

Supplementary Materials

Identification and expression characterisation of *SbERECTA* family genes in *Sorghum bicolor*

Jia Cheng Zheng^{A,B,†}, Jie Yu^{†B}, Ting Liu^{†B}, Xin Wang^B, Qiu Wen Zhan^B, Jie Qin Li^B, Zhao Shi Xu^{A,C} and You Zhi Ma^{A,C}

^AInstitute of Crop Science, Chinese Academy of Agricultural Sciences (CAAS)/National Key Facility for Crop Gene Resources and Genetic Improvement, Key Laboratory of Biology and Genetic Improvement of Triticeae Crops, Ministry of Agriculture, Beijing 10081, P.R. China.

^BCollege of Agronomy, Anhui Science and Technology University, Fengyang, Anhui, 233100, P.R. China.

^CCorresponding authors. Email: xuzhaoshi@caas.cn; mayouzhi@caas.cn

†These authors contributed equally to this work.

Table S1. The URLs list of biological database

No.	Website	URLs of database
1	Phytozome v12.1	https://phytozome.jgi.doe.gov/pz/portal.html
2	NCBI	https://www.ncbi.nlm.nih.gov
3	PROSITE	https://prosite.expasy.org
4	GSDS	http://gsds.cbi.pku.edu.cn
5	Plant-mPLoc	http://www.csbio.sjtu.edu.cn/bioinf/plant-multi/
6	PlantCARE	http://bioinformatics.psb.ugent.be/webtools/plantcare/html/
7	SMART	http://smart.embl-heidelberg.de

Table S2. The sequence of primers related to PCR amplification

Name	Pair of forward and reverse primers (5'-3')	Tm	Size	Restriction Enzyme	Note
<i>SbER10_XI-2F</i>	GCGAGAACCCACGAAACC AACC	64°C	3376 bp	—	<i>SbER10_XI</i> gene isolation
<i>SbER10_XI-2R</i>	CCCATG ATTCCAACACGGCAA				
<i>GFP-SbER10_XI-F1</i>	<u>GGGGACGAGCTCGGTACCATGCCTGTCCGCAGCTCA</u>	62°C	2982 bp	<i>Kpn</i> I	Subcellular localization
<i>GFP-SbER10-XI-R1</i>	<u>CATGGTGTGACTCTAGACTCCATGTTCTGCGAGATGGC</u>			<i>Xba</i> I	
<i>SbER10-KF</i>	<u>CGGGATCCCCTTCATATGAACATGGCCC</u>	61°C	965 bp	<i>BamH</i> I	Prokaryotic expression
<i>SbER10-KR</i>	<u>CCCAAGCTTTGATGTAGCTTGGCACGGC</u>			<i>Hind</i> III	
<i>SbER10-2QF</i>	GCCAGCCGTGCCAAGCTACATC	62°C	179 bp	—	qRT-PCR
<i>SbER10-2QR</i>	CTCCCTGCATCCTTCAACTCCC				
<i>SbActin-F</i>	ACGGCCTGGAT GGCGACGTACATG	62°C	185 bp	—	
<i>SbActin-R</i>	GCAGAAGGACGCCTACGTTGGTGAC				

Table S3. Name and sources of 32 sweet sorghum collection

Code	No. of national genebank	Origin (province)
<u>1</u>	357	Hei long jiang
<u>2</u>	360	Hei long jiang
<u>3</u>	2349	Shan xi
<u>4</u>	2350	Shan xi
<u>5</u>	7322	Nei meng gu
<u>6</u>	7323	Nei meng gu
<u>7</u>	7324	Liao ning
<u>8</u>	7325	Liao ning
<u>9</u>	7329	Hei long jiang
<u>10</u>	7342	An hui
<u>11</u>	7343	An hui
<u>12</u>	7357	Hu bei
<u>13</u>	7358	Hu bei
<u>14</u>	10246	Ji lin
<u>15</u>	10264	Shaan xi
<u>16</u>	10265	Shaan xi
17	10288	Yun nan
18	10289	Yun nan
19	12484	Bei jing
20	12520	Bei jing
21	12630	Bei jing
22	13241	Si chuan
23	13242	Si chuan
24	13361	Gui zhou

Code	No. of national genebank	Origin (province)
25	13428	Hu bei
26	13443	Hu bei
27	13543	Hai nan
28	13812	Guangxi
29	13854	Gui zhou
30	14181	He nan
31	M81	American
32	Yuxitian	An hui

Table S4. NCBI accession numbers of proteins in phylogenetic tree

Code	ER family proteins	Plant Species	NCBI accession number	Classification
Group I				
1	SbER10_X1	<i>Sorghum bicolor</i> L.	XP_002438023.1	monocots
2	SbER10_X2	<i>Sorghum bicolor</i> L.	XP_021304405.1	monocots
3	SbER10_X3	<i>Sorghum bicolor</i> L.	XP_021304406.1	monocots
4	SbER4_X1	<i>Sorghum bicolor</i> L.	XP_021314268.1	monocots
5	SbER4_X2	<i>Sorghum bicolor</i> L.	XP_021314269.1	monocots
6	ZmERECTA_1	<i>Zea mays</i> L.	AQK83004.1	monocots
7	ZmERECTA_2	<i>Zea mays</i> L.	AQK83013.1	monocots
8	ZmERECTA_3	<i>Zea mays</i> L.	AQK83012.1	monocots
9	PamERECTA_X2	<i>Panicum miliaceum</i> L.	RLN13615.1	monocots
10	SiERECTA_X2	<i>Setaria italica</i> L.	XP_004964942.1	monocots
11	SiERECTA_X1	<i>Setaria italica</i> L.	XP_004964941.1	monocots
12	OtERECTA	<i>Oropetium thomaeum</i> L.	AYU56584.1	monocots
13	BdERECTA_X1	<i>Brachypodium distachyon</i> L.	XP_003564133.1	monocots
14	BdERECTA_X2	<i>Brachypodium distachyon</i> L.	XP_010227884.1	monocots
15	HvERECTA	<i>Hordeum vulgare</i> subsp. <i>vulgare</i> L.	AKU38976.1	monocots
20	HvERL_1	<i>Hordeum vulgare</i> L.	KAE8808234.1	monocots
17	AetERECTA_X1	<i>Aegilops tauschii</i> subsp. <i>t.auschii</i> L.	XP_020193416.1	monocots
19	AetERECTA_X2	<i>Aegilops tauschii</i> subsp. <i>tauschii</i> L.	XP_020193417.1	monocots
16	TaERECTA_2	<i>Triticum aestivum</i> L.	AFJ38187.2	monocots
18	TaERECTA_1	<i>Triticum aestivum</i> L.	AFJ38186.2	monocots
21	ObERECTA_X1	<i>Oryza brachyantha</i> L.	XP_006649092.1	monocots
22	ObERECTA_X2	<i>Oryza brachyantha</i> L.	XP_015689543.1	monocots
23	OsERECTA_X1	<i>Oryza sativa Japonica Group</i> L.	XP_015623966.1	monocots

Code	ER family proteins	Plant Species	NCBI accession number	Classification
24	<i>OsERL</i>	<i>Oryza sativa Japonica Group</i> L.	AFJ14786.1	monocots
25	<i>OsERECTA_X2</i>	<i>Oryza sativa Japonica Group</i>	XP_015623967.1	monocots
Group II				
26	<i>AcERECTA_X2</i>	<i>Ananas comosus</i> L.	XP_020081626.1	monocots
27	<i>AcERECTA_X1</i>	<i>Ananas comosus</i> L.	XP_020081619.1	monocots
28	<i>PeERECTA</i>	<i>Phalaenopsis equestris</i> L.	XP_020588182.1	monocots
29	<i>EgERECTA</i>	<i>Elaeis guineensis</i> L.	XP_010917192.1	monocots
30	<i>MaERECTA_X2</i>	<i>Musa acuminata subsp. malaccensis</i> L.	XP_018684789.1	monocots
31	<i>MaERECTA_X4</i>	<i>Musa acuminata subsp. malaccensis</i> L.	XP_009408412.1	monocots
32	<i>MaERECTA_X3</i>	<i>Musa acuminata subsp. malaccensis</i> L.	XP_018684790.1	monocots
33	<i>MaERECTA_X1</i>	<i>Musa acuminata subsp. malaccensis</i> L.	XP_018684788.1	monocots
Group III				
34	<i>SeiERECTA</i>	<i>Sesamum indicum</i> L.	XP_011069519.1	dicots
35	<i>PoERECTA1</i>	<i>Populus trichocarpa</i> L.	XP_024458576.1	dicots
36	<i>PoERECTA2</i>	<i>Populus euphratica</i> L.	XP_011019342.1	dicots
38	<i>GmERECTA2</i>	<i>Glycine max</i> L.	XP_014632764.1	dicots
39	<i>GmERECTA_X1</i>	<i>Glycine max</i> L.	XP_006578096	dicots
40	<i>GmERECTA1</i>	<i>Glycine max</i> L.	XP_003544548.1	dicots
41	<i>PaERECTA_X2</i>	<i>Prunus avium</i> L.	XP_021805236.1	dicots
42	<i>PaERECTA_X1</i>	<i>Prunus avium</i> L.	XP_021805235.1	dicots
43	<i>PmERECTA</i>	<i>Prunus mume</i> L.	XP_008220576.1	dicots
44	<i>VvERECTA1</i>	<i>Vitis vinifera</i> L.	XP_002280069.2	dicots
45	<i>VvERECTA2</i>	<i>Vitis vinifera</i> L.	RVX17310.1	dicots
46	<i>TcERECTA1</i>	<i>Theobroma cacao</i> L.	XP_017983493.1	dicots
47	<i>TcERECTA2</i>	<i>Theobroma cacao</i> L.	EOY30050.1	dicots
48	<i>DzERECTA</i>	<i>Durio zibethinus</i> L.	XP_022754146.1	dicots

Code	ER family proteins	Plant Species	NCBI accession number	Classification
49	CsERECTA	<i>Citrus sinensis</i> L.	XP_006474503.1	dicots
50	JcERECTA	<i>Jatropha curcas</i> L.	XP_012077188.1	dicots
51	MeERECTA	<i>Manihot esculenta</i> L.	XP_021598132.1	dicots
52	HbERECTA	<i>Hevea brasiliensis</i> L.	XP_021676852.1	dicots
37	AtERECTA	<i>Arabidopsis thaliana</i> L.	NP_180201.1	dicots
Group IV				
53	AtERL_1	<i>Arabidopsis thaliana</i> L.	NP_001190595.1	dicots
54	AtERL_2	<i>Arabidopsis thaliana</i> L.	NP_001331523.1	dicots

Table S5. NCBI accession numbers of *ER* genes in phylogenetic tree

No.	ER family genes	Plant species	NCBI accession number	Classification
Group I				
1	<i>AetERECTA</i>	<i>Aegilops tauschii</i> subsp. <i>tauschii</i> L.	XM_020337827.1	monocots
2	<i>TaERECTA</i>	<i>Triticum aestivum</i> L.	JQ599261.2	monocots
3	<i>HvERECTA</i>	<i>Hordeum vulgare</i> subsp. <i>vulgare</i> L.	KP284445.1	monocots
4	<i>BdERECTA</i>	<i>Brachypodium distachyon</i> L.	XM_003564085.4	monocots
5	<i>SiERECTA</i>	<i>Setaria italica</i> L.	XM_004964885.3	monocots
6	<i>SbER1_X1</i>	<i>Sorghum bicolor</i> L.	XM_002437978.2	monocots
7	<i>ZmERECTA</i>	<i>Zea mays</i> L.	NM_001358834.1	monocots
8	<i>OsERECTA</i>	<i>Oryza sativa Japonica Group</i> L.	XM_015768480.2	monocots
Group II				
9	<i>NtERECTA</i>	<i>Nicotiana tabacum</i> L.	XM_016632277.1	dicots
10	<i>AtERECTA</i>	<i>Arabidopsis thaliana</i> L.	NM_128190.3	dicots
11	<i>GmERECTA</i>	<i>Glycine max</i> L.	XM_003544500.4	dicots
12	<i>PoERECTA</i>	<i>Populus trichocarpa</i> L.	XM_024602808.1	dicots
13	<i>VvERECTA</i>	<i>Vitis vinifera</i> L.	XM_002280033.3	dicots
Group III				
14	<i>AtERL_1</i>	<i>Arabidopsis thaliana</i> L.	NM_001203666.1	dicots
15	<i>AtERL_2</i>	<i>Arabidopsis thaliana</i> L.	NM_001342930.1	dicots

Fig. S1. The putative protein transcripts of SbER10_X1 subcellular localization.

DNAMAN File: DNAMAN1_Translation

SbER10_X1 (subcellular localization)

Total amino acid number: 981, MW=107187
Max ORF starts at AA pos 1(may be DNA pos 1) for 981 AA(2943 bases), MW=107187

	10	20	30	40	50	60
1	ATGCCTGTCGCCAGCTCAGTGGCCATGACGACGACGGCCGCCGTGCTCTCGTCGCCCTC					
1	M P V R S S V A M T T T A A R A L V A L					
	70	80	90	100	110	120
61	CTCCTCGTCGCCGTGCCGACGATGGGCGACGCTGGAGATCAAGAAAGTCC					
21	L L V A V A V A D D G A T L V E I K K S					
	130	140	150	160	170	180
121	TTCCGCAACGTCGGCAACGTACTGTACGATTGGGCCGGCACGACTACTGCTCCTGGCC					
41	F R N V G N V L Y D W A G D D Y C S W R					
	190	200	210	220	230	240
181	GGCGCTCTGTGCCAACGTCACATTGCCGTGCCGCTCACACCTCTGGCCCTAAC					
61	G V L C D N V T F A V A A L N L S G L N					
	250	260	270	280	290	300
241	CTTGAGGGCGAGATCTCTCAGCGTGGCAGCCTCAAGAGCCTCGTCTCCATCGATCTG					
81	L E G E I S P A V G S L K S L V S I D L					
	310	320	330	340	350	360
301	AAGTCAAATGGGCTATCCGGGAGATCCCTGATGGAGATTGGTATTGTTCATCACTTAGG					
101	K S N G L S G Q I P D E I G D C S S L R					
	370	380	390	400	410	420
361	ACGCTGGACTTTCTTCAACAACATTGGATGGCAGACATACCAATTCTATATCAAAGCTG					
121	T L D F S F N N L D G D I P F S I S K L					
	430	440	450	460	470	480
421	AAGCACCTGGAGAACATTGATATTGAAGAACACCAACAGCTGATTGGTGCATCCATCAACA					
141	K H L E N L I L K N N Q L I G A I P S T					
	490	500	510	520	530	540
481	TTGTCACAGCTCCAAATTGAGATTGGATTTGGCACAAAACAAACTGACTGGGAG					
161	L S Q L P N L K I L D L A Q N K L T G E					
	550	560	570	580	590	600
541	ATACCAAGGCTTACTGGAATGAGGTTCTCAATATCTGGCTTACGGGCAATCAT					
181	I P R L I Y W N E V L Q Y L G L R G N H					
	610	620	630	640	650	660
601	TTAGAAGGAAGCCTCTCCTGATATGCTGCCAGCTGACTGGCCTTGGTACTTGATGTG					
201	L E G S L S P D M C Q L T G L W Y F D V					
	670	680	690	700	710	720
661	AAGACAATAGCTTGACGGGGTGATACCAGACACATTGGAACTGTACAAGTTCAA					
221	K N N S L T G V I P D T I G N C T S F Q					
	730	740	750	760	770	780
721	GTCTTGGATTGTCTTACAACCGCTTTACTGGACCAATCCCATTCAACATTGGTTCTA					
241	V L D L S Y N R F T G P I P F N I G F L					
	790	800	810	820	830	840
781	CAAGTGGTACACTATCCCTGCAAGGAAACAAGTTACCGGCTCAATTCTTCAATT					
261	Q V A T L S L Q G N K F T G P I P S V I					
	850	860	870	880	890	900
841	GGCTTATGCAGGCTCTCGCTCTAGATCTGAGTTACAACCAATTATCTGGCTCTATA					
281	G L M Q A L A V L D L S Y N Q L S G P I					
	910	920	930	940	950	960
901	CCATCAATACTAGGCAACTTGACATACACTGAGAACGCTGACATCCAAGGCAATAAGTTA					
301	P S I L G N L T Y T E K L Y I Q G N K L					
	970	980	990	1000	1010	1020
961	ACTGGGTCGATACCAACCCAGAGTTAGGAAATATGCAACACTCATTACCTAGAAC					
321	T G S I P P E L G N M S T L H Y L E L N					
	1030	1040	1050	1060	1070	1080
1021	GATAATCAACTTACTGGGTCATTCCACCAAGAGCTTGGAAAGGCTAACAGGCTTGTGAC					

DNAMAN File: DNAMAN1 Translation

341	D N Q L T G S I P P E L G R L T G L F D
1081	1090 1100 1110 1120 1130 1140
361	CTGAACCTTGCATAACCACCTGGAGGACCAATTCTGACACACCTAAGTTCATGTGTG L N L A N N H L E G P I P D N L S S C V
1141	1150 1160 1170 1180 1190 1200
381	AATCTCAATAGCTTCATGCTTATGCCAACAGTTAAATGGGACCATTCCTCGTTG N L N S F N A Y G N K L N G T I P R S L
1201	1210 1220 1230 1240 1250 1260
401	CGGAAACTTGAAAGCATGACCTATTTAAATCTGTCATCAAACCTCATAAAGTGGCTCTATT R K L E S M T Y L N L S S N F I S G S I
1261	1270 1280 1290 1300 1310 1320
421	CCTATTGAGTTATCAAGGATCAACAATTGGACACGCTGGATTATCCTGTAACATGATG P I E L S R I N N L D T L D L S C N M M
1321	1330 1340 1350 1360 1370 1380
441	ACTGGTCCAATTCCATCATCAATTGGCAGCCTAGAGCATCTATTGAGACTTAACTTGAGC T G P I P S S I G S L E H L L R L N L S
1381	1390 1400 1410 1420 1430 1440
461	AAGAATGGTCTAGTTGGATTATCCCCGGAGTTGGTAATTGAGGAGTGTACATGGAG K N G L V G F I P A E F G N L R S V M E
1441	1450 1460 1470 1480 1490 1500
481	ATTGATTATCCTATAATCACCTGGCTGGCTGATTCTCAAGAACTTGAAATGCTGCAA I D L S Y N H L G G L I P Q E L E M L Q
1501	1510 1520 1530 1540 1550 1560
501	AACCTGATGTTGCTAAACACTGGAAAATAACAATATAACTGGTATCTGCTCTGATG N L M L L K L E N N N I T G D L S S L M
1561	1570 1580 1590 1600 1610 1620
521	AACTGCTTCAGCCCTCAATATCTTAAATGTGTCGTCACAATAATTGGCTGGTGTG N C F S L N I L N V S Y N N L A G V V P
1621	1630 1640 1650 1660 1670 1680
541	GCTGACAACAACCTCACACGGTTTACCTGACAGCTTTAGGTAATCCTGACTCTG A D N N F T R F S P D S F L G N P G L C
1681	1690 1700 1710 1720 1730 1740
561	GGATACTGGCTTGGTTCTGTCGTTCCACTGGCACCGAGAAACCGCTATCTCA G Y W L G S S C R S T G H H E K P P I S
1741	1750 1760 1770 1780 1790 1800
581	AAGGCTGCCATAATTGGTGTGGTGACTGGTGTATCCTCTGATGATCTTAGTA K A A I I G V A V G G L V I L L M I L V
1801	1810 1820 1830 1840 1850 1860
601	GCTGTTTCAGGCCACATCGTCCACCTGCTTTAAAGATGTCAGTAAAGCAAGCCAGTG A V C R P H R P P A F K D V T V S K P V
1861	1870 1880 1890 1900 1910 1920
621	AGAAAATGCTCCCCCAAGCTGGTATCCATATGAACATGCCCTCATGTATACGAT R N A P P K L V I L H M N M A L H V Y D
1921	1930 1940 1950 1960 1970 1980
641	GACATAATGAGGATGACTGAGAACCTTGAGTGAGAAATACATCATTGGATACGGGGCGTCA D I M R M T E N L S E K Y I I G Y G A S
1981	1990 2000 2010 2020 2030 2040
661	ACTACAGTTATAATGTGCTAAAGAATTGCAAACGGGTGCCAATAAAAAGCTGTAT S T V Y K C V L K N C K P V A I K K L Y
2041	2050 2060 2070 2080 2090 2100
681	GCCCACCTACCCACAGAGCCTTAAGGAATTGAAACTGAGCTTGAGACTGTTGGTAGCATC A H Y P Q S L K E F E T E L E T V G S I
2101	2110 2120 2130 2140 2150 2160
701	AAGCACCGGAATCTAGTCAGCCTCAAGGGTACTCATTACACCTGTTGGAACCTCCTC K H R N L V S L Q G Y S L S P V G N L L

— DNAMAN File: DNAMAN1 Translation —
2161 2170 2180 2190 2200 2210 2220
721 TTTTATGATTATGGAAATGTGGCAGCTTATGGATTTACATGAAGGTTCATCCAAG
F Y D Y M E C G S L W D V L H E G S S K
2221 2230 2240 2250 2260 2270 2280
741 AAGAAAAAAACTTGACTGGGAGACTCGCCTACGGATTGCTCTTGGTGCAGCTCAAGGCCCT
K K K L D W E T R L R I A L G A A Q G L
2281 2290 2300 2310 2320 2330 2340
761 GCTTACCTTCACCATGACTGCAGTCCACGGATAATTACATGGGATGTAATCAAAGAAT
A Y L H D C S P R I I H R D V K S K N
2341 2350 2360 2370 2380 2390 2400
781 ATACTCCTTGACAAAGATTATGAGGCCATCTACAGACTTTGGAATTGCTAACAGAGCTTA
I L L D K D Y E A H L T D F G I A K S L
2401 2410 2420 2430 2440 2450 2460
801 TGTTGTCCTAAAACACACATCAACCTATGTCATGGAACTATTGGCTACATTGATCCT
C V S K T H T S T Y V M G T I G Y I D P
2461 2470 2480 2490 2500 2510 2520
821 GAGTACGCCGCACCTCCCGTCTCAACGAAAGCTGATGTCTACAGCTATGGCATTGTT
E Y A R T S R L N E K S D V Y S Y G I V
2521 2530 2540 2550 2560 2570 2580
841 CTGCTGGAGCTGCTGACTGGCAAGAACGAAAGCCAGTGGACAACGAGTGCAATCTCCATCACTTG
L L E L L T G K K P V D N E C N L H H L
2581 2590 2600 2610 2620 2630 2640
861 ATCTATCGAACGCGAACGAAACGAGGTATGGATACCCTGACATCGGGAC
I L S K T A S N E V M D T V D P D I G D
2641 2650 2660 2670 2680 2690 2700
881 ACCTGCAAGGACCTCGGCGAGGTGAAGAACGACTGGCTCCAGCTGGCGCTCCTTGACACCAAG
T C K P L G E V K K L F Q L A L L C T K
2701 2710 2720 2730 2740 2750 2760
901 CGGCAACCCCTCGGACCGACCGAGATGCACGAGGTGGTGCCTGGACTGGCTGGTG
R Q P S D R P T M H E V V R V L D C L V
2761 2770 2780 2790 2800 2810 2820
921 AACCCGGACCCGGCCAAAGCCGTCGGCGCACCGTGCCGAGCCCTCGCCAGCCCTG
N P D P P P K P S A H Q L P Q P S P A V
2821 2830 2840 2850 2860 2870 2880
941 CCAAGCTACATCAACGAGTACGTCAGCCTGCAGGGCACCGCGCTCTCGCCAGCCAAAC
P S Y I N E Y V S L R G T G A L S C A N
2881 2890 2900 2910 2920 2930 2940
961 TCGACCAGCACCTCGGACGCCAGCTGTTCCCTAAGTTGGCGAGGCCATCTCGCAGAAC
S T S T S D A E L F L K F G E A I S Q N
2941 ATGGAG
981 M E

Fig. S2. The growth and development stages of sorghum.

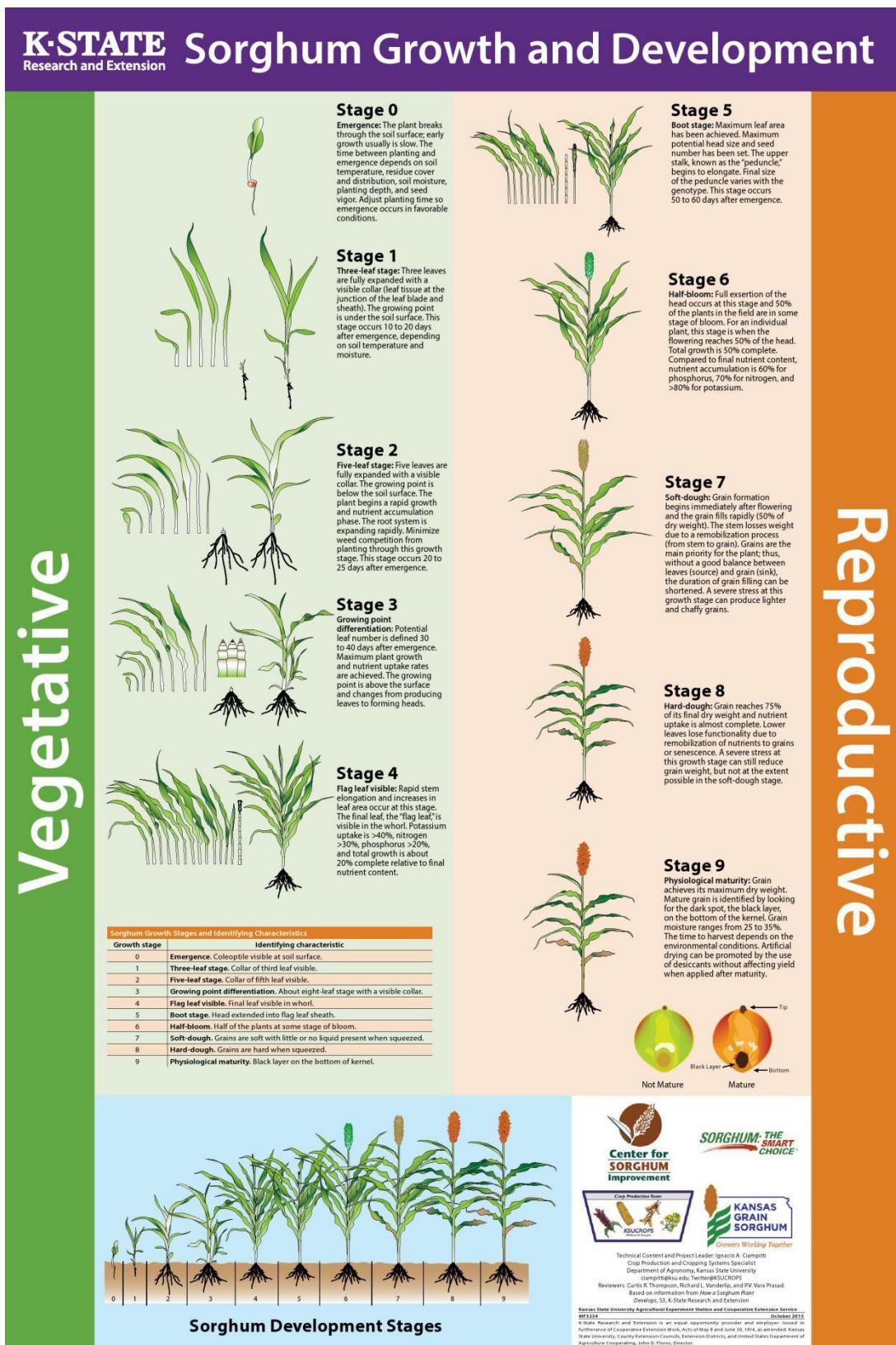
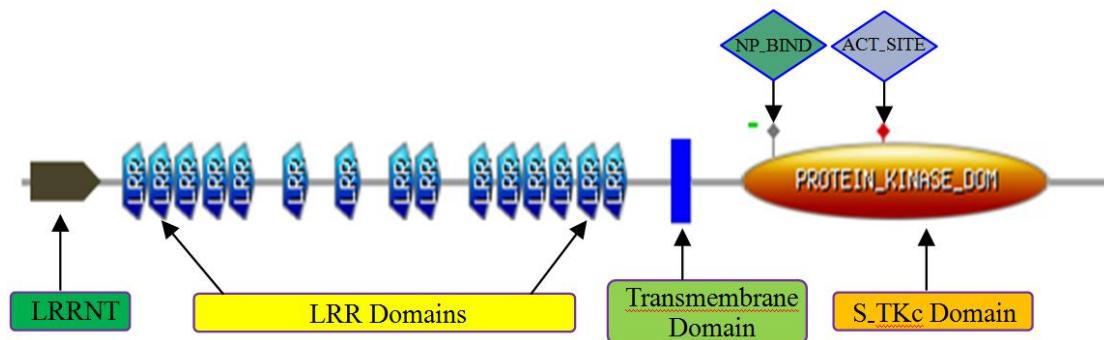
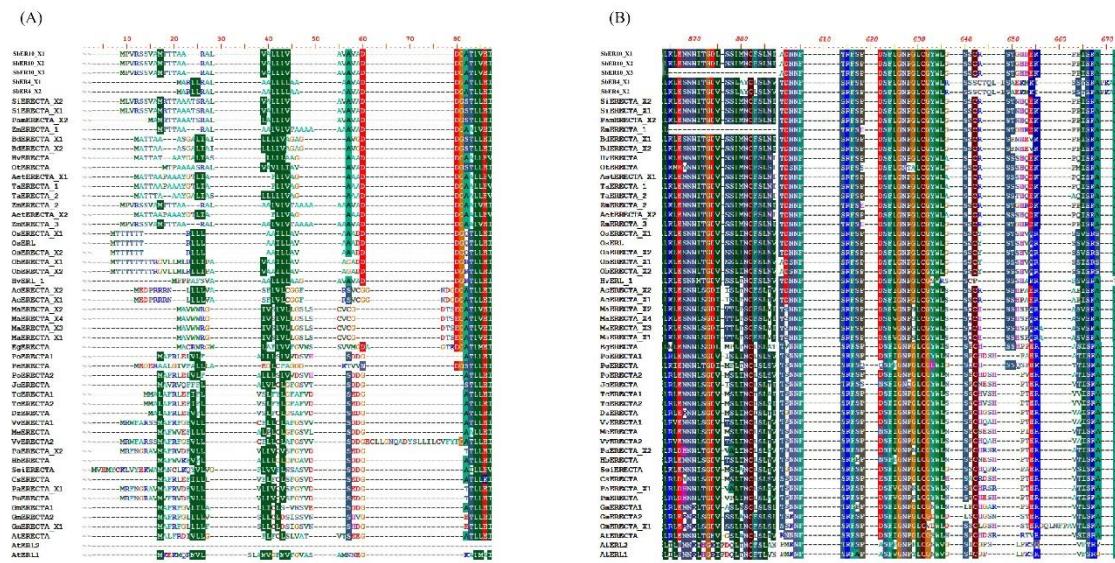


Fig. S3. Analysis of conserved domains of SbER family.



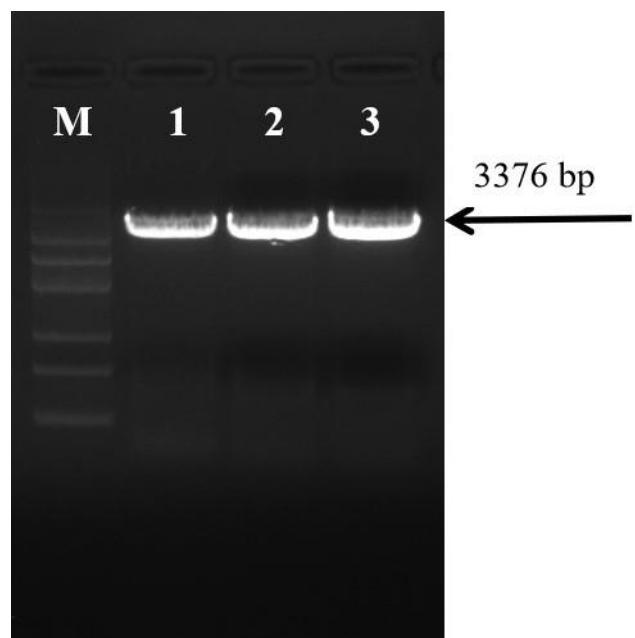
LRRNT was the N-terminal signal peptide region, LRR Domain was the leucine repeat region, S_TKc was the Serine/threonine protein kinase region, catalytic region, NP_BIND was the ATP binding site (rich in lysine residues), ACT_SITE was the proton accepting site (rich in aspartic acid residues).

Fig. S4. Amino acid sequence alignment of ER family members in monocotyledonous and dicotyledonous plants.



(A) was the N-terminal signal peptide region; (B) was the transmembrane region.

Fig. S5. *SbER10_XI* gene isolation.



M was the Maker III (Tian gen, China), lane 1, 2 and 3 were the fragment of *SbER10_XI* gene