



牡丹SPL基因家族的分子特征及表达分析
Molecular characterization and expression analysis
of the *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE* gene
family in *Paeonia suffruticosa*

汇报人: 崔宇通

中国农业科学研究院蔬菜花卉研究所

G6: 崔宇通 徐生阳 梁智玲 王奕婷

目录

CONTENTS

01

研究背景

02

*PsSPL*基因的鉴定

03

*PsSPL*基因保守结构域及系统进化分析

04

蛋白序列motif预测

05

*PsSPL*基因的表达分析

研究背景



四大名品



姚黄



魏紫



豆绿



赵粉

九大色系



➤ 白色



➤ 绿色



➤ 黑色



➤ 粉色



➤ 黄色



➤ 蓝色



➤ 红色



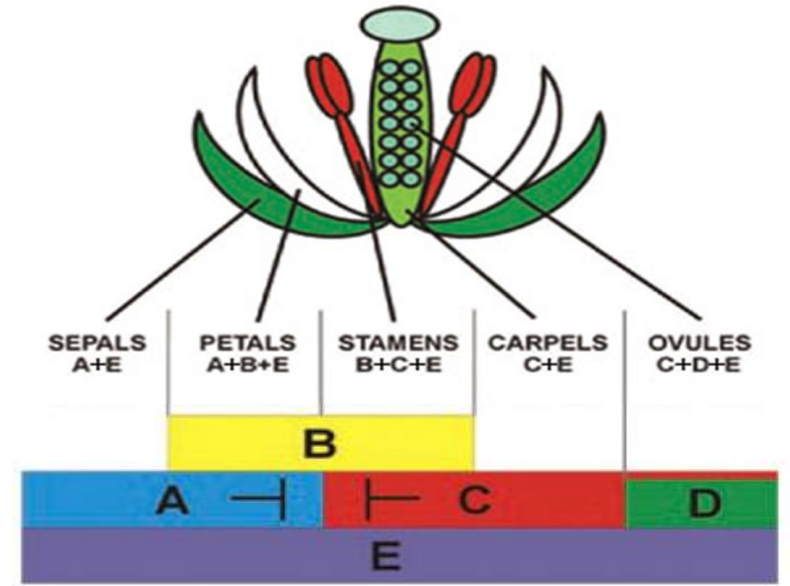
➤ 紫色



➤ 复色

花器官发育

1 单瓣型	2 荷花型	3 菊花型	4 蔷薇型
5 金蕊型	6 托挂型	7 金环型	8 皇冠型
9 绣球型		10 台阁型	



从外向内:

第1轮生组织: 萼片 (sepals) A

第2轮生组织: 花瓣 (petals) A+B

第3轮生组织: 雄蕊 (stamen) B+C

第4轮生组织: 心皮 (carpel) C

D的功能是促进胚珠形成。

E基因与其他基因结合以促进花的形成。

PsSPL基因的鉴定

1. 已知物种基因家族成员序列获取

拟南芥数据库TAIR (<https://www.arabidopsis.org/>)

The screenshot shows the TAIR website interface. The 'Gene Families' section is highlighted with a red circle. Below it, a table lists various gene families with columns for 'Gene Family Name', 'Family Count/ Gene Count', and 'Submitter'. Red handwritten text '家族信息' is above the 'Family Count/ Gene Count' column, and '上传者' is above the 'Submitter' column. The 'Gene Family Name' column is also circled in red.

Gene Family Name	Family Count/ Gene Count	Submitter
14-3-3 family	1 family 13 members	R Ferl Laboratory
ABC Superfamily	Subfamilies 8 Members 136	Paul Verrier Freddie Theodoulou Angus Murphy
ABI3VP1 Transcription Factor Family	1 family 11 members	AGRIS
AGC Kinase Gene Family	1 family 39 members	Laszlo Bogre Laszlo Okresz
Aldehyde Dehydrogenase Gene Family	9 families 14 members	Dorothea Bartels Hans-Hubert Kirch
Amino Acid/Auxin Permease AAP Family	1 family 43 members	John Ward

Arabidopsis SBP-box Gene Family

Source: Cardon G, Hohmann S, Klein J, Nettesheim K, Saedler H, Huijser P. (1999) Molecular characterisation of the Arabidopsis SBP-box genes Gene 237:91-104
Functional analysis of the SBP-domain protein family in Arabidopsis thaliana

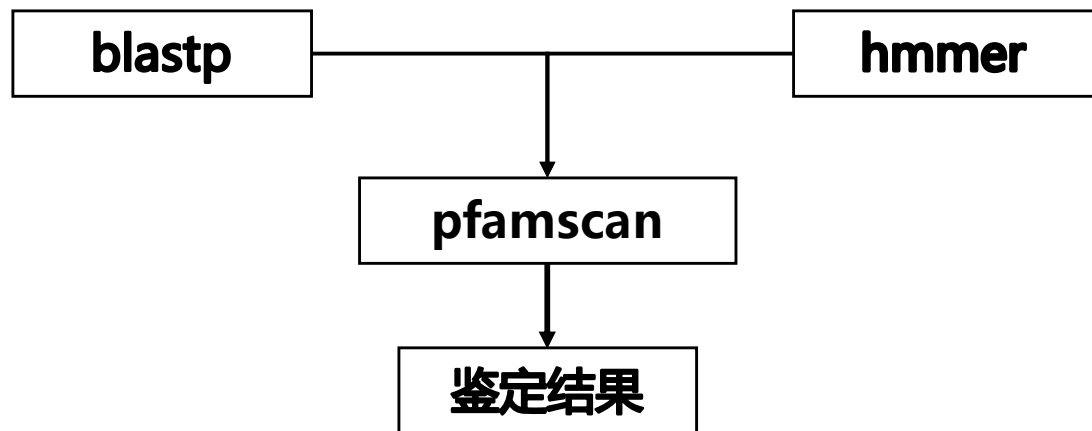
Gene Family Criteria: Defined by the presence of the DNA-binding SBP domain (Pfam profile PF03110)

Contact: Peter Huijser

Gene Family Name:	Protein Name: Alternate Name:	BAC Locus: Genomic Locus:	Accession:	TIGR Annotation Comment:
SBP-box family of putative plant transcription factors	AtSPL1	F14M4_9 At2g47070	NP_850468	squamosa promoter-binding protein-like 1 (SPL1)
	AtSPL2	MNL12_9 At5g43270	NP_851122	squamosa promoter binding protein-like 2 (SPL2)
	AtSPL3	T1B8_11 At2g33810	NP_565771	squamosa promoter binding protein-like 3 (SPL3)
	AtSPL4	F12M16_2 At1g53160	NP_175723	squamosa promoter-binding protein-like 4 (SPL4)
	AtSPL5	K7L4_7 At3g15270	NP_188145	squamosa promoter binding protein-like 5 (SPL5)
	AtSPL6	F23O10_24 At1g69170	NP_177077	squamosa promoter binding protein-like 6 (SPL6)
	AtSPL7	F17K4_80 At5g18830	NP_850850	squamosa promoter binding protein-like 7 (SPL7)
	AtSPL8	T7I23_32 At1g02065	NP_683267	squamosa promoter-binding protein-like 8 (SPL8)
	AtSPL9	T24P15_11 At2g42200	NP_850365	squamosa promoter binding protein-like 9 (SPL9)
	AtSPL10	F17L21_15 At1g27370	NP_174057	squamosa promoter binding protein-like 10 (SPL10)
	AtSPL11	F17L21_14 At1g27360	NP_564280	squamosa promoter binding protein-like 11 (SPL11)
	AtSPL12	T2O9_10 At3g60030	NP_191562	squamosa promoter binding protein-like 12 (SPL12)
	AtSPL13A	MBA10_13 At5g50570	NP_851161	"squamosa promoter-binding protein, putative"
	AtSPL13B	MFB16_6 At5g50670	NP_851161	"squamosa promoter-binding protein, putative"
	AtSPL14 SPL1-Related2	F9H16_3 At1g20980	NP_173522	SPL1-Related2 protein (SPL1R2)
	AtSPL15	T10K17_130 At3g57920	NP_191351	"squamosa promoter-binding protein, putative"
AtSPL16 SPL1-Related3	F14G6_18 At1g76580	NP_177784	SPL1-Related3 protein (SPL1R3)	

2. *PsSPL*基因家族成员的鉴定

方法: blast+hmmer+pfamscan



BLAST: 根据已知序列寻找与该序列同源性较高的序列。

Hmmer: 根据蛋白保守结构域来寻找与该结构域相似的序列，也可以对蛋白质序列进行功能注释。

PfamScan: Pfam是蛋白质家族的集合，每个蛋白家族由多序列比对和Hmm模型的形式表示。

3. 家族成员基因结构信息和理化性质信息统计

基因结构信息统计

根据基因结构注释文件（GFF3）统计基因在染色体上的位置、转录本、外显子、内含子数目。

理化性质信息统计：

网址：<https://web.expasy.org/protparam/>

ProtParam tool

ProtParam (References / Documentation) is a tool which allows the computation of various physical and chemical parameters for a given protein sequence. The computed parameters include the molecular weight, theoretical pI, amino acid composition, atomic composition, instability index, aliphatic index and grand average of hydropathicity (GRAVY) (Disclaimer).

Please note that you may only fill out **one** of the following fields at a time.

Enter a Swiss-Prot/TrEMBL accession number (AC) (for example **P05130**) or a sequence identifier (ID) (for example **KPC1_DROME**):

Or you can paste your own amino acid sequence (in one-letter code) in the box below:

粘贴序列

RESET Compute parameters

ProtParam

User-provided sequence:

```
MSQEEQPKRP QEPPTVGVDF EYSGELAKP IAFEDAMNG AAETRVFGHT QKGGAAAYRQ
SAATANERG FVHPGDTIDL AERGVYVAQ IDVPGARVTT EFGGQVPGQ YVEFPVYATA
AAMEAEVWGL SLQSAITIGE ALEATVQTAG HRFYDQSDAA AIQAAEVRAC GTHVIAPGGI
AASASAAHH NATIDRDENK IKLIDVLAGA TSKLAADKAY TRQDABGVYS AELRNHPNLS
THPGGYAASI TAAARLNERA DI
```

References and documentation are available.

Number of amino acids: 262
Molecular weight: 26743.70
Theoretical pI: 4.73

Amino acid composition: CSV format

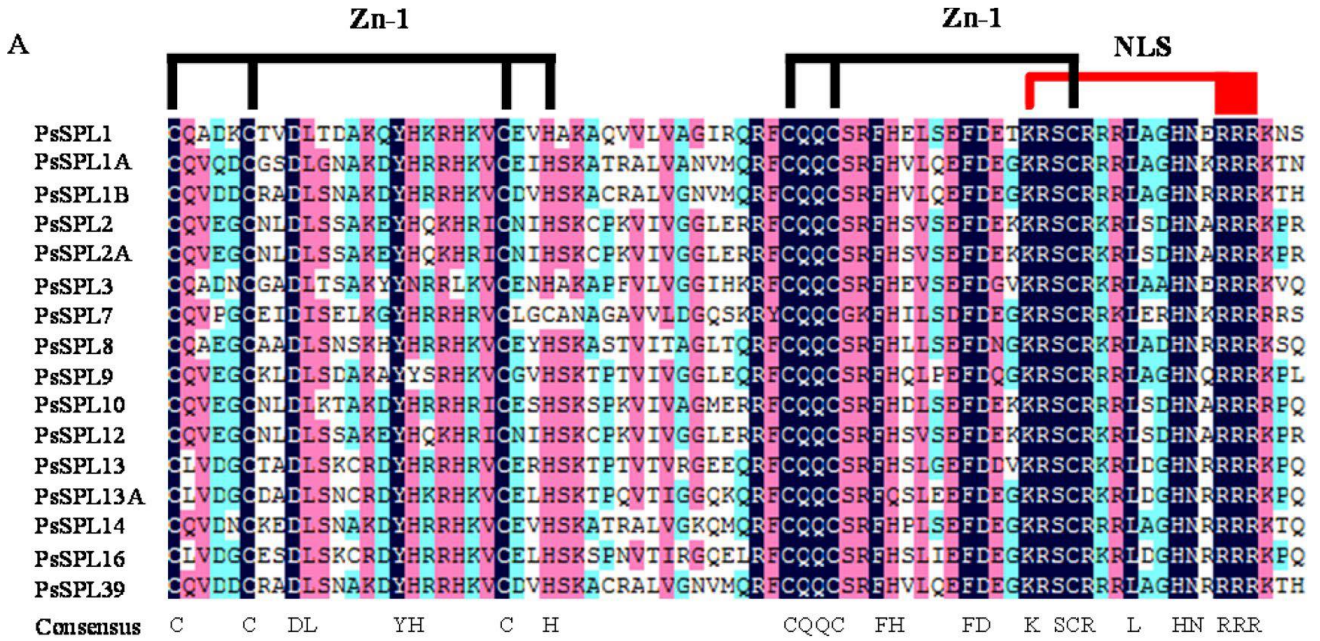
Ala (A)	60	19.1%
Arg (R)	12	4.6%

Table 1 Gene information of *SPLs* genes in *Paeonia suffruticosa*

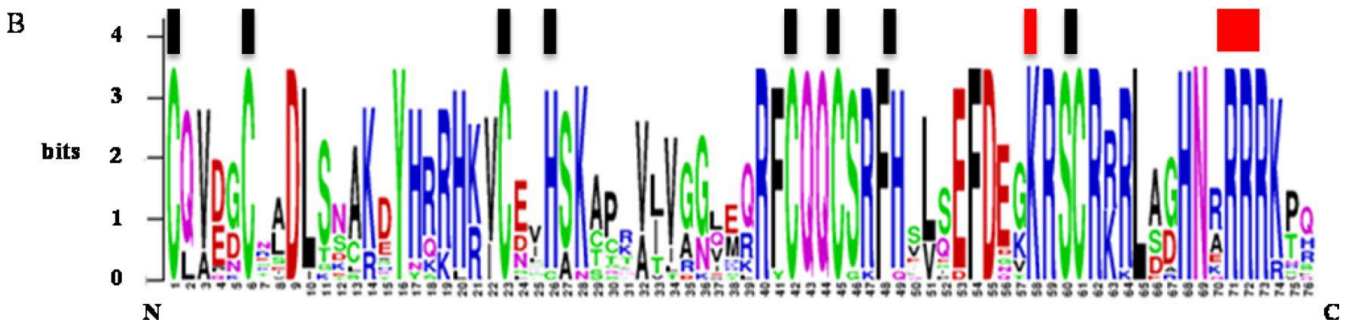
Gene name	Accession No	Gene length (bp)	CDS (bp)	No. of amino acid (aa)	Mw (Da)	<i>PI</i>	Upstream (bp)	Down-stream (bp)	Group	miR156 target site
<i>SPL1</i>	MT239458	1020	549	182	20,363.42	9.64	125	345	4	TTGCTCTCTCTTCTGTAC*
<i>SPL1A</i>	MT239459	3731	2991	996	110,753.06	6.28	432	307	3	
<i>SPL1B</i>	MT239460	4424	3042	1013	112,994.24	6.62	748	633	3	
<i>SPL2</i>	MT239461	2084	1398	465	51,332.33	8.91	376	309	6	ATGCTCCCTCTCTTCTGTCAA
<i>SPL2A</i>	MT239462	2030	1440	479	52,967.25	8.82	577	14	6	ATGCTCCCTCTCTTCTGTCAA
<i>SPL3</i>	MT239463	1467	1224	407	46,855.12	4.99	122	120	4	
<i>SPL7</i>	MT239464	2642	2324	773	86,477.83	6.53	113	206	1	
<i>SPL8</i>	MT239465	1218	939	312	34,834.42	8.55	178	100	2	
<i>SPL9</i>	MT239466	1636	1125	374	40,158.71	9.24	275	235	6	GTGCTCTCTCTTCTGTCAA
<i>SPL10</i>	MT239467	1745	1398	465	51,630.54	8.43	89	257	6	GTGCTCTCTCTTCTGTCAA
<i>SPL12</i>	MT239468	1815	1398	465	51,403.37	8.91	252	166	6	ATGCTCCCTCTCTTCTGTCAA
<i>SPL13</i>	MT239469	2123	1086	361	40,025.22	9.38	833	203	5	GTGCTCTCTCTTCTGTCAA
<i>SPL13A</i>	MT239470	1386	957	318	35,455.53	8.94	325	103	5	GTGCTCTCTCTTCTGTCA
<i>SPL14</i>	MT239471	4112	3267	1088	119,898.06	8.78	232	612	3	
<i>SPL16</i>	MT239472	1789	963	320	35,647.65	8.32	558	267	5	GTGCTCTCTCTTCTGTCA
<i>SPL39</i>	MT239473	3366	2700	899	100,528.64	7.81	437	228	3	

Asterisks represents miR156 target site was distributed at 3'-UTR region of *PsSPL1*, and the other target sites were all distributed at CDS region of *PsSPLs*. Red nucleotide represents unmatched nucleotide with nucleotide of miR156

PsSPL 基因保守结构域及系统进化分析



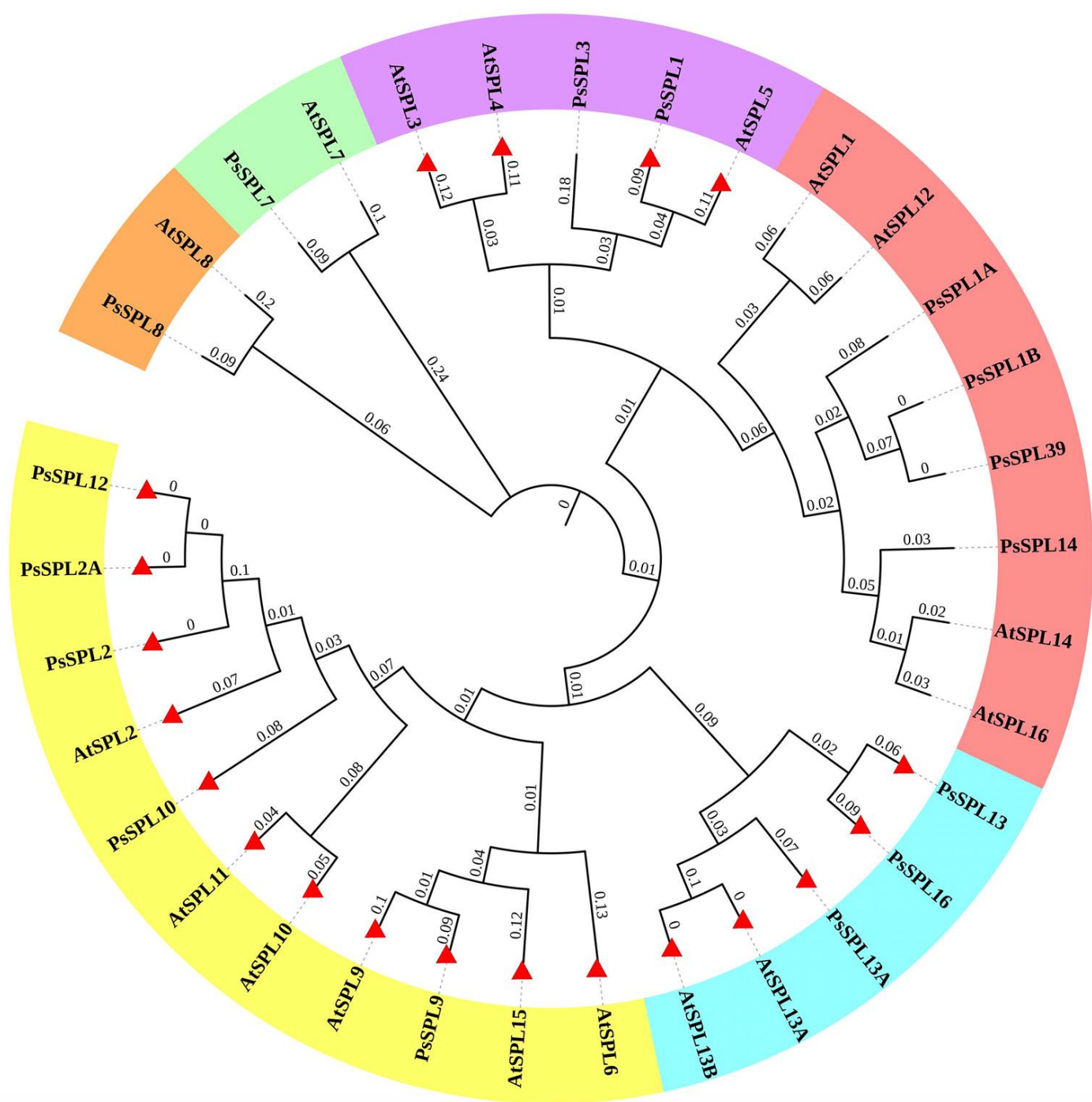
用DNAMAN软件进行多序列比对，获得*PsSPL*基因家族成员的保守结构域的：2个保守的锌指结构(Zn-1, Zn-2)和核苷酸定位信号(NLS)。



*PsSPL*基因家族成员的序列保守性 logo图。每个堆叠的总高度代表该位置的保守程度，而每个堆叠中字母的高度表示相应氨基酸的相对频率。

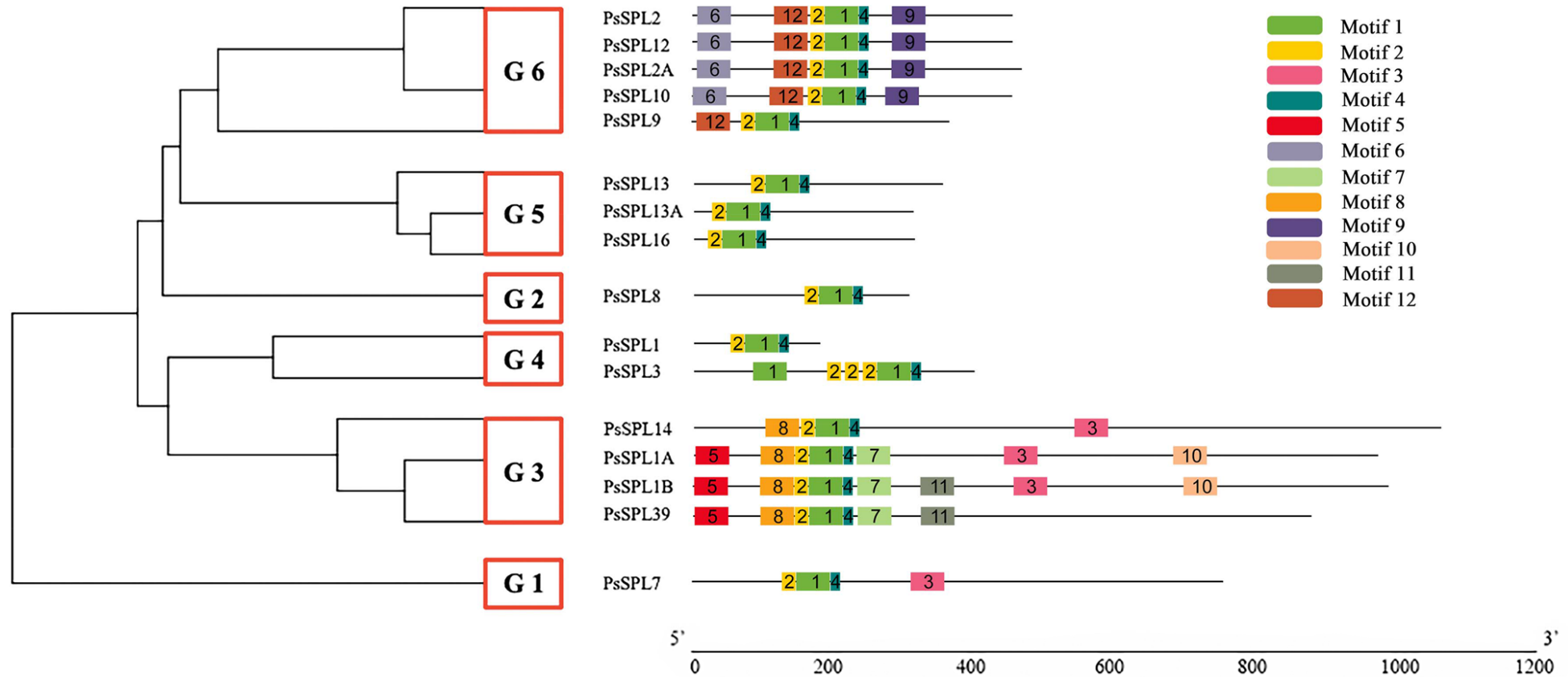
Groups

- Group 1
- Group 2
- Group 3
- Group 4
- Group 5
- Group 6



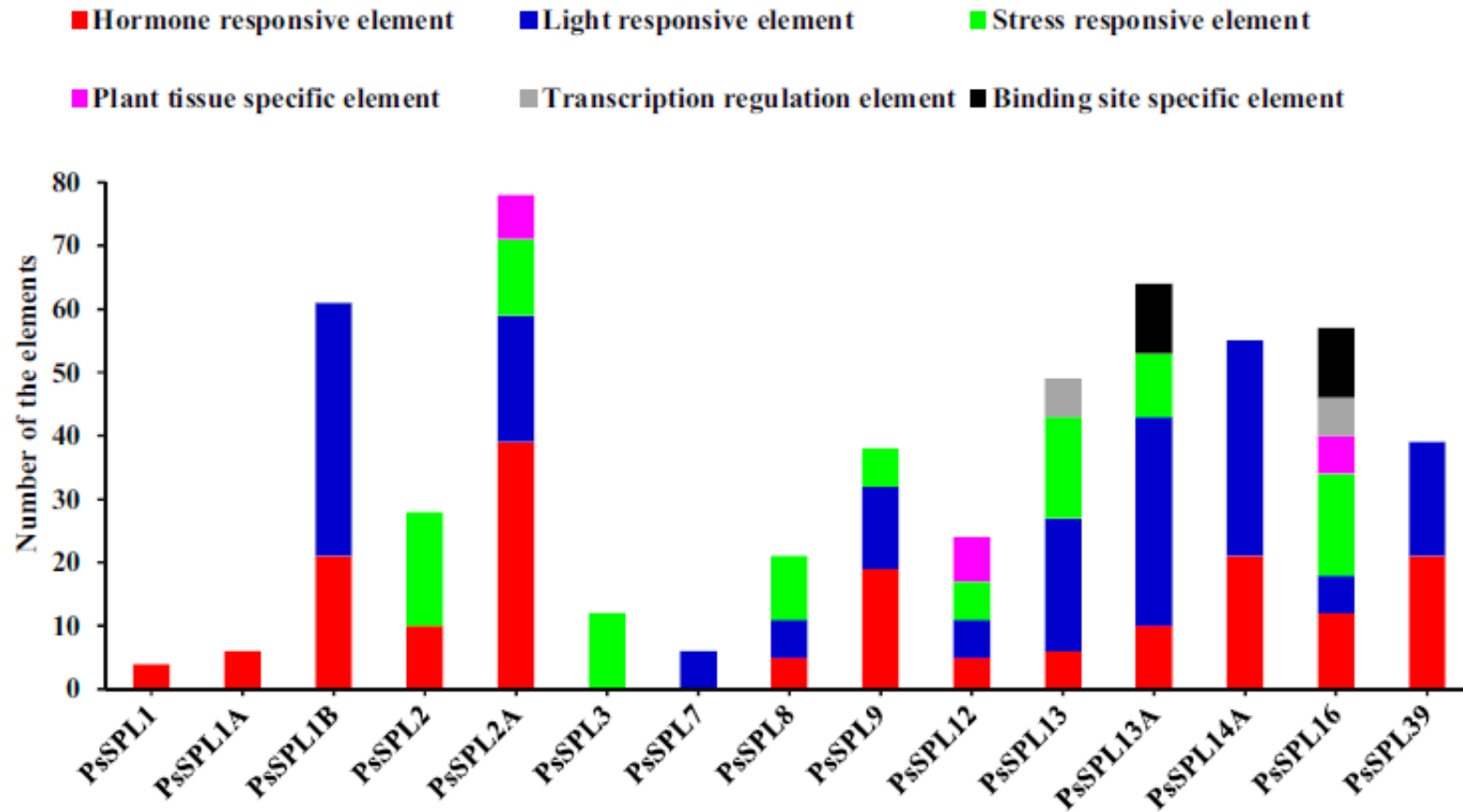
蛋白序列motif预测

MEME (<https://meme.sdsc.edu/meme/>)

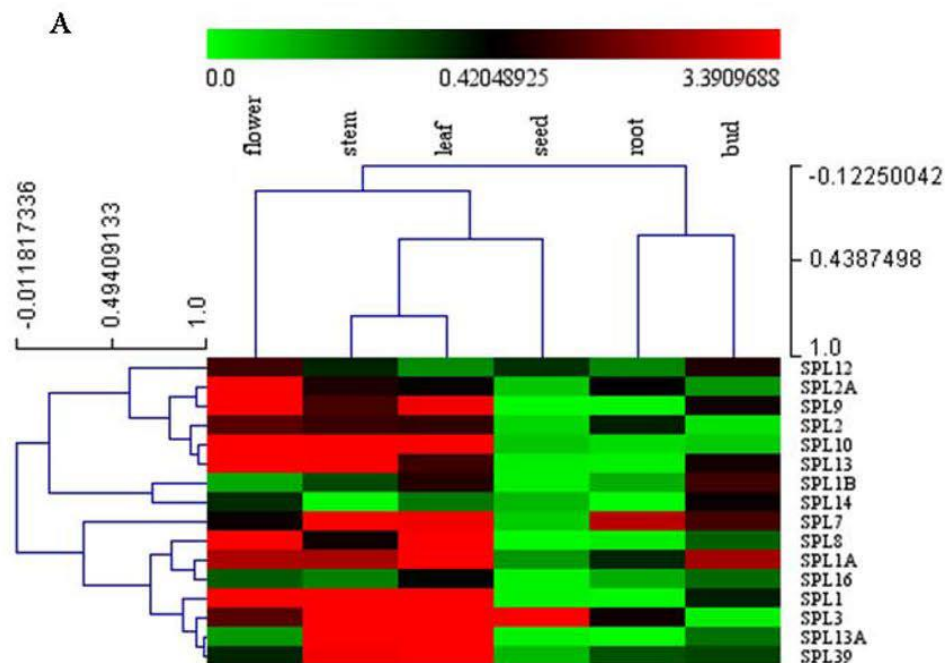
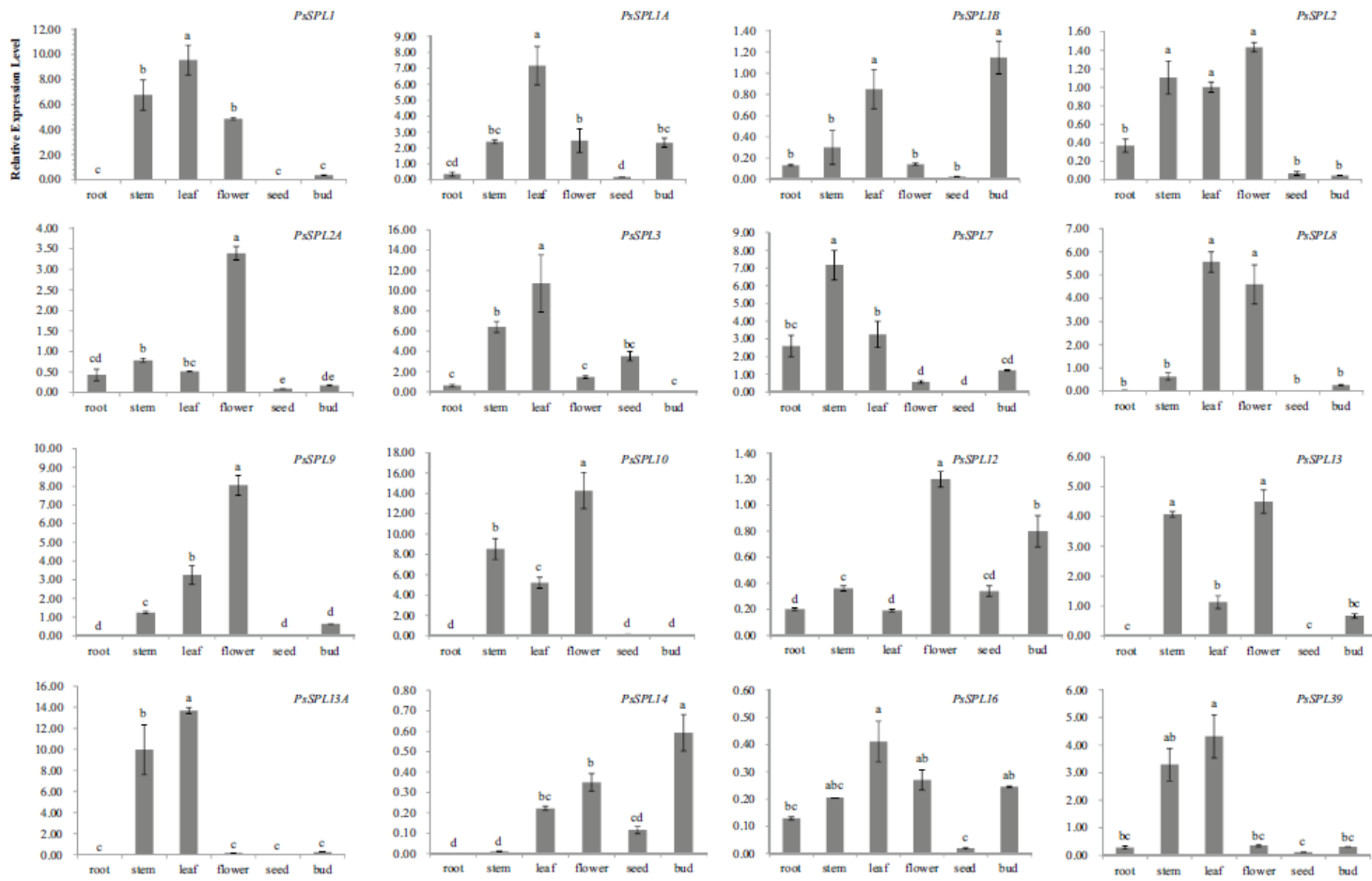


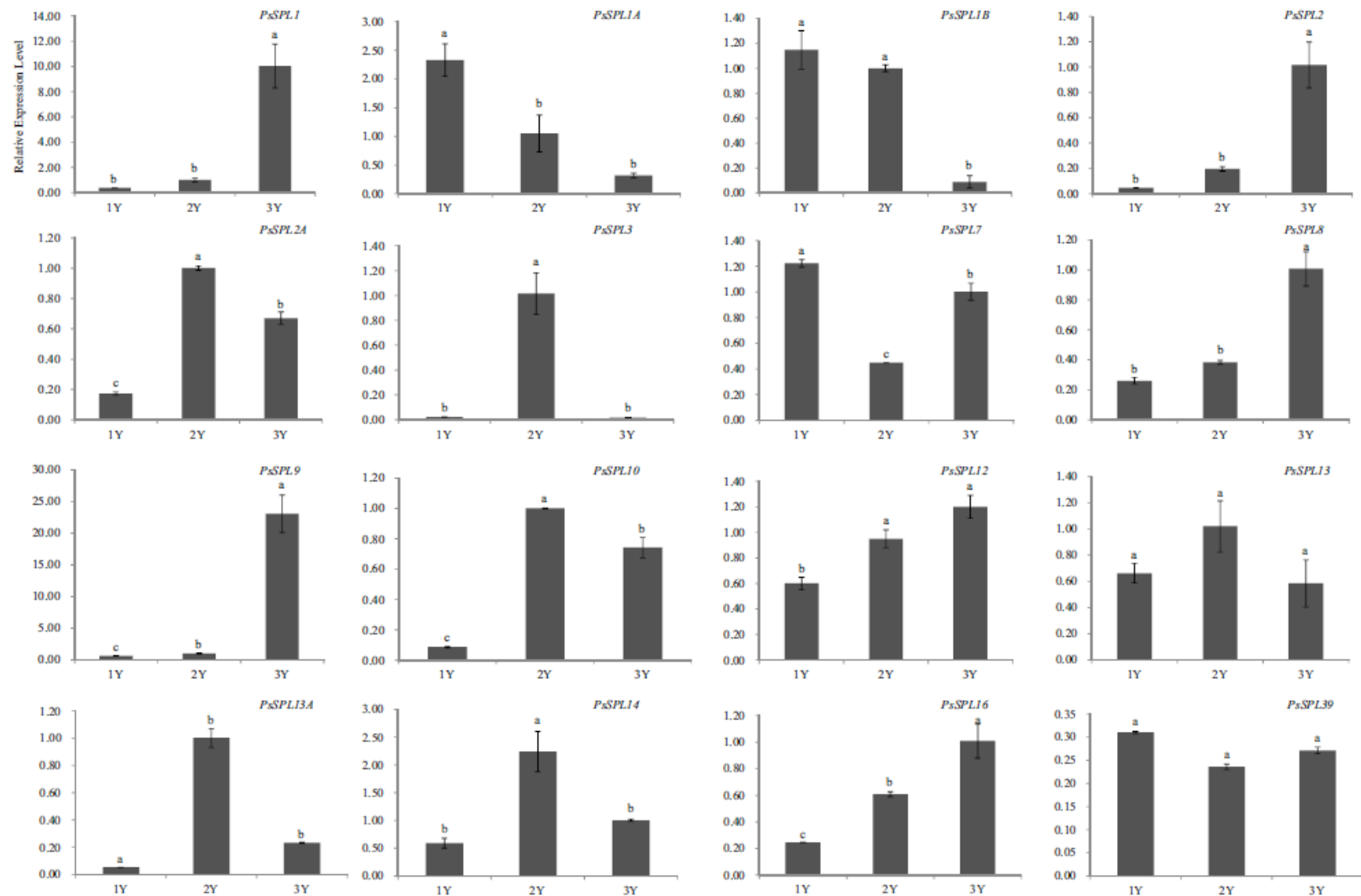
PsSPL基因启动子区域的顺式元件

PlantCARE database (<https://bioinformatics.psb.ugent.be/webtools/plantcare/html/>).

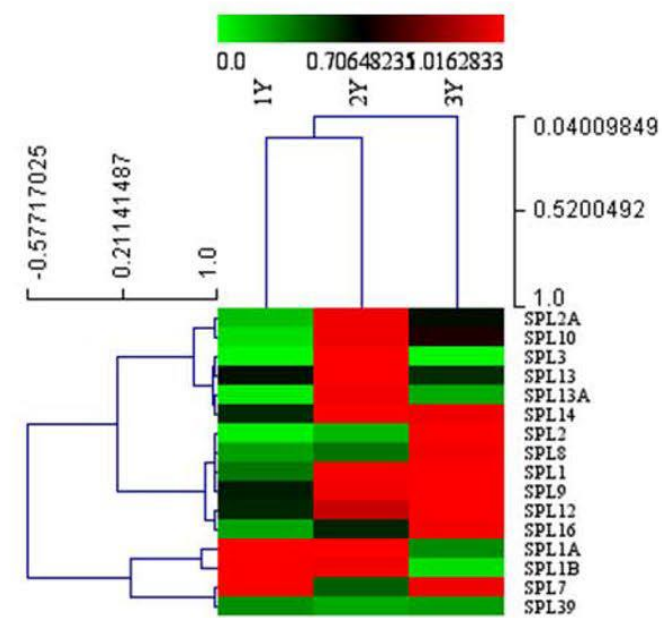


PsSPL基因的表达分析

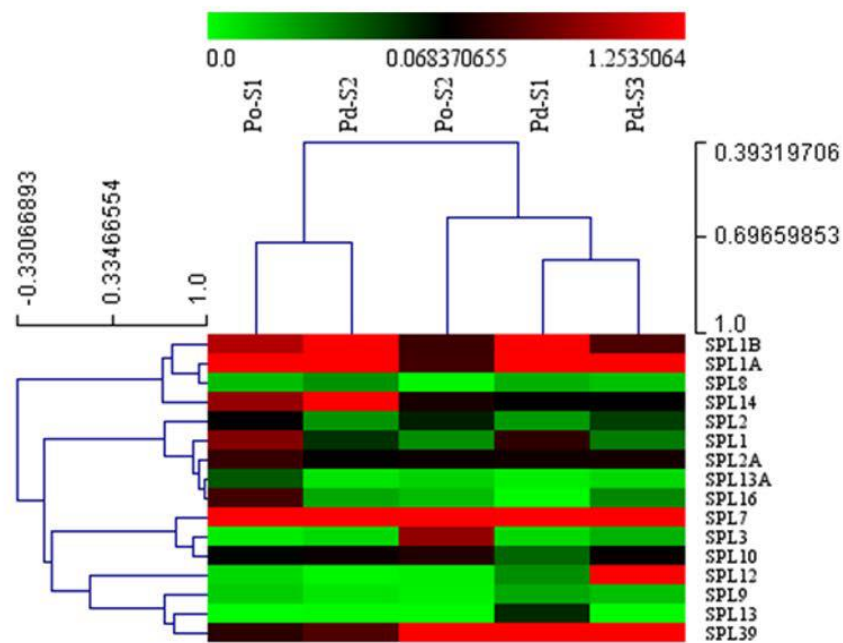
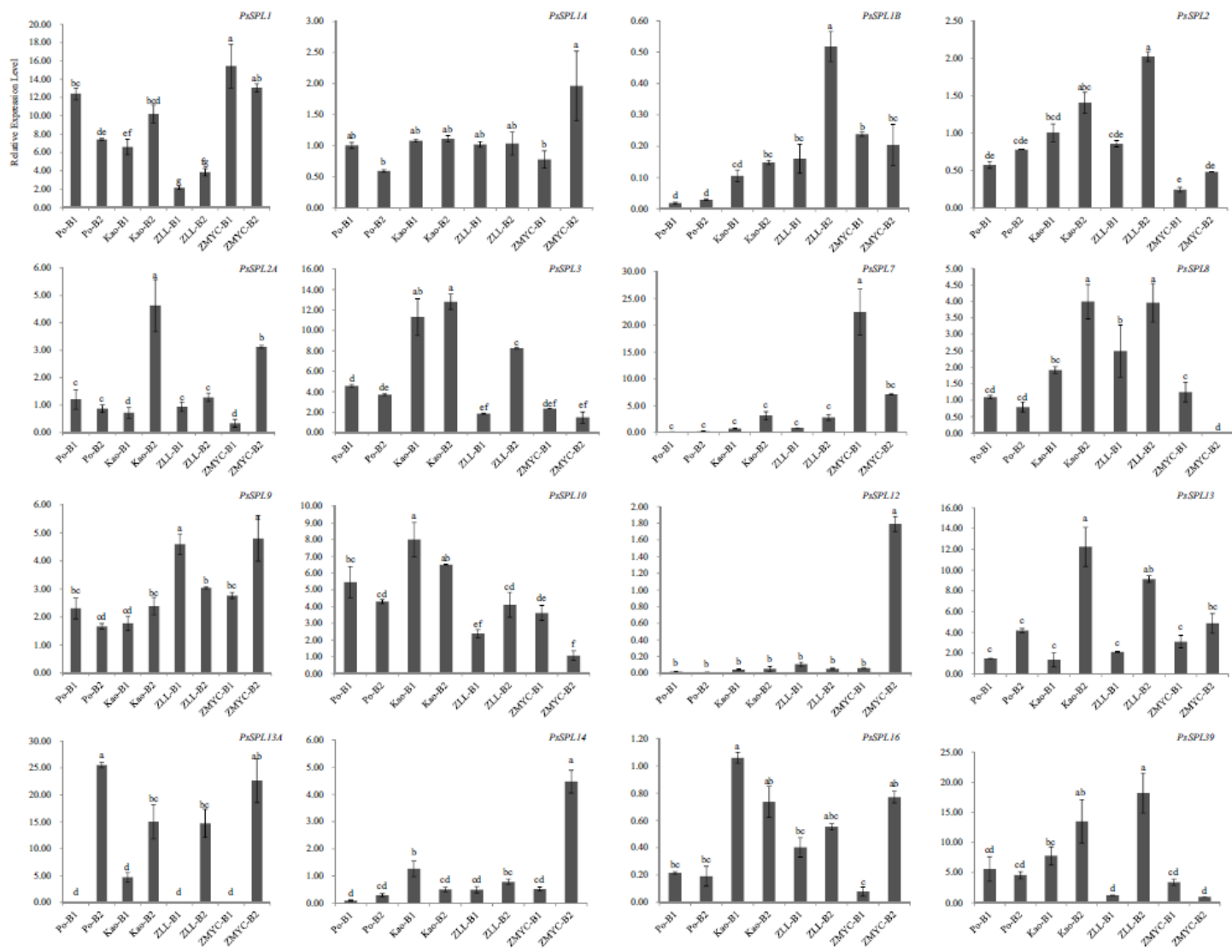




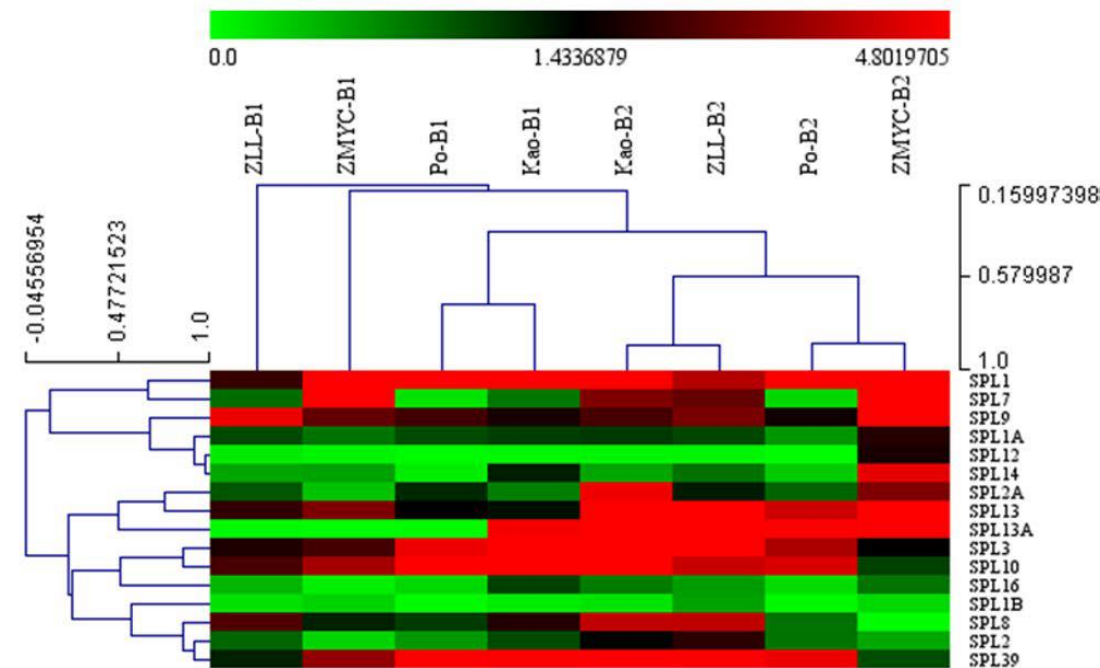
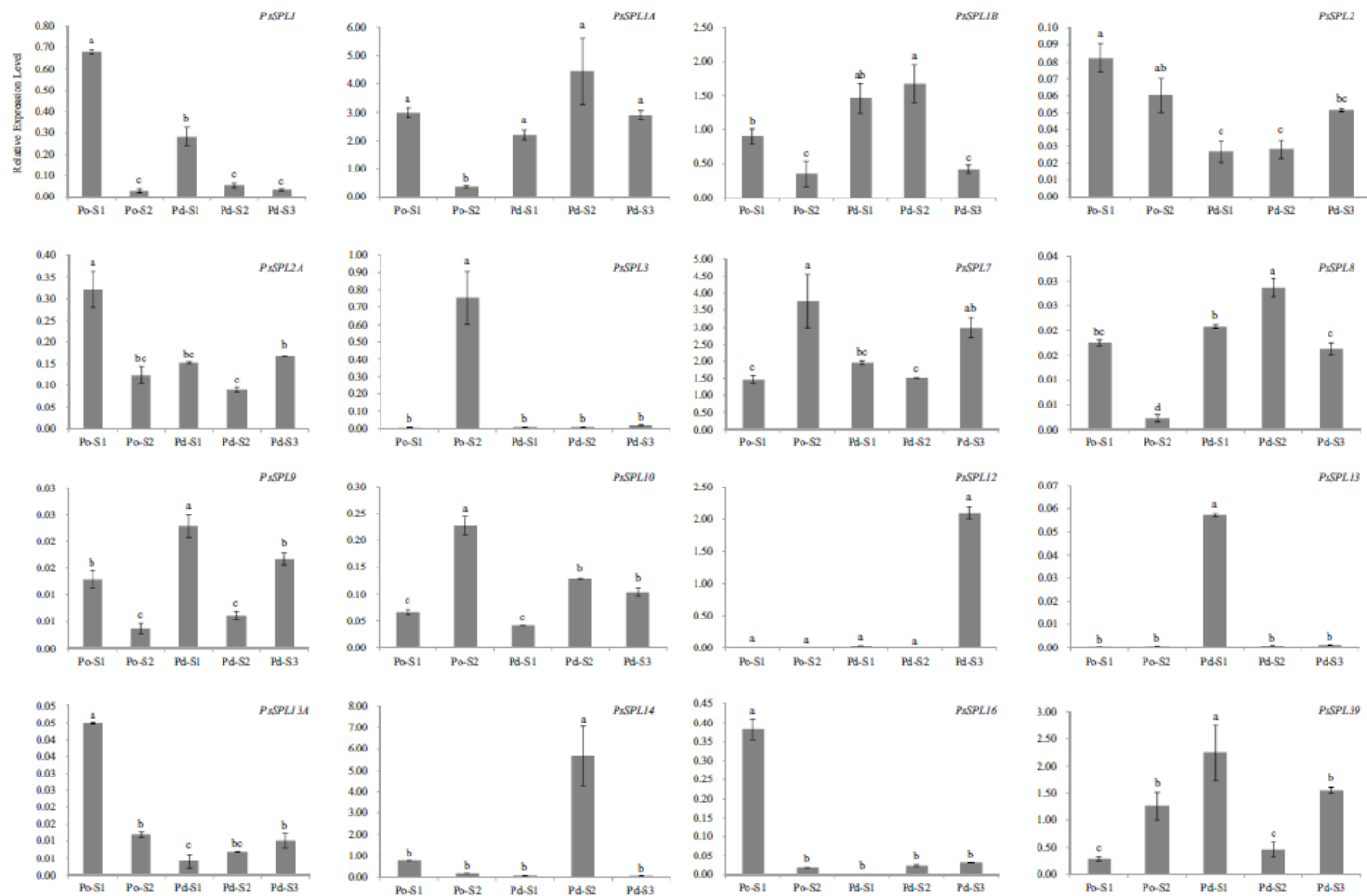
B



不同株龄PsSPL基因的表达模式



不同花期的花蕾中PsSPL基因的表达模式



不同发育阶段PsSPL基因的表达模式

入门很简单，而精通却很难，有一万种方法教你入门，而精通只有靠自己!!!

