

**Genome-wide identification of the tea plant bHLH
transcription factor family and discovery of candidate
regulator of trichome formation
'Supplemental Information'**

Renjian Liu, Yuyuan Wang, Song Tang, Jiarong Cai, Shaoqun Liu, Peng Zheng, Binmei Sun

March 14th, 2021

Contents

Figure S1. The multiple alignment of CsbHLH conserved domain.	3
Figure S2. Phylogenetic tree of the bHLH family in <i>Camellia sinensis</i> L., <i>Arabidopsis thaliana</i> , <i>Actinidia chinensis</i> and <i>Oryza sativa</i> L..	4
Figure S3. Conserved motifs and gene structure distributions of CsbHLH gene family.	5
Figure S4. The expression pattern of CsbHLHs of cluster 1, cluster 3 and cluster 5 in different tissues.	6

	basic	helix1	loop	helix2
CsHL1001	LPAK	NLM	AER	RRR
CsHL1046	LPAK	NLM	AER	RRR
CsHL1030	LPAK	NLM	AER	RRR
CsHL1051	QPSK	NLM	AER	RRR
CsHL1126	QPSK	NLM	AER	RRR
CsHL1075	QPSK	NLM	AER	RRR
CsHL1058	QPSK	NLM	AER	RRR
CsHL1061	AASK	NIV	SERN	RRR
CsHL1114	AASK	NIV	SERN	RRR
CsHL1004	A-QDH	VVA	EKR	RRR
CsHL1060	V-QDR	VMA	EKR	RRR
CsHL1017	SNQD	HIA	EKR	RRR
CsHL1128	NNQD	HIA	EKR	RRR
CsHL1038	A-RNH	ILA	EKR	RRR
CsHL1091	A-ENR	ILA	EKR	RRR
CsHL1054	A-KNH	VVA	EKR	RRR
CsHL1005	-PLN	HVE	AER	RRR
CsHL1013	-PLN	HVE	AER	RRR
CsHL1062	PPVNH	VQAE	RRE	RRR
CsHL1066	PPVNH	VQAE	RRE	RRR
CsHL1132	PPVNH	VQAE	RRE	RRR
CsHL1009	TPINH	VQAE	RRE	RRR
CsHL1028	-PLN	HVE	AER	RRR
CsHL1086	-PLN	HVE	AER	RRR
CsHL1041	-PLN	HVE	AER	RRR
CsHL1078	-PLN	HVE	AER	RRR
CsHL1045	-PLN	HVE	AER	RRR
CsHL1125	-PLN	HVE	AER	RRR
CsHL1024	DHETH	IWT	ERE	RRR
CsHL1133	IYVQ	CAF	PEK	RRR
CsHL1033	LSANH	VLA	EER	RRR
CsHL1112	LSANH	VLA	EER	RRR
CsHL1034	ALKSH	SEA	RRR	RRR
CsHL1129	ALKSH	SEA	RRR	RRR
CsHL1070	GRMTH	I	AVERN	RRR
CsHL1062	ASKSH	SEA	RRR	RRR
CsHL1103	ASKSH	SEA	RRR	RRR
CsHL1099	ASKSH	SEA	RRR	RRR
CsHL1052	ASKNH	KEA	EKR	RRR
CsHL1029	QRMTH	I	AVERN	RRR
CsHL1048	VVG	-G	KST	DR
CsHL1032	QRMTH	I	AVERN	RRR
CsHL1053	QRMTH	I	AVERN	RRR
CsHL1011	QRMTH	I	AVERN	RRR
CsHL1069	QRMTH	I	AVERN	RRR
CsHL1116	GRMTH	I	AVERN	RRR
CsHL1036	QRMTH	I	AVERN	RRR
CsHL1107	QKVS	HIT	VERN	RRR
CsHL1109	QKVS	HIT	VERN	RRR
CsHL1096	QMSH	I	AVERN	RRR
CsHL1029	DHE	I	HWT	ERE
CsHL1089	DHETH	I	WT	ERE
CsHL1008	NDRKT	I	EKNR	-RNQ
CsHL1050	TDRKT	I	EKNR	-RNH
CsHL1064	LDRKL	V	ERNR	-RIH
CsHL1094	LDRKT	V	ERNR	-RIH
CsHL1074	LDRNE	I	EKS	R
CsHL1026	IYVH	S	I	EKR
CsHL1119	IMHRD	I	ERL	R
CsHL1027	INRS	D	I	EKR
CsHL1071	LQ-HL	I	DHNR	SS
CsHL1121	LPPH	T	DQNR	SS
CsHL1073	VAAE	G	G	K
CsHL1084	VVG	-G	K	S
CsHL1003	- - - -	V	K	S
CsHL1035	- - - -	K	H	C
CsHL1067	- - - -	I	N	E
CsHL1025	- - - -	K	A	R
CsHL1031	- - - -	K	A	R
CsHL1111	- - - -	K	A	R
CsHL1068	- - - -	K	A	R
CsHL1076	- - - -	K	A	R
CsHL1002	AMDS	H	L	A
CsHL1063	ATDS	H	L	A
CsHL1019	ATDS	H	L	A
CsHL1079	ATDS	H	L	A
CsHL1106	ATDS	H	L	A
CsHL1130	ATDS	H	L	A
CsHL1043	ATDK	H	L	A
CsHL1044	ATDS	H	L	A
CsHL1007	ATDS	H	L	A
CsHL1012	ATDS	H	L	A
CsHL1042	ATDN	H	L	A
CsHL1101	ATDS	H	L	A
CsHL1095	- - - -	A	R	E
CsHL1104	LV	- - - -	V	C
CsHL1015	- - - -	A	R	E
CsHL1098	ATNS	H	L	A
CsHL1127	ATNS	H	L	A
CsHL1081	ATNS	H	L	A
CsHL1023	SKDP	S	V	A
CsHL1118	SKDP	S	V	A
CsHL1014	SDDP	S	V	A
CsHL1093	STDP	S	V	A
CsHL1117	STDP	S	V	A
CsHL1072	ATDP	S	L	A
CsHL1105	ATDP	S	L	A
CsHL1022	ATHP	S	I	A
CsHL1083	ATHP	S	I	A
CsHL1115	ATHP	S	I	A
CsHL1090	ATHP	S	I	A
CsHL1020	AAEV	H	N	L
CsHL1134	AAEA	H	N	L
CsHL1085	RSSK	R	S	R
CsHL1021	AAEV	H	N	L
CsHL1131	AAEV	H	N	L
CsHL1037	AAVH	N	Q	S
CsHL1097	AAAN	H	N	Q
CsHL1088	AAVH	N	Q	S
CsHL1124	- - - -	Q	K	R
CsHL1018	ATDP	S	I	A
CsHL1122	ATDP	S	I	A
CsHL1049	ATDP	S	I	A
CsHL1056	ATDP	S	I	A
CsHL1077	KKKS	H	S	E
CsHL1120	GGGF	K	I	O
CsHL1047	GRYS	K	G	V
CsHL1102	ATDS	H	L	A
CsHL1087	NINS	H	S	O
CsHL1100	ATDS	H	L	A
CsHL1123	ATNS	H	L	A
CsHL1092	ATHP	S	I	A
CsHL1039	PQSK	N	L	V
CsHL1080	YQSK	N	L	V
CsHL1108	SKDP	S	V	A
CsHL1059	DRS	R	T	L
CsHL1113	SRO	R	A	T
CsHL1016	KKL	N	H	N
CsHL1006	GTT	K	H	C
CsHL1057	VVMP	O	S	T
CsHL1110	CL	S	A	O
CsHL1040	TE	I	E	N
CsHL1082	RSY	R	H	M
CsHL1055	VQDC	N	E	T

Figure S1. The multiple alignment of CsbHLH conserved domain. The conserved amino acids with more than 33% identify are shown in blue. The conservation ratio of them increased with deepening colour.

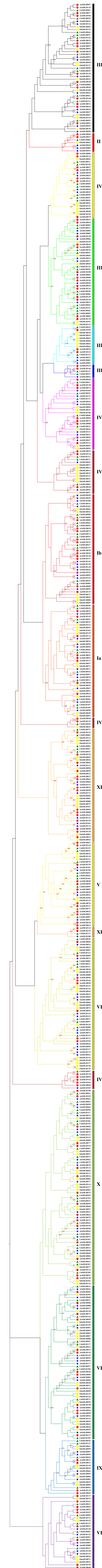


Figure S2. Phylogenetic tree of the bHLH family in *Camellia sinensis* L., *Arabidopsis thaliana*, *Actinidia chinensis* and *Oryza sativa* L.. The different-coloured branches indicated different groups of bHLH proteins. Red circle, green triangle, yellow rect and blue star represented bHLH proteins in *Actinidia chinensis*, *Camellia sinensis* L., *Oryza sativa* L. and *Arabidopsis thaliana*, respectively.

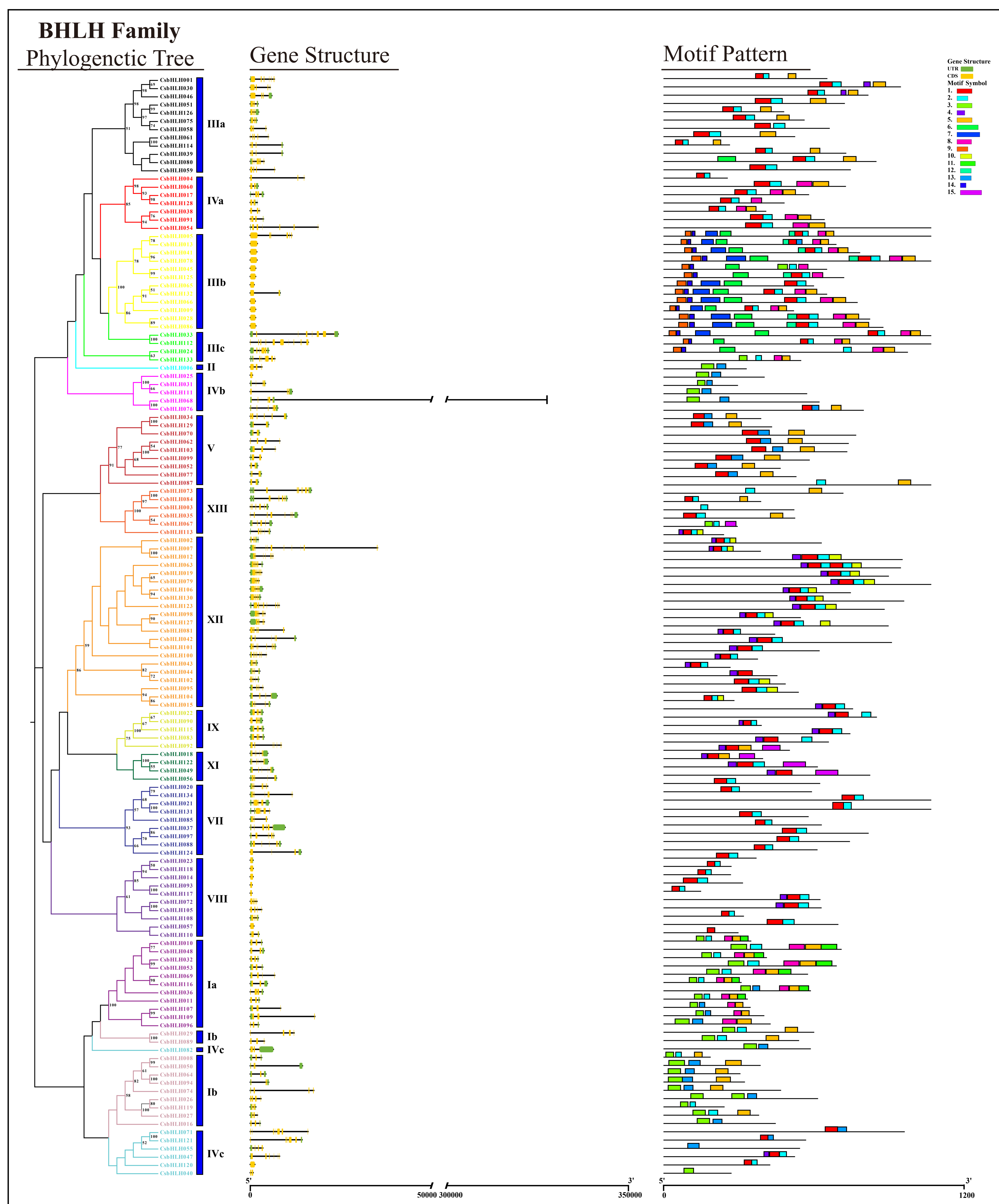


Figure S3. Conserved motifs and gene structure distributions of CsbHLH gene family. The phylogenetic tree included all identified CsbHLHs was constructed using Neighbor-joining method. The yellow and green boxes indicated CDS and UTR of *CsbHLHs* (Left). Different coloured boxes represented different motifs (Right).

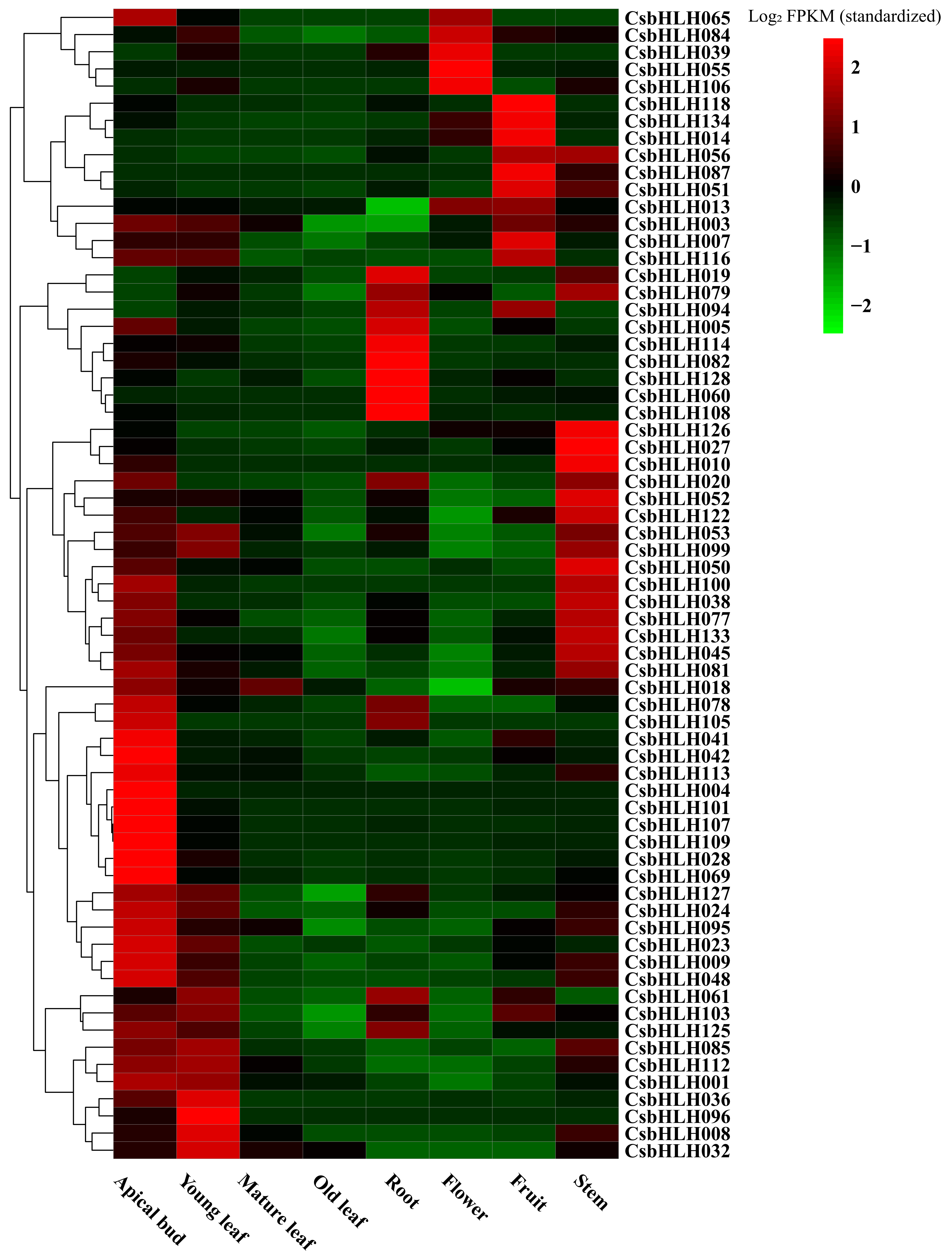


Figure S4. The expression pattern of CsbHLHs of cluster 1, cluster 3 and cluster 5 in different tissues. The RPKM values were downloaded from the Tea Plant Information Archive (TPIA) and used to construct the heat map.