

## Supplementary data:

# Characterization and phylogenetic analysis of $\alpha$ -gliadin gene sequences reveals significant genomic divergence in Triticeae species

Guang-Rong Li, Tao Lang, En-Nian Yang, Cheng Liu and Zu-Jun Yang

*J. Genet.* **93**, 725–731

**Table 1.** The  $\alpha$ - gliadin gene sequences in tested Triticeae species.

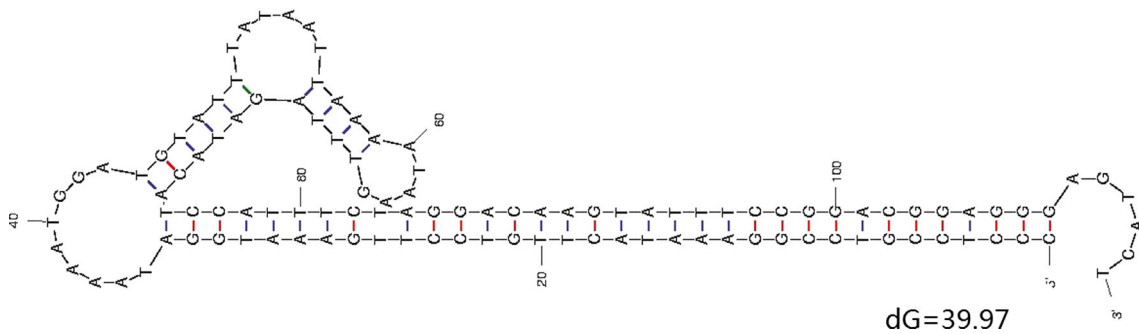
Species	USDA accession	Genome formula	Putatively functional (GenBank no.)	Pseudogene (GenBank no.)	Total
<i>S. sylvestre</i>	CN41954	R	HQ141102–HQ141111		10
<i>A. retrofractum</i>	PI553013	W	EU263386–EU263395	EU263396–EU263401	18
<i>A. desertum</i>	PI380621	P	EU026388–EU026396	EU026397–EU026403	16
<i>E. bonaepartis</i>	PI401345	F	HM452979–HM452986	HM452987–HM452994	16
<i>D. breviaristatum</i>	99008-17	V <sup>b</sup>	GU168578–GU168585	GU168586–GU168590	14
<i>D. villosum</i>	AS1365	V	EU325688–EU325697	EU325713–EU325719	16
<i>P. spicata</i>	PI232131	St	HM452964–HM452978 EU849647–EU849648	HM453022–HM453029 EU849653–EU849657	28
<i>L. elongatum</i>	PI153179	E	HM452951–HM452961 EU849641–EU849646	HM453003–HM453006 EU849651–EU849652	20
<i>T. bessaradicum</i>	PI431711	J	HM452962–HM452963	HM453007–HM453021	17
<i>T. intermedium</i>	PI440043	StJJ <sup>s</sup>	HM452946–HM452950 EU849650–EU849651	HM452995–HM452002 EU849658–EU849662	20
<i>D. breviaristatum</i>	PI546317	V <sup>b</sup> V <sup>b</sup>	EU186102–EU186108 EU325699–EU325707	EU325708–EU325719	28

```

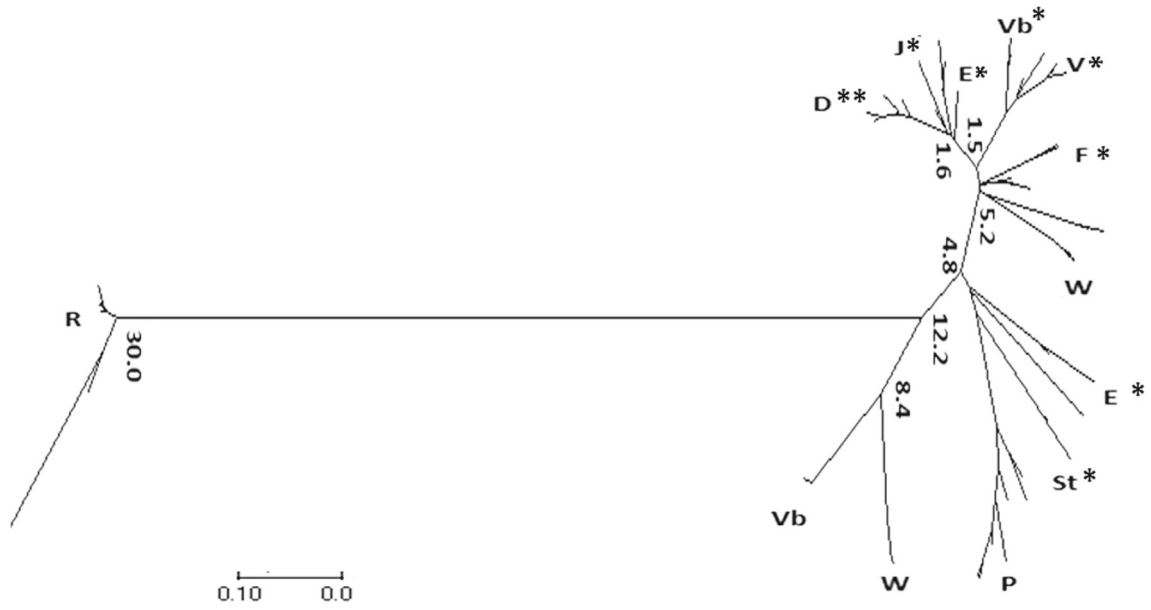
ACT TGT AAG TAG TGC CAC CAA ACA CAA CAT ACC AAA TTA TGA TTT CCG GAA GCA TCC AAG CAC TTT CCA GAC AAG CAA ATG CCA ATT GTG   90
AAA GAG ATC ATG CCA TGG CAG CTA TAA ATA GGC CCG TAG CAT GAT GAT CAT CCT TCC TCA TCC ATC ATT CTC ATA AGT AGA GCT CAT CAT   180
TCG AGC CAA GCA AGC AGT GGT CAA TAC AAA TCC ACC ATG AAG ACC TTT CTC ATC CTT GCC CTC CTT GCT ATC GTG GCG ACC ACC GCC ACA   270
                                     M K T F L I L A L L A I V A T T A T   18
ACC GCA GTT GGA GTT CCA GTG CCA CAA TTG CAG CCG CAA AAT CCA TCT CAG CAA CAA CCA CAA GAA CAA GTT CCA TTG ATG CAA CAA CAA   360
T A V G V P V P Q L Q P Q N P S Q Q Q P Q E Q V P L M Q Q Q   48
CAA TTT CCA GGG CAG CAA CAA CCA TTT TCA CCA CAA CAG CCA TAT CCG CAG CCG CAA CCA TTT CCA TCA CAA CAA CCA TAT CCG CAG CCG   450
Q F P G Q Q Q P F S P Q Q P Y P Q P Q P F P S Q Q P Y P Q P   78
CAA CCA TTT CCA CCA CAA CAA CTA TTT CCG CAG CCG CAG CCA TTT CCG CCA CAA CTA CCA TAT CCG CAG CCG CAG CCA TTT CCA CCA CAA   540
Q P F P P Q Q L F P Q P P P P Q L P Y P Q P P P P Q   108
CAA CCA TAT CCA CAA CCG CAA CAA CAG TAT CCG CAA CCA CAG CCA ATT TCG CAG CAA CAA GAA CAA CAA ATC CTT CAA CAA CTT CTG   630
Q P Y P Q P Q Q Q Y P Q P Q Q P I S Q Q Q E Q Q I L Q Q L L   138
CAA CAA CCG CTG ATT CCA TGC AGG CAT GTC GTC TTG CAA CAA CAC AAC ATA GCG CAT GGA AGC TCA CAA GTA TTG CAA CAA AGT AGT TAC   720
Q Q R L I P C R H V V L Q Q H N I A H G S S Q V L Q Q S S Y   168
CAG GAG TTG CAA CAA TTG TGT TGT CAG CAG CTG TGG CAG ATC CCT GAG CAG TCG CCG TGC CAA GCC ATC CAC AGT GTC GTT CAC GCT ATT   810
Q E L Q Q L C C Q Q L W Q I P E Q S R C Q A I H S V V H A I   198
ATT CTG CAT CAA CAA CAA CAA CCG TCA AGC CAG GGC TCC TTC CAA CAG CCT CAA CAA CAA TAT CCG TCA GGC CAG GGC TCC TTC CAG CCA   900
I L H Q Q Q Q P S S Q G S F Q Q P Q Q Q Y P S G Q G S F Q P   228
TCT CAG CAA AAC CCA CAG GGC CAG AGC TCT GTC CAG CCT CAA CAA CTA CCC CAG TTC GAG GAA ATA AGG AGG TTA GCG CTG CAG ACG CTA   990
S Q Q N P Q G Q S S V Q P Q Q L P Q F E E I R R L A L Q T L   258
CCA ACA ATG TGC AAT GTC TAC GTC CCT CCA TAT TGC TCG ACC ACC ATC GCG CCA TTT AGT AGC ATC GGT ATT AAC TGA GAA GAG AAG AAA   1080
P T M C N V Y V P P Y C S T T I A P F S S I G I N
TCT AGT ACT CCC TCC GTC CCG AAA TAC TTG TCC TTG AAA TGG ATA AAA TGG ATG TAT TTA TAA TTA AAA TAA GTT TAG ATA CAT CCA TTT   1170
CTA GGA CAA GTA TTT CCG GAC GGA GGG AGT ACT AGA TAT ATG GAA CAC CAT TTT TAG TAC ATG GTT TGG TCA TTG TAG CCG TGA AAA AAT   1260
AAA GTG ACA TGC ACT ATC ATG TAA GAA CCC GAC CTA TAC TAG TTC AAA CAT GGG   1314

```

**Figure 1.** The sequence HM452949 from *Thinopyrum intermedium*. The MITE insertion at the 3'UTR is boxed.



**Figure 2.** The secondary structure of MITE insertion in HM452949.



**Figure 3.** Average distance of  $\alpha$ -gliadin gene sequences among the Triticeae genomes. For details of genome symbols for each diploid accession see table 1. \* Represents the existence of Gli $\alpha$ - $\alpha$ , \*\* show the presence of four epitopes.