Barley Bmy1 gene intron III inter- and intra-haplotype variability

Sjakste T.G., Zhuk A.F.

in

Molina-Cano J.L. (ed.), Christou P. (ed.), Graner A. (ed.), Hammer K. (ed.), Jouve N. (ed.), Keller B. (ed.), Lasa J.M. (ed.), Powell W. (ed.), Royo C. (ed.), Shewry P. (ed.), Stanca A.M. (ed.).
Cereal science and technology for feeding ten billion people: genomics era and beyond

Zaragoza : CIHEAM / IRTA
Options Méditerranéennes : Série A. Séminaires Méditerranéens; n. 81

2008
pages 49-52

Barley *Bmy1* gene intron III inter- and intra-haplotype variability

T.G. Sjakste and A.F. Zhuk
Institute of Biology of the University of Latvia, Miera str. 3, LV 2169, Salaspils, Latvia
e-mail: tanja@email.lubi.edu.lv

Introduction

The highly polymorphic barley endosperm specific β-amylase gene (*Bmy1*) provides an excellent model for comparative genomic analysis of the allelic variants and evaluation of the mechanism of the gene structural rearrangements. Simple alignment of the available sequences of four allelic variants of the structural gene (Haruna Nijo: GenBank D4999; Adorra: GenBank AF061203; *H.spontaneum* strain PI 296897: GenBank AF061204; and the Finnish landrace line HA52: GenBank AJ301645), suggests huge number of polymorphic loci in coding and noncoding regions of the gene. Some of them could influence enzymatic properties or/and could be used for more precise variety discrimination (Eglinton *et al*., 1998; Erkkilä *et al*., 1998; Erkkilä and Ahokas, 2001; Kaneko *et al*., 2000; Ma *et al*., 2001; Paris *et al*., 2002; Polakova *et al*., 2003). Interesting data were accumulated on the intron III polymorphisms. 126bp insertion/deletion events (indel) in 5' region is associated with allelic variants of the gene encoding enzymes of low/high thermostability correspondingly (Erkkilä and Ahokas, 2001). Genotyping of the 3' region of the intron in 55 North European barley varieties revealed linkage between the 1+6 bp indel (S/L allele) and $C^{698}\rightarrow T$ (V233A) allelic variations of exon IV (Sjakste and Röder, 2004) responsible for the highest and the lowest thermostability of the enzyme correspondingly (Paris *et al*., 2002). High intron III microsatellite (MS) length polymorphism was identified in the varieties from the entire world (Sjakste and Röder, 2004, Malisheva *et al*., 2004). We described several linkage blocks on three polymorphic events (1+6 bp indel, MS size and $C^{698}\rightarrow T$ variation) that are transmitted through the generations of different independent pedigrees (Sjakste and Röder, 2004). Our analysis suggested the existence of certain associations between MS locus variability and other *Bmy1* polymorphisms. Therefore in our following work (Sjakste and Zhuk, 2006) we tried to obtain more general view on *Bmy1* intron III polymorphisms in different barley varieties and to develop approaches for more precise haplotype classification and characterization of inter- and intra-haplotype variability. Here we would like to present short up to date review of our data.

Genotyping of polymorphisms and haplotype classification

Data on *Bmy1* intron III genomic sequence polymorphism were generated among 20 Latvian and one Danish variety Maja, compared by sequence alignment to previously published data on the corresponding region of the gene of four allelic variants (Haruna Nijo, Adorra, *H. spontaneum* strain PI 296897, and Finnish landrace line HA52) and summarized in the Tables 1 and Table 2 (Sjakste and Zhuk, 2006) and other materials published in (Sjakste and Zhuk, 2006). Data on new β-amylase structural gene intron III nucleotide sequences and are available in the GenBank database (http://www.ncbi.nlm.nih.gov/) under accession numbers DQ316895 – DQ316905.

Our finding revealed two variable components of the repeated portion of the intron. Variable 5'-MS component of MS could be presented either by $(TG)_m$ (Haruna Nijo , *H.spontaneum* strain PI 296897, Maja, Latvijas Vietejie, and Abava-like Latvian accessions), either by the restructured motifs of Finnish landrace HA52, Adorra and Adorra-like Latvian accessions. The 5'-$(TG)_m$ repeat was shown to be followed immediately by a $(G)_n$ repeat in genes of Maja and Latvijas Vietejie similarly to Haruna Nijo, and *H. spontaneum* strain PI 296897. The 5'-TG-reach region is separated from the 3'-$(G)_n$ repeat by a TT motif in Finnish landrace HA 52, Adorra, and all Latvian accessions besides Latvijas Vietejie (Tables 1 and 2). High level of the 3'-$(G)_n$ repeat polymorphism as well as several single nucleotide polymorphisms (SNPs) in the 5'- and the 3'-MS flanks were revealed in accessions analyzed (Table 1).

Table 1. Sequence information (5' → 3') of the *Bmy1* MS region and formulas of the MS motif of five *Bmy1* intron III haplotypes. The numbering of the MS position is relative to the Haruna Nijo sequence (GenBank D49999). SNPs determined the differences in MS motif compared to HN haplotype are underlined

Haplotype	MS sequence of the accession featured the haplotype	Formula of the MS motif
Haruna Nijo like/HN	GA^{1964}TGTGTGTGTGTGTGTGTGGGGGGGGGGGGGGGGTG A^{2002}TG	$(TG)_m(G)_nTG$
H. *spontaneum* PI 296897like/HS	GA TGTGTGTGTGTGTGTG—GGGGGGGGGGGGGGGGGGTTA–G	$(TG)_8(G)_{19}TT$
HA52 like/HA	GA TGTGCCATTTTGTT————————GGGGGGGGGGGGGGGGGG ATG	$(TG)_2CCA(T)_3TGTT(G)_{17}$
Abava like/AB	GA TGTGTGTGTGTGTT———— GGGGGGGGGGGGGGGGGG ATG	$(TG)_6TT(G)_n$
Adorra like/AD	GA TGGGTGTGTGTGTT———————— GGGGGGGGGGGGGGATG	$TG(G)_2(TG)_4TT(G)_n$

Table 2. Description of the accessions including year of release, origin/breeding station, form of spike, β-amylase intron III haplotype, and details on MS polymorphism. Abbreviations LV and LLU indicates correspondingly Latvian origin and Latvian University of Agriculture. Data on MS size are indicated according to Sjakste and Röder (2004)

Cultivar	Year of release	Origin/breeding station	Form	Haplotype	MS region sequence data	MS size (bp)
Maja	1934	Denmark	2	HN	$(TG)_9(G)_{16}TG$	150
Latvijas Vietejie	-	LV	2	HN	$(TG)_{10}(G)_{18}TG$	154
Ansis	1995	LV/Stende	2	AD	$TG(G)_2(TG)_4TT(G)_{12}$	140
Balga	1990	LV/Priekuli	2	AD	$TG(G)_2(TG)_4TT(G)_{12}$	140
Kombainiers	1950	LV/Stende	2	AD	$TG(G)_2(TG)_4TT(G)_{12}$	140
Linga	1985	LV/Priekuli	2	AD	$TG(G)_2(TG)_4TT(G)_{12}$	140
Rasa	1991	LV/Stende	2	AD	$TG(G)_2(TG)_4TT(G)_{12}$	140
Sencis	1994	LV/Stende	2	AD	$TG(G)_2(TG)_4TT(G)_{14}$	140
Druvis	1999	LV/Stende	6	AD	$TG(G)_2(TG)_4TT(G)_{15}$	143
Malva	2001	LV/LLU	2	AD	$TG(G)_2(TG)_4TT(G)_{14}$	143
Idumeja	1998	LV/Priekuli	2	AD	$TG(G)_2(TG)_4TT(G)_{18}$	146
Imula	1985	LV/Stende	2	AD	$TG(G)_2(TG)_4TT(G)_{18}$	146
Priekulu 1	1954	LV/Priekuli	2	AD	$TG(G)_2(TG)_4TT(G)_{18}$	146
Ilga	1983	LV/Priekuli	2	AB	$(TG)_6TT(G)_{14}$	140
Klinta	1993	LV/LLU	2	AB	$(TG)_6TT(G)_{15}$	140
Abava	1978	LV/Stende	2	AB	$(TG)_6TT(G)_{16}$	143
Gate	1995	LV/Priekuli	2	AB	$(TG)_6TT(G)_{15}$	143
Dzintars	1930	LV/Stende	6	AB	$(TG)_6TT(G)_{16}$	143
Ruja	1992	LV/Priekuli	2	AB	$(TG)_6TT(G)_{16}$	143
Vairogs	1930	LV/Priekuli	6	AB	$(TG)_6TT(G)_{16}$	143
Agra	1984	LV/Priekuli	6	AB	$(TG)_6TT(G)_{18}$	146

Summarizing the data on MS motif polymorphism (in contrast to MS repeat number), we describe here all variations of the region by five formulas and classify them as HN (Haruna Nijo-like), HS (*H. spontaneum* strain PI 296897-like), HA (Finnish landrace HA52-like), AB (Latvian cultivar Abava-like), and AD (Adorra-like) MS variants (Table 1).

Five *Bmy 1* intron III haplotypes were classified according the MS motif as HN, HS, HA, AB, and AD haplotypes. Other polymorphisms detected including 16 indels, 38 SNPs, 3 double SNPs (dSNPs), one fragment substitution were used in precise haplotype characterization.

Based on polymorphisms of 59 loci and the specificity of the microsatellite motif, eleven Latvian

varieties turned out to have haplotype similar to cultivar Adorra, one – to Haruna Nijo, and eight – to the newly described Abava β-amylase intron III haplotype (Tables 1 and 2).

High level of repeat polymorphisms of $(TG)_m$ as well as $(G)_n$ component of the microsatellite repeat was revealed for all the haplotypes studied.

Interhaplotype variability

Each of intron III haplotypes analyzed possesses a unique spectrum of polymorphic loci allelic variants that could be used in haplotype discrimination. Simultaneously, some polymorphisms are common for several haplotypes. MS sequence motif and its flanking sequences are among the main features discriminated in all the analyzed haplotypes. Sequence polymorphisms of the 3' end of MS in contrast to MS repeat number was used earlier to differentiate two α-amylase gene families in crustaceans (Van Wormhould and Sellos, 2003). High variability of MS and its flanking sequences may be the common mechanism of the evolution of at least some amylase genes in different taxes.

Obviously, interhaplotype variability of intron III is the consequence of sequence structural reorganization during both natural evolution (HS compared to cultivated barley) and breeding process. Interestingly, Latvian barley accessions analyzed belong only to two haplotypes AB and AD. Including of the different allelic variants in the pedigrees of the Latvian barley cultivars did not result in its β-amylase intron III reorganization. These findings reflect stability of the transmittance of this genomic portion through generations in different crosses during approximately 70 years of breeding process in Latvia. It should be underlined that many fragments of intron III sequence of different length are conserved between all haplotypes studied, and may have more ancient origin compared to polymorphic loci.

Intrahaplotype variability and validity of MS markers

Variations in the number of repeat units of the (TG)m or/and (G)n components of MS revealed for HN, AB, and AD-like sequences result in apparent similarity in the size of repeated portions of the different haplotypes. This can give rise of mimicry of different haplotypes when exclusively the MS fragment size analysis is used for genotyping studies. As it was shown (Sjakste and Zhuk, 2006), approximately half of the Latvian varieties studied belong to AB-like haplotype, others to AD haplotype. Spectrum of the MS sizes of the corresponding sequences was similar in both haplotypes and can be considered as intrahaplotype variability of repeated portion of intron III. High level of intahaplotype variability in number of repeats was shown previously for MS of intron VI of the *Bmy1* gene of a crustacean *Litopenaeus vannamei* (Van Wormhoudt and Sellos, 2003). Our previous results presented in (Sjakste and Röder, 2004) provide a good example of mimicry of different haplotypes by the intrahaplotype variably in number of the repeated units. The precise MS region sequence analysis enabled us to find out real haplotype variants. Conclusion on the absence of the association between *Bmy1* intron III MS repeat number and the C^{698} mutation in exon IV (Sjakste and Röder, 2004) was confirmed by the sequencing results. Indeed, not MS length (MS repeat number) but MS sequence motif itself is the main locus trait that correlates with mutations in coding region of the *Bmy 1* gene (Sjakste and Zhuk, 2006).

Taken into account intensive application of microsatellite markers in basic and applied research, we stress the necessity to test the applicability of MS marker size as sole parameter of genomic region polymorphism in every particular association and linkage disequilibrium study.

References

Eglinton, J.K., Langridge, P. and Evans, D.E. (1998). *J Cereal Sci*., 28: 301-309.
Erkkilä, M.J. and Ahokas, H. (2001). *Hereditas,* 134: 91-95.
Erkkilä, M.J., Leah, R., Ahokas, H. and Cameron-Mills, V. (1998). *Plant Physiol*., 117: 679–685.
Kaneko, T., Kihara, M. and Ito, K. (2000). *Plant Breed.,* 119: 197-201.
Ma, Y.F., Evans, D.E., Logue, S.J. and Langridge, P. (2001). *Mol. Genet. Genomics,* 266: 345-352.

Malisheva, L., Ganal, M.W. and Röder, M.S. (2004). *Plant Breed.,* 123: 128-131.
Paris, M., Jones, M.G.K. and Eglinton, J.K. (2002). *Plant Mol. Biol. Rep.,* 20: 149-159.
Polakova, K., Laurie, D., Vaculova, K. and Ovesna, J. (2003). *Plant Mol. Biol. Rep*., 21: 439-447.
Sjakste, T. and Röder, M. (2004). *Hereditas,* 141: 39-45.
Sjakste, T.G. and Zhuk, A.F. (2006). *Theor. Appl. Genet*., 113: 1063-1079.
Van Wormhoudt, A. and Sellos, D. (2003). *J Mol. Evol*., 57: 659–761.