




**Bioinformatical Analysis of
SCR gene in Brassicaceae
species**

生物信息14小组

刘基生、刘东明、刘锦、张学铭

- 
- ❖ **Introduction**
 - ❖ **Methods**
 - ❖ **Discussion**
 - ❖ **Acknowledgement**

Introduction

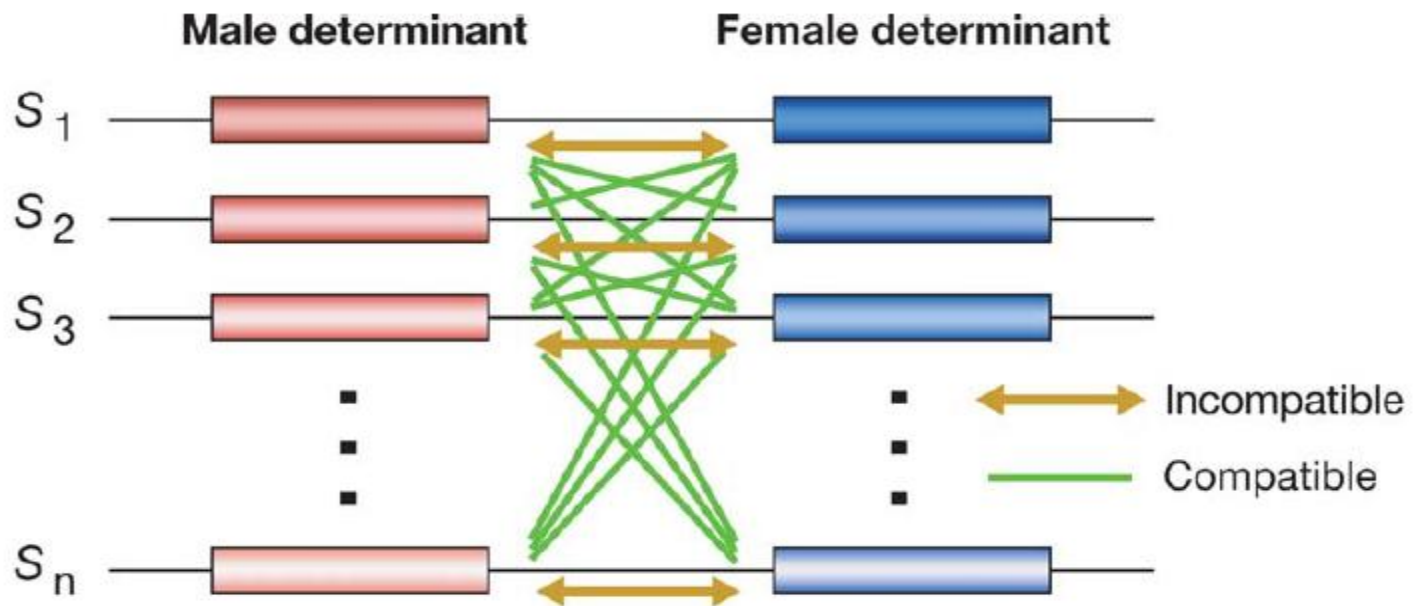
- ❖ **Self-incompatibility (SI) is an important mechanism used by many angiosperm species for preventing self-fertilization.**



- ❖ **SI is often controlled by a single polymorphic locus, S locus. In Brassicaceae, the S locus contains two specificity-determinant genes.**

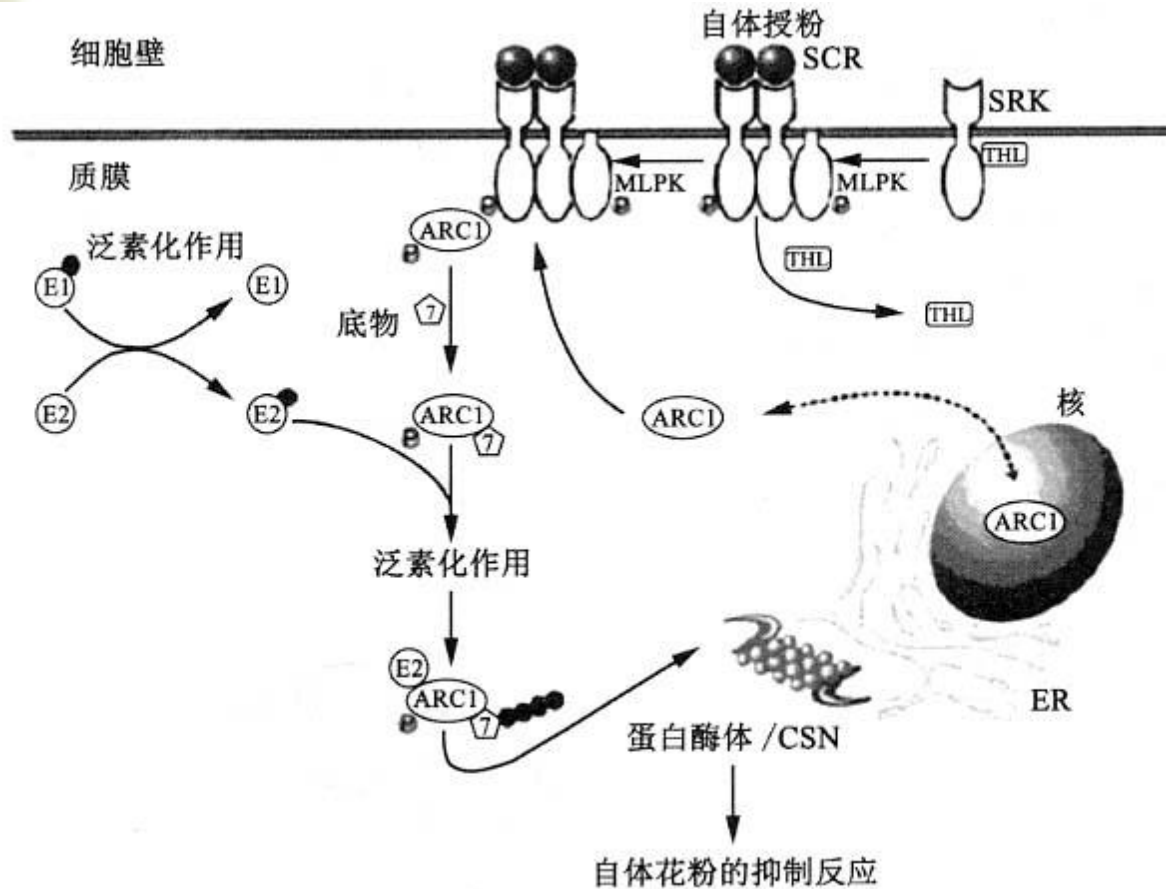
- ❖ ***SRK*** encodes the stigmatic receptor kinase.
- ❖ ***SCR*** encodes a small cys-rich protein localized in the pollen coat, which is the ligand for the ***SRK*** receptor (Nasrallah, 2002).
- ❖ Pollen inhibition occurs when the same **S**-locus specificity is expressed by both pollen and pistil (Nasrallah, 2002).

- ❖ **Classic genetic studies unraveled two distinct forms of SI, the gametophytic (GSI) and the sporophytic (SSI).**



Seiji Takayama et al ,2005

Signal transduction of Self-incompatibility reaction



A

```
ATGAGATGTGTTGTTTTGTTTTATGGTTTCTTGTCTTCTCATAGTTCCTCCTTATAAACCATTTTGAAGGT
M R C V V L F M V S C L L I V L L I N H F E

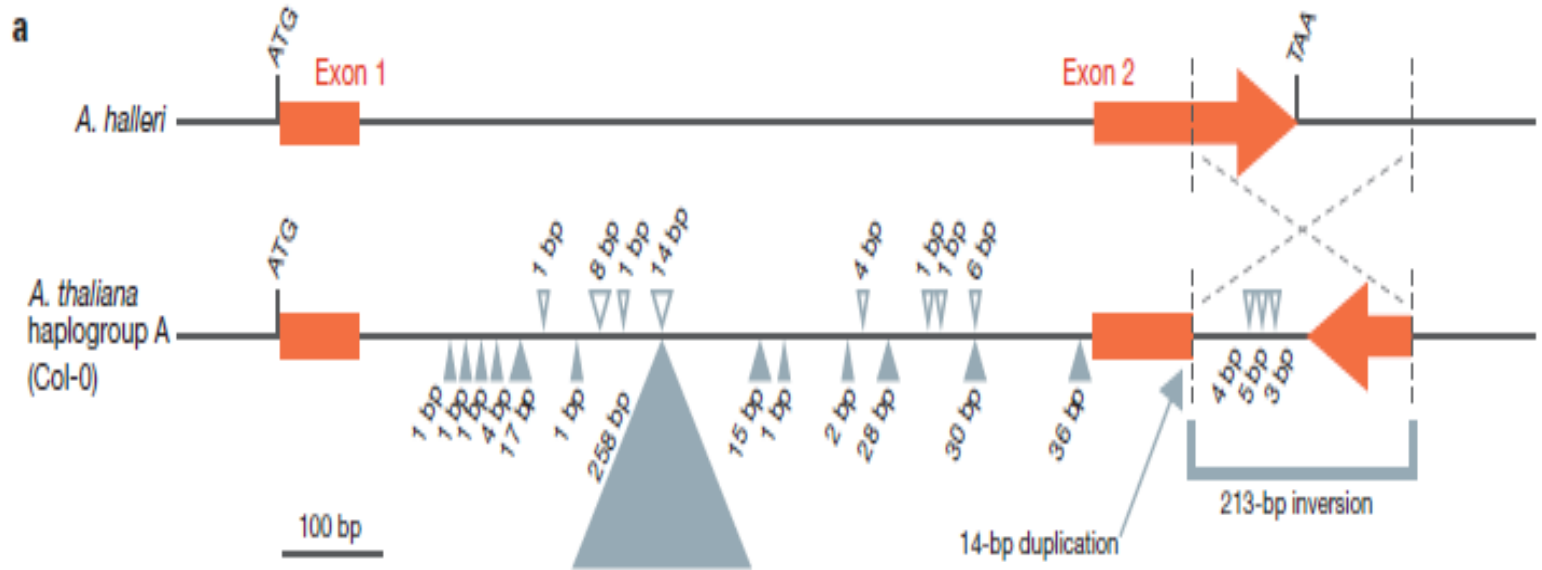
ATTTTACTATTTTCTTTTAACTACAGTGATTGTTATTGGTTCAAGAAATTGGTGTTTTCTACAATAACAT
GCAATATTCGAGTTTGAATTCATGTGGTCACAGAACAATTACAGTATACAAAGTACTAGTATGCATGGT
CTGTAAAAATAGTATTATGATCAAAGTAATTTCTCAATGATAATTGATGAGACATAAACATTATATTTT
TCAATTAGTTCATTAAAAATAAAAATTTCTAGCAAATGTTTGATTTTTTTATTTGTTTTATAAGATAATTTCT
GGACCTGGTAAAGAGGCATAATTTTAATTTCAAAGTGATTAATTGATCGAATTAATATTACACAACATA
AGATATTTAGAGATATCTACTATATATCTACTTAAGACGAGAATTATCATGATTAATATTAATACCTAT
CTTTAGCTATATCATAAGAATATTATAATATTCTAAATTAATAAAAAATTATTTTATTTCCCTTTTACAGA
GTTAATTAAGAACAATATACGTCACACCATTAGTAAATACTAAATAGTGAACAATATAATATATAAGTTG
ACCATCATAAATAGCTTCCTTTTATCATGTTCTTAAATGTGGAAGGTGTGGTTACTCTCTGTATATGTG
TATATAACCCTCCTTCAAAGTTATCTTCGAAATAATATATATGTGTATAAAACTCTCTAACTATTACTAA
TCGCAAAATGCAGTATTGCGACAAGAAAAAAATAGTTTATATTCTAAAAATTTACTGCATTATCAAGT
ATTTATCTAAGTATTCTAAAGTAATAACAGAAGTGGGAAGTTCAGAAGTGGGAACAAGTGCTTTCTTAGGG
      G V E A Q K W N K C F L R

ACATTTTCCCTGGGAAATGTGAACATGACGCAAACGCAAACACTACGATGCAAAGAAGACGATGCAAAGA
D I F P G K C E H D A N A K L R C K E D -----
AGACGATGCAAAGAAGACATTAGCGTAAAATACAGCTTTAAAATGAATATTTTGTGATAAATTTGTTTT
-----
TCTTCAATTAAATTTGTGCTTTAATTATCACTGTTTCATAGGTATATAGATTTTGGTTTTTGGAGTTGTTA
-----
AACCAAGCATTTTTACAATAGCAAATTCACCTTTATCAAAGTTTGACAATTGCATTCAAAGGGCG
V L C K K C Y C I G G K D F T Q C N C E F P R
AGAGGGTCTGAAATTCCTTAGCAATGTCTTCTTT
S P R F N K A I D E K
```

B

```
AlSCRa CNLEGNFPGRCVGNGDEE--QCKRDLTE--DGNNPSKCRCR-FRAGRRHCRCIYCEVFGM.
AlSCR37 CNLRDIFPGKCEHDANAKLRCKEDIAKNFRPSRPFECDCQTFDQGR-ICYCKCLV.
.....
AtSCR1 CFLRDIFPGKCEHDANAKLRCKED.VLCKKCYCIGGKDFTQCNCEFPRSPRFNKAI
```

Boggs et al,2009



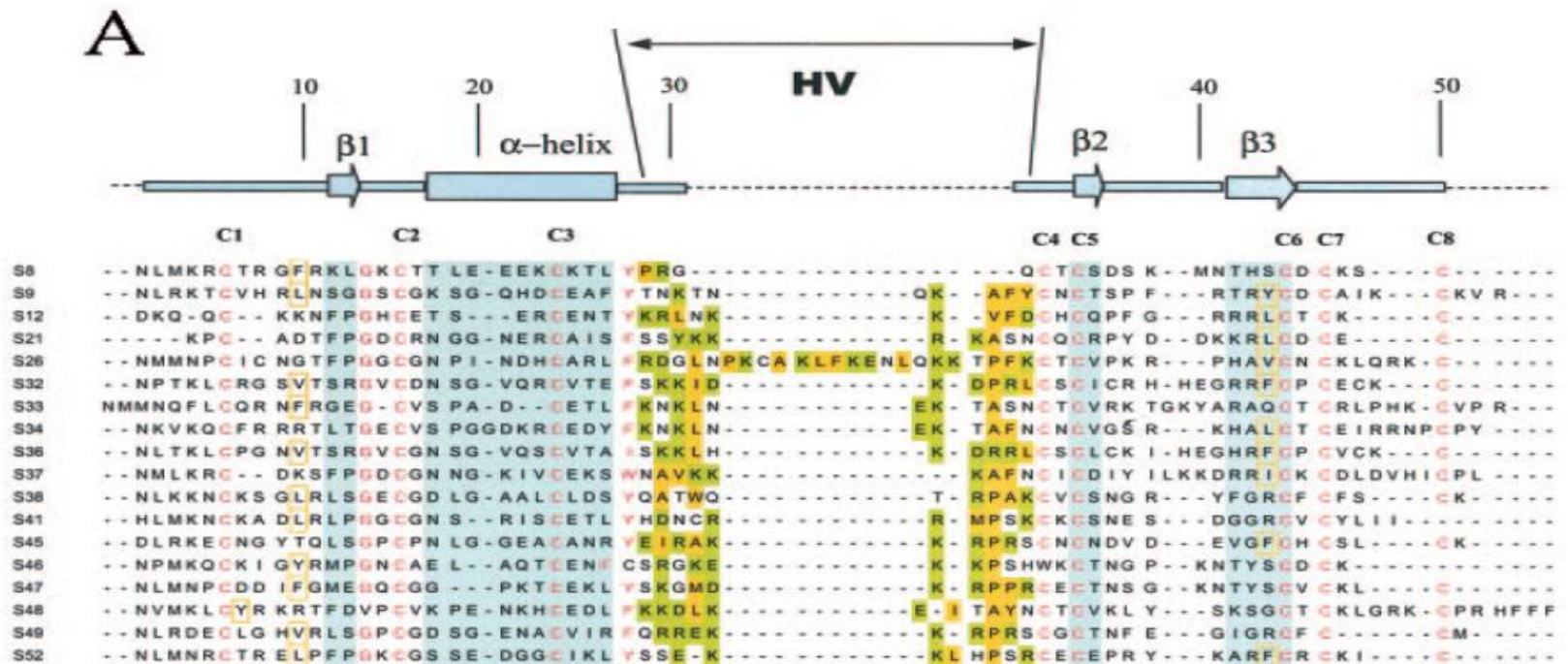
b

	10	20	30	40	50	60	70	80	90													
SCR-A- <i>Ahalleri</i>	MRCVVLFMV	SCLLIVLL	INHFEVEVAQK	WKE-CNLRD	I	PP-GKCE	BHDANAKLR	CKEDI	AKNFRPSR	PFECDCQ	TFDQG-R	ICYCKK	CLV*									
SCR-A- <i>Col-0</i>	MRCVVLFMV	SCLLIVLL	INHFEVEVAQK	WKNK-CPLRD	I	PP-GKCE	BHDANAKLR	CKEDDA	KKTLA*													
SCR-A- <i>Col-0-restored</i>	MRCVVLFMV	SCLLIVLL	INHFEVEVAQK	WKNK-CPLRD	I	PP-GKCE	BHDANAKLR	CKEDI	AKNFRPSR	PFECNCQ	TFDKG-G	ICYCKK	CLV*									
SCR-B- <i>Cvi-0</i>	MKCAVSP	FMVSCFL	IVFFTRHI	KELEA	QKWA-CLIK	QICP-G	SCRTDGY--	IRCKND	ITKNGKH-	RPFECK	CKD	VD-GDRL	CFYKCLVLR	RASSDLTT*								
SCRa- <i>Alyrata</i>	MRCVLFV	VVS	VIMSL	LISHV	QGMED	QKWKV	CNLEGN	FPP-G	RCV	GVNG	DEQ--	CKRDL	TEDGNN--	PSKCR	C-RFR	AGRRH	CR	CIY	CEV	FGM*		
SCRb- <i>Alyrata</i>	MRNATFF	FIVF	YVFI	SLVLS	SNVQ	VDVTA	QK-NK-C	MRSE	MPPT	GPCGN	NGRET--	CKKDF	KNI	YRT--	PIQ	CKCL	D	KYDF	ARL	CD	CR	FC*

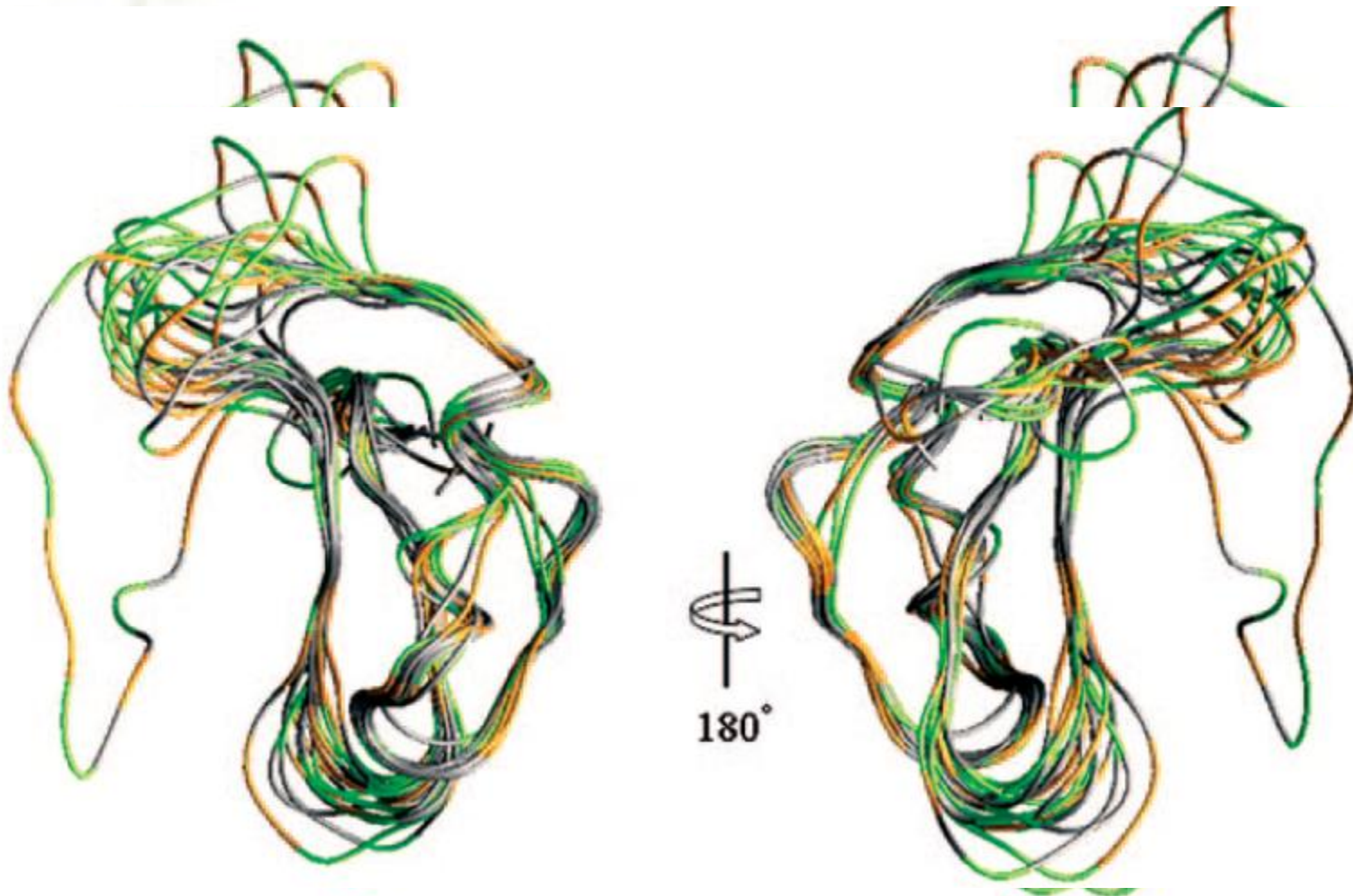
Takashi et al, 2010

Methods

Solution Structure of S₈-SP11



Mishima, et al, 2003



Mishima et al,2003

Signal Peptide Prediction

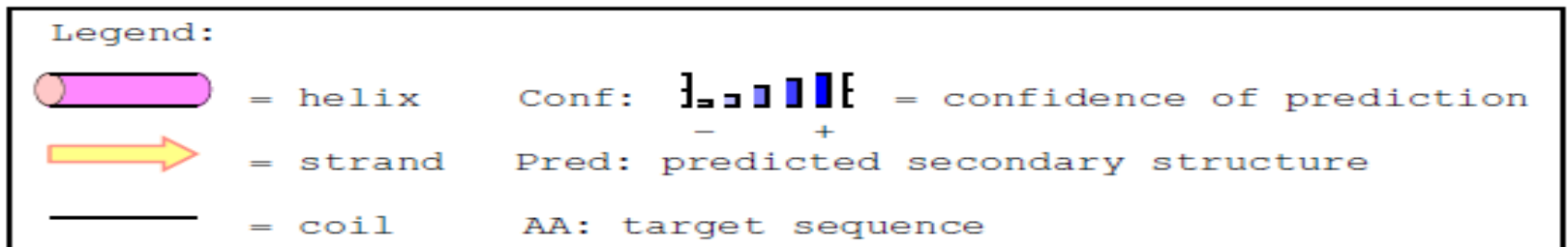
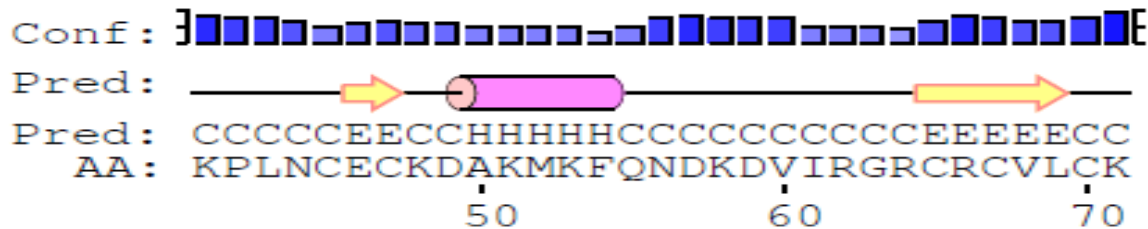
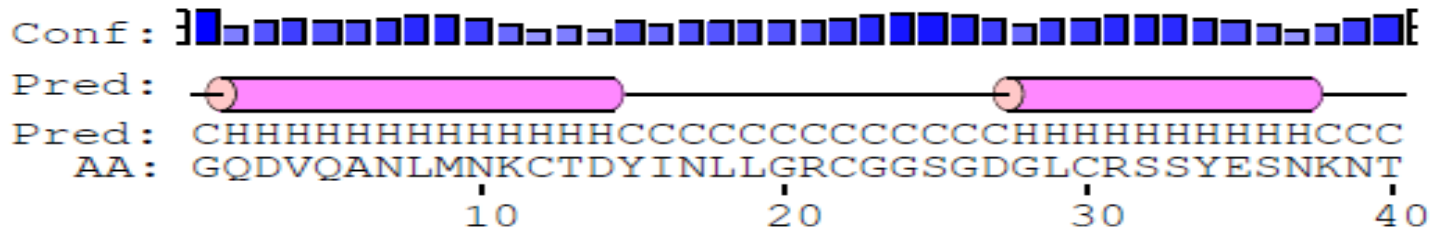
gene name	Cut Site		Cut Seq
BrSP11-8	24	25	LEA-NL
BrSP11-40	29	30	VQA-LD
BrSP11-29	29	30	VQA-LD
BoSP11-2	29	30	VQA-LD
BoSP11-5	29	30	VQA-LD
BoSP11-13	24	25	VEA-NL
BoSP11-8			
SCR-A-Col_0	26	27	VEA-QK
SCRa-Alyrata	23	24	VQG-ME
SCRa-Ahalleri	26	27	VEA-QK

A

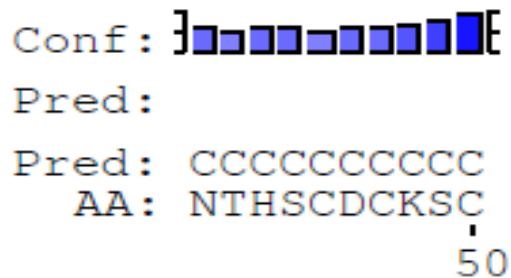
		C1	C2	C3	C4C5	C6C7	C8				
SP11-8	IFIV-SGHIQE---LEANLMKR--	CTRGFRKLGKCTTLEEEK		CKTLYPRGQ-----	CTCSDSKMNTHS--	CDCKS----	C				
SP11-9	IFIV-SSHVQE---VEANLRKT--	CVHRLNSGGSCGKSGQHD		CEAFYTNKTNQK-	AFYCNCTSPFRTRY--	CDCAIK---	CKVR				
SP11-12	IFIILS-RSQELTEVGADKQQ---	CKKN--	FPGHC---	ETSERCENTYKRLNKKVFD--	CHCQPFGRRL---	CTCK-----	C				
SP11-21	IFIISSHFQ----VEAKP-----	CADTF--	PGDCRNGGN-	ERCAISFSSYKKRK-	ASNQCQRPYDDKKRL--	CDCE-----	C				
SP11-32	IFVV--SIHVQG--VEANPTKL--	CRGSVTSRQVCDNSGVQ-	RCVTEF	SKKIDKDPRL-	CSCICRHHEGRRF-	CPCE-----	CKC				
SP11-33	IFVI-SSHFQE---VGANMNMQFLCQRNFRGEG-	CVSPAD--	CETLF	KNKLNEKTASNCT	CVRKTGKYARAQCT	CRLPHK-	CVPR				
SP11-34	LFII-SSHSQE---VEANKVKQ--	CFRRRTLGTGECVSPGGDKR	CEDYF	KNKLNEKTAFNCN	CVGSRKHAL---	CTCEIRRN	PCPY				
SP11-36	LFVV-SSHVQG---VEANLTKL--	CPGNVTSRQVCGNSGVQS-	CVTAISK	LHKDRRL-	CSCLC	CKIHEGHRF-	CPCV-----	CKC			
SP11-37	IIIIAS-HFQ-----EANMLKR--	CDKSF--	PGDCGNGNKIV-	CEKSN	NAVKKK--	AFNCICDIYILKKDRRI	CKCDLDVHICPL				
SP11-38	IFIV-SSHGQE---VEANLKKN--	CKSGLRLSGECGDLGAAL-	CLDSY	QATWQTRPAK-	CVC	SNGRYFGR---	CF	CFS----	CK		
SP11-41	IFIV-SSHGQE---LEAHLMKN--	CKADLRLPGGCGNSRIS-	CETLY	HDNCR	RRMPSK-	CKCSNESD	GGR---	CV	CYLII		
SP11-45	MFII-SSHGQG---VEADLRKE--	CNGYTQLSGPCPNLGGEA-	CANRY	EIRAKKR	PRS-	CNCNDV	DEVGf---	CH	CSL----	CK	
SP11-46	IFIV-SGHIQE---VEANPMKQ--	CKIGYRMPGNCAELAQT-	CENF	CSR	GKEKKPSH-	WKCTNG	PKNTYS--	CD	CK		
SP11-47	IFIV-SGHIQE---VEANLMNP--	CDDIFGMEGQCGGPKT---	CEKLY	SKGMD	KRPPR-	CECTNS	GKNTYS--	CV	CKL----	C	
SP11-48	IFII-SSHFQE---VEANVMKL--	CYRKRTFDVPCVKPENKH-	CEDLF	KKDLKEITAYN	CTCVKLY	SKSG---	CT	CKLGRK-	CPRHFF		
SP11-49	IFIV-SSHCQG---VEANLRDE--	CLGHVRLSGPCGDSGENA-	CVIR	FQRRE	KKRPRS-	CGCTN	FEIGR---	CF	C-----	CM	
SP11-52	IFIV-SSHAQD---VEANLMNR--	CTRELPPFGKCGSSEDGG-	CIKLY	SSE	KKLHPSR-	CECE	PRYKARF---	CR	CKI----	C	
SCR6	IFLV-SSHGQE---VEANLKKN--	CVGKTRLPGPCGDSGASS-	CRDLY	NQTE	KTMPVS-	CRC---	VPTGR---	CF	CSL----	CK	
SCR13	IFIV-SGHIQE---VEANLMMP--	C--	GSFMFGNCRNIGARE-	CEKL-	NSPG	KRKPSH-	CKCTDT	QMGTYs--	CD	CKL----	C

Watanabe et al,2000


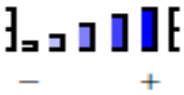
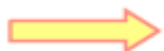

Secondary Structure prediction



BoSP11-8 secondary struction

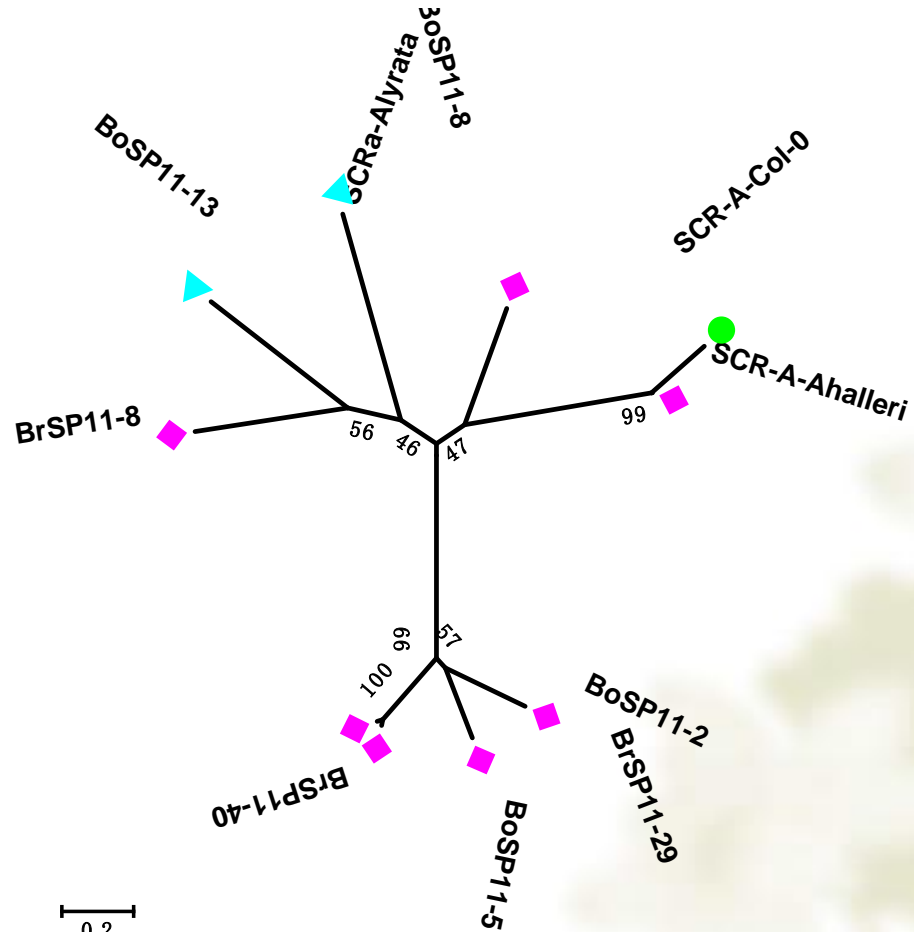


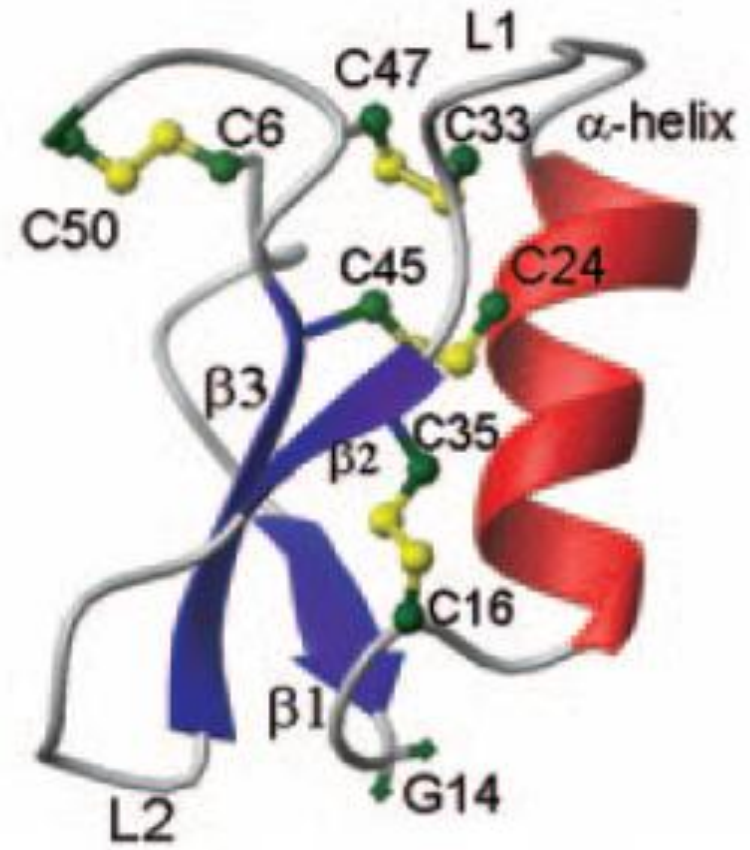
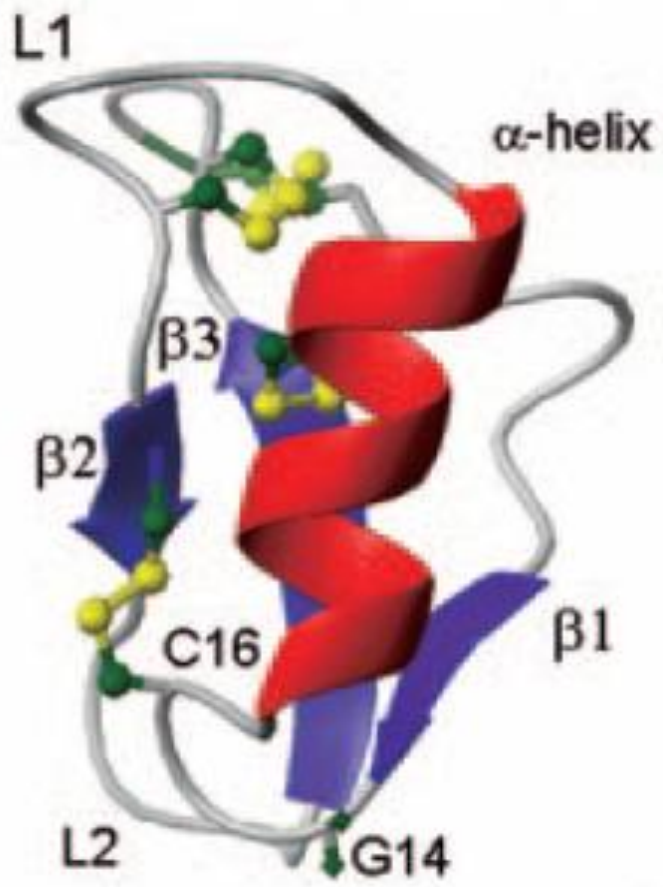
Legend:

	= helix	Conf: 	= confidence of prediction
	= strand	Pred:	predicted secondary structure
	= coil	AA:	target sequence

BoSP11-8 secondary structure

Phylogenetic tree of SP11s





\curvearrowright 90°

Mishima et al,2003

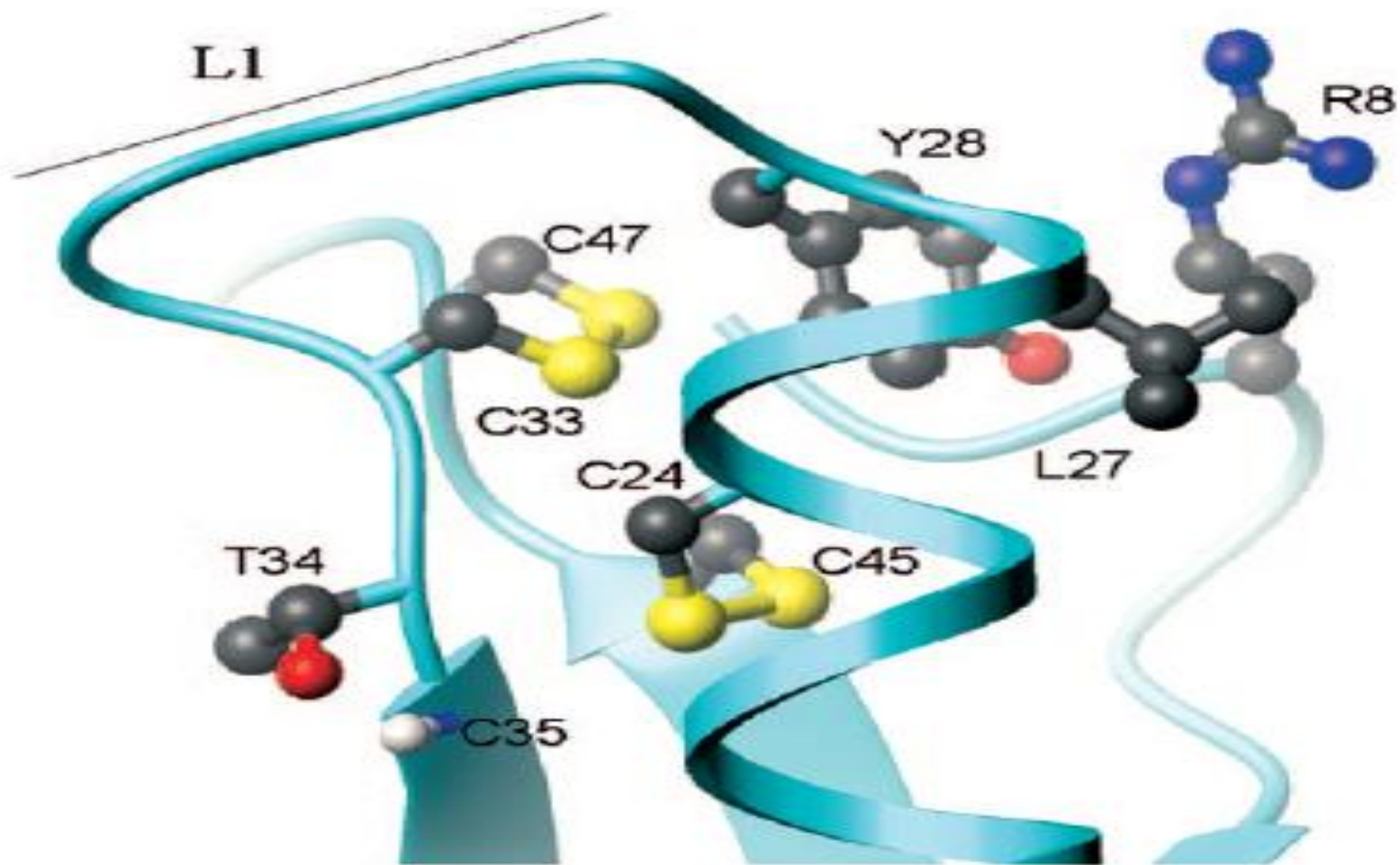


FIG. 3. Protein-core forming residues and loop L1. The *ribbon representation* and the *ball-and-stick models* of the side chain around the protein core. The *ball-and-stick model* of the backbone amide group of Cys-35 is also depicted. The L1 loop, referred to as HV, is guided by the *solid line*.

Discussion

- ❖ Here, we align different SP11 genes in *Brassica rapa*, *Brassica oleracea*, *Arabidopsis thaliana* and *Arabidopsis lyrata*, with a finding that 8 conserved cysteine residues exist in every SP11. But no hyper-variable region (HV) is found.
- ❖ As a result of too low identity between SP11-8 and other SP11s, we can't do homology modeling. But based on 8 conserved cysteine, we presume that other SP11s share the similar structure.

Acknowledgement

感谢罗老师孜孜不倦的教诲！

感谢生物信息14小组的各位组员！

感谢生物信息课程寒假班的所有同学！