

Phylogenetics and Conserved motifs analysis of cryptochrome(CRY)

CRY的系统发生及保守结构域分析

PKU 16F
Group 05
韩雪 王瑞璞
樊德 李丛冉

Background:

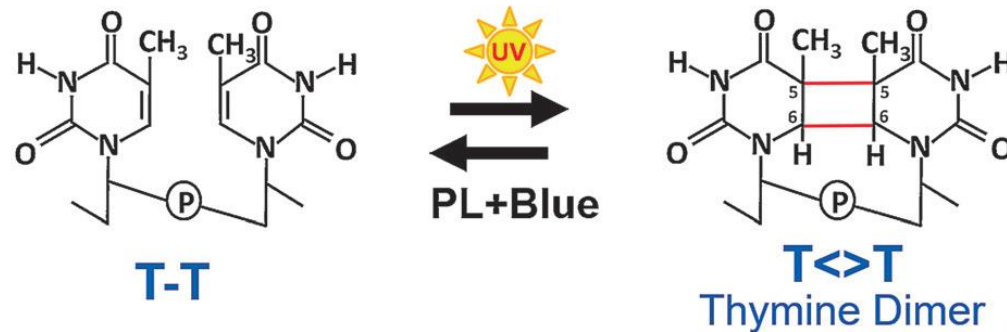
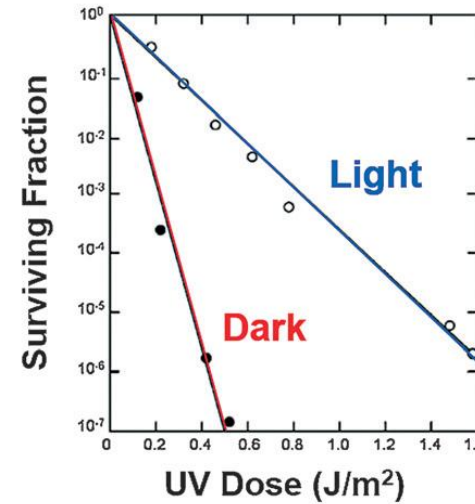
- 1、 Photolyase–cryptochrome connection
- 2、 The function of cryptochrome in different species
- 3、 The evolution of cryptochrome

1、 Photolyase–cryptochrome connection

(1) Photoreactivation and photolyase



Rupert and Sancar, UT Dallas, 2009



UV induces the formation of a cyclobutane thymine dimer (T< >T), photolyase binds to the dimer, absorbs a blue light photon, and converts the dimer to two canonical thymines.

(2) Photolyase–cryptochrome connection

- Humans do not have photolyase

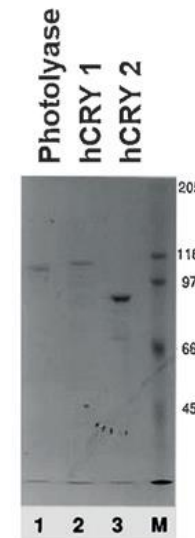
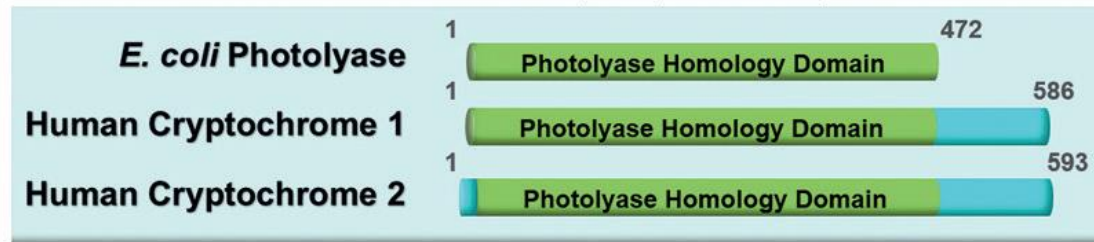
Li YF, *et al* (1993) *PNAS* 90:4389-93

- Humans have a photolyase homolog

Adams MD, *et al* (1995) *Nature* 377:3-174

- Humans have **2** photolyase paralogs

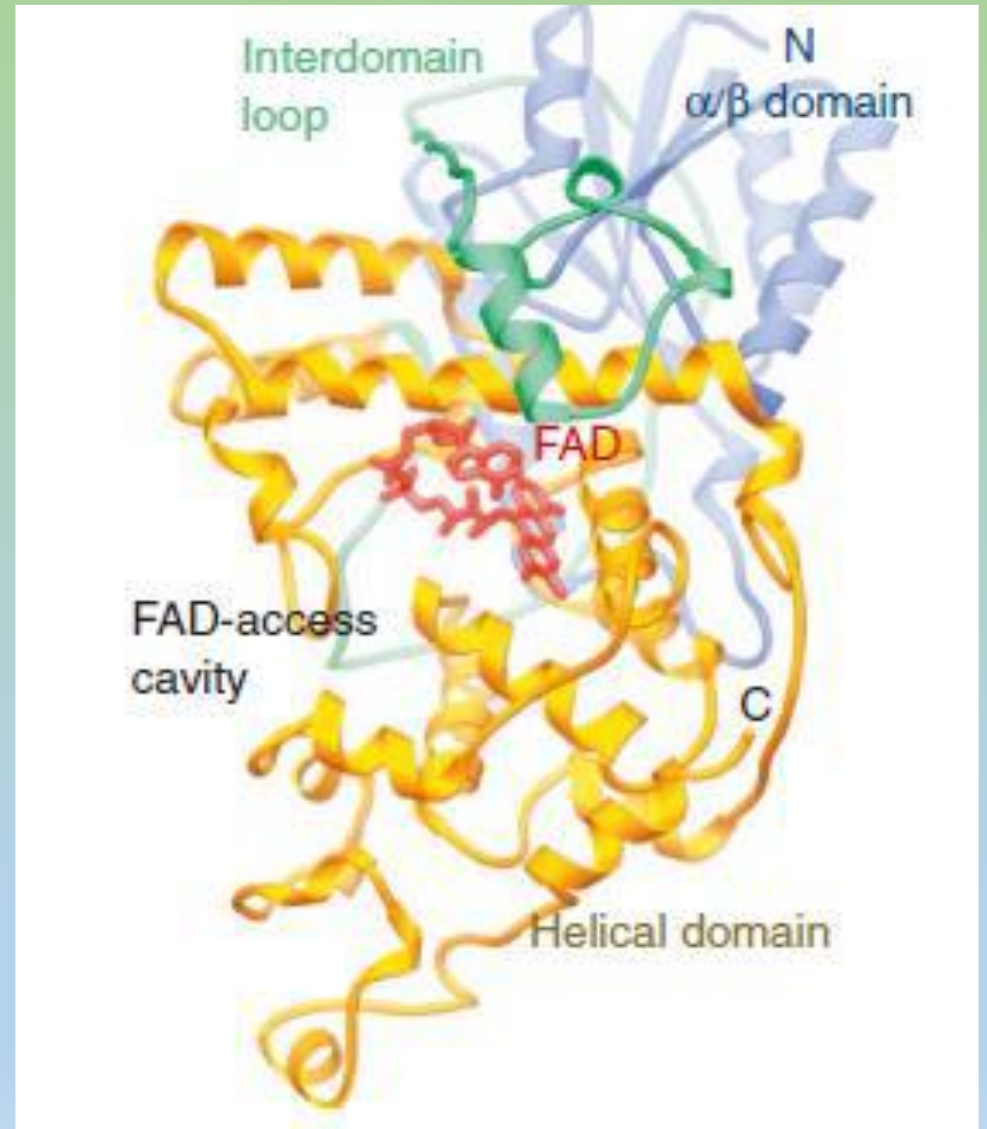
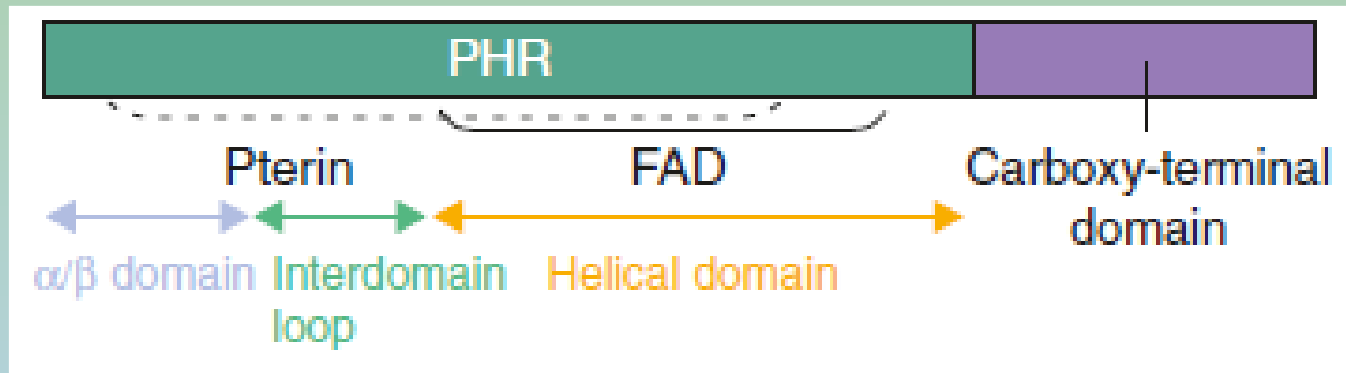
Hsu DS, *et al* (1996) *Biochemistry* 35:13871-77



Photolyase Cryptochrome

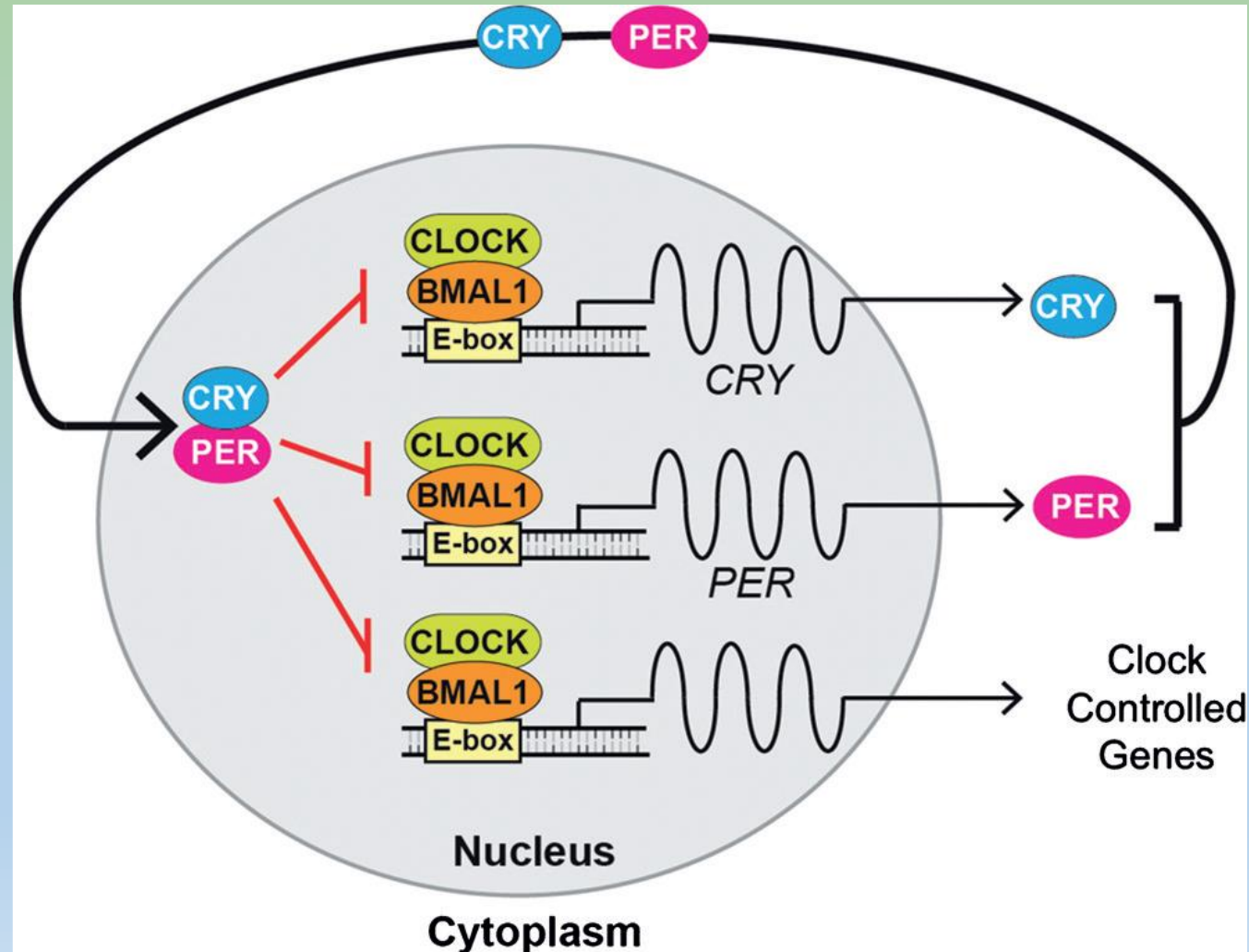
C. A. Brautigam *et al.*, *Proc. Natl. Acad. Sci. USA*
2004, 101, 12142–12147.

(3) The structure of cryptochrome

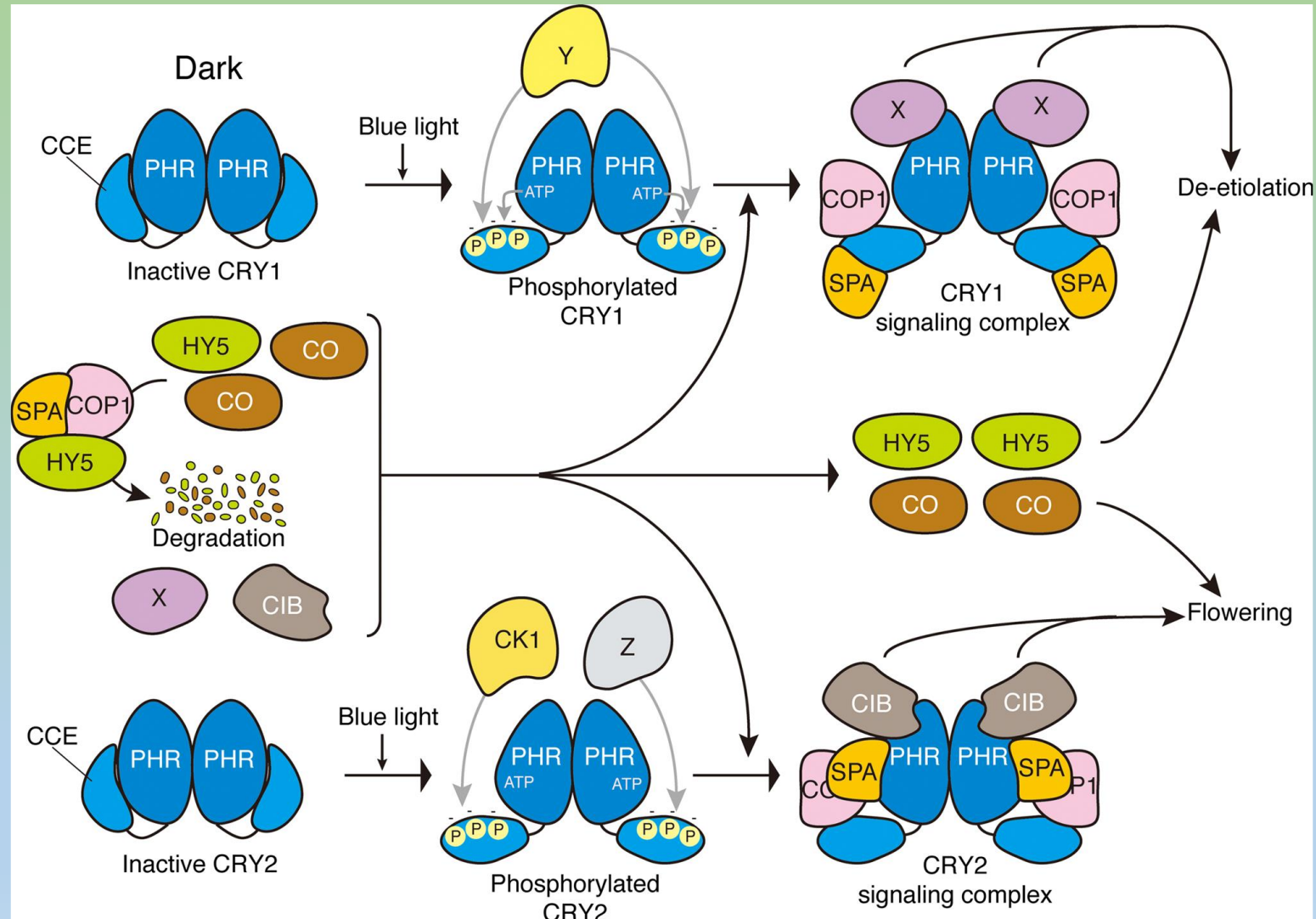


2、 The function of cryptochrome in different species

(1) Molecular mechanism of the mammalian circadian clock

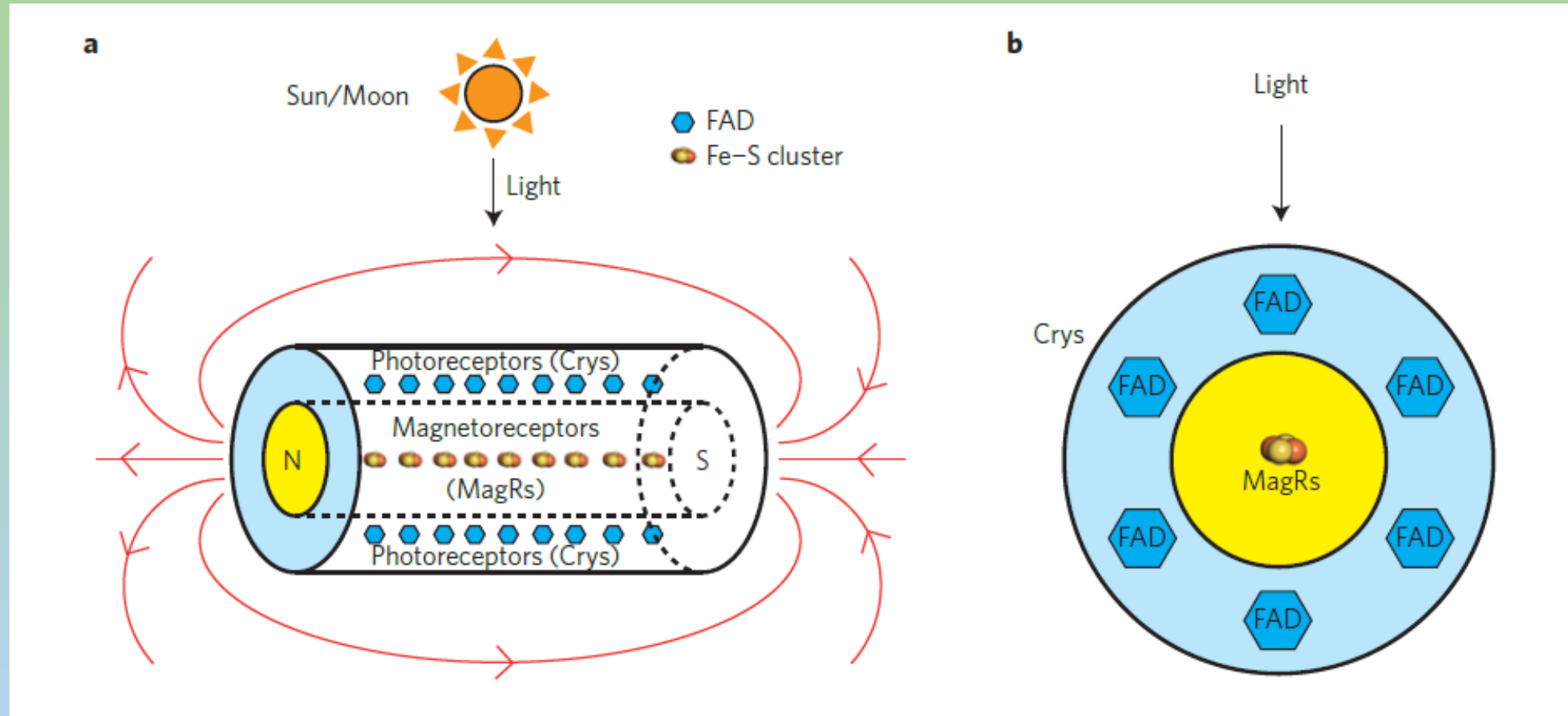


(2) The model of cryptochrome signal transduction mechanisms in Arabidopsis.



(Yoshito Oka, et al, 2016)

(3) The biocompass model of animal magnetoreception and navigation

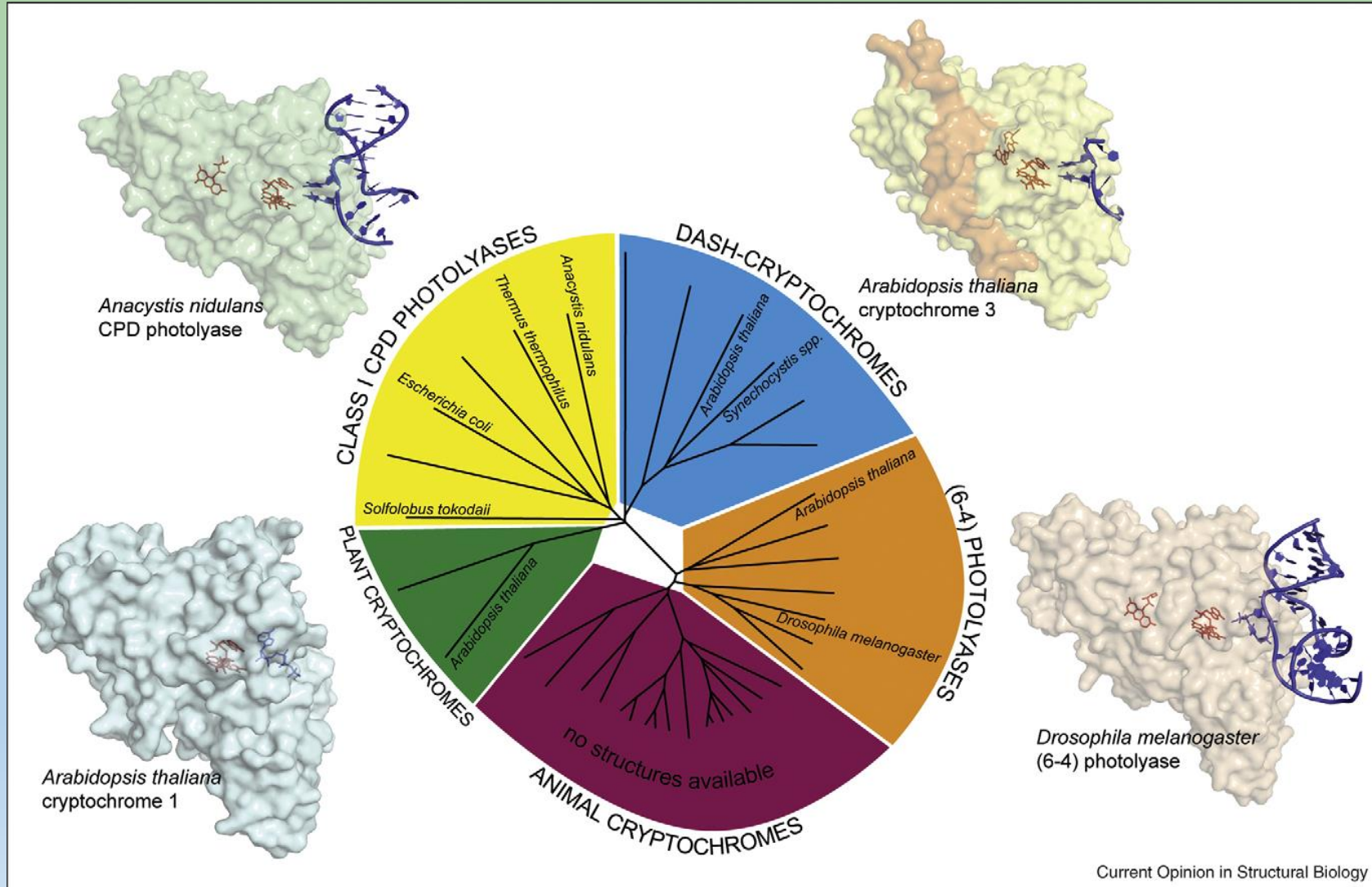


(Can Xie, et al,2015)

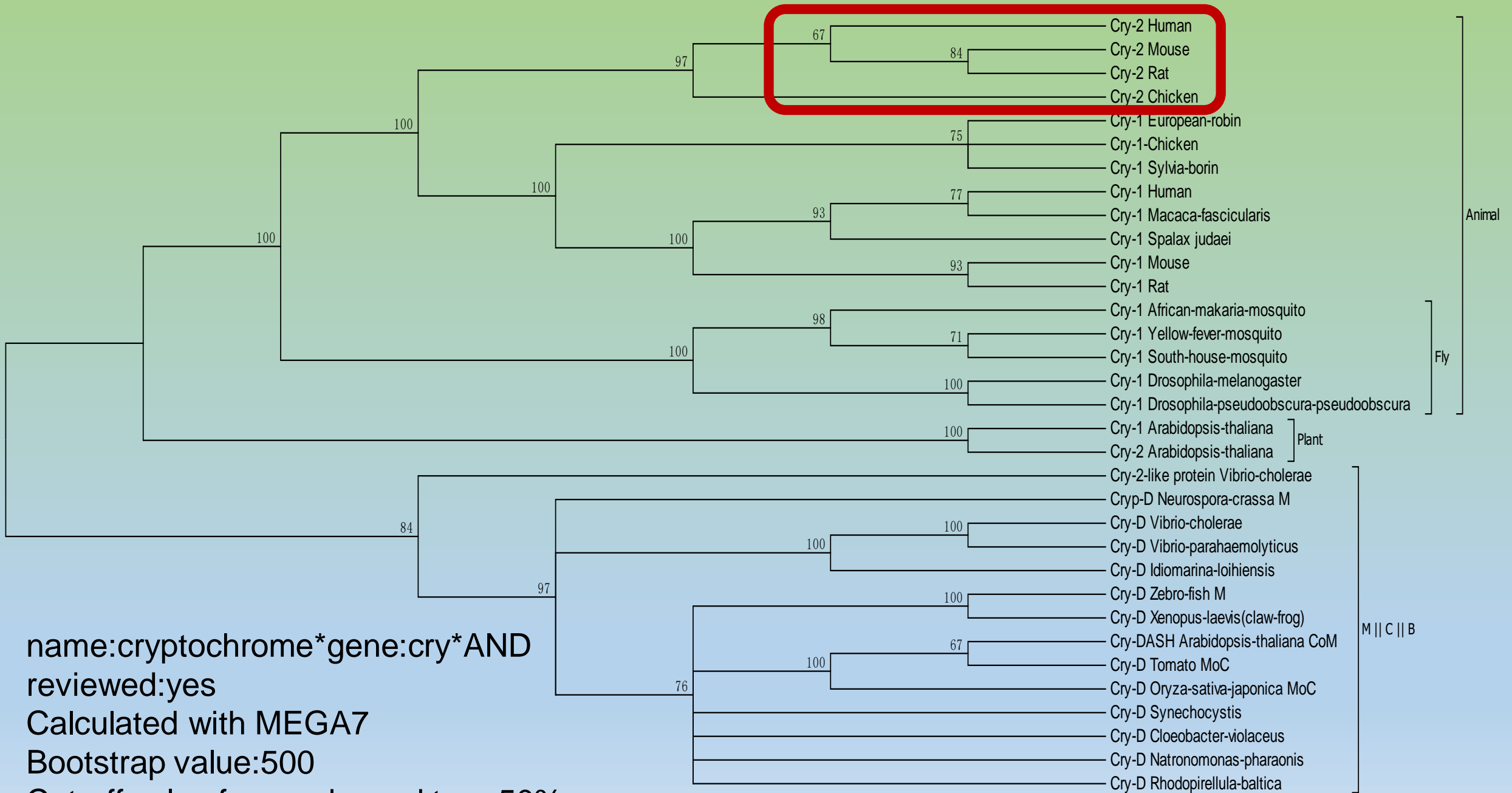
A nanoscale Cry/MagR magnetosensor complex with intrinsic magnetic polarity acts as a light-dependent biocompass

3、 The evolution of cryptochrome

Phylogenetic tree of representative members of the cryptochrome/photolyase family, calculated with ClustalX

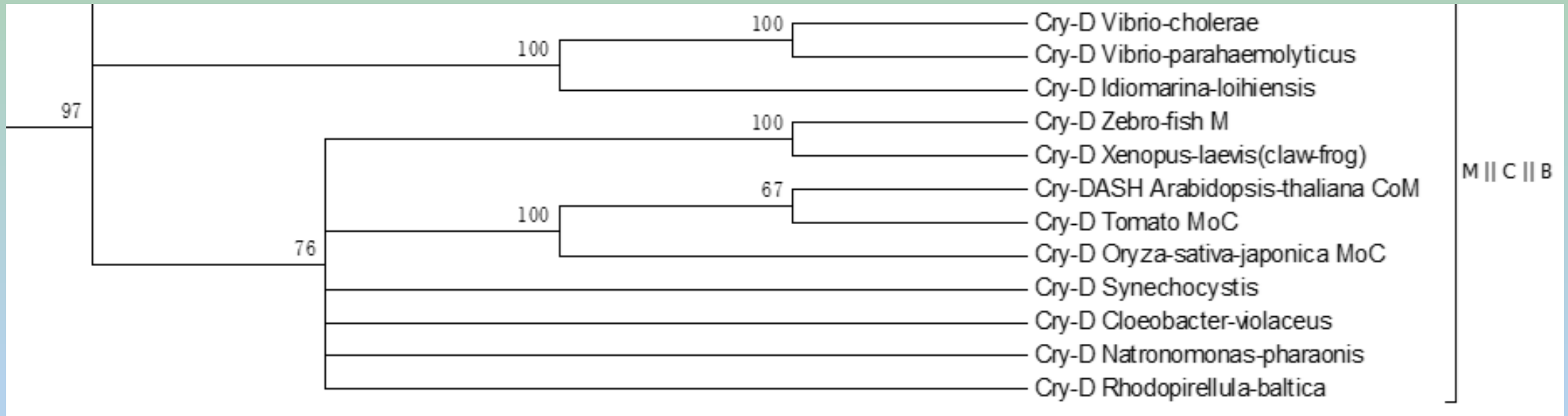


Phylogenetics analysis of Cry



name:cryptochrome*gene:cry*AND
 reviewed:yes
 Calculated with MEGA7
 Bootstrap value:500
 Cut-off value for condensed tree:50%

Endosymbiosis 内共生



- a. Unreviewed sequences
- b. The plants we focused on
- c. Similar sequence length

Filter byⁱ



Reviewed (33)

Swiss-Prot

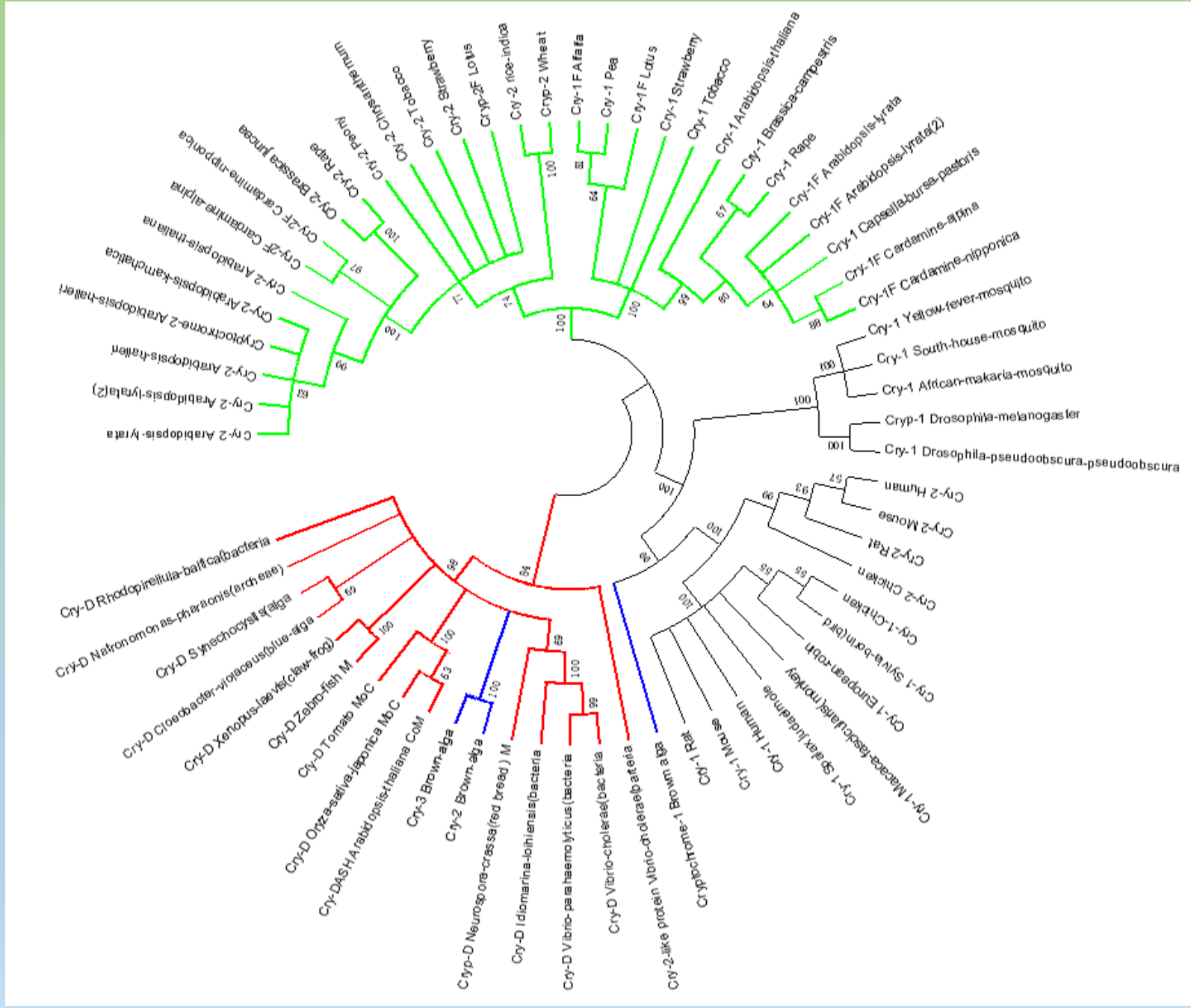


Unreviewed (718)

TrEMBL

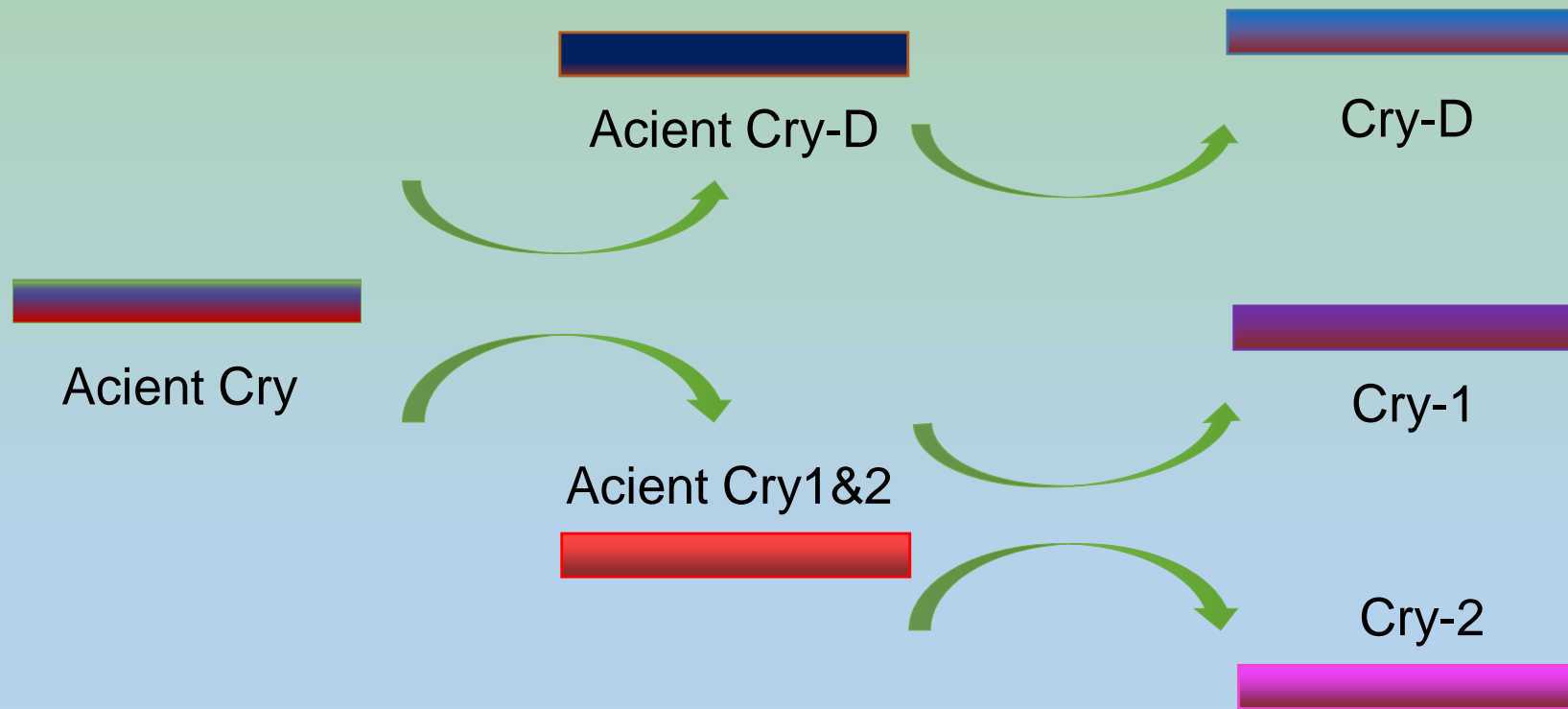
Entry	Entry name	Protein names	Gene names	Organism	Length
A0A1B3TPC4	A0A1B3TPC4_DEPLA	Cryptochrome	CRY2	Deparia lancea (False spleenwort) (Asplenium lanceum)	8
A0A1A8EXX6	A0A1A8EXX6_9TELE	Cryptochrome 1 (Photolyase-like)	CRY1	Nothobranchius korthausae	10
A0A1B3TPC9	A0A1B3TPC9_9MONI	Cryptochrome	CRY2	Deparia prolifera	10
A0A1A8VG70	A0A1A8VG70_NOTFU	Cryptochrome 3	CRY3	Nothobranchius furzeri (Turquoise killifish)	13
A0A1A7ZRM0	A0A1A7ZRM0_NOTFU	Cryptochrome 1 (Photolyase-like)	CRY1	Nothobranchius furzeri (Turquoise killifish)	15
A0A0A7NVJ8	A0A0A7NVJ8_TOMSC	Cryptochrome-1	CRY1	Tomistoma schlegelii (False gharial)	19
A0A0A7P0F9	A0A0A7P0F9_9SAUR	Cryptochrome-1	CRY1	Caiman yacare (yacare caiman)	19
A0A0A7NVM3	A0A0A7NVM3_9SAUR	Cryptochrome-1	CRY1	Mecistops cataphractus (African slender-snouted crocodile)	19
A0A0A7NVQ7	A0A0A7NVQ7_OSTTE	Cryptochrome-1	CRY1	Osteolaemus tetraspis (African dwarf crocodile)	19

Calculated with
MEGA7
Bootstrap
value:500
Cut-off value for
condensed
tree:50%

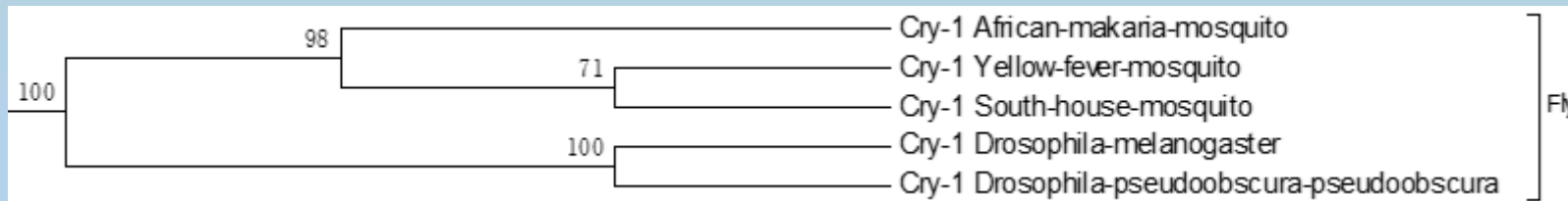


Plants
Animals
Cry-D
Brown-alga 褐藻

The predicted phylogenetic model of Cry



Cry-2 in the FLY: absent gene? or undiscovered gene ?



Search the potential Cry-2 in the *Drosophila melanogaster* by BlastP

Max target sequences 100
Select the maximum number of aligned sequences to display

Short queries Automatically adjust parameters for short input sequences

Expect threshold 0.001

Word size 6

Max matches in a query range 0

Scoring Parameters

Matrix BLOSUM62

Gap Costs Existence: 11 Extension: 1

Compositional adjustments Conditional compositional score matrix adjustment

Domazet-Loso and Tautz,
Mol. Biol. Evol., 2008

Sequences producing significant alignments:

Select: All None Selected: 0

Alignments Download GenPept Graphics Distance tree of results Multiple alignment

Description	Max score	Total score	Query cover	E value	Ident	Accession
Chain A, Crystal Structure Of Drosophila Cryprochrome	1129	1129	100%	0.0	100%	4K03_A
cryptochrome [Drosophila melanogaster]	1128	1128	100%	0.0	100%	NP_732407.1
Chain A, Structure Of Full-length Drosophila Cryptochrome	1123	1123	99%	0.0	100%	4GU5_A
blue light photoreceptor [Drosophila melanogaster]	1122	1122	100%	0.0	99%	BAA35000.1
Chain A, Crystal Structures Of Drosophila Cryptochrome	1109	1109	99%	0.0	99%	4JZY_A
(6-4)-photolase, isoform B [Drosophila melanogaster]	335	335	93%	1e-108	38%	NP_477188.1
RE11660p [Drosophila melanogaster]	335	335	93%	1e-108	38%	AAL90322.1
Chain A, Drosophila Melanogaster (6-4) Photolase Bound To Ds Dna With A T-T (6-4) Photolesion	335	335	93%	2e-108	38%	3CVU_A
Chain A, Drosophila Melanogaster (6-4) Photolase Bound To Ds Dna With A T-T (6-4) Photolesion	334	334	93%	4e-108	38%	3CVX_A
photolase [Drosophila melanogaster]	333	333	93%	7e-108	38%	BAA12067.1
Chain A, Drosophila Melanogaster (6-4) Photolase Bound To Ds Dna With A T-T (6-4) Photolesion And Cofactor F0	332	332	93%	2e-107	38%	3CVW_A

BlastP with Cry-1_Drome

Sequences producing significant alignments:

Select: All None Selected: 0

Alignments Download GenPept Graphics Distance tree of results Multiple alignment


Description	Max score	Total score	Query cover	E value	Ident	Accession
photolase [Drosophila melanogaster]	170	170	77%	1e-45	29%	BAA12067.1
Chain A, Drosophila Melanogaster (6-4) Photolase H365n Mutant Bound To Ds Dna With A T-T (6-4) Photolesion And Cofactor F0	169	169	77%	2e-45	29%	3CVW_A
Chain A, Drosophila Melanogaster (6-4) Photolase H369m Mutant Bound To Ds Dna With A T-T (6-4) Photolesion	169	169	77%	3e-45	29%	3CVX_A

BlastP with Cry-2_Arado

Cry-1 and Photolyase are different gene in *Drosophila melanogaster*

ACTIONS	QUERY	SCORE	START	END	QSIZE	IDENTITY	CHRO	STRAND	START	END	SPAN
browser details	Cryp-1_Drosophila-melanogaster	623	1	542	542	100.0%	3R	++	19212341	19215359	3019


chr2L:20,641,965-20,647,013 5,049 bp.



chr2L (38D1-38D2)

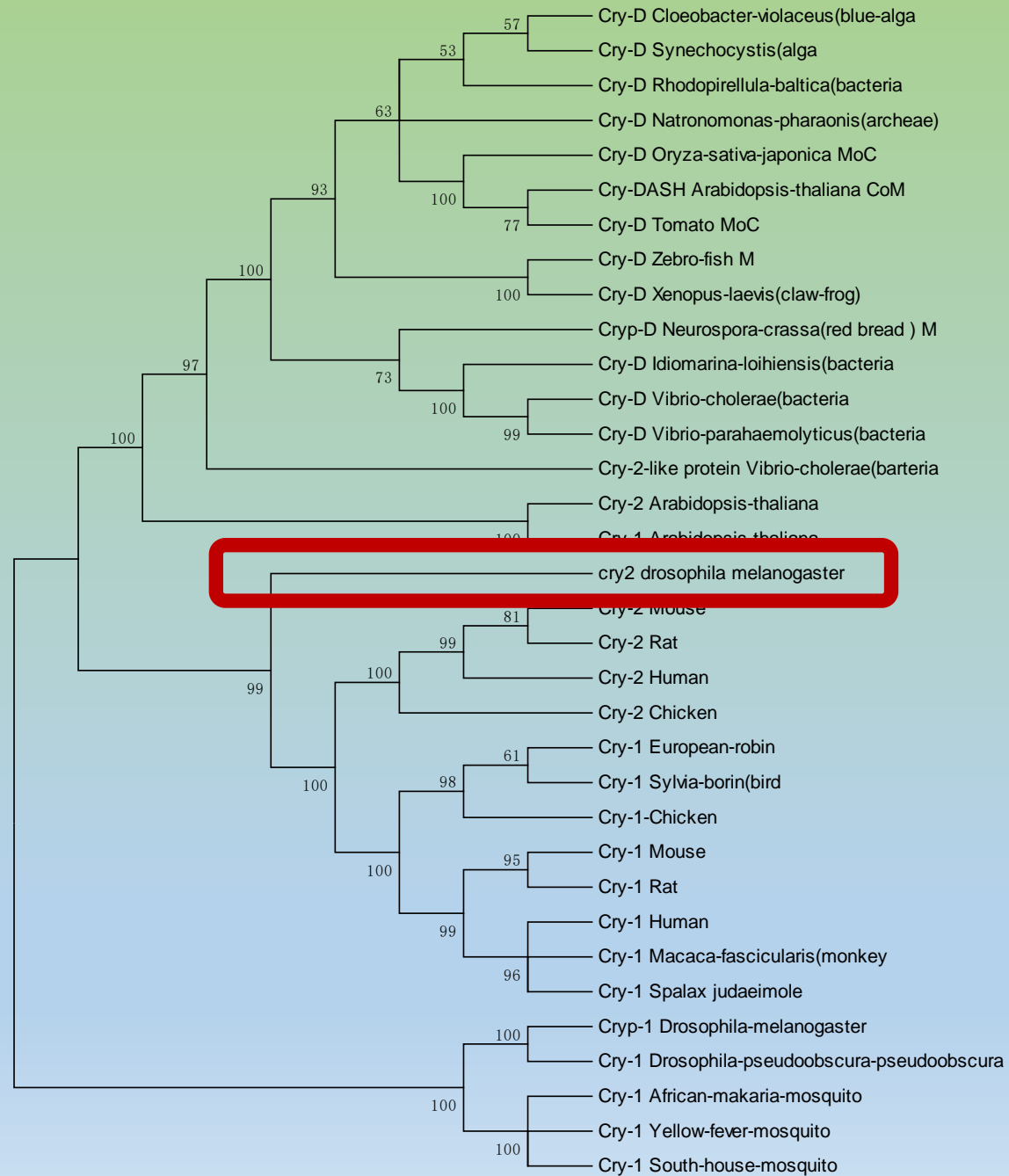
ACTIONS	QUERY	SCORE	START	END	QSIZE	IDENTITY	CHRO	STRAND	START	END	SPAN
browser details	BAA12067.1	1595	1	540	540	99.3%	2L	+ -	20643648	20645330	1683

chr3R:19,212,341-19,215,359 3,019 bp.



chr3R (91F11)

Calculated with MEGA7
Bootstrap value:500
Cut-off value for
condensed tree:50%



Search the potential Cry in the *Amphimedon queenslandica* by BlastP

Sequences producing significant alignments:

Select: [All](#) [None](#) Selected:0

[Alignments](#) [Download](#) [GenPept](#) [Graphics](#) [Distance tree of results](#) [Multiple alignment](#)

	Description	Max score	Total score	Query cover	E value	Ident	Accession
<input type="checkbox"/>	PREDICTED: cryptochrome-2-like [Amphimedon queenslandica]	142	142	67%	3e-36	27%	XP_003386582.1
<input type="checkbox"/>	PREDICTED: cryptochrome-2-like [Amphimedon queenslandica]	142	142	70%	3e-36	27%	XP_003386569.1
<input type="checkbox"/>	PREDICTED: deoxyribodipyrimidine photo-lyase-like [Amphimedon queenslandica]	33.9	33.9	8%	0.20	32%	XP_003390874.1
<input type="checkbox"/>	PREDICTED: deoxyribodipyrimidine photo-lyase-like [Amphimedon queenslandica]	33.5	33.5	8%	0.22	30%	XP_003390970.2
<input type="checkbox"/>	PREDICTED: glutamate carboxypeptidase 2-like [Amphimedon queenslandica]	27.3	27.3	3%	6.3	44%	XP_003392075.1

BlastP with Cry-1_Arado

Sequences producing significant alignments:

Select: [All](#) [None](#) Selected:0

[Alignments](#) [Download](#) [GenPept](#) [Graphics](#) [Distance tree of results](#) [Multiple alignment](#)

	Description	Max score	Total score	Query cover	E value	Ident	Accession
<input type="checkbox"/>	PREDICTED: cryptochrome-2-like [Amphimedon queenslandica]	157	157	76%	1e-41	28%	XP_003386569.1
<input type="checkbox"/>	PREDICTED: cryptochrome-2-like [Amphimedon queenslandica]	129	129	77%	4e-32	24%	XP_003386582.1
<input type="checkbox"/>	PREDICTED: deoxyribodipyrimidine photo-lyase-like [Amphimedon queenslandica]	40.4	40.4	24%	0.001	25%	XP_003390970.2

BlastP with Cry-2_Arado

Conclusion one

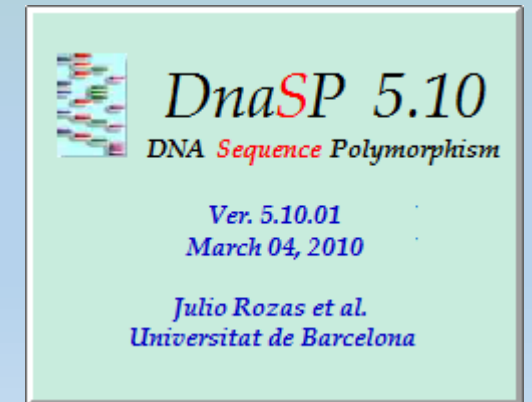
- Giving a predicted phylogenetic model of the Cry
- The differentiation of Cry-1 and Cry-2 in animal happened after the appearance of the *Drosophila*

K_A/K_S of Cry

- K_A :非同义氨基酸替代位点的核酸平均替代数
- K_S :同义氨基酸替代位点的核酸平均替代数
- 根据Kimura提出的中性演化理论，同义替代位点和非同义替代位点上发生的替代被固定的概率相同，即 $K_A/K_S=1$ 。
- $K_A/K_S < 1$ ，则造成氨基酸改变的替代数少，基因受到负选择；
 $K_A/K_S > 1$ ，则造成氨基酸改变的替代数多，基因受到正选择。



Seq 1	Seq 2	SynDif	SynPos	K_S	NSynDif	NSynPos	K_A
Cry1-Danio	Cry1-Mus	48.00	362.83	0.1455	306.00	1185.17	0.3165
Cry1-Danio	Cry1-Droso	187.92	366.83	0.8617	606.08	1181.17	0.8644
Cry1-Danio	Cry1-Homo	47.83	352.50	0.1497	328.17	1195.50	0.3418
Cry1-Mus	Cry1-Droso	170.08	355.17	0.7631	630.92	1192.83	0.9162
Cry1-Mus	Cry1-Homo	10.50	340.83	0.0315	142.50	1207.17	0.1284
Cry1-Droso	Cry1-Homo	171.58	344.83	0.8167	663.42	1203.17	0.9966



Conserved motifs analysis

Search the conserved motifs by the similarity of the sequences

M7: Alignment Explorer

Data Edit Search Alignment Web Sequencer Display Help

Protein Sequences

Species/Abbrv	Gr
3. Cry-2_Arabidopsis-thaliana	
4. Cry-1_Arabidopsis-thaliana	
5. Cry-DASH_Arabidopsis-thaliana_CoM	
6. Cry-1_South-house-mosquito	
7. Cry-D_Zebro-fish_M	
8. Cry-1_Drosophila-melanogaster	
9. Cry-1_Drosophila-pseudoobscura-pseudoobscura	
10. Cry-1_European-robin	
11. Cry-1_Chicken	
12. Cry-2_Chicken	
13. Cry-D_Cloeobacter-violaceus(blue-alga	
14. Cry-1_Human	
15. Cry-2_Human	
16. Cry-D_Idiomarina-loihiensis(bacteria	
17. Cry-1_Macaca-fascicularis(monkey	
18. Cry-1_Mouse	

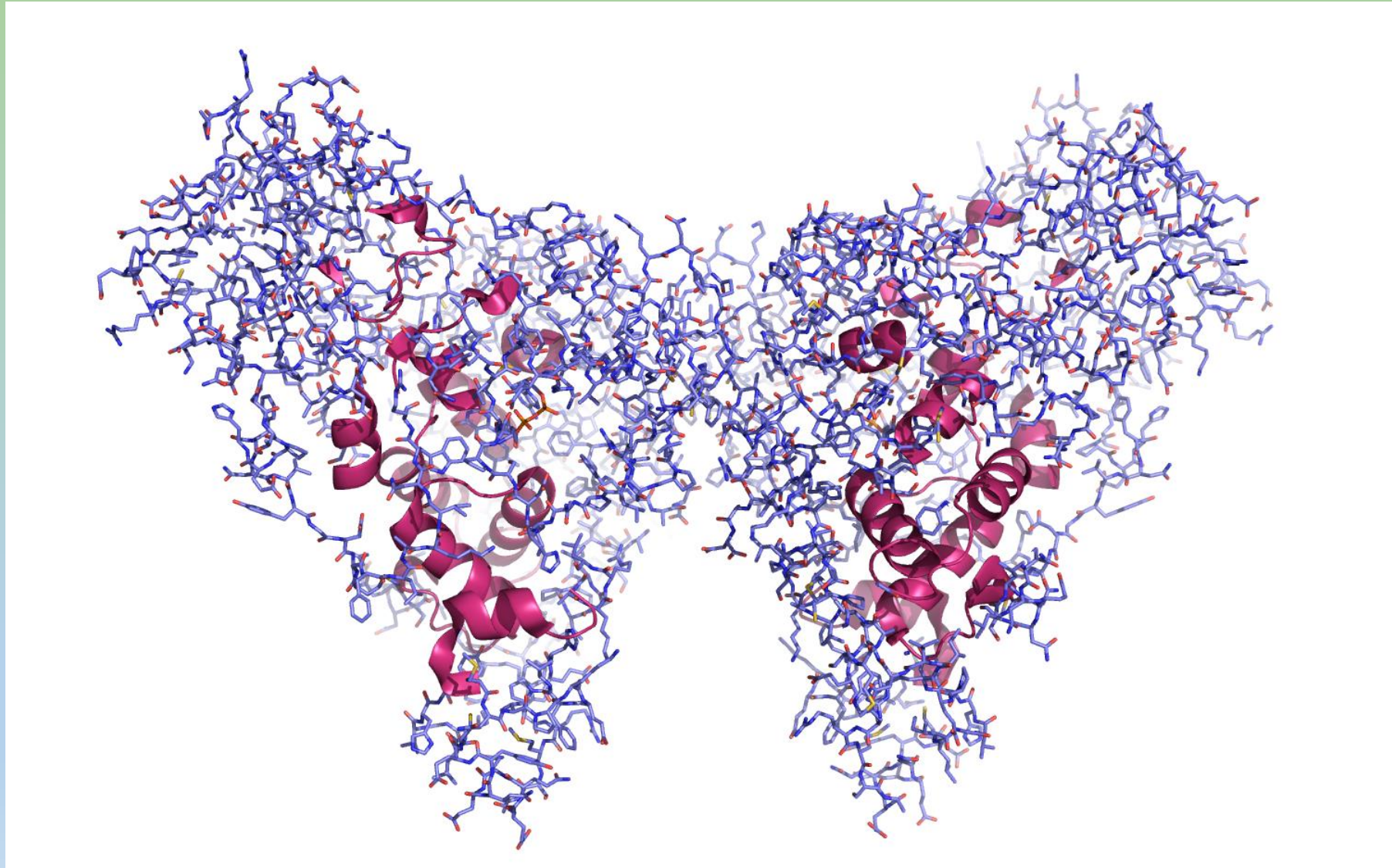
Tool Bars
Columns
Use Color
Background Color
Toggle Conserved Sites
Font

at 50% Level
at 60% Level
at 70% Level
at 80% Level
at 90% Level
at 100% Level
none

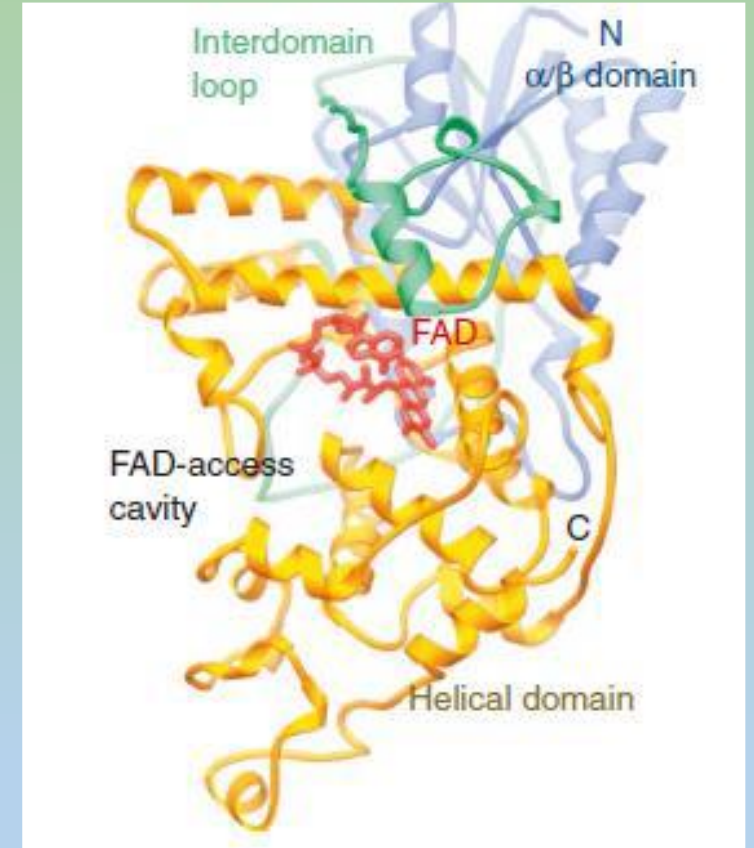
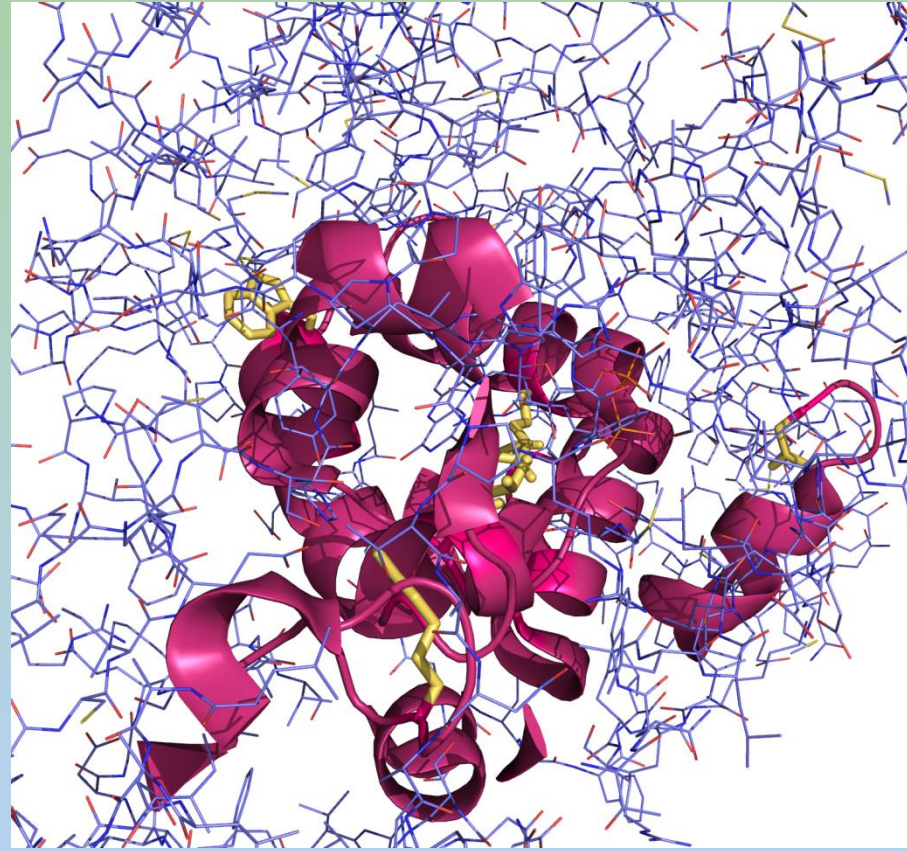
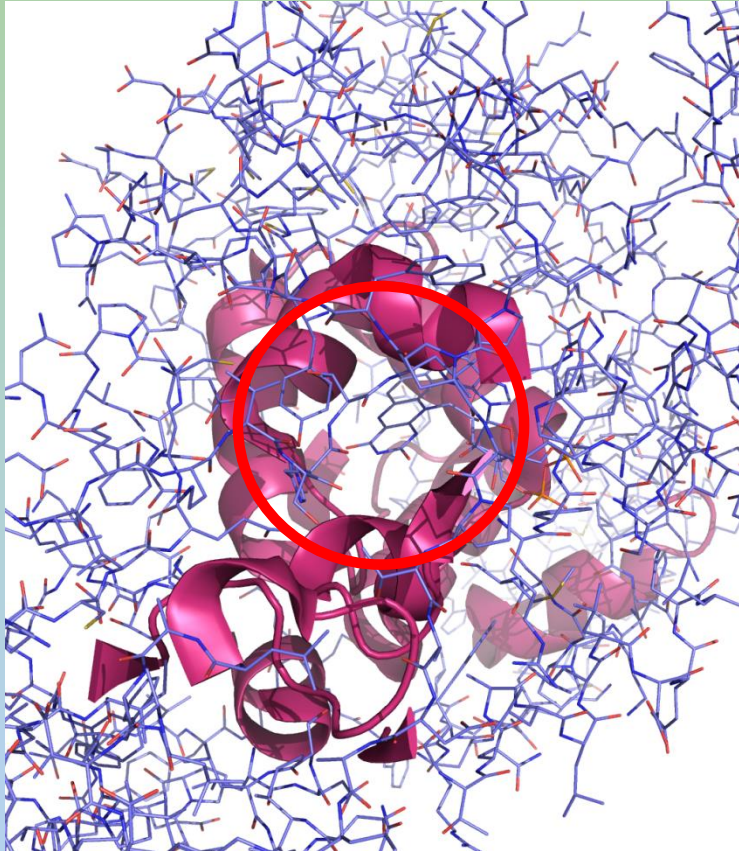
>Cryp-1_Drosophila-melanogaster

MATRGANVI-10-WFRHGLRLHDNPAL-24-LAALADKDQGIALIPVFIF
 DGESAGTKNV-53-GYNRMRF-61-LDSLQDIDDQLQAATDGRGRLLV
 FEGEPAYIFRRLHEQVRLHRICIEQDCEPIWNERDES-121-IRSLCRELNI
 DFVEKVSHTLWDPQLVIETNGGIPPLTYQMFLHTVQIIGLPPRPTADAR
 L-181-EDATFVELDPEFCRSLKLFELPTPEHFNVYGDNMGFLAKINW
 R-225-GGE-228-TQALLLLDERLKVEQHAFERGFYLPNQALPNIHDSPK
 SM-267-SAHLRFGLSVR-279-RFYWSVHDLFKNVQLRACVRGVQM
 TGGAHITGQ-312-LIWREYF-319-YTMSVNNPNYDRMEGNDICLSIP
 WAKPNENLLQS-353-WRLGQTGFPLIDGAMRQLLAEGWLHHTLRNT
 VATFLTRGGLWQSWEHGLQHFLKYLLDADWSVCAGNW-421-MWVS
 SSAFERLLDSSLVTCVALAKRL-448-DPDGTYIKQYVPEL-462-MNVPK
 EHVHEPWRMSAEQQ-481-EQYECLIGVHYPERIIDLSMAVKRNMLAM
 KSLRNSLITPPPHCRPSNEEEVRQFFWLADV-541-VV

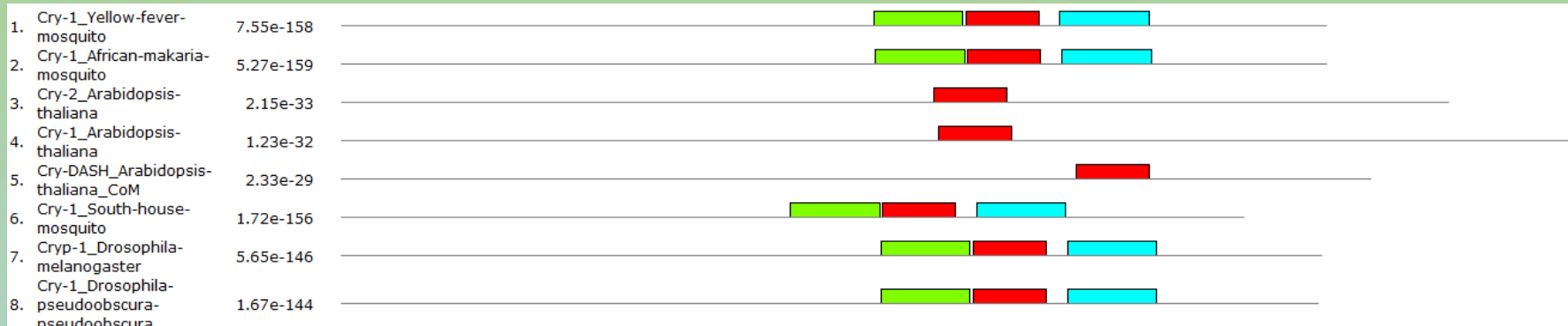
The conserved motifs of CRY-1 in *Drosophila*



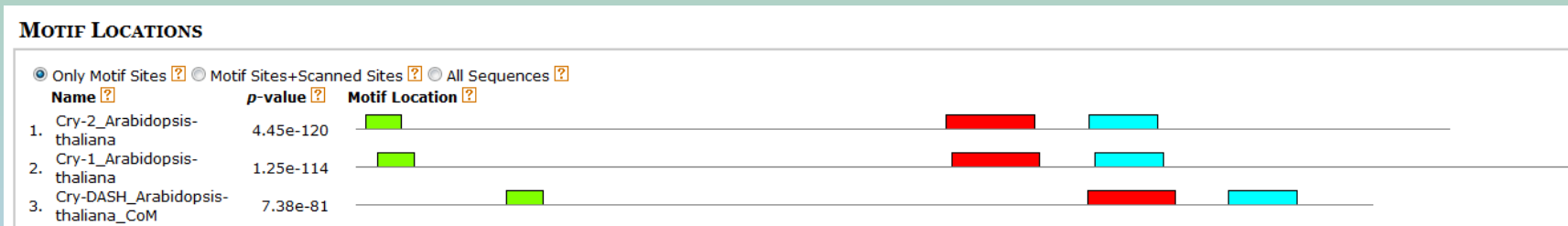
The conserved motif of CRY-1 in *Drosophila*



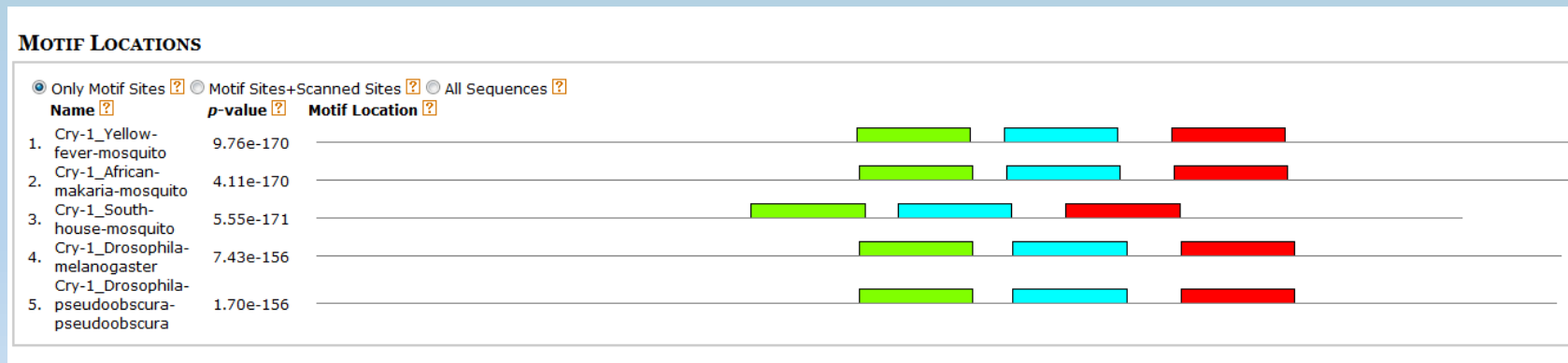
Search the conserved motifs by MEME



Enriched motifs with Cry of *Drosophila* and *Arabidopsis*



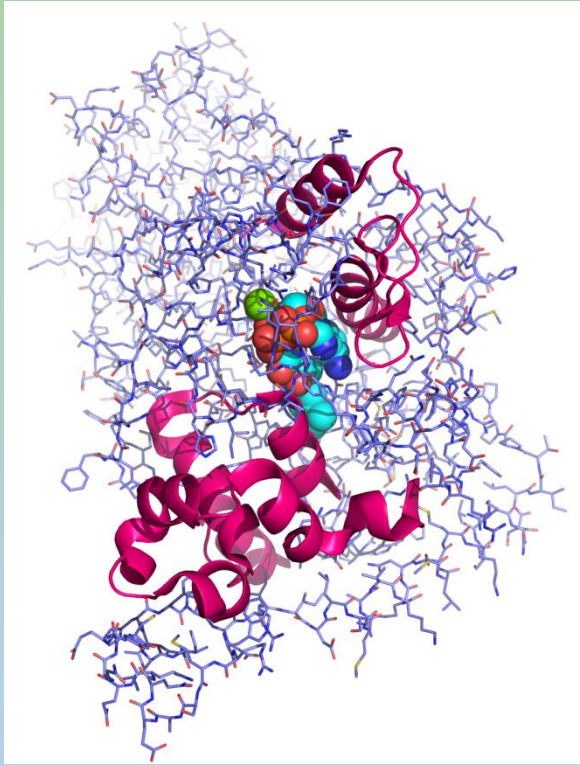
Enriched motifs with Cry of *Arabidopsis*



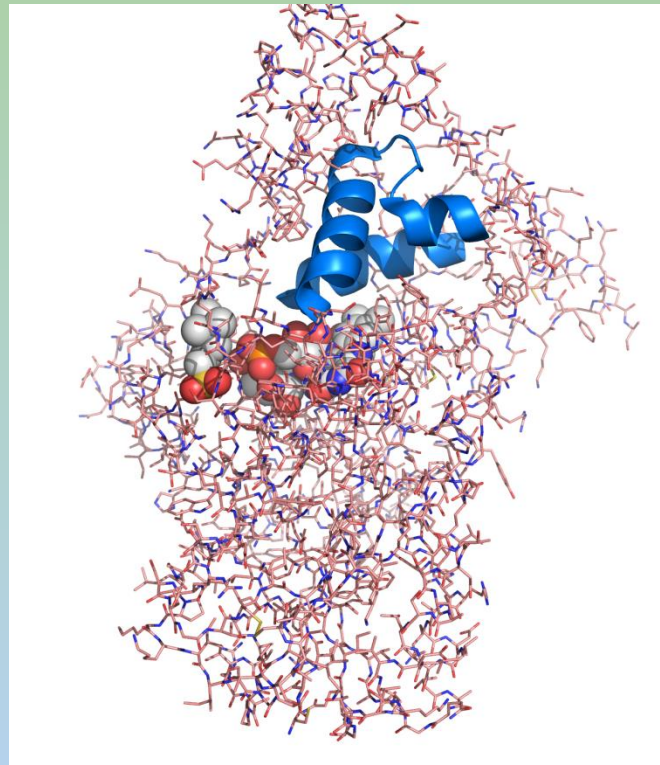
Enriched motifs with Cry of *Drosophila*

MEME:
<http://meme-suite.org/tools/meme>

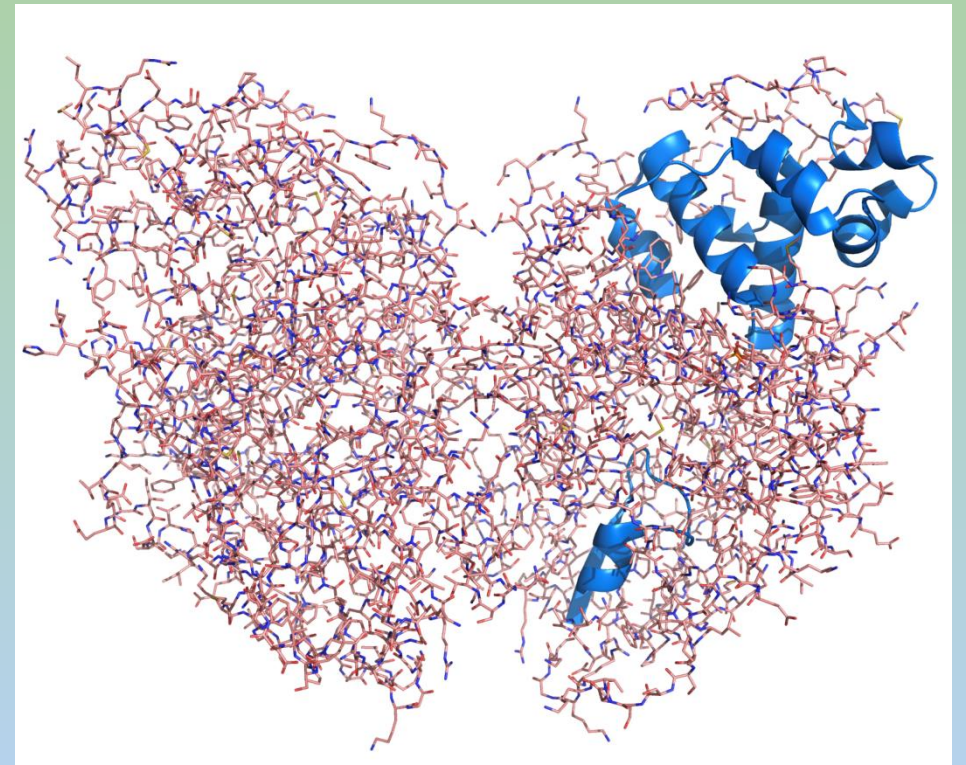
The conserved motif of CRY-1 in *Drosophila* and *Arabidopsis*



Drosophila



Arabidopsis



Arabidopsis

Conclusion two

- The conserved motif of 33 Crys sequences may have a common function by binding FAD.
- The specific motif of *Arabidopsis* may contribute to the functions that other species do not have.

致谢

- 罗静初老师、柯岚师姐
- 邓兴旺老师、范六民老师、朱丹萌老师、何航老师
- 郝柏林老师、龙漫远老师、顾红雅老师
- 樊德、韩雪、王瑞璞、王尧
- 所有参与ABC课程的同学