

Supplementary material

Table S1. Genes and pseudogenes on the *T. monococcum* and *T. durum* contigs.

Gene ID ^a	<i>T. monococcum</i> (AY146588)	<i>T. durum</i> (AY146587)	Predicted function ^b	Supporting evidence ^c
1 (-), 3 (-), 9 (-)	<i>TmGluA3-2</i> (266981-267264) <i>TmGluA3-3</i> (274690-274791)	<i>TdGluA3-1</i> (64353-65519)	<u>TmGluA3-2</u> : LMW glutenin protein <i>T. aestivum</i> P93794 (E=e-53) <u>TmGluA3-3</u> : LMW glutenin protein <i>T. aestivum</i> P93794 (E=e-51) <u>TdGluA3-1</u> : LMW glutenin protein <i>T. durum</i> P16315 (E=4e-44)	<u>TmGluA3-2</u> : BU607215 <i>T. aestivum</i> (E=e-163) <u>TmGluA3-3</u> : BU607215 <i>T. aestivum</i> (E=e-151) <u>TdGluA3-1</u> : BF293077 <i>T. durum</i> (E=4e-177)
2 (-)	<i>TmHG-1</i> (108502-110998)	-	Unknown protein	RiceGAAS annotation system, BE492296 <i>T. monococcum</i> (E=4e-40)
4 (+)	<i>TmHG-2</i> (269228-271588)	-	Unknown protein	RiceGAAS annotation system, BJ285833 <i>T. aestivum</i> (E=e-147)
5 (+)	<i>TmPIK-1</i> ^ψ (266981-267264, 274690-274791)	-	85% similarity to a soybean phosphatidyl-inositol 3-kinase from (amino acid positions 502–608) P42347 (E= 2e-9)	-
6 (+),14 (+)	<i>TmRGL-1</i> ^ψ (276394-280055)	<i>TdRGL-1</i> (140783-141799)	<u>TmRGL-1</u> : Resistance complex protein I2C-2 <i>L. esculentum</i> O24016 (E=5e-68) <u>TdRGL-1</u> : A320-4 protein <i>O. sativa</i> Q9SXR0 (E=e-38)	<u>TmRGL-1</u> : CA501286 <i>T. aestivum</i> (E=e-113) <u>TdRGL-1</u> : -
7 (+)	<i>TmSTF-1</i> (284366-285442)	-	Sulfotransferase like gene <i>A. thaliana</i> Q9M1V2 (E=e-107)	BE493623 <i>T. durum</i> (E=e-107)
8 (+)	-	<i>TdHbox-1</i> ^{ψ*} (29674-29802)	Similarity with Glabra2 like-1 protein <i>A. thaliana</i> Q9LFW5 (E=5e-8)	-
10 (+)	-	<i>TdLRR1</i> ^ψ (77451-78390)	Similarity with Disease resistance protein like <i>A. thaliana</i> Q9FFS6 (E= 6 e-6)	BE591368 <i>T. aestivum</i> (E=e-168)
11 (+)	-	<i>TdHG-1</i> (102386-105314)	Unknown protein	RiceGAAS annotation system, BE497896 <i>T. aestivum</i> (E=6e-58)
12 (-)	-	<i>TdHG-2</i> (107822-109161)	Unknown protein	RiceGAAS annotation system, BQ788997 <i>T. aestivum</i> (E=2e-12)
13 (-)	-	<i>TdHG-3</i> (109559-116377)	Unknown protein	RiceGAAS annotation system, CA728026 <i>T. aestivum</i> (E=e-135)

^a Genes and pseudogenes numbered as in Figure 3 and their strand orientation.

^b Protein function is based on BLASTX searches. E value are given in parenthesis.

^c Evidence supporting the existence of the gene : Identification of identical or similar sequence in *Triticeae* ESTs (E value are given in parenthesis) or prediction by the RiceGAAS annotation system (Sakata et al., 2002).

^ψ Sequence determined as pseudogene.

* *TdHbox-1* corresponds to a sequence of 111bp with 70% homology to the protein sequence encoded by the last exon of the *Glabra2* like-1 protein. It has been found inside a gypsy retrotransposon (Fatima_107G22-2, Wicker et al., 2001). It is likely that it has been acquired by Fatima_107G22-2. Acquisition of segments of cellular genes by LTR retrotransposons has been previously described (Jin and Bennetzen, 1994; Elrouby and Bureau, 2001).

Table S2. Chronology of retrotransposon insertions in the LMW *Glu-A3* region

Element	bp aligned	Substitutions	MYA ^a	SD ^b
<i>WIS_426K20-1</i>	1713	22	0.99	0.21
<i>Angela_426K20-4</i>	1715	45	2.05	0.31
<i>WIS_107G22-2</i>	1440	39	2.12	0.34
<i>Angela_107G22-1</i>	1731	60	2.73	0.36
<i>Angela_426K20-5</i>	1730	65	2.97	0.37
<i>WIS_107G22-1</i>	1718	69	3.18	0.39
<i>Sabrina_107G22-3</i>	1216	55	3.60	0.49
<i>WHAM-1_453N11-1</i>	1324	65	3.92	0.49
<i>Sabrina_18B1-2</i>	1556	80	4.10	0.46
<i>Sabrina_453N11-2</i>	1564	97	4.98	0.51
<i>Wilma_426K20-3</i>	1386	94	5.49	0.57