

Supplementary data:

Characterization of low-molecular-weight glutenin subunit genes of *Aegilops* section *Sitopsis* and comparative analysis with those of wheat (*Triticum aestivum* L.) and some *Aegilops* species

Zhuo Huang, Hai Long, Xiao-Fang Yu, Yu-Ming Wei, Ze-Hong Yan and You-Liang Zheng

J. Genet. **94**, 497–501

Table 1. The accessions of *Sitopsis* species used in this study.

Species	Accession no. (symbol for clones)
<i>Ae. bicornis</i> (S ^b S ^b)	CIae 47 (Sb47)
<i>Ae. longissima</i> (S ^l S ^l)	PI 604105 (SL105), PI 604106 (SL106), PI 604124 (SL123) PI 604131 (SL131), PI 604133 (SL133)
<i>Ae. searsii</i> (S ^s S ^s)	PI 599138 (SS138), PI 599139 (SS139), PI 599151 (SS151)
<i>Ae. sharonesis</i> (S ^{sh} S ^{sh})	PI 584347 (Sh47), PI 584349 (Sh49), PI 584357 (Sh57)
<i>Ae. speltoides</i> (SS)	PI 560527 (S27)

Table 2. LMW-GS genes obtained in this study.

Species	Accession no.	Clone	Length (bp)	Sequence similarity
<i>Ae. bicornis</i> (S ^b S ^b)	CIae 47	Sb47-1	1053	EU189089 96% <i>T. aestivum</i>
		Sb47-2	1053	EU189089 96% <i>T. aestivum</i>
		Sb47-3	1053	EU305550 97% <i>Ae. longissima</i>
		Sb47-4	1053	EU189089 96% <i>T. aestivum</i>
		Sb47-5	1053	EU189089 96% <i>T. aestivum</i>
		Sb47-6	1053	EU189089 96% <i>T. aestivum</i>
<i>Ae. longissima</i> (S ^l S ^l)	PI 604103	SL103-1 ^a	1062	EU189089 95% <i>T. aestivum</i>
		SL103-2 ^a	972	EU036695 96% <i>T. aestivum</i>
	PI 604105	SL105-1	903	EU189093 93% <i>T. aestivum</i>
		SL105-2	903	DQ457417 93% <i>T. aestivum</i>
		SL105-3	903	EU189093 93% <i>T. aestivum</i>
	PI 604106	SL106-1	903	DQ457417 93% <i>T. aestivum</i>
		SL106-2	903	EU189093 93% <i>T. aestivum</i>
		SL106-3	903	EU189093 93% <i>T. aestivum</i>
	PI 604126	SL126-1 ^a	903	DQ457417 93% <i>T. aestivum</i>
	PI 604129	SL129-1 ^a	837	EF371505 99% <i>T. turgidum</i>
		SL129-3 ^a	903	EU189093 93% <i>T. aestivum</i>
		SL129-4 ^{a*}	903	EU189093 93% <i>T. aestivum</i>
	PI 604130	SL130-1	903	DQ457417 93% <i>T. aestivum</i>
		SL130-2	903	DQ457417 93% <i>T. aestivum</i>
		SL130-3	903	DQ457417 93% <i>T. aestivum</i>
		SL130-4*	969	AB062859 92% <i>T. aestivum</i>
		SL130-5*	969	AB062859 92% <i>T. aestivum</i>
		SL130-6	903	EU189093 93% <i>T. aestivum</i>
	PI 604131	SL131-1*	1047	EU189089 89% <i>T. aestivum</i>
		SL131-2*	999	EU036695 89% <i>T. aestivum</i>
		SL131-3*	867	EF371505 91% <i>T. aestivum</i>
	PI 604133	SL133-1*	1002	EU036695 89% <i>T. aestivum</i>
		SL133-2*	918	AB062864 89% <i>T. aestivum</i>
SL133-3*		999	EU036695 89% <i>T. aestivum</i>	
<i>Ae. searsii</i> (S ^s S ^s)	PI 599138	SS138-1*	903	AJ293099 94% <i>T. durum</i>
		SS138-2*	903	AJ293099 94% <i>T. durum</i>
		SS138-3*	903	AJ293099 94% <i>T. durum</i>
		SS138-4*	903	AJ293099 94% <i>T. durum</i>
	PI 599139	SS139-1*	903	AJ293099 94% <i>T. durum</i>
		SS139-2*	903	AJ293099 94% <i>T. durum</i>
	PI 599151	SS151-1	903	EU189093 93% <i>T. aestivum</i>
		SS151-2	903	EU189093 93% <i>T. durum</i>
SS151-3	903	AJ293099 93% <i>T. durum</i>		
<i>Ae. sharonesis</i> (S ^{sh} S ^{sh})	PI 584347	Sh47-1	903	EU189093 93% <i>T. durum</i>
		Sh47-2	903	EU189093 93% <i>T. durum</i>
		Sh47-3	903	DQ457417 94% <i>T. aestivum</i>
		Sh47-4	903	DQ457417 93% <i>T. aestivum</i>
		Sh47-5*	903	DQ457417 93% <i>T. aestivum</i>
		Sh47-6	813	EF371505 95% <i>T. turgidum</i>
		Sh47-7	1101	AB062851 95% <i>T. aestivum</i>
	PI 584349	Sh49-1*	960	AB062859 93% <i>T. aestivum</i>
		Sh49-2*	953	EU189096 95% <i>T. aestivum</i>
	PI 584357	Sh57-1	903	EU189093 93% <i>T. durum</i>
		Sh57-3*	903	EU189093 93% <i>T. durum</i>
		Sh57-4	903	EU189093 93% <i>T. durum</i>
<i>Ae. speltoides</i> (SS)	PI 560527	S27-1	903	EU305553 97% <i>Ae. longissima</i>
		S27-2*	903	AJ293099 93% <i>T. durum</i>
		S27-3	903	AJ293099 93% <i>T. durum</i>
		S27-4	903	AJ293099 93% <i>T. durum</i>
		S27-5	903	AJ293099 93% <i>T. durum</i>

^aGenes obtained in our previous study (Huang *et al.* 2010); *pseudogenes; ^agene obtained by our previous study (Huang *et al.* 2010)

Table 3. N-terminal sequences of 109 LMW-GS genes analysed in this study.

Gene	N-terminal sequence
GQ389629.1_Ae._cylindrica	METRCVPLERPW
AB062851_METSHIP-D	METSHIPGLEKPS
GQ389634.1_Ae._speltoides	IENSHIPGLEKPS
EU594335.1_Ae._comosa	IPISQQQQPPFSQQQQ
GQ870243.1_Ae._uniaristata	IPISQQQQPPFSQQQQ
GU942731.1_Ae._uniaristata	IPISQQQQPPFSQQQQ
GU942732.1_Ae._uniaristata	IPISQQQQPPFSQQQQ
AB062876-ISQQQ-A	ISQQQQPPFSQQQQ
EU594336.1_Ae._comosa	ISQQQQPPFSQQQQ
EU594337.1_Ae._comosa	ISQQQQPPFSQQQQ
GQ870242.1_Ae._umbellulata	ISQQQQPPFSQQQQ
GQ870244.1_Ae._uniaristata	ISQQQQPPFSQQQQ
GQ870245.1_Ae._markgrafii	ISQQQQPPFSQQQQ
EU594331.1_Ae._comosa	LPISQQQQPPFSQQQQ
EU594332.1_Ae._comosa	LPISQQQQPPFSQQQQ
EU594333.1_Ae._comosa	LPISQQQQPPFSQQQQ
EU594334.1_Ae._comosa	LPISQQQQPPFSQQQQ
EU594338.1_Ae._comosa	MATSCIPGSESPW
AB062868_MDT-A	MDTSCIPGLERPW
SL130-4	MENSHIPGLEKPS
SL130-5	MENSHIPGLEKPS
SL131-2	MENSHIPGLEKPS
SL133-1	MENSHIPGLEKPS
SL133-2	MENSHIPGLEKPS
SL133-3	MENSHIPGLEKPS
AB062862_MEN-D	MENSHIPGLERPS
Sh49-1	MENSHIPGLERPS
SL131-1	MENSHIPGLERPS
X13306_METRC-D	METRCIPGLERPW
EU329425.1_Ae._markgrafii	METSCIPGLERPW
EU571720.1_Ae._markgrafii	METSCIPGLERPW
EU571721.1_Ae._speltoides	METSCIPGLERPW
EU571722.1_Ae._triuncialis	METSCIPGLERPW
EU571723.1_Ae._triuncialis	METSCIPGLERPW
EU571724.1_Ae._uniaristata	METSCIPGLERPW
EU571725.1_Ae._umbellulata	METSCIPGLERPW
EU594339.1_Ae._comosa	METSCIPGLERPW
EU594340.1_Ae._comosa	METSCIPGLERPW
GQ870235.1_Ae._triuncialis	METSCIPGLERPW
GQ870236.1_Ae._triuncialis	METSCIPGLERPW
GQ870237.1_Ae._triuncialis	METSCIPGLERPW
GQ870238.1_Ae._triuncialis	METSCIPGLERPW
GQ870239.1_Ae._triuncialis	METSCIPGLERPW
GQ870240.1_Ae._umbellulata	METSCIPGLERPW
GQ870241.1_Ae._umbellulata	METSCIPGLERPW
GQ870250.1_Ae._markgrafii	METSCIPGLERPW
GQ980034.1_Ae._umbellulata	METSCIPGLERPW
GQ980035.1_Ae._umbellulata	METSCIPGLERPW
HQ220188.1_Ae._speltoides	METSCIPGLERPW
X51759_METSC-A	METSCIPGLERPW
S27-1	METSCIPSLERPW
S27-2	METSCIPSLERPW
S27-3	METSCIPSLERPW
S27-4	METSCIPSLERPW
S27-5	METSCIPSLERPW
Sh47-1	METSCIPSLERPW
Sh47-2	METSCIPSLERPW
Sh47-3	METSCIPSLERPW
Sh47-4	METSCIPSLERPW
Sh47-5	METSCIPSLERPW
Sh57-1	METSCIPSLERPW

Table 3 (*contd*)

Genes	N-terminla sequences
Sh57-3	METSCIPSLERPW
Sh57-4	METSCIPSLERPW
SL105-1	METSCIPSLERPW
SL105-2	METSCIPSLERPW
SL105-3	METSCIPSLERPW
SL106-1	METSCIPSLERPW
SL106-2	METSCIPSLERPW
SL106-3	METSCIPSLERPW
SL126-1	METSCIPSLERPW
SL129-3	METSCIPSLERPW
SL129-4	METSCIPSLERPW
SL130-1	METSCIPSLERPW
SL130-2	METSCIPSLERPW
SL130-3	METSCIPSLERPW
SL130-6	METSCIPSLERPW
SS138-1	METSCIPSLERPW
SS138-2	METSCIPSLERPW
SS138-3	METSCIPSLERPW
SS138-4	METSCIPSLERPW
SS139-1	METSCIPSLERPW
SS139-2	METSCIPSLERPW
SS151-1	METSCIPSLERPW
SS151-2	METSCIPSLERPW
SS151-3	METSCIPSLERPW
B062872_METSC_D	METSCISGLERPW
GQ389628.1_Ae._cylindrica	METSCISGLERPW
JQ726549.1_Ae._crassa	METSCISGLERPW
JQ726550.1_Ae._cylindrica	METSCISGLERPW
GQ389635.1_Ae._speltoides	METSHILSLEKPL
Sb47-1	METSHILSLEKPL
Sb47-2	METSHILSLEKPL
Sb47-3	METSHILSLEKPL
Sb47-4	METSHILSLEKPL
Sb47-5	METSHILSLEKPL
Sb47-6	METSHILSLEKPL
Sh47-6	METSHILSLEKPL
Sh47-7	METSHILSLEKPL
Sh49-2	METSHILSLEKPL
SL103-1	METSHILSLEKPL
SL103-2	METSHILSLEKPL
SL129-1	METSHILSLEKPL
SL131-3	METSHILSLEKPL
GQ389630.1_Ae._cylindrica	METSHIPGLEKPS
FJ824794.1_Ae._speltoides	METSHIPGLENPS
FJ824795.1_Ae._speltoides	METSHIPGLENPS
Y14104_METSHIP-B	METSHIPSLEKPL
AB062865_METSRV-D	METSRVPGLEKPW
GQ389633.1_Ae._markgrafii	METSRVPGLEKPW

Table 4. Alignment of amino acids of 15 representing typical N-terminal sequences.

	1	2	3	4	5	6	7	8	9	10	11	12	13
METSCIPGLERPW	M	E	T	S	C	I	P	G	L	E	R	P	W
MDTSCIPGLERPW	-	D	-	-	-	-	-	-	-	-	-	-	-
METRCIPGLERPW	-	-	-	R	-	-	-	-	-	-	-	-	-
METSCISGLERPW	-	-	-	-	-	-	S	-	-	-	-	-	-
METSCIPSLERPW	-	-	-	-	-	-	-	S	-	-	-	-	-
METRCVPGLERPW	-	-	-	R	-	V	-	-	-	-	-	-	-
MATSCIPGSESPW	-	A	-	-	-	-	-	-	S	-	S	-	-
METSHIPGLEKPS	-	-	-	-	H	-	-	-	-	-	K	-	S
METSHIPGLENPS	-	-	-	-	H	-	-	-	-	-	N	-	S
METSRVPGLEKPW	-	-	-	-	R	V	-	-	-	-	K	-	-
MENSHIPGLERPS	-	-	N	-	H	-	-	-	-	-	-	-	S
MENSHIPGLEKPS	-	-	N	-	H	-	-	-	-	-	K	-	S
IENSHIPGLEKPS	I	-	N	-	H	-	-	-	-	-	K	-	S
METSHILSLEKPL	-	-	-	-	H	-	L	S	-	-	K	-	L
METSHIPSLEKPL	-	-	-	-	H	-	-	S	-	-	K	-	L

Table 5. Cysteines in different LMW-GS.

	Sig	N-	Rep.	CI	CII	CIII	Total
MATSCIPGSESPW		1	0	5	1	1	8
METRCVPGLERPW		1	0	5	1	1	8
METRCIPGLERPW		1	0	5	1	1	8
METSCIPGLERPW ^a		1	0	5	1	1/2	8/9
METSCISGLERPW		1	0	5	1	1	8
METSCIPSLERPW		1	0	5	1	1	8
MDTSCIPGLERPW		1	0	5	1	1	8
METSRVPGLEKPW		0	1	5	1	1	8
METSHIPGLEKPS		0	1	5	1	1	8
METSHIPGLENPS		0	1	5	1	1	8
METSHILSLEKPL		0	1	5	1	1	8
METSHIPSLEKPL		0	1	5	1	1	8
MENSHIPGLEKPS ^a		0	1	5	1/2	1	8/9
IENSHIPGLEKPS		0	1	5	1	1	8
ISQQQ-A		0	0	6	1	1	8
(L)IPISQQQ-	1	0	1	5	1	1	9

^aGenes contain the 9th cysteine were found in the C-terminal domain of five clones of SL131-2, SL131-1, SL131-3, SL133-3, and SL133-1. Sig, signal peptide; N-, N-terminal domain; rep., repetitive domain. CI, CII and CIII indicate the CI, CII and CIII regions of C-terminal domain.

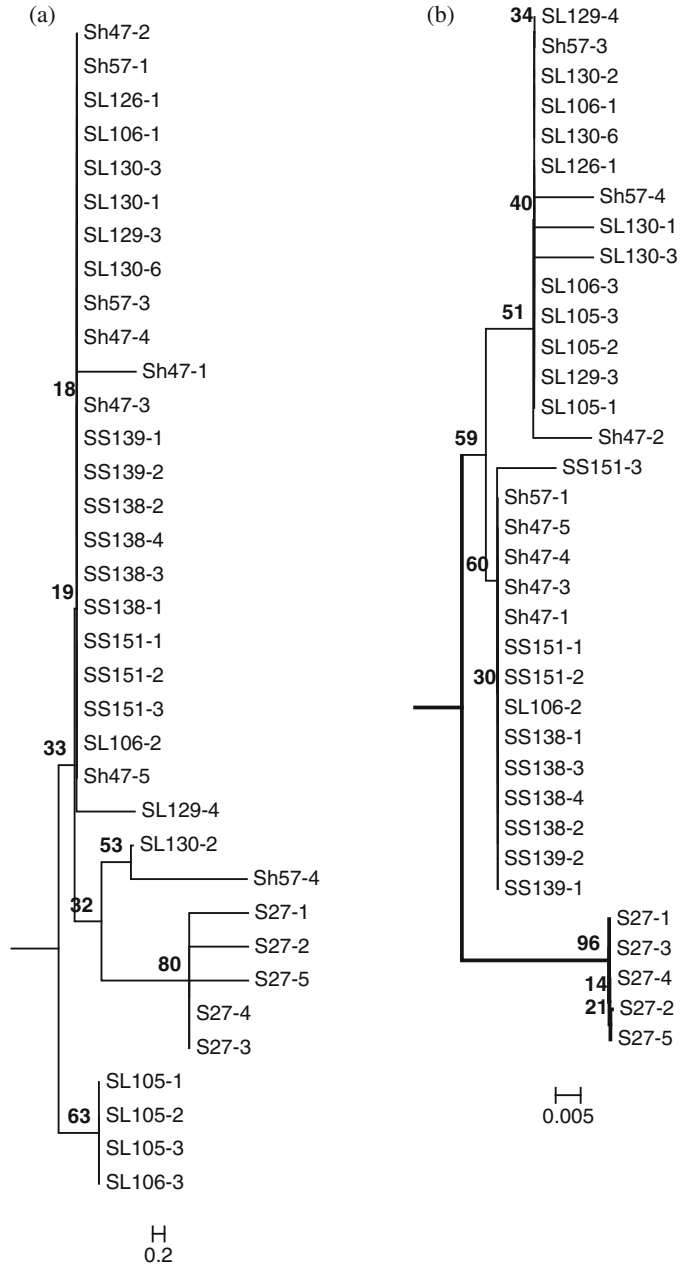


Figure 2. Subtrees correspond to those compressed in figures 1a and 3b.

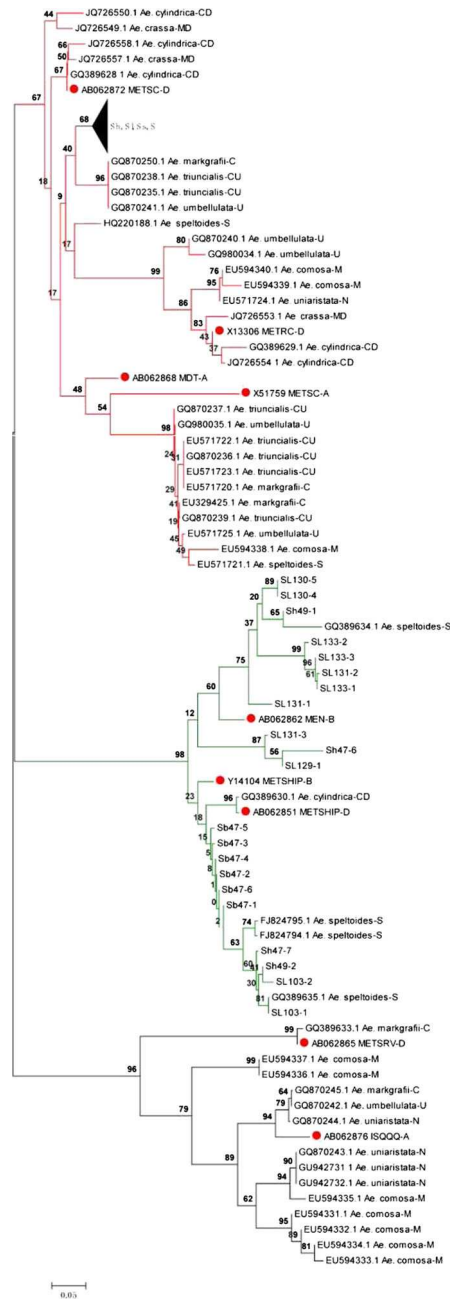


Figure 3. Phylogenetic analysis of LMW-GS based on deduced amino acids of repetitive domain. The representing typical LMW-GS in specific genomes of wheat are indicated by red dot. The compressed branch contains genes obtained in this study and is shown in figure 2. The scale bar indicates the distance scale.