AFLP and SSR Markers Linked to the Yellow Seed Colour Gene in *Brassica juncea* L.

Zhen HUANG, Yu ZHANG, Hui Qiang LI, Li YANG, Yuan Yuan BAN, Ai Xia XU and En Shi XIAO

College of Agronomy, Northwest A&F University, Yangling, Shaanxi, China

Abstract: Yellow mustard, cultivated in northern Shaanxi of China, is a valuable germplasm of *Brassica juncea* with low erucic acid content. Its yellow seed colour is controlled by a recessive allele of a single gene, whose dominant allele conditions brown seed colour. To map the yellow seed colour allele, amplified fragment length polymorphism (AFLP) and simple sequence repeat (SSR) technologies were used to identify markers linked to the recessive allele. The analysis was done on 386 F_2 plants, segregating for seed colour, from the cross Wuqi yellow mustard × Wugong mustard. The plants were selfed to determine their seed colour genotype. Twenty AFLP markers and eight SSR markers were identified from 256 AFLP primer combinations and 624 pairs of SSR primers, respectively. Blast analysis indicated that the sequences of four closely linked AFLP and SSR markers showed good collinearity with those of *Arabidopsis* chromosome 3, and the homologue of the yellow seed colour allele was located between *At3g14190* and *At3g32130*. Sequence information of the region between the two genes of *Arabidopsis* could be used to develop more closely linked markers to narrow down the homologue of the yellow seed colour allele. These results would accelerate the procedure of yellow seed colour gene cloning and marker-assisted selection for yellow mustard

Keywords: *Brassica juncea* L.; comparative analysis; fine mapping; yellow seed colour allele

High oil content is a very important goal for rapeseed production. Compared with brown-seeded rapeseed, yellow-seeded rapeseed has some advantages, such as higher oil and protein contents, lower fibre content and good quality characteristics. Yellow-seeded rapeseed is considered to be useful for improving the oil content, protein content and quality of rapeseed oil (Liu *et al*. 1991). However, stable lines of yellow-seeded rapeseed are rare in *Brassica napus* L. and the yellow-seeded rapeseed germplasm is very limited worldwide as well. Yellow seed colour genes with genetic stability in *B. rapa* L. are not stably expressed, once transferred to *B. napus* (Zhang *et al*. 2009b). In China, there are many yellow-seeded types of *B. juncea* L. whose trait of yellow seedness is genetically stable. Therefore, they are good materials for studying the inheritance of yellow seedness trait in yellow-seeded rapeseed. Meanwhile, *B. juncea* has a number of valuable agronomic characteristics, such as drought tolerance, high temperature resistance, tolerance to lean soil, shattering resistance, disease resistance, heavy metal tolerance, early maturity and late sowing tolerance (Downey 1990; Woods *et al*. 1991). Besides, yellow mustard is becoming one of the major oil crops in the western arid region of China. In recent years, mustard has been transformed into the type of canola and is becoming increasingly popular in North America and Australia (Burton *et al*. 2003; Rakow 2003).

Found in northern Shaanxi, yellow mustard is a type of *B. juncea* widely distributed in west-

ern China. Previous studies showed that yellow seedness is controlled by a recessive allele of a single gene. Yellow mustard has specific fatty acid composition of about 25% erucic acid and high oil content of 48% (Huang *et al*. 1999; Xu *et al*. 2010), and some well-performing varieties have been developed by now.

Molecular markers are used to map the yellow seed colour genes in different species. In *B. rapa*, Rahman *et al*. (2007) developed a SRAP marker and successfully converted it into single nucleotide polymorphism (SNP) marker and sequence characterized amplified region (SCAR) markers. Zhang *et al*. (2009a) successfully mapped and cloned one yellow seed colour gene in *B. rapa* and found that *TTG1* in *Arabidopsis thaliana* L. (Heynh.) was one candidate of the yellow seed colour gene. In *B. juncea*, Xu *et al*. (2010) identified 15 AFLP markers linked to the brown seed colour gene in *B. juncea*, and converted four of them into SCAR markers. Negi *et al*. (2000) converted an AFLP marker into a co-dominant SCAR marker (SCM08) linked to the seed colour gene in *B. juncea*, and amplified three polymorphic bands: 0.5 and 1.2 kb bands, which were linked to the yellow seed colour allele, and a 1.1 kb band, which was linked to the brown seed colour allele. Recently, some additional molecular markers linked to yellow seed colour gene were reported in *B. juncea* (Lionneton *et al*. 2004; Lakshmi *et al*. 2005). In *B. napus*, Liu *et al*. (2005) and Xiao *et al*. (2007) identified some AFLP and SSR markers linked to the yellow seed colour gene in resynthesized *B. napus* line No. 2127-17, and mapped the gene to chromosome A9/N9 of *B. napus*.

Many genetic maps of yellow seed colour genes are constructed, but there are very few reports on yellow seed colour gene cloning. A high-resolution genetic map is the foundation for gene cloning, and genomic information on *Arabidopsis thaliana* and *Brassica* is helpful for such gene mapping. Recently, Xu *et al*. (2010) reported markers linked to a dominant brown seed colour allele in yellow mustard distributed in northern Shaanxi, however, the marker density of the reported map was not high enough to clone the gene underlying this trait. The aim of this study was to identify some other molecular markers linked to the yellow seed colour allele in yellow mustard, construct its high-resolution genetic map, and identify an *Arabidopsis* collinear region that probably carries the homologue of the yellow seed colour gene.

MATERIALS AND METHODS

Plant materials and population construction

Wuqi yellow mustard, a *B. juncea* landrace with light-yellow seed that grows on the loess plateau in northern Shaanxi, was crossed with Wugong mustard, a brown-seeded landrace. An F_2 population was obtained by selfing a single F_1 plant; each F_2 individual was selfed to generate an $F_{2:3}$ population. The genotype of each F_2 individual was judged in terms of the phenotypes of the F_3 population. Seeds were harvested from self-pollinated plants and their colours were observed and recorded. Seeds presenting pure yellow were judged as yellow seeds, and seeds presenting other or mixed colours were judged as brown seeds. There were 386 individuals in the $F_{2:3}$ population which were used to identify the markers linked to the yellow seed colour allele.

DNA extraction and bulked segregant analysis

Genomic DNA was extracted by the CTAB method (Doyle & Doyle 1990). Young leaves of F_2 individuals were taken at the seedling stage, and equal amounts of DNA from ten yellow-seeded F_2 plants and ten brown-seeded F_2 plants were pooled to form the yellow seed colour bulk (the genotype was *bb*) and the brown seed colour bulk (the genotype was *BB*), respectively. The DNA concentrations were adjusted to 50 ng/ul.

AFLP and SSR analysis

Genomic DNA samples were digested with *EcoRI*/*MseI*. The adaptor ligation and two successive PCR reactions were followed by the AFLP method described by Vos *et al*. (1995). The sequences of pre-amplified primers and the selective amplified primers could be seen in Huang *et al*. (2007). The SSR amplification was performed as described by Lowe *et al*. (2002). Sequences of all the SSR markers were obtained from public sources: http://ukcrop.net/perl/ace/search/BrassicaDB (Lowe *et al*. 2004), http://www. brassica.info/ssr/SSRinfo.htm (prefixes: Ra, Na, BN, and BRMS), the primer pairs with the prefixes BRAS and CB were from the electronic supplementary material of Piquemal *et al*. (2005). PCR products were detected by silver staining (Lu *et al*. 2001).

Linkage analysis

The specific AFLP fragments that were present only in yellow-seeded individuals, but absent in brown-seeded individuals, were regarded as recessive allele linked markers. The individual marker data and phenotypes were processed with MAPMAKER/EXP 3.0 (Lander *et al*. 1987; Lincoln *et al*. 1992). A minimum LOD score of 3.0 was adopted for the map construction. The map distances were calculated using Kosambi's (1944) mapping function.

Sequence similarity analysis with *Arabidopsis* genome

In order to identify a putative syntenic region around the yellow seed colour allele in the *Arabidopsis* genome, the six closest markers (CB10022, EA10MC09, EA03MC03, EA14MC03, EA07MC06-100 and EA09MC11) from the yellow seed colour allele were cloned and sequenced according to the method described by Cho *et al*. (1996). The sequence similarity between molecular markers and *Arabidopsis* genome was analysed using the BLAST programs of the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/). Then the linear arrangement relationship of these molecular markers in the *B. juncea* genome was compared with those of their homologues in the *Arabidopsis* genome to determine whether synteny occurs around the seed colour genes in these two related species.

Table 1. Primers and fragment sizes of AFLP markers linked to the yellow seed colour allele in *Brassica juncea*

AFLP markers	Primers	Size of markers (bp)
EA02MC10	EA+AT/MC+CT	150
EA03MC03	EA+AC/MC+AC	100
EA04MC15	EA+AG/MC+GC	180
EA04MC15	EA+AG/MC+GC	200
EA05MC04	EA+TA/MC+AG	300
EA06MC09	EA+TT/MC+CA	90
EA06MC11	EA+TT/MC+CC	150
EA07MC06	EA+TC/MC+TT	100
EA07MC06	EA+TC/MC+TT	200
EA08MC12	EA+TG/MC+CG	200
EA09MC03	EA+CA/MC+AC	90
EA09MC11	EA+CA/MC+AC	100
EA10MC09	EA+CT/MC+CA	200
EA11MC08	EA+CC/MC+TG	100
EA12MC08	EA+CG/MC+TG	100
EA12MC11	EA+CG/MC+CC	100
EA13MC12	EA+GA/MC+CG	200
EA14MC03	EA+GT/MC+AC	300
EA14MC12	EA+GT/MC+CG	150
EA15MC01	EA+GC/MC+AA	100

RESULTS

Genetic analysis of the yellow-seeded gene

The F_2 population consisted of 386 individuals, which included 101 homozygous brown-seeded plants, 193 heterozygous brown-seeded plants and 92 yellow-seeded plants. The segregation ratio of the three genotypes was consistent with the expected ratio of 1:2:1 ($\chi^2 = 0.42$, and $P > 0.05$), confirming the hypothesis that only one seed colour gene segregated in the F_2 population.

Identification of molecular markers linked to the yellow-seeded gene

In order to identify the markers linked to the yellow seed colour allele, 256 pairs of selective primer combinations of EA + 2/MC + 2 and 624 pairs of SSR primers were screened in the two bulks, i.e. the yellow seed colour bulk and the brown seed colour bulk. The primers that showed reproducible polymorphism were used to test twenty individuals of the two bulks. As a result, 20 positive AFLP markers and eight SSR markers present in the yellow seed colour individuals but absent in the brown seed colour individuals were identified (Tables 1 and 2).

Linkage analysis

All of the markers identified were tested in the F_2 population composed of 386 individuals. These markers were able to detect polymorphism in the

Table 2. Oligonucleotide sequences and fragment sizes of SSR markers linked to the yellow seed colour allele in *Brassica juncea*

SSR marker	Sequence (5'–3')	Size (bp)
CB10299	TACAGGTTCCTTGCGATG ATGGACGAGACAACATGG	100
CB10065	CGGCAATAATGGACCACTGG CGGCTTTCACGCAGACTTCG	170
CB10022	AACAACCAAACATAGTCCC GTTGACTTTGACCTTGACTT	230
CB10501	GTACCAGCCGGTTATCAA CGATGGAGTGGAAGTGAG	230
CB10336	CAAAACACCCAATTCTCG GTGGTTGGTTCAGCTTTG	180
CB10504	GGTGTCCCAACTGTTGAA CATTGGCATAGGAACAGG	100
CB10330	AGGCGAGTTTACGAGGAT ACCTGCACCAGTCATTTG	100
CB10415	GAACTCGTCGCGGTAGTA TCTCTTTCCTCGCAGATG	500

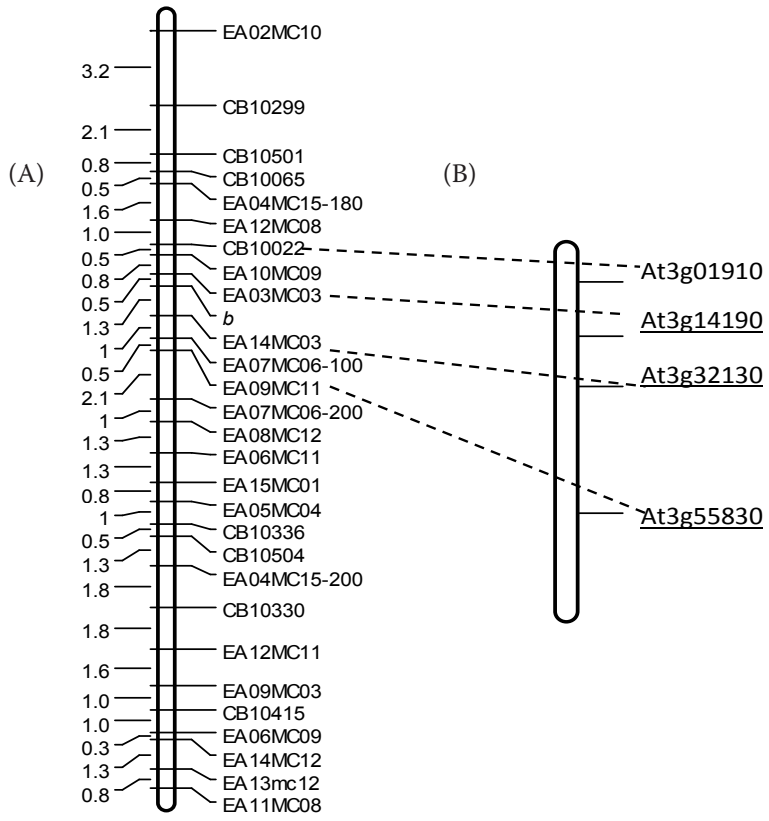

Figure 1. (A) A linkage map of the region surrounding the yellow seed colour allele (*b*) from the $F_{2:3}$ population of *Brassica juncea*; (B) A partial physical map of chromosome 3 of *Arabidopsis thaliana* showing the homologues of the mapped marker sequences; the dotted lines indicate the relationship of the two maps

F_2 population, and all was linked to the gene controlling yellow seeds. Twenty AFLP and eight SSR markers were mapped to a region of 32.7 cM around the yellow seed colour allele and their average distance from the gene was 1.1 cM. EA03MC03 and EA14MC03 were the two closest markers to the gene, which were on either side of the gene with their distances of 0.5 and 1.3 cM, respectively (Figure 1).

Sequence similarity analysis with *Arabidopsis* genome

The six most closely linked markers on either side of the gene were cloned and sequenced. Their sequences were submitted to the NCBI website for BLAST search. The sequences of four of them (CB10022, EA03MC03, EA14MC03 and EA09MC11) were found to have highly conserved homologues in the *Arabidopsis* genome. They were distributed in a large region from *At3g14190* to *At3g55830* on chromosome 3 of *Arabidopsis*. Furthermore, the arrangement order of the four molecular markers linked to the yellow seed colour allele was consistent with that of their homologues in *Arabidopsis* (Figure 1). The homologues of the two closely linked markers in *Arabidopsis* genome were *At3g14190* and *At3g32130*, respectively, which indicated that the homologue of the gene was between *At3g14190* and *At3g32130*.

DISCUSSION

In *B. napus*, the inheritance of yellow seedness as a trait is complex due to intra- and inter-genomic actions concerned, and this creates a barrier to study the yellow seedness mechanism. It is an effective way to employ yellow-seeded materials with a simple genetic basis to study a seed colour mechanism. Yellow seedness as a trait of yellow mustard found in northern Shaanxi is controlled by a single gene and the genetic effect of the gene is very stable, so the map-based cloning strategy can be adopted to map the gene. Xu *et al*. (2010) constructed a primary genetic map around the dominant brown seed colour allele, but the density of genetic map markers was not very high. More information is needed to finely map the gene. The brown seed colour allele and yellow seed colour allele have the same chromosomal location. When the number of markers linked to the brown seed colour allele is limited, we can develop yellow seed colour allele linked markers. So it is a promising strategy to use markers linked to these two alleles to increase the genetic map density, and to narrow the distance between markers of the target gene. In fact, there are many reports on a successful use of this method (Huang *et al*. 2007; Wang *et al*. 2007; He *et al*. 2008).

Although many results of mapping yellow seed colour genes have been reported, few such genes have been cloned up to now. One reason for this is the genetic complexity of the yellow seed colour trait and the other reason probably is that insufficient genome information has been available. The collinearity between *Brassica* and *Arabidopsis* genomes in many regions is very perfect (Parkin *et al*. 2005; Babula *et al*. 2003). So we could first determine the homologous region of the yellow seed colour allele in *Brassica* genome or *Arabidopsis* genome, and then we used the sequence information on the candidate region to develop closer markers linked to the yellow seeded gene, and to gradually narrow down the homologous region of the gene. In our research, the homologous region carrying the yellow seed colour allele in *Arabidopsis* was identified and it is located between *At3g14190* and *At3g32130*. We can make use of the sequences of this homologous region to develop new markers to finely map the gene. The closely linked markers obtained in our research will also facilitate the yellow-seeded mustard breeding.

Acknowledgments. The study was supported by Natural Science Foundation of Shaanxi (2009JQ3003), National Program 863 (2009AA101105), Tang Zhongying Breeding Foundation of the Northwest A&F University (A212020905, A212020523), Public Service Sectors (agriculture) Special Funds (200903002-04), Earmarked Fund for China Agriculture Research System (CARS-13) and Special Funds for Basic Research of Northwest A&F University (QN2011085).

References

Babula D., Kaczmarek M., Barakat A., Delseny M., Quiros C.F., Sadowski J. (2003): Chromosomal mapping of *Brassica oleracea* based on ESTs from *Arabidopsis thaliana*: complexity of the comparative map. Molecular Genetics and Genomics, **268**: 656–665.

Burton W., Salisbury P., Potts D. (2003): The potential of canola quality *Brassica juncea* as an oilseed crop

for Australia. In: Proc. 11th Int. Rapeseed Congress. Volume 1. Copenhagen, 5–7.

Cho Y.G., Blair W.M., Panaud O., McCouch S.R. (1996): Cloning and mapping of variety-specific rice genomic DNA sequences: amplified fragment length polymorphisms (AFLP) from silver stained polyacrylamide gels. Genome, **39**: 373–378.

Downey R.K. (1990): *Brassica* oilseed breeding – achievements and opportunities. Plant Breeding, **60**: 1165–1170.

Doyle J.J., Doyle J.L. (1990): Isolation of plant DNA from fresh tissue. Focus, **12**: 13–15.

He J.P., Ke L.P., Hong D.F., Xie Y.Z., Wang G.C., Liu P.W., Yang G.S. (2008): Fine mapping of a recessive genic male sterility gene (*Bnms3*) in rapeseed (*Brassica napus*) with AFLP- and *Arabidopsis*-derived PCR markers. Theoretical and Applied Genetics, **117**: 11–18.

Huang J.Y., Xu A.X., Ma C.Z., Wang S.Z. (1999): Study on the yellow mustard and the ecologic feature of its distribution zone in North Shaanxi. Journal of Northwest Sci-Tech University of Agriculture and Forestry, **6**: 33–37.

Huang Z., Chen Y.F., Yi B., Xiao L., Ma C.Z., Tu J.X.. Fu T.D. (2007): Fine mapping of the recessive genic male sterility gene (*BnMs3*) in *Brassica napus* L. Theoretical and Applied Genetics, **115**: 113–118.

Kosambi D.D. (1944): The estimation of map distances from recombination values. Annals of Eugenics, **12**: 172–175.

Lakshmi P.K., Arumugam N., Gupta V., Mukhopadhyay A., Sodhi Y.S., Pental D., Pradhan A.K. (2005): Mapping and tagging of seed coat colour and the identification of microsatellite markers for marker-assisted manipulation of the trait in *Brassica juncea*. Theoretical and Applied Genetics, **111**: 8–14.

Lander E.S., Green P., Abrahamson J., Barlow A., Daly M.J., Lincoln S., Newburg L. (1987): MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics, **1**: 174–181.

Lincoln S., Daly M., Lander E. (1992): Constructing genetic maps with MAPMAKER/EXP 3.0. In: Whitehead Technical Institute Report. 3rd Ed., Whitehead Technical Institute, Cambridge.

Lionneton E., Aubert G., Ochatt S., Merah O. (2004): Genetic analysis of agronomic and quality traits in mustard (*Brassica juncea*). Theoretical and Applied Genetics, **109**: 792–799.

Liu H.L., Han J.X., Hu X.J. (1991): Studies on the inheritance of seed coat colour and other related characteristics of yellow seeded *Brassica napus*. In: Proc. 8th Int. Rapeseed Congress. Volume 5. Saskatoon, 1438–1444.

Liu Z.W., Fu T.D., Wang Y., Tu J.X., Chen B.Y. (2005): Inheritance of seed colour and identification of RAPD and AFLP markers linked to the seed colour gene in rapeseed (*Brassica napus* L.). Theoretical and Applied Genetics, **110**: 303–310.

Lowe A.J., Jones A.E., Raybould A.F., Trick M., Moule C.J., Edwards K.J. (2002): Transferability and genome specificity of a new set of microsatellite primers among *Brassica* species of the U triangle. Molecular Ecology Notes, **2**: 7–11.

Lowe A., Moule C., Trick M., Edwards K. (2004): Efficient large-scale development of microsatellites for marker and mapping applications in *Brassica* crop species. Theoretical and Applied Genetics, **108**: 1103–1112.

Lu G.Y., Yang G.S., Fu T.D. (2001): Silver stained AFLP – a novel assay for DNA fingerprinting in *Brassica napus*. Journal of Huazhong Agricultural University, **20**: 413–415. (in Chinese)

Negi M.S., Devic M., Delseny M., Lakshimikumaran M. (2000): Identification of AFLP fragments linked to seed coat colour in *Brassica juncea* and conversion to a SCAR marker for rapid selection. Theoretical and Applied Genetics, **101**: 146–152.

Parkin I.A.P., Gulden S.M., Sharpe A.G., Lukens L., Trick M., Osborn T.C., Lydiate D.J. (2005): Segmental structure of the *Brassica napus* genome based on comparative analysis with *Arabidopsis thaliana*. Genetics, **171**: 765–781.

Piquemal J., Cinquin E., Couton F. (2005): Construction of an oilseed rape (*Brassica napus* L.) genetic map with SSR markers. Theoretical and Applied Genetics, **111**: 1514–1523.

Rahman M., McVetty P.B.E., Li G.Y. (2007): Development of SRAP, SNP and multiplexed SCAR molecular markers for the major seed coat color gene in *Brassica rapa* L. Theoretical and Applied Genetics, **115**: 1101–1107.

Rakow G. (2003): The potential of *Brassica juncea*, *B. carinata* and *Sinapis alba* as new oilseed crops in Canada. In: Proc. 11th Int. Rapeseed Congress.

Vos P., Hogers R., Bleeker M., Reijians M., Vande Lee T., Hornes M., Frijters A., Pot J., Peleman J., Kuiper M., Zabeau M. (1995): AFLP: a new technique for DNA fingerprinting. Nucleic Acids Research, **23**: 4407–4414.

Wang G.C., He J.P., Hong D.F., Xie Y.Z., Xu Z.H., Liu P.W., Yang G.S. (2007): Development of AFLP and SCAR markers linked to a recessive genic male sterile gene (*Bnms3*) in rapeseed for marker-assisted selection. Korean Journal of Genetics, **29**: 481–487.

Woods D.L., Capcara J.J., Downey R.K. (1991): The potential of mustard (*Brassica juncea*) as an edible

crop on the Canadian prairies. Canadian Journal of Plant Science, **71**: 195–198.

Xiao S.S., Xu J.S., Li Y., Zhang L., Shi S.J., Shi S.W., Wu J.S., Liu K.D. (2007): Generation and mapping of SCAR and CAPS markers linked to the seed coat color gene in *Brassica napus* using a genome-walking technique. Genome, **50**: 611–618.

Xu A.X., Huang Z., Ma C.Z., Xiao E.S., Tian G.W., Zhang X.S., Tu J.X., Fu T.D.., Zhang G.S. (2010): Inheritance of seed colour and molecular markers linked to the seed colour gene in *Brassica juncea*. Molecular Breeding, **25**: 57–65.

Zhang J.F, Lu Y., Yuan Y.X., Zhang X.W., Geng J.F., Chen Y., Cloutier S., McVetty P.B.E., Li G.Y. (2009a): Map-based cloning and characterization of a gene controlling hairiness and seed coat color traits in *Brassica rapa*. Plant Molecular Biology, **69**: 553–563.

Zhang Y., Li X., Ma C.Z., Shen J.X., Chen B.Y., Tu J.X., Fu T.D. (2009b): The inheritance of seed color in a resynthesized *Brassica napus* line No. 2127-17 including a new epistatic locus. Genes & Genomics, **31**: 413–419.

Received for publication September 1, 2011
Accepted after corrections November 29, 2011

Corresponding author:

Ai Xia Xu, PhD., College of Agronomy, Northwest A&F University, Yangling, Shaanxi 712100, China
e-mail: xuaixia64@yahoo.com.cn