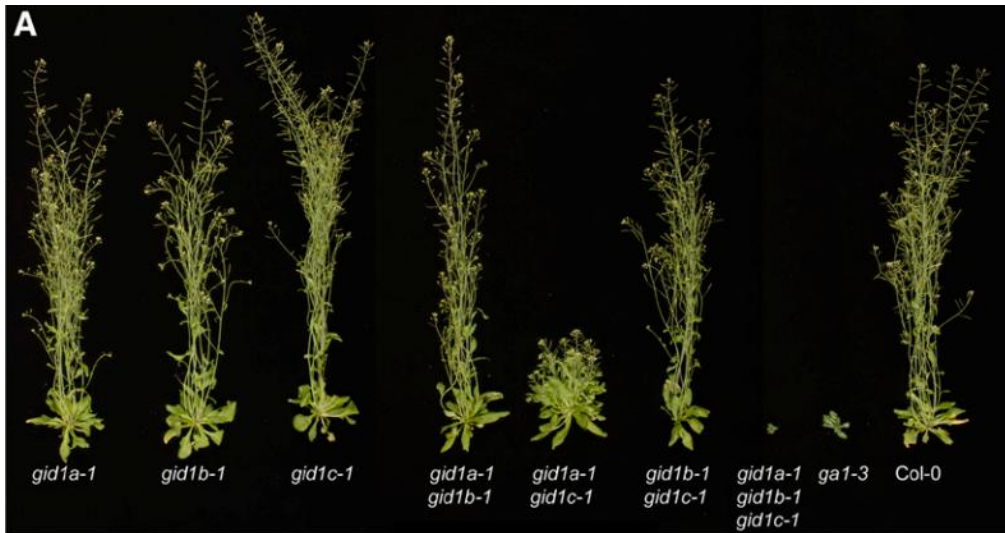
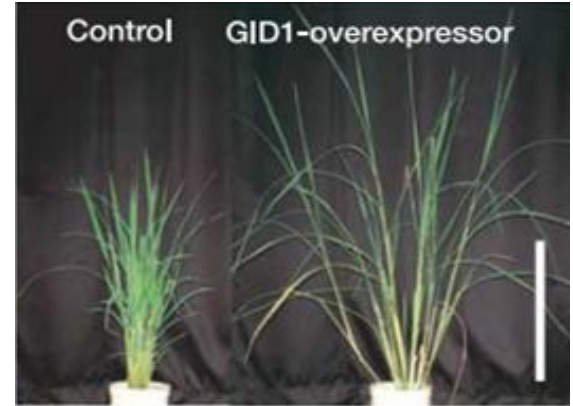
4. 易于调制干草。刈割的鲜草在阳光下晒15d以上叶片也不脱落^[1]。
5. 饲喂利用期比苜蓿早。东方山羊豆的第一次刈割时间要早于新疆大叶苜蓿20-25d左右^[1]。
6. 青饲反刍家畜不得臃胀病，并具有促进奶牛泌乳等功能^[1]。
7. 长寿多年生牧草。在同一地块上可生长10~15年或更长时间^[1]。

[1]张清斌、刘万胜, 东方山羊豆引种研究初报. 中国草地, 2001. **23**(4): p. 17-20.

赤霉素受体，首次于2005年，从水稻一个赤霉素不敏感矮化(gibberellin insensitive dwarf, GID1)突变体中克隆出来。

随后，在拟南芥、棉花中分别克隆到三个 (AtGID1a, AtGID1b, AtGID1c)、二个 (GhGID1a、GhGID1b) 赤霉素受体；在大麦、杨树等物种中也发现赤霉素受体。

功能:



What does gene encode?

```
      10      20      30      40      50      60
MTGSNEVNLS ESKSVVPLNT WVLISNFKLA YNLLRRADGT FNRELAEFLD RKVPANTIPV

      70      80      90     100     110     120
DGVFSEFDHVD RNTGLFSRVY QPAENVITW GIIELEKPLS TTEIVFVIIF FHGGSFSSHSS

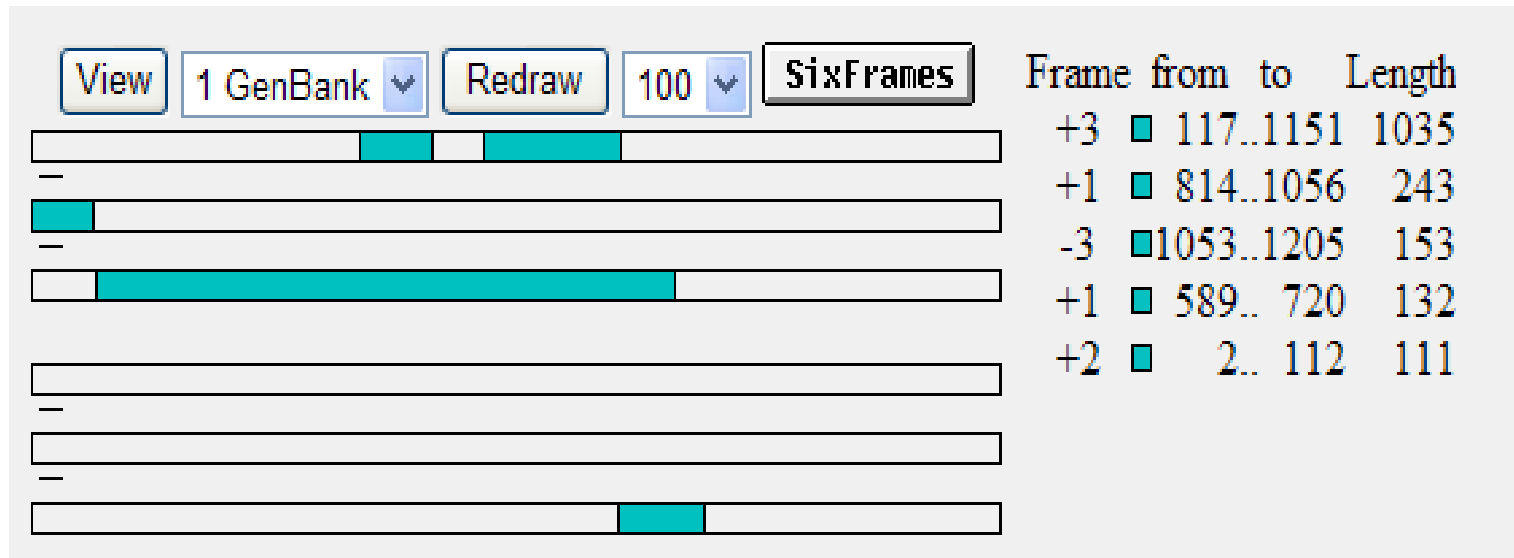
     130     140     150     160     170     180
ANSAIYDTPC RRLVSMCKAV VVSVNYRASP EHRYPAYED GWNALQWVKS RTWLQSGKDS

     190     200     210     220     230     240
KVYVYMAGDS SGGNIAHHVA VRAAEEDVEV LGNILLHPLF GGERRTESER KLDGKYFVRL

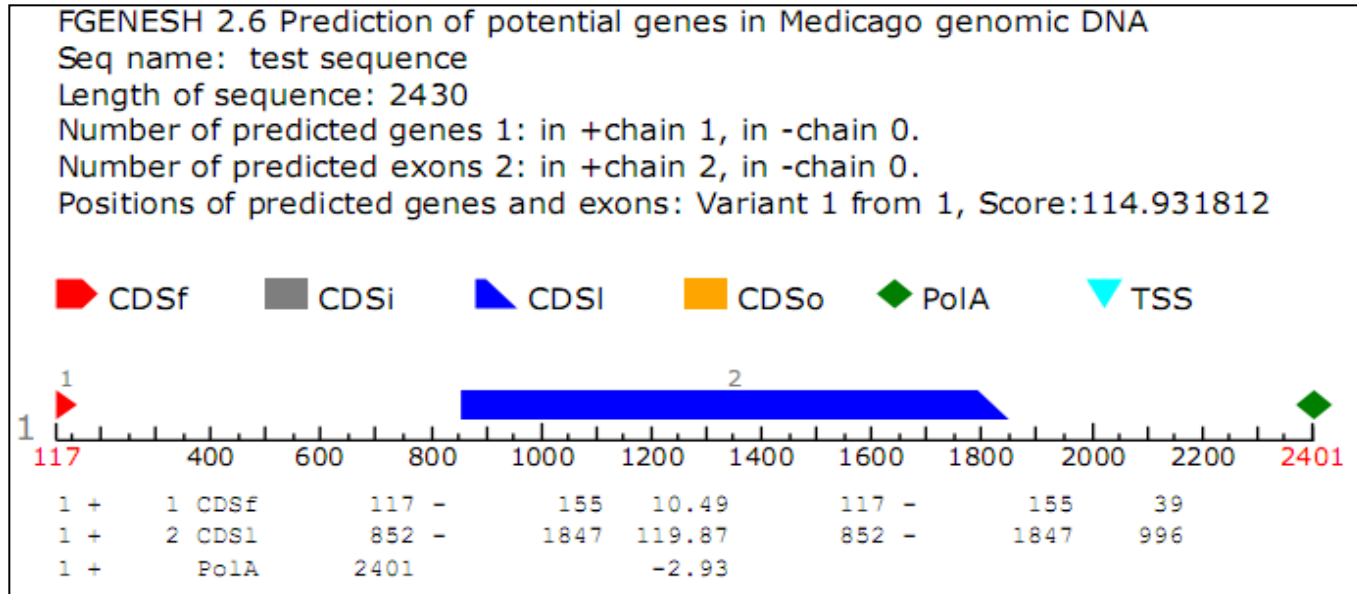
     250     260     270     280     290     300
QDRDWYWRAP LPEGEDRDHP ACNPFPGPKGK SLAGLKFAKS LVCVAGLDLL QDWQLEYVEG

     310     320     330     340
LKSFDQDVRL LYLKEATIGF YFLFNNDHFY CLFNEINTEV HPNC
```

东方山羊豆赤霉素受体mRNA（HM989010）和DNA全长序列（未登录）。



ORF查找，该mRNA的开放阅读框为：117-1151bp，长度为1035bp，编码344个氨基酸。



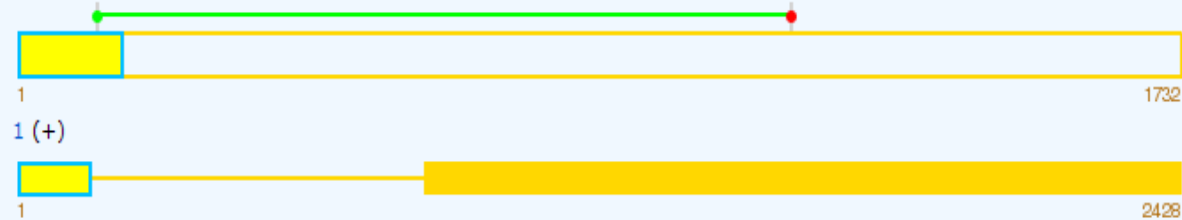
FGENESH软件，基因结构预测，可知：有一个696bp内含子，两个外显子。

#	Query	Subject	Span(bp)	Coverage(%)	Overall(%)	Exon(%)	CDS(%)	In-frame(%)
1	HM989010(+)	1(+)	1-2428	100.00	100.00	100.00	100.00	100.00

[Graphics](#) | [Text](#)

Model 1	Coverage	100.00%	CDS	100.00%	Mismatches and indels	0
	Overall	100.00%	In-frame	100.00%	Exons (min/max/ave), bp	155 / 1577 / 866
	Exon	100.00%	Primary transcript	1732 bp	Introns (min/max/ave), bp	697 / 697 / 697

HM989010 (+) *Galega orientalis* gibberellin receptor (GID) gene, complete cds



splign工具对cDNA、DNA进行比对，从下图可知：有一个内含子，位于155-852bp之间。

Search in

Protein Knowledgebase (UniProtKB)



Query

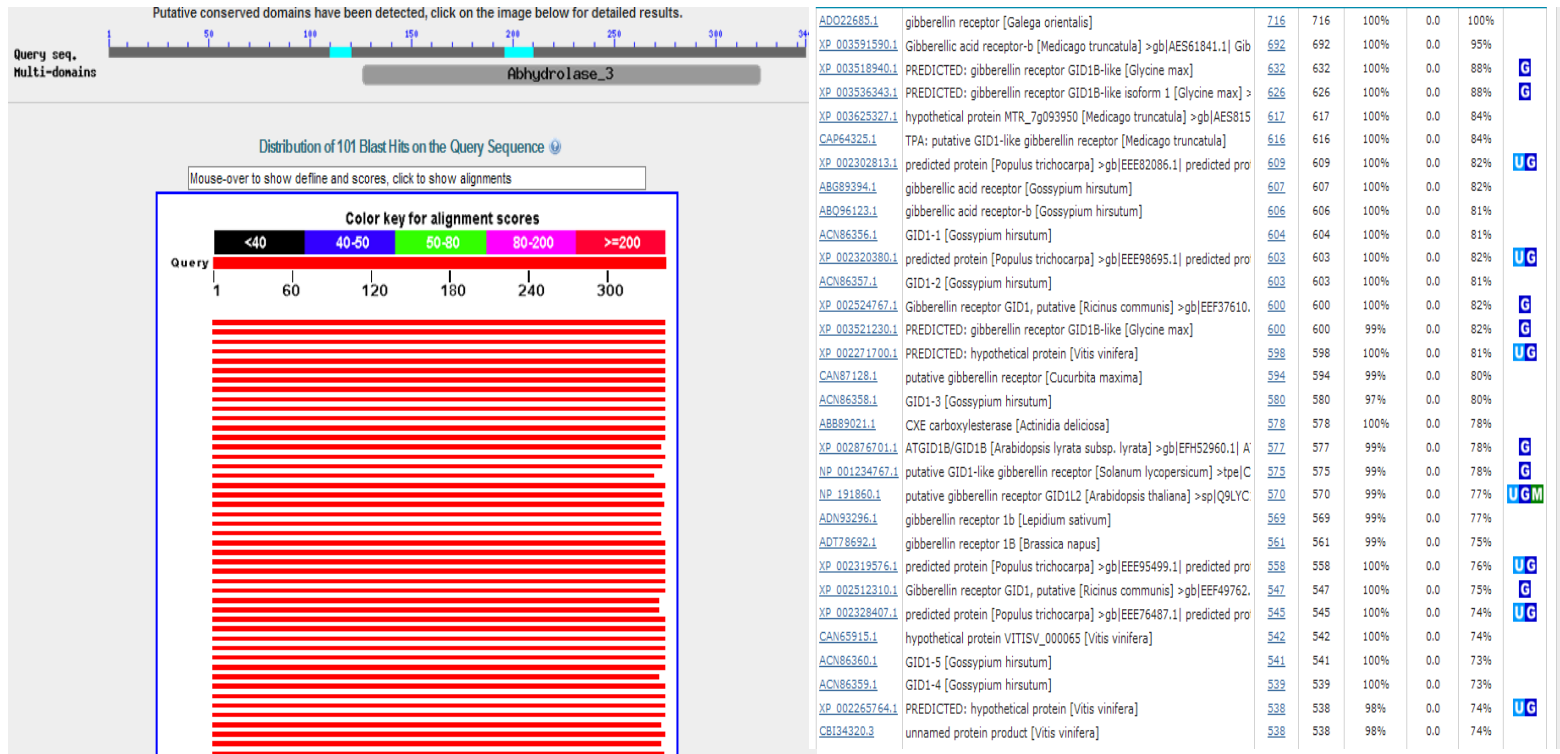
name:"gibberellin receptor"

66 results

- Physcomitrella patens (2)
- Tracheophyta (64)
 - Selaginella (3)
 - Spermatophyta (61)
 - Magnoliophyta (59)
 - Eudicotyledons (29)
 - Aquilegia formosa x Aquilegia pubescens (1)
 - Core eudicotyledons (28)
 - Rosids (27)
 - Brassicaceae (7)
 - Arabidopsis thaliana (Mouse-ear cress) (3)
 - Brassica napus (Rape) (1)
 - Lepidium sativum (Garden cress) (3)
 - Fabids (20)
 - Castanea mollissima (Chinese chestnut) (1)
 - Cucurbita maxima (Pumpkin) (Winter squash) (2)
 - Papilionoideae (3)
 - Galega orientalis (1)
 - Medicago truncatula (Barrel medic) (Medicago tribuloides) (2)
 - Ricinus communis (Castor bean) (14)
 - Solanum lycopersicum (Tomato) (Lycopersicon esculentum) (1)

- Liliopsida (30)
 - Allium cepa (Onion) (2)
 - Poaceae (28)
 - Andropogoneae (25)
 - Saccharum officinarum (Sugarcane) (1)
 - Sorghum bicolor (Sorghum) (Sorghum vulgare) (1)
 - Zea mays (Maize) (23)
 - BEP clade (3)
 - Oryza sativa subsp. japonica (Rice) (1)
 - Triticeae (2)
 - Hordeum vulgare (Barley) (1)
 - Triticum aestivum (Wheat) (1)
- Pinaceae (2)

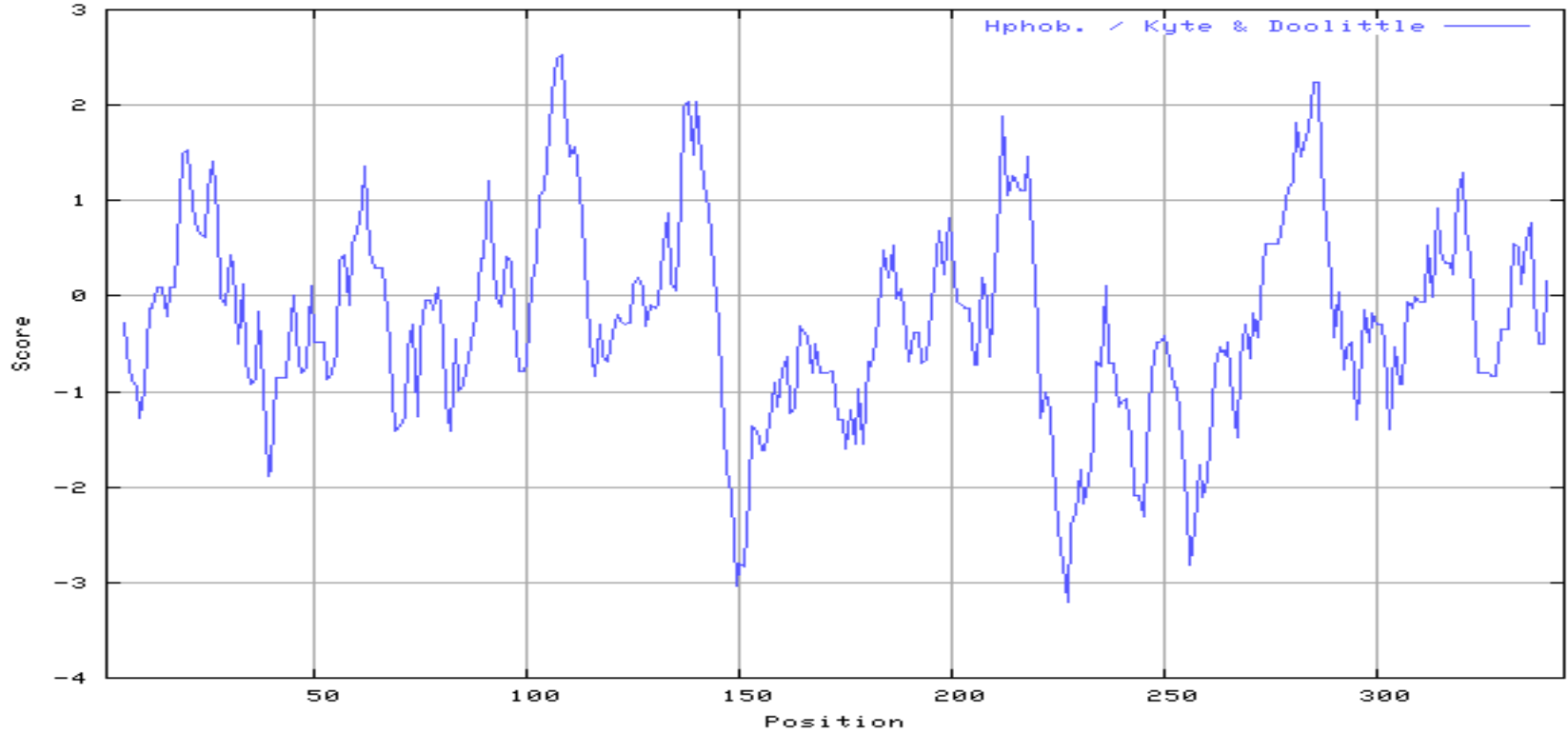
BLAST



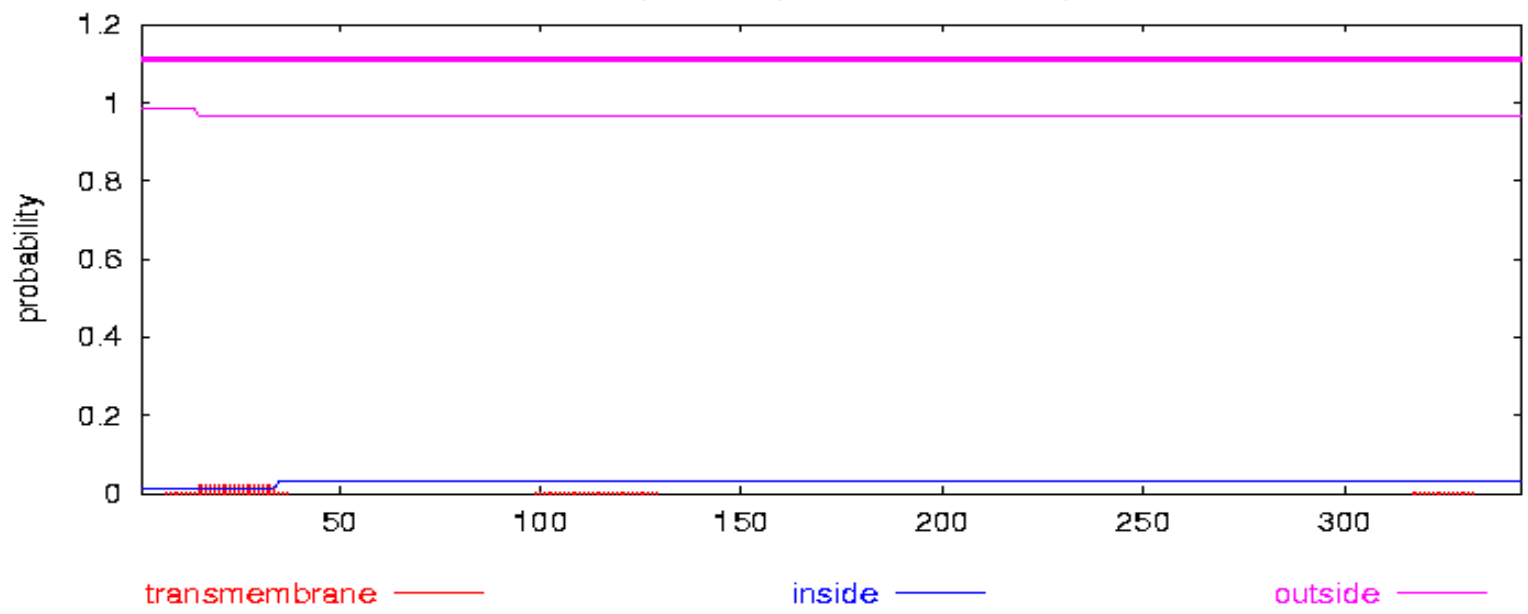
Putative GA receptor

蛋白的等电点为6.34，pH 7.0时的电荷为-3.603，分子量为39.085kDa，碱性氨基酸（K、R）37个，酸性氨基酸（D、E）42个，疏水氨基酸（A、I、F、W、V）126个，极性氨基酸（N、C、Q、S、T、Y）88个。

ProtScale output for user sequence



TMHMM posterior probabilities for Sequence



hits by patterns: [2 hits (by 2 distinct patterns) on 1 sequence]

E3VJP7
(E3VJP7_9FABA)  (344 aa)

SubName: Full=Gibberellin receptor;. *Galega orientalis*

PS01173 LIPASE_GDXG_HIS *Lipolytic enzymes :*

108 - 124: IIfFHGGSFshsSanSA

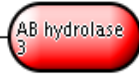
PS01174 LIPASE_GDXG_SER *Lipolytic enzymes :*

184 - 196: VyMAGDSSGGnIA

Predicted feature:

ACT_SITE 190 By similarity [condition: none]

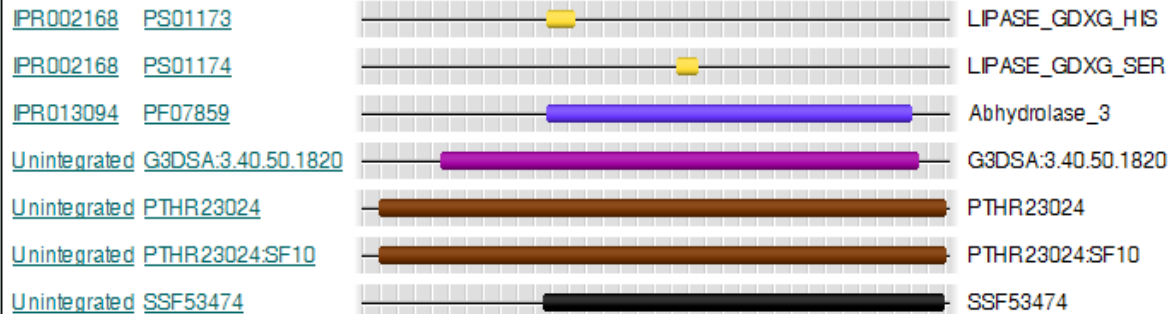
Gibberellin receptor



IPR002168	IPR002168	IPR013094	Unintegrated	Unintegrated	Unintegrated	Unintegrated
PS01173	PS01174	PF07859	G3DSA:3.40.50.1820	PTHR23024	PTHR23024:SF10	SSF53474
108 - 124	184 - 196	108 - 321	46 - 325	10 - 341	10 - 341	106 - 340

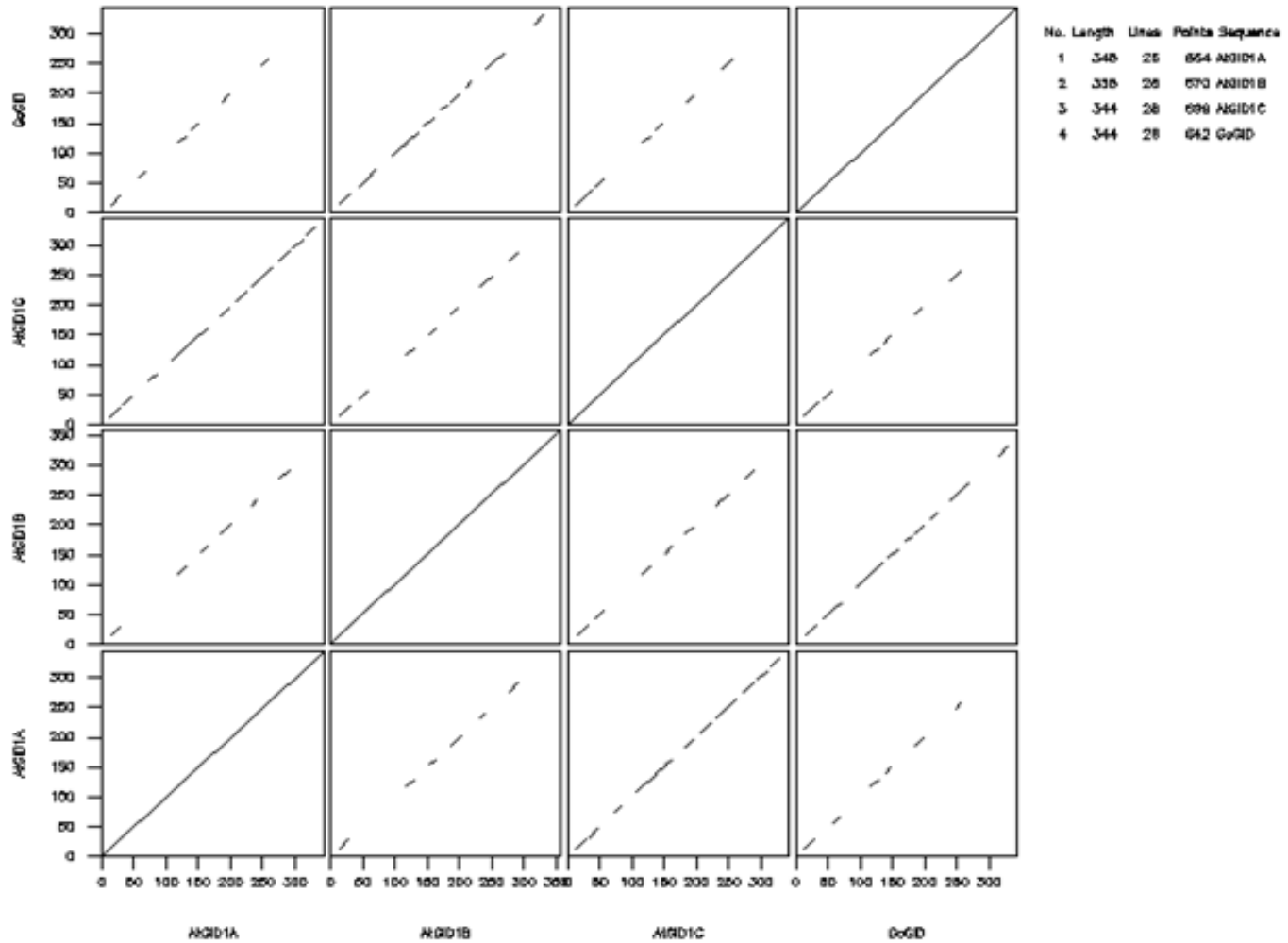
[UniProtKB/TrEMBL : E3VJP7](#)
 Scale:10aa
E3VJP7_9FABA
[Signature Matches in BioMart](#)
[Supermatches in BioMart](#)
[Dasty](#)
[GO!](#)
[Galega orientalis](#)

InterPro Signatures



Poly dotplot of 493603

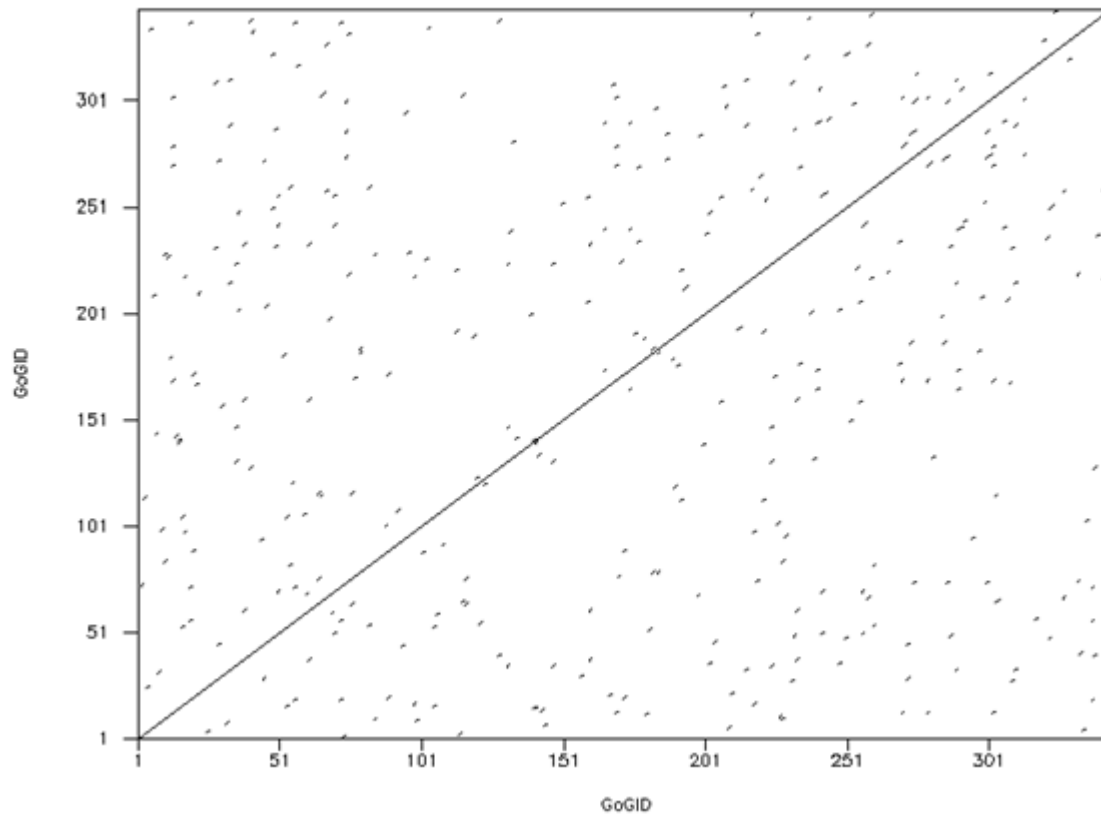
Sun 25 Dec 2011 03:10:02



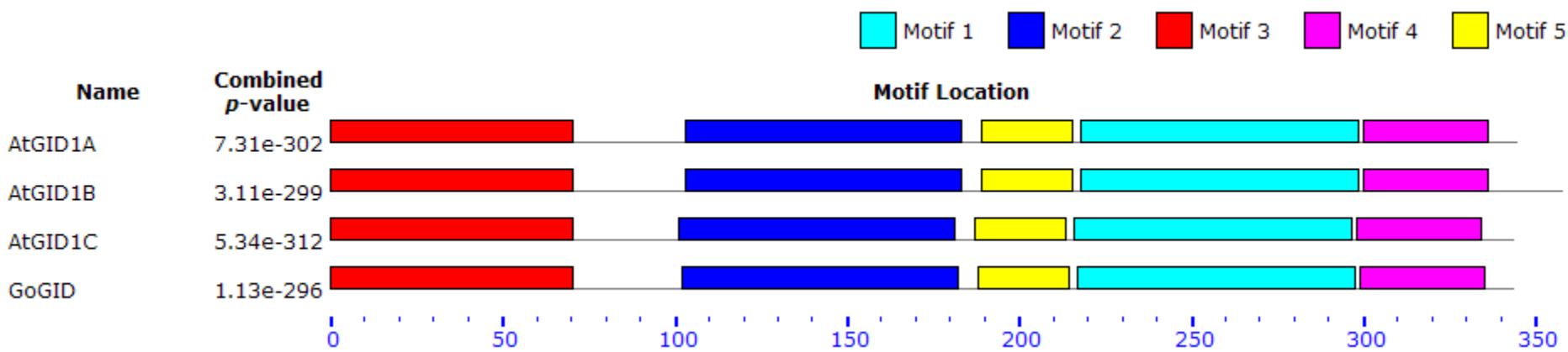
Word size 8

Dottup: fasta::491441:GoGID vs fasta::491441:GoGID

Sun 25 Dec 2011 03:05:07



Word size 1



Motif 1

- 4.8e-129
- 4 sites



Motif 2

- 5.8e-127
- 4 sites



Motif 3

- 3.3e-091
- 4 sites



Motif 4

- 6.8e-045
- 4 sites



Motif 5

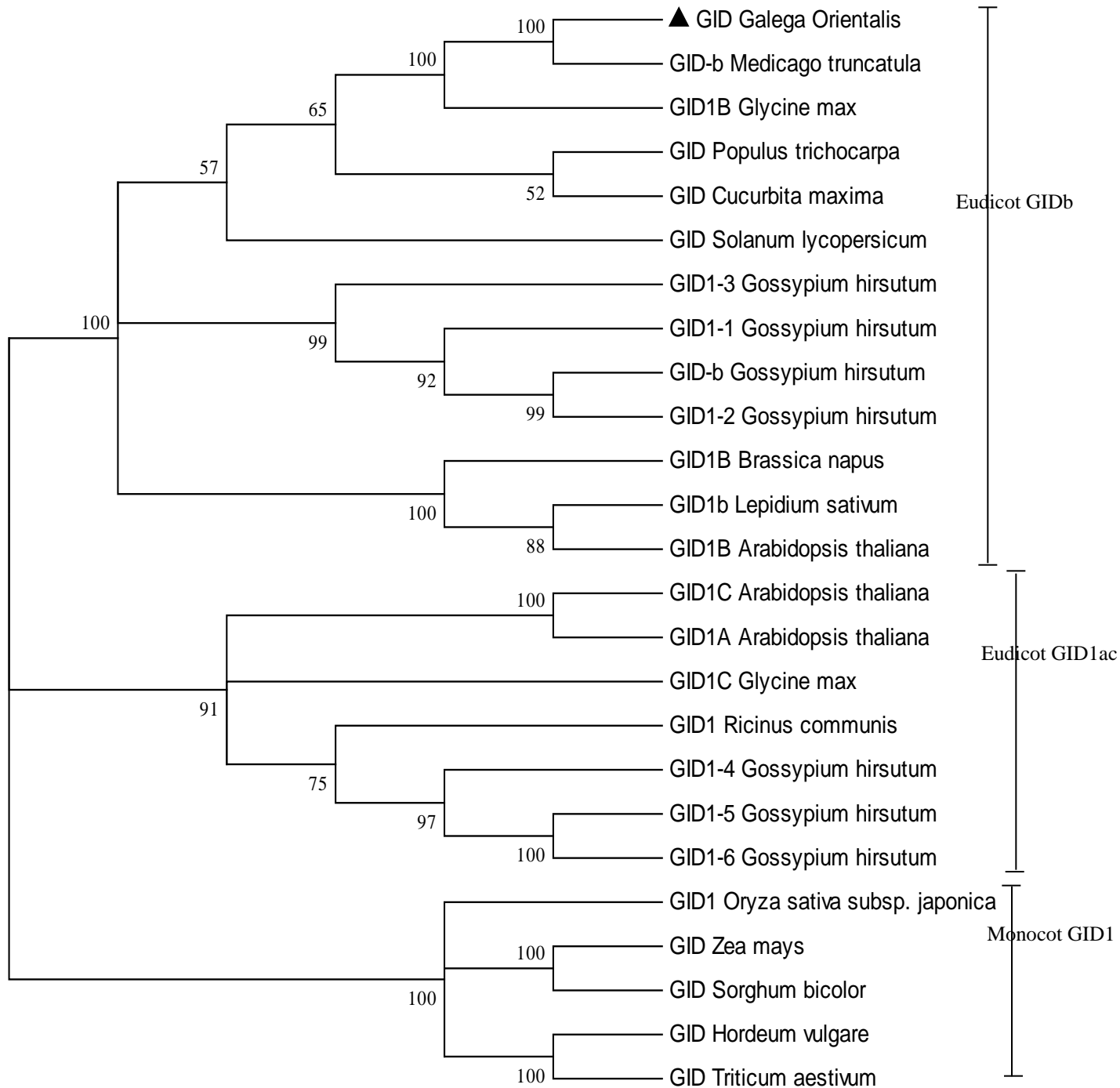
- 3.9e-025
- 4 sites



Minimum 10, maximum 80, motif 5

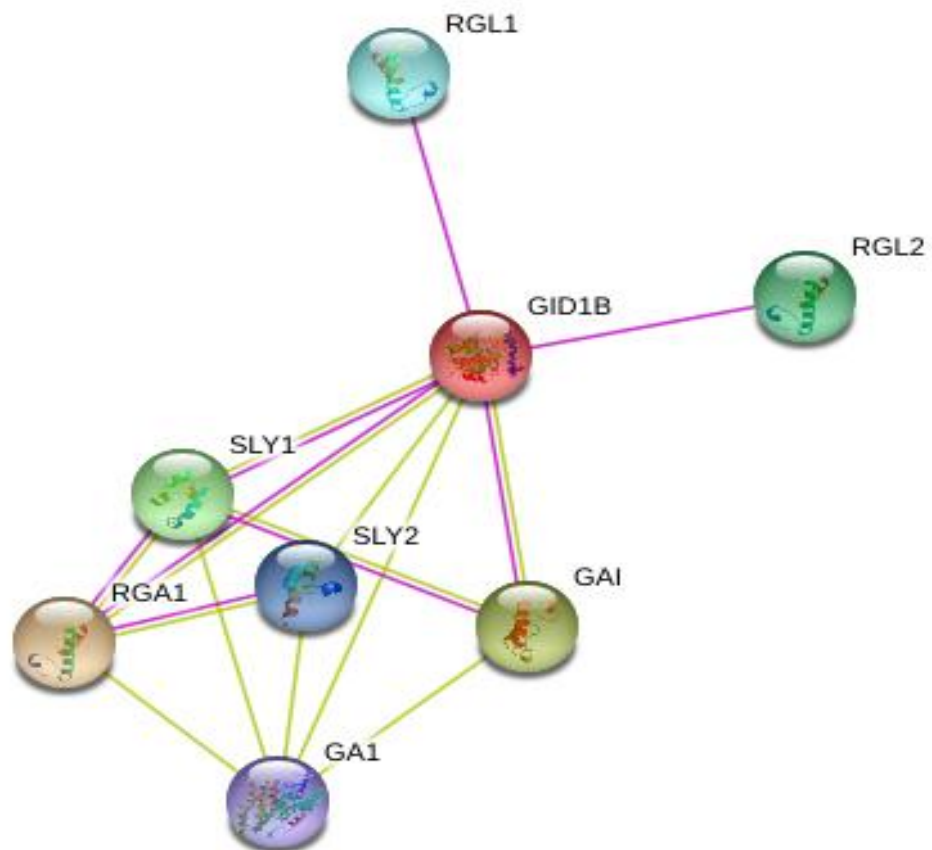
GoGID	MTGSNEVNLSESKSVVPLNTWVLI	SNFKLAYNLLRRADGTFNRHLAEFLDRKWPANTI	PVDGVVFSFD	HVDRNTGLFSRVYQPASE	NV	87																																																																																							
AtGID1a	MAASDEVNLI	ESRTVVPLNTWVLI	SNFKVAVNI	LRRPDGTFNRHLAEFLDRKWTANANP	VDGVVFSFDVLI	DRRI	NLLSRVYRPAYAD	QE	89																																																																																				
AtGID1b	MAGNEVNLNECKRI	VPLNTWVLI	SNFKLAYKVLRRPDGTFNRDLAEFLDRKWPANS	FPLDGVVFSFD	HVDS	TTNLLTRI	YQPASLL	HQ	88																																																																																				
AtGID1c	MAGSEVNLI	ESKTVVPLNTWVLI	SNFKLAYNLLRRPDGTFNRHLAEFLDRKWPANANP	VNGVFSFDVI	I	DRQTNLLS	RVYRPADAG		87																																																																																				
OsGID1	MAGSDEVNRNECKTVVPLHTWVLI	SNFKLSYNI	LRRADGTFERDLGEFLDRRWPANARPLE	GVS	SFDHI	I	DQSVGLEVRI	YRAAAEGDAEEGAAA		95																																																																																				
GhGID1a	MAGSNEVNLNESKRVVPLNTWVLI	SNFKLAYNLQRRPDGTFNRDLSEFLDRRWPANI	NPVDGVVFSFD	HVDGATGL	LNRVYQPS	SL		NE	87																																																																																				
GhGID1b	MAGSNEVNLNESKRVVPLNTWVLI	SNFKLSYNLQRRPDGTFNRDLSEFLDRRWPANI	NPVDGVVFSFD	HVDGATGL	LNRVYQPS	PK		NE	87																																																																																				
HGG																																																																																														
GoGID	TTWGI	I	ELEKPLS	TTEI	VPVI	I	FFHGGS	FS	HS	SANS	AI	YDTFCRRLWS	MCKAWVVS	VNYRRS	PEHRYP	CAYEDGWNAL	QWVKS	RTWLOS	GKDS	KV	182																																																																									
AtGID1a	QPPSI	LDLEK	VDG	DI	VPVI	I	FFHGGS	FAHS	SANS	AI	YDTLCRRLWGL	CKCQVVS	VNYRRAPENP	YPCAY	DDGWI	AL	NWVNS	RS	WLK	KKDS	KV	183																																																																								
AtGID1b	TRHGT	LELT	KPLS	TTEI	VPVLI	I	FFHGGS	FTHS	SANS	AI	YDTFCRRLVTI	CGVWVVS	VDYRRS	PEHRYP	CAYDDGWNAL	NWVKS	RVWLOS	GKDS	NV		183																																																																									
AtGID1c	TSPSI	TDLQN	VDG	EI	VPVI	I	VFFHGGS	FAHS	SANS	AI	YDTLCRRLWGL	CGAVVVS	VNYRRAPENRYPCAY	DDGWAVL	KWVNS	S	WLRS	KKDS	KV		181																																																																									
OsGID1	VTRPI	LEFL	TDAP	AAEP	FPVI	I	FFHGGS	FVHS	SANS	TI	YDSL	CRRFV	KL	SKGVVVS	VNYRRAPENRYPCAY	DDGWT	AL	KWVNS	QPF	MRS	GG	LAQA	190																																																																							
GhGID1a	AQWGM	VDLEK	PLS	TTEI	VPVI	I	VFFHGGS	FTHS	SANS	AI	YDTFCRRLWSL	CKAVVVS	VNYRRS	PEHRYP	CAYDDGWAAL	KWVKS	RTWLOS	GKDS	NV		182																																																																									
GhGID1b	AQWGI	VDLEK	PLS	TTEI	VPVI	I	VFFHGGS	FTHS	SANS	AI	YDTFCRRLWNI	CKAVVVS	VNYRRS	PEHRYP	CAYDDGWAAL	KWVKS	RTWLOS	GKDS	KV		182																																																																									
GXSXG																																																																																														
GoGID	YVYNA	GDSS	GGNI	AHHVAVRAAEE	DVEVL	GNL	LLHP	LF	GG	RRTE	SE	KKLDGKYF	VRLQDRD	DWY	WRAFL	PEGEDRDHP	ACNPF	FG	PKGKS	LA	GLKF	277																																																																								
AtGID1a	HI	F	L	A	G	D	S	S	G	G	N	I	A	H	N	V	A	R	A	E	E	D	V	E	V	L	G	N	L	L	H	P	M	F	G	G	N	E	R	T	E	S	E	K	S	L	D	G	K	Y	F	V	T	V	R	D	R	D	W	Y	W	K	A	F	L	P	E	G	E	D	R	H	P	A	C	N	P	F	S	P	R	G	K	S	L	E	G	W	S	F	278			
AtGID1b	YVYLA	GDSS	GGNI	AHNVA	VRATNE	GVKVL	GNL	LLHP	M	F	GG	Q	R	T	E	S	E	K	L	D	G	K	Y	F	V	T	I	Q	D	R	D	W	Y	W	R	A	I	L	P	E	G	E	D	R	H	P	A	C	N	P	F	G	P	R	G	Q	S	L	K	G	W	N	F	278																														
AtGID1c	RI	F	L	A	G	D	S	S	G	G	N	I	V	H	N	V	A	R	A	V	E	S	R	I	D	V	L	G	N	L	L	H	P	M	F	G	G	T	E	R	T	E	S	E	K	R	L	D	G	K	Y	F	V	T	V	R	D	R	D	W	Y	W	R	A	F	L	P	E	G	E	D	R	H	P	A	C	S	P	F	G	P	R	S	K	S	L	E	G	L	S	F	276		
OsGID1	R	V	F	L	S	G	D	S	S	G	G	N	I	A	H	H	V	A	V	R	A	A	E	D	G	V	K	V	C	G	N	L	L	N	A	M	F	G	G	T	E	R	T	E	S	E	R	R	L	D	G	K	Y	F	V	T	L	Q	D	R	D	W	Y	W	K	A	Y	L	P	E	D	A	D	R	H	P	A	C	N	P	F	G	P	N	G	R	R	L	G	L	P	F	285	
GhGID1a	H	V	Y	L	A	G	D	S	S	G	G	N	I	A	H	H	V	A	V	R	A	A	E	A	D	V	E	V	L	G	N	L	L	H	P	M	F	G	G	Q	R	T	E	S	E	K	R	L	D	G	K	Y	F	V	T	L	H	D	R	D	W	Y	W	R	A	Y	L	P	E	G	E	D	R	H	P	A	C	N	P	F	G	P	R	G	R	S	L	E	G	L	K	F	277	
GhGID1b	H	V	Y	L	A	G	D	S	S	G	G	N	I	A	H	H	V	A	V	R	A	A	E	A	D	V	E	V	L	G	N	L	L	H	P	M	F	G	G	Q	M	R	T	E	S	E	K	R	L	D	G	K	Y	F	V	T	L	H	D	R	D	W	Y	W	R	A	Y	L	P	E	G	E	D	R	H	P	A	C	N	P	F	G	P	R	G	R	T	L	E	G	L	K	F	277
GoGID	AKS	L	V	W	A	G	L	D	L	I	Q	D	W	L	A	Y	V	E	G	L	K	S	F	D	Q	D	V	K	L	L	Y	L	K	E	A	T	I	G	F	Y	F	L	P	N	N	H	F	Y	C	L	F	N	E	I	N	T	F	V	H	P	N	C		344																													
AtGID1a	PKS	L	V	W	A	G	L	D	L	I	R	D	W	L	A	Y	A	E	G	L	K	K	A	G	Q	E	V	K	L	M	H	L	E	K	A	T	V	G	F	Y	L	P	N	N	H	F	H	N	V	M	D	E	I	S	A	F	V	N	A	E	C		345																														
AtGID1b	PKS	L	V	W	A	G	L	D	L	I	Q	D	W	L	A	Y	V	D	G	L	K	T	G	L	E	V	N	L	L	Y	L	K	Q	A	T	I	G	F	Y	F	L	P	N	N	H	F	H	C	L	M	E	L	N	K	F	V	H	S	I	E	D	S	Q	S	K	S	P	V	L	L	T		357																				
AtGID1c	PKS	L	V	W	A	G	L	D	L	I	Q	D	W	L	K	Y	A	E	G	L	K	K	A	G	Q	E	V	K	L	L	Y	L	E	Q	A	T	I	G	F	Y	L	P	N	N	H	F	H	T	V	M	D	E	I	A	A	F	V	N	A	E	C		344																														
OsGID1	AKS	L	I	I	W	S	G	L	D	L	T	C	D	R	L	A	Y	A	D	A	L	R	E	D	G	H	H	V	K	V	Q	C	E	N	A	T	V	G	F	Y	L	P	N	T	V	H	Y	H	E	V	M	E	E	I	S	D	F	L	N	A	N	L	Y		354																												
GhGID1a	PKS	L	V	W	A	G	L	D	L	I	Q	D	W	L	A	Y	V	E	G	L	K	K	S	G	Q	E	V	N	L	L	F	L	E	K	A	T	I	G	F	Y	F	L	P	N	N	H	F	Y	C	L	M	E	E	I	K	N	F	V	N	P	N	C		344																													
GhGID1b	PKS	L	V	W	A	G	L	D	L	I	Q	D	W	L	A	Y	V	E	G	L	K	K	S	G	Q	E	V	K	L	L	F	L	E	K	A	T	I	G	F	Y	F	L	P	N	N	H	F	Y	R	L	M	E	E	M	N	F	V	H	S	N	C		344																														

通过DNAMAN6.0(采用默认参数)对GoGID、OsGID1、AtGID1s和GhGID1s进行序列比对。比对的氨基酸序列为：GID1a(Q9MAA7)，tGID1b(Q9LYC1)，AtGID1c(Q940G6)，OsGID1(Q6L545)，GhGID1a(ABG89394)，GhGID1b(ABQ96123)。黑色阴影部分代表100%相同。HGG、GXSXG基序为激素敏感脂肪酶中保守序列；箭头所指氨基酸（Gly-196、Arg-251，在水稻中对GA结合活性所必需）同样在GoGID中保守。



Function	<p>Functions as soluble gibberellin (GA) receptor. GA is an essential hormone that regulates growth and development in plants. Binds with high affinity the biologically active gibberellin GA4, but has no affinity for the biologically inactive GAs. In response to GA, interacts with specific DELLA proteins, known as repressors of GA-induced growth, and targets them for degradation via proteasome. Seems to be required for GA signaling that controls root growth, seed germination and flower development. May function as a dominant GA receptor at low GA concentrations in germination. Partially redundant with GID1A and GID1C. Ref.6</p> <p>Ref.7 Ref.8</p>
Subunit structure	<p>Interacts with the DELLA proteins GAI, RGA, RGL1, RGL2 and RGL3 in a GA-dependent manner. Ref.6 Ref.7 Ref.9</p>
Subcellular location	<p>Nucleus By similarity.</p>
Tissue specificity	<p>Widely expressed. Ref.5</p>
Disruption phenotype	<p>No visible phenotype under normal growth condition. Ref.6 Ref.8</p>
Sequence similarities	<p>Belongs to the 'GDXG' lipolytic enzyme family.</p>

AtGID1b

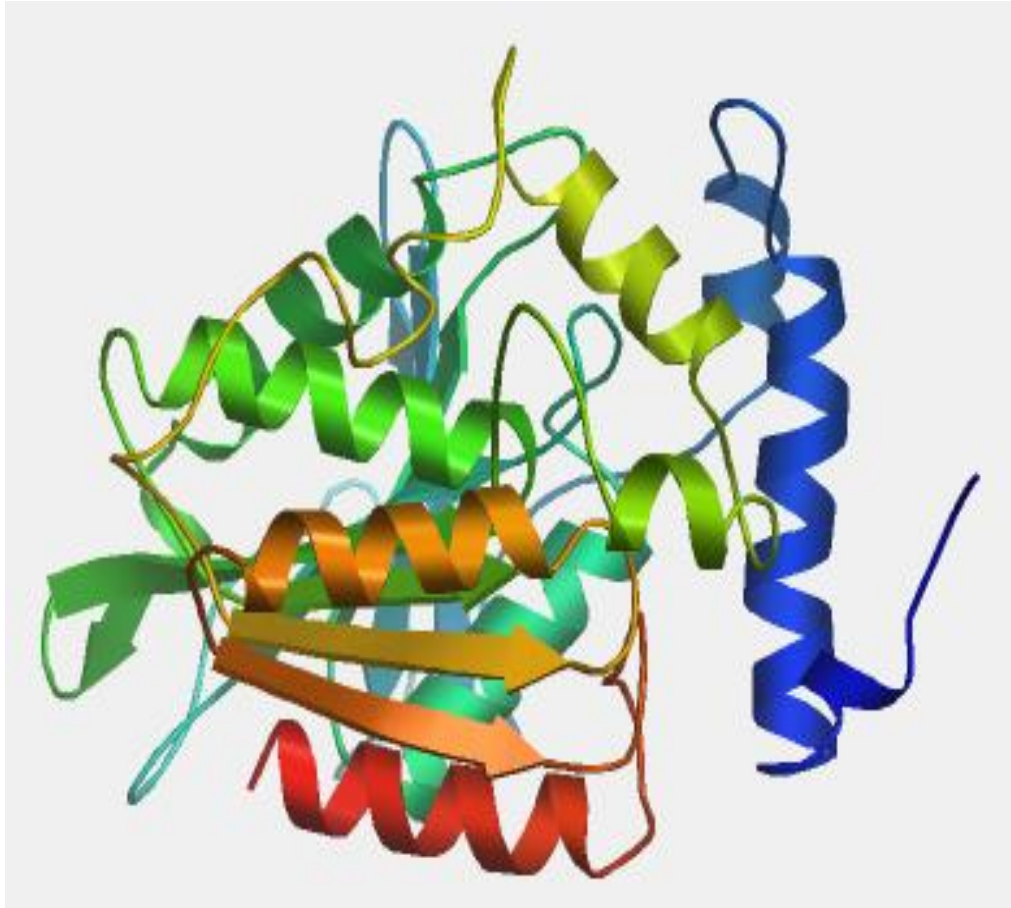


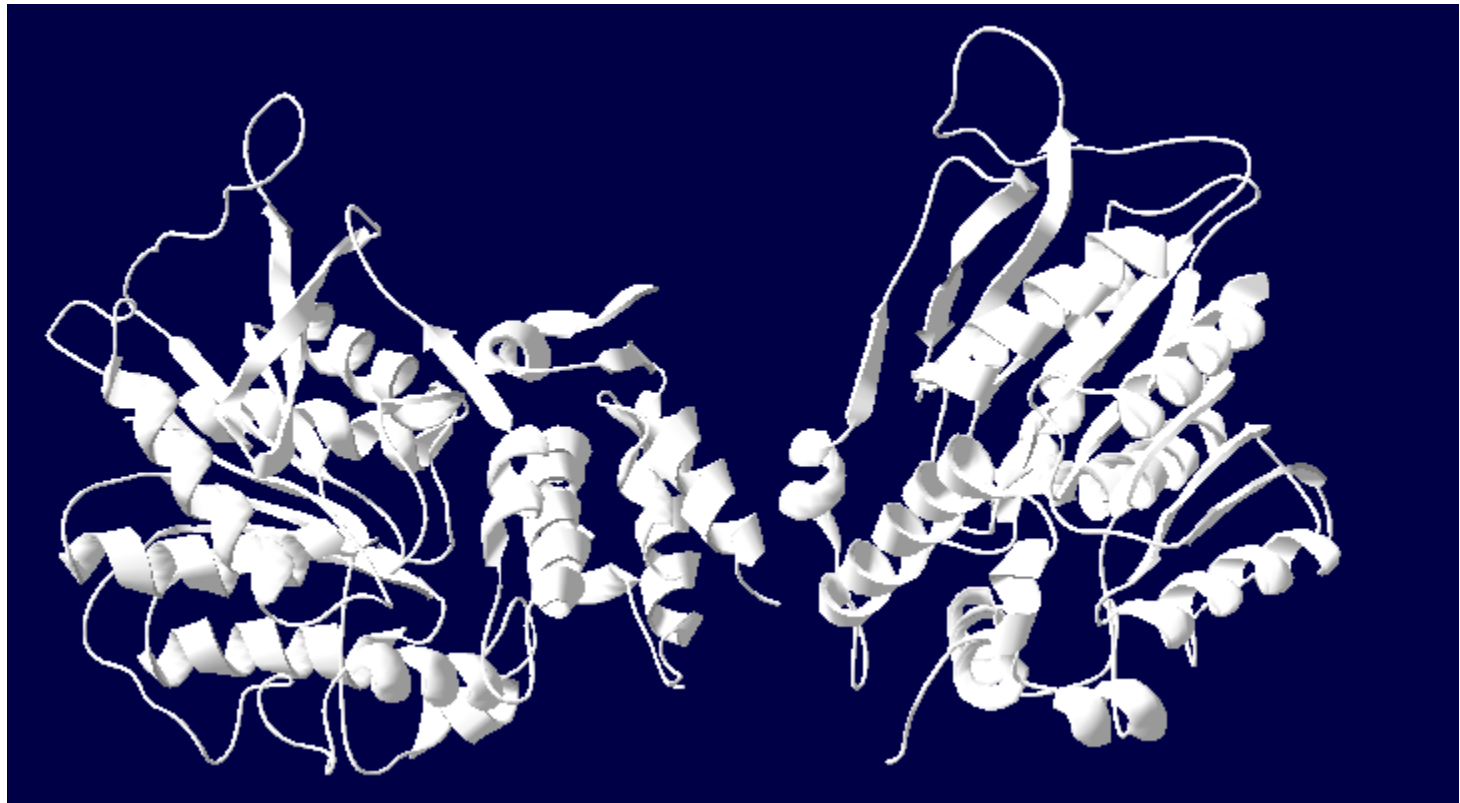
Your Input:

- GID1B GID1B (GA INSENSITIVE DWARF1B); hydrolase; Encodes a gibberellin (GA) receptor ortholog of the rice GA receptor gene (OsGID1). Has GA-binding activity, showing higher affinity to GA4. Interacts with DELLA proteins in vivo in the presence of GA4. ; Probable soluble gibberellin (GA) receptor. GA is an essential hormone that regulates growth and development in plants (By similarity) (358 aa)
(Arabidopsis thaliana)

Predicted Functional Partners:

		Neighborhood	Gene Fusion	Cooccurrence	Coexpression	Experiments	Databases	Textmining	[Homology]	Score
●	RGA1	RGA1 (REPRESSOR OF GA1-3 1); protein binding / transcription factor; Member of the VHIID/DELLA [...]				●			●	0.999
●	GAI	GAI (GIBBERELIC ACID INSENSITIVE); transcription factor; Similar to a putative transcription f [...]				●			●	0.998
●	SLY1	SLY1 (SLEEPY1); F-box protein that is involved in GA signaling. Regulates seed germination. Com [...]				●			●	0.994
●	RGL2	RGL2 (RGA-LIKE 2); transcription factor; Encodes a DELLA protein, a member of the GRAS superfam [...]				●				0.951
●	RGL1	RGL1 (RGA-LIKE 1); transcription factor; Negative regulator of GA responses, member of GRAS fam [...]				●				0.951
●	SLY2	SLY2 (SLEEPY2); encodes an F-box protein whose protein sequence is similar to SLY1, which belon [...]							●	0.562
●	GA1	GA1 (GA REQUIRING 1); ent-copalyl diphosphate synthase/ magnesium ion binding; Catalyzes the co [...]							●	0.440





GoGID







感谢大家参与

请各位老师同学批评、指教