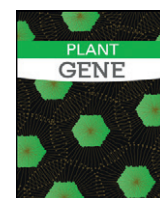


Contents lists available at [ScienceDirect](http://ScienceDirect.com)

Plant Gene

journal homepage: www.elsevier.com/locate/plant-gene

Genetic distance of inbred lines of Chinese cabbage and its relationship to heterosis



Kazutaka Kawamura^{a,1}, Takahiro Kawanabe^{b,1}, Motoki Shimizu^a, Atsushi J. Nagano^{c,d,e}, Natsumi Saeki^a, Keiichi Okazaki^a, Makoto Kaji^f, Elizabeth S. Dennis^g, Kenji Osabe^h, Ryo Fujimoto^{b,d,*}

^a Graduate School of Science and Technology, Niigata University, Ikarashi-nincho, Niigata 950-2181, Japan

^b Graduate School of Agricultural Science, Kobe University, Rokkodai, Nada-ku, Kobe 657-8501, Japan

^c Faculty of Agriculture, Ryukoku University, Yokatani 1-5, Seta Ohe-cho, Otsu-shi, Shiga 520-2194, Japan

^d JST PRESTO, Honcho 4-1-8, Kawaguchi, Saitama 332-0012, Japan

^e Center for Ecological Research, Kyoto University, Hirano 509-3-2, Otsu, Shiga 520-2113, Japan

^f Watanabe seed Co., Ltd, Machiyashiki, Misato-cho, Miyagi 987-0003, Japan

^g CSIRO Agriculture, Canberra, ACT 2601, Australia

^h Plant Epigenetics Unit, Okinawa, Institute of Science and Technology Graduate University, Onna-son, Okinawa 904-0495, Japan

ARTICLE INFO

Article history:

Received 8 August 2015

Received in revised form 15 October 2015

Accepted 23 October 2015

Available online 10 November 2015

Keywords:

Brassica rapa

F₁ hybrid cultivar

Marker assisted selection

Heterosis

Genetic distance

ABSTRACT

Chinese cabbage (*Brassica rapa* L. var. *pekinensis*) is an important vegetable in Asia. Most Japanese commercial cultivars of Chinese cabbage use an F₁ hybrid seed production system because of the high yielding cultivars produced. An efficient method for predicting hybrid performance in the parental generations is desired, and genetic distance between parental lines might be a good indicator of the level of hybrid vigor in a cross. Information concerning the genetic relationships among parental candidate inbred lines is useful for variety protection. The number of DNA markers available that can be used to assess the purity of inbred lines is limited in *B. rapa*. The aim of this study is to use DNA markers to assess the genetic distance between inbred lines to examine early developmental and yield heterosis so as to develop methods for selecting the best parental lines for the production of hybrids. We screened highly polymorphic SSR and CAPS markers to assess the genetic uniformity of inbred lines and characterize their genetic relationship. We examined the early size and yield heterosis in 32 F₁ hybrids of Chinese cabbage. There was a moderate correlation in mid-parent heterosis between leaf size at 21 days after sowing and harvested biomass but not in best-parent heterosis. In contrast there was no correlation between genetic distance and mid-parent or best-parent heterosis, indicating that genetic distance does not predict the heterosis phenotype.

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Brassica rapa L. comprises commercially important vegetable crops consumed worldwide such as leafy vegetables including Chinese cabbage (var. *pekinensis*), pak choi (var. *chinensis*), and komatsuna (var. *perviridis*), root vegetables including turnip (var. *rapa*), and oilseed (var. *oleifera*). Chinese cabbage forms a head with large pale-green colored leaves and wide white midribs and is an important vegetable in Asia. As the genome sequence of Chinese cabbage (Chiifu-401-42) has been released and *B. rapa* is related to the model plant *Arabidopsis thaliana* (The *Brassica rapa* Genome Sequencing Project Consortium., 2011), detailed genetic and evolutionary studies have become possible.

In Japan, most commercial cultivars of Chinese cabbage are F₁ hybrids because of their agronomic benefits such as high yield, stress tolerance, disease resistance, and uniform phenotype. Hybrid breeding came from the discovery of heterosis or hybrid vigor, which is defined as the superior performance of hybrid plants over the parents (Crow, 1998). When breeding F₁ hybrid cultivars, breeders developed elite pure lines (inbred lines) as parents for hybrid production. About five to seven generations of selfing and selection based on traits concerned with the breeding objective are required for developing inbred lines as parental candidates. The level of heterosis of crosses of all possible combinations of the inbred lines is used to identify suitable parents for F₁ hybrid generation. An efficient method for predicting hybrid performance in the parental generations is desired as hybrid production can be expensive, time consuming, and labor intensive. Genetic distance between parental lines might be a good predictor, though the relationship between genetic distance and heterosis is controversial (Barth et al., 2003; Dreisigacker et al., 2005; Flint-Garcia et al., 2009; Geleta

* Corresponding author at: Graduate School of Agricultural Science, Kobe University, Rokkodai, Nada-ku, Kobe 657-8501, Japan.

E-mail address: leo@people.kobe-u.ac.jp (R. Fujimoto).

¹ These authors equally contributed to this work.

et al., 2004; Meyer et al., 2004; Yu et al., 2005). Information concerning the genetic relationships among parental candidate inbred lines is also useful for variety protection.

There are various types of DNA markers such as cleaved amplified polymorphic sequences (CAPS), amplified fragment length polymorphisms (AFLP), randomly amplified polymorphic DNA (RAPD), simple sequence repeats (SSRs), single nucleotide polymorphisms (SNPs), and insertion/deletion polymorphism (InDel) markers. SSR markers have been widely used because of high polymorphism, reproducibility, co-dominant inheritance, and genome-wide coverage. In addition, SSR markers require only small amounts of DNA for PCR, and can be used for high-throughput analysis. SSR markers have been widely used for detecting genetic diversity and making genetic linkage maps, and many SSR markers are available for the genus *Brassica* (Guo et al., 2014; Hatakeyama et al., 2010; Lowe et al., 2004; Pino Del Carpio et al., 2011; Ramchiary et al., 2011; Suwabe et al., 2002; Suwabe et al., 2006). Sequencing technology enables us to identify SNPs easily, and SNPs are wide spread in the *B. rapa* genome (Metzker, 2010; Rafalski, 2002). SNPs detected by RNA-sequencing (RNA-seq) in coding regions are used for developing gene-based markers (Paritosh et al., 2013). Restriction-site associated DNA sequencing (RAD-seq) where the flanking region is sequenced from a specific restriction site, is useful for developing DNA markers and high-throughput genotyping (Baird et al., 2008).

The purpose of this study is to examine the possibility of using DNA markers to enhance breeding in *B. rapa*. We focused on the relationship between genetic distance and heterosis. We identified highly polymorphic DNA markers for calculating genetic distance or assessment of homozygosity in inbred lines of Chinese cabbage and examined the relationship between heterosis and genetic distance calculated by DNA markers. The information obtained in this study will be useful for breeding in *Brassica*.

2. Materials and methods

2.1. Plant materials and DNA extraction

The commercial F₁ hybrid cultivars of Chinese cabbage (Table S1) and F₂ populations produced by bud pollination of these F₁ hybrid cultivars were used as plant materials. A commercial cultivar of Chinese cabbage 'Chiifu', 31 Chinese cabbage inbred lines (RJKB-T01–T20, -T22–T24, -T26–T28, and -T30–T35), 3 turnip doubled haploid (DH) lines (Atsumi, Kisobeni, and BRA2209), 2 komatsuna inbred lines (YBCG-T01 and -T02) and 2 komatsuna DH lines (OSD2 and Wase Maruba) were also used in this study. We developed 31 F₁ hybrids by crossing between seed (RJKB-T01–T03, T07–T20, and -T26–T28) and pollen (RJKB-T30–T35) parents. Parental combinations were defined with regards to flowering time; RJKB-T33 and -T34 were early flowering, RJKB-T31 and -T32 were intermediate flowering, and RJKB-T30 and -T35 were late flowering.

Seeds were sown on soil and plants were grown in growth chambers under a 16-h/8-h light/dark cycle at 22 °C. Young leaves harvested from the F₁ and F₂ seedlings were used for genomic DNA extraction. Total genomic DNA was isolated by the Cetyl trimethyl ammonium bromide method (Murray and Thompson, 1980).

2.2. Evaluation of heterosis phenotype

For examining the heterosis phenotype of early developmental stages, plants were grown in plastic dishes containing Murashige and Skoog (MS) agar medium supplemented with 1.0% sucrose (pH 5.7) in growth chambers under a 16-h/8-h light/dark cycle at 22 °C. Cotyledons at 6 days after sowing (DAS), and 1st and 2nd leaves at 14 DAS were fixed in a formalin/acetic acid/alcohol solution (ethanol:acetic acid:formalin = 16:1:1). The image of the whole cotyledon or leaf was photographed under a stereoscopic microscope, and sizes were determined with Image-J software (<http://rsb.info.nih.gov/ij/>).

For examining the yield under field conditions, seeds were sown on multi-cell trays on 21st August 2014 and grown in a greenhouse. At 6 DAS, cotyledons were photographed, and the area of the cotyledons was determined with Image-J software (<http://rsb.info.nih.gov/ij/>). On 3rd September 2014, seedlings were transplanted to the field at Osaki, Miyagi, Japan (38°57'N, 141°00'E). Thirty plants per plot were transplanted and plot size was 13.5 × 0.7 m. Row spacing is 70 cm and planting distance is 40 cm. At 21 DAS, leaf lengths and widths of the first and second largest leaves were measured. On 13th and 14th November 2014, plants were harvested. Statistical comparisons of cotyledon area, leaf size, fresh weight of total biomass and harvested biomass were performed using Student's *t*-test (*p* < 0.05).

The ratio of heterosis performance between F₁ and mid parent value (MPV) (termed rMPV) is calculated as follows, $rMPV = F_1(\text{mean})/MPV$ (mean of two parents). The ratio of heterosis performance between F₁ and better parent value (BPV) (termed rBPV) is calculated as follows, $rBPV = F_1(\text{mean})/BPV$ (mean of better parent).

2.3. Detection of DNA polymorphism with SSR markers

A total of 321 SSR markers, "BRAS", "BRMS", "BnGMS", "CB", "KBr", "Na", "Ni" and "OI", were used to screen for polymorphisms among F₂ individual plants derived from the F₁ hybrid cultivar 'W77' (Table S2). The PCR reaction was performed using the following conditions; 1 cycle of 94 °C for 3 min, 40 cycles of 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 1 min, and final extension at 72 °C for 3 min. The PCR products were electrophoresed on 10% or 13% polyacrylamide gel using NA-1040 or NA-1118 (NIHON EIDO, Japan). The gel was stained with Gelstar solution (0.1 µl/10 ml; Takara Bio Inc., Japan). Primer sequences used in this study are shown in Table S3.

2.4. Detection of DNA polymorphism with CAPS markers

A total of 38 CAPS markers were used for examining the genetic distances among Chinese cabbage inbred lines. The PCR reaction was performed using the following conditions; 1 cycle of 94 °C for 3 min, 35 cycles of 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 1 min, and final extension at 72 °C for 3 min. Amplified DNA digested by *Afa* I, *Hae* III, *Hha* I, *Hinf* I, *Mbo* I, *Msp* I, or *Hae* III restriction enzymes were electrophoresed on 13% polyacrylamide gel. The gel was stained with Gelstar solution (0.1 µl/10 ml; Takara Bio Inc., Japan). Primer sequences used in this study are shown in Table S3.

2.5. RAD-seq

Genomic DNA was digested using two restriction enzymes, *Bgl* II and *Eco* RI. The digested DNA fragments and two adapters (*Bgl* II adapter and *Eco* RI adapter) were ligated. The digestion and ligation were simultaneously performed at 37 °C for 16 h. The reaction mixture consisted of 20 ng of genomic DNA, 5 units of *Bgl* II (NEB), 10 units of *Eco* RI-HF (NEB), 1 × NEB buffer2 (NEB), 1 × BSA (NEB), 0.2 µM *Bgl* II adapter, 0.2 µM *Eco* RI adapter, 1 mM ATP (Takara), 300 units of T4 DNA ligase (Enzymatics). The ligation product was purified by the AMPureXP (Beckman coulter) according to manufacturer's instructions. One tenth of the purified DNA was used in the PCR enrichment with the KAPA HiFi HS ReadyMix (KAPA biosystems). Sequences of adaptors and primers used in this study are shown in Table S4. Approximately 350 bp fragments of the PCR product was selected by the E-Gel size select 2% (Life technologies). Single end 50 bp and index sequence of the library was sequenced by the HiSeq2500 (Illumina) with the TruSeq v3 chemistry. Preprocessing of the sequence data was performed by the trimmomatic-0.32 with the following parameters: ILLUMINACLIP TruSeq3-SE.fa:2:30:10 LEADING:19 TRAILING:19 SLIDINGWINDOW:30:20 AVGQUAL:20 MINLEN:51 (Bolger et al., 2014). The preprocessed sequences were analyzed by the Stacks program with default parameters (Catchen et al., 2013). We selected 288

positions of a nucleotide sequence with less than 3 missing data per site among 30 lines.

2.6. Phylogenetic analysis

Sixty-six SSR markers, 38 CAPS markers, and RAD-seq data from 288 positions were used for screening polymorphic loci among 22 inbred lines (Table S2). The software Populations version 1.2.31 (<http://bioinformatics.org/~tryphon/populations/>) was used to calculate the genetic distance among inbred lines and to construct a neighbor-joining phylogenetic tree with 1000 bootstraps (Nei et al., 1983). The tree was depicted using TreeView version 1.6.6 (<http://tree.bio.ed.ac.uk/software/figtree/>).

3. Results

3.1. Screening of polymorphic DNA markers between the inbred lines of Chinese cabbage

SSR markers can detect nucleotide sequence differences between F_2 individuals and determine segregation of the parental genomes in the F_2 population. To find SSR markers that can efficiently detect polymorphisms among Chinese cabbage genotypes, we screened SSR markers using an F_2 population derived from an F_1 hybrid cultivar. We assessed 321 SSR markers in a commercial F_1 hybrid plant ('W77') and 6 individual F_2 plants of 'W77' (Tables S1, S2). Fifty-nine DNA markers (18%) detected polymorphism among these 7 plants (Table S2). We selected 11 co-dominant DNA markers confirming their Mendelian segregation using 96 F_2 plants derived from 'W77' (Table S5).

We selected 7 (BRMS007, BRMS026, BRMS027, BRMS040, BRMS163, BRMS226, and BRMS276) of the 59 SSR markers because the BRMS series are highly polymorphic markers in Brassica vegetables (Dr. Satoru Matsumoto, personal communication). To examine whether these 7 SSR markers that can detect polymorphisms in parental lines of 'W77' are applicable to other parental combinations of F_1 hybrid cultivars, we assessed 13 F_2 populations derived from commercial F_1 hybrid cultivars. These SSR markers showed 38% (5 of 13 F_2 populations) to 92% (12 of 13 F_2 populations) polymorphism rate among 6 individual F_2 plants, of which BRMS007 and BRMS027 showed the highest polymorphism rate (Table S6). More than 3 DNA markers (43%) were found to detect parental polymorphisms of F_1 hybrid cultivars (Table S6).

RNA-sequencing (RNA-seq) analysis in leaves using two Chinese cabbage inbred lines, RJKB-T23 and RJKB-T24, was used for SNP analysis (Shimizu et al., 2014). More than three CAPS markers in each chromosome were developed, and their marker locations were spread over the reference genome (Fig. S1). Totally 38 CAPS markers with co-dominance were developed, and their Mendelian segregation was confirmed in 12 of 38 CAPS markers using 96 F_2 plants between RJKB-T23 and RJKB-T24 (Table S5).

Genetically uniform lines are desired for both laboratory experiments and F_1 hybrid breeding. To confirm genetic uniformity of these inbred lines, we used 7 DNA markers (BRMS007, BRMS026, BRMS027, BRMS040, BRMS163, BRMS226, and BRMS276) on 8 individual plants for each inbred line. These SSR markers did not show any polymorphism among the 8 plants in any inbred line (data not shown). We further tested the genetic uniformity by RAD-seq. The numbers of raw sequenced reads per inbred line ranged from 526,146 to 5,386,735 (mean = 1,774,162). The heterozygous genotypes in 22 inbred lines were less than 1.42% (mean = 0.54%) in 288 positions of a nucleotide sequences determined by RAD-seq (Fig. S1, Table S7). These results indicate that these 22 inbred lines are highly homozygous.

3.2. Genetic distance of inbred lines in Chinese cabbage

The genetic distance (GD) among 22 genotypes of Chinese cabbage inbred lines was examined for use in F_1 hybrid cultivars. The genotypes

of one commercial cultivar, Chiifu, 3 turnip DH lines, and 4 komatsuna inbred or DH lines were also examined. We selected 66 SSR markers to calculate the GD, and a total of 163 alleles were identified among all 30 lines. There are 1 to 7 alleles (mean = 5.1 alleles) per marker among 30 lines. The average scores of GD among Chinese cabbage, turnip, komatsuna, and all lines were 0.38, 0.48, 0.45, and 0.44, respectively (Table S8). Scores of GD between Chinese cabbage and turnip lines, between Chinese cabbage and komatsuna lines, and between turnip and komatsuna lines were 0.53, 0.51, and 0.46, respectively.

Next we calculated the GD using 38 CAPS markers, and a total of 143 alleles were identified in all 30 lines. There are 1 to 8 alleles (mean = 4.8 alleles) per marker among all 30 lines. The average scores of genetic distance among Chinese cabbage, turnip, komatsuna, and all lines were 0.42, 0.58, 0.59, and 0.50, respectively (Table S9). Scores of genetic distances between Chinese cabbage and turnip lines, between Chinese cabbage and komatsuna lines, and between turnip and komatsuna lines were 0.56, 0.63, and 0.59, respectively.

We selected 288 nucleotide positions determined by RAD-seq. The average scores of genetic distance among Chinese cabbage, turnip, komatsuna, and all lines were 0.07, 0.20, 0.16, and 0.10, respectively (Table S10). Scores of genetic distances between Chinese cabbage and turnip lines, between Chinese cabbage and komatsuna lines, and between turnip and komatsuna lines were 0.16, 0.15, and 0.19, respectively.

GD calculated by the three methods was compared, and the correlation coefficient between SSR and CAPS, between SSR and RAD-seq, and between CAPS and RAD-seq were 0.65 ($p < 0.01$), 0.68 ($p < 0.01$), and 0.73 ($p < 0.01$) (Fig. 1), respectively, indicating that there is a high correlation among three methods. Using all genotype information among 22 inbred lines of Chinese cabbage, the genetic distance between RJKB-T05 and -T14 was lowest, and highest between RJKB-T07 and -T19. A dendrogram based on cluster analysis in Neighbor-joining placed all lines into 5 major groups. Four groups (I, III, IV, and V) were comprised of Chinese cabbage inbred lines, and turnip and komatsuna lines were in group-II (Fig. S2).

3.3. Evaluation of heterosis phenotype

Considering the future breeding cultivars we used 3 new inbred lines, RJKB-T26–T28, and 18 of the 22 inbred lines for seed parents and 6 new inbred lines, RJKB-T30–T35, for pollen parents. 32 F_1 hybrids were obtained. We examined the cotyledon area at 6 days after sowing (DAS) and leaf area at 14 DAS when grown on MS medium in a growth chamber of 12 F_1 hybrids to confirm that both heterotic and non-heterotic hybrids are included in this sample. In cotyledon area, the scores of rMPVs (the ratio of heterosis performance between F_1 and mid parent value (MPV)) ranged from 0.94 to 1.96, and the average scores of rMPV was 1.34. The scores of rBPV (the ratio of heterosis performance between F_1 and better parent value (BPV)) ranged from 0.90 to 1.90, and the average scores of rBPV was 1.20. In the 1st and 2nd leaf areas, scores of the rMPV and rBPV ranged from 0.52 to 2.94 and from 0.33 to 2.29, respectively, and the average scores of rMPV and rBPV were 1.41 and 1.16, respectively (Table S11).

Next we examined the cotyledon and leaf size at an early developmental time and at the final stage in the field. The average scores of the rMPV and rBPV in the cotyledon area at 6 DAS were 1.48 and 1.27, respectively (Table S12). The average scores of rMPV and rBPV in leaf length of largest leaves were 1.12 and 1.06, respectively, and the average scores of rMPV and rBPV in leaf width of largest leaves were 1.13 and 1.06, respectively. The average scores of rMPV and rBPV of the product of leaf length \times width in the largest leaves were 1.27 and 1.14, respectively (Table S13). Finally total biomass (whole above-ground), harvested biomass (in which the outer leaves were stripped for marketing), and height and circumference of the harvested plants were examined (Table S14). The average scores of rMPV and rBPV of the harvested biomass were 1.28 and 1.17, respectively. The plant circumference in F_1

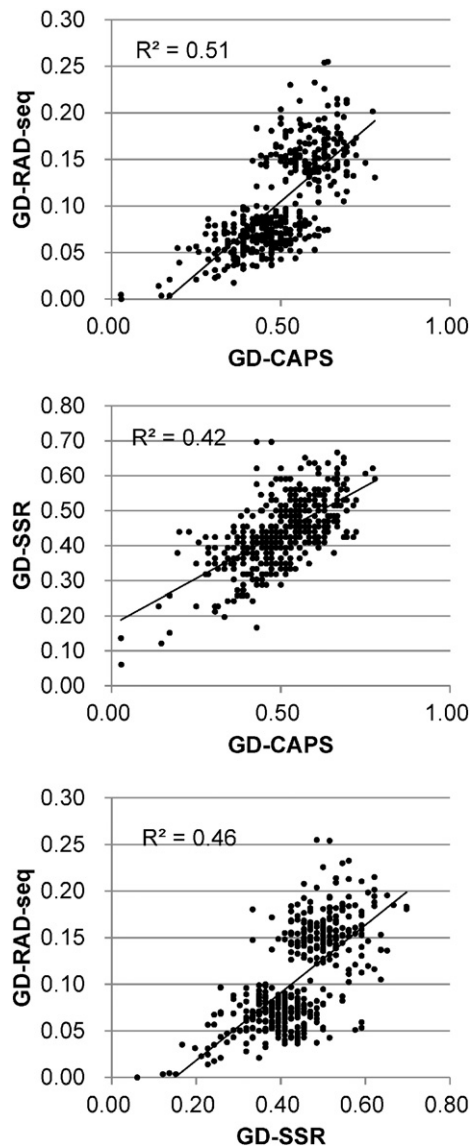


Fig. 1. Relationship between genetic distance (GD) among three types of DNA markers.

tended to be larger than parental lines, while there was little difference in the plant height between F_1 and parental lines (Table S14). The F_1 hybrids using RJKB-T35 as a pollen parent tended to show higher rMPV through 12 traits (Figs. S3, S4, Tables S12–S14).

We compared the rMPV and rBPV of 6 day cotyledon area and the 1st and 2nd leaf sizes between plants grown on MS medium in a growth chamber and plants grown on soil in the greenhouse or field. In these two comparisons, there was a significant correlation between the two experiments except for the rBPV of the 1st and 2nd leaf sizes (cotyledon, rMPV: $r = 0.86^{**}$, rBPV: $r = 0.83^{**}$; leaf size, rMPV: $r = 0.71^*$, rBPV: $r = 0.39$).

We examined the relationships of rMPV and rBPV among 12 traits at three developmental stages. There was no correlation between the rMPV/rBPV-cotyledon area at 6 DAS and rMPV/rBPV-leaf area at 21 DAS or between rMPV/rBPV-cotyledon area at 6 DAS and rMPV/rBPV-harvested biomass (Table 1). There was a moderate correlation between rMPV-leaf area at 21 DAS and rMPV-harvested biomass, while there was no correlation between rBPV-leaf area at 21 DAS and rBPV-harvested biomass (Table 1).

RAD-seq was performed on another 9 inbred lines (RJKB-T26–T28, T30–T35) and combined with the previous RAD-seq data (Table S15) to examine the relationship between rMPV and GD or between rBPV

and GD. Of 12 traits, no trait showed a positive correlation between rMPV and GD or between rBPV and GD. The leaf length of the second largest leaf at 21 DAS and plant height of the harvested plants showed a moderate negative correlation between rBPV and GD (Table 2, Fig. 2), indicating that it is difficult to predict the level of heterosis from the genetic distance between parental lines.

4. Discussion

4.1. Screening of polymorphic SSR markers

We screened SSR markers that can distinguish between the parental alleles of commercial F_1 hybrid cultivars of Chinese cabbage. Fifty-nine of 321 SSR markers detected a polymorphism between the parental alleles of a commercial F_1 hybrid cultivar, 'W77'. We assessed 7 of the 59 SSR markers on 13 parental combinations of F_1 hybrid cultivars, and more than 3 SSR markers detected polymorphisms of parental alleles, suggesting that these highly polymorphic SSR markers can be applied to assess the genetic uniformity of inbred lines. We also developed 38 CAPS markers using the SNP information detected by RNA-seq. As the putative chromosomal positions of these CAPS markers were able to be predicted, these CAPS markers enable us to avoid any bias of DNA marker positions relative to SSR markers. When breeders develop inbred lines, five to seven generations of selfing are performed, but it is not clear how many generations are sufficient for generating genetically uniform inbred lines. Completion of the creation of inbred lines is evaluated by the uniformity of the traits in a field test, but this has the risk of being affected by the environment and is time consuming and expensive. The assessment of genetic uniformity of inbred lines by DNA markers becomes a possibility, though it is not clear how many markers are sufficient. We tested the same 7 SSR markers on 8 individual offspring of 22 inbred lines, and did not find any polymorphisms among them. Genetic uniformity was also confirmed by RAD-seq analysis, which is a more powerful method of assessing the genetic uniformity. The highly polymorphic 7 SSR markers might be useful as a first test, and the use of additional DNA markers would increase the reliability of the data. Thus, the 59 SSR and 38 CAPS markers identified in this study are useful for assessing the genetic uniformity of inbred lines in Chinese cabbage, and the combination with RAD-seq gives a more precise analysis. As there are some reports of interspecific transferability of SSR and CAPS markers in closely related species (Márquez-Lema et al., 2010), these marker sets may also apply to other varieties of *B. rapa* species such as turnip or komatsuna or to related species such as cabbage (*B. oleracea* L.).

4.2. Characterization of genetic distance among inbred lines of Chinese cabbage using SSR markers

We generated neighbor-joining trees using 23 Chinese cabbage, 3 turnip, and 4 komatsuna lines. Five major clusters were constructed and 4 of them included only the Chinese cabbage lines. Turnip and komatsuna lines were clustered into the same group, but using more turnip and komatsuna lines will separate the two varieties into different groups. When *B. rapa* collections such as Chinese cabbage, turnip, and pak-choi, are examined, Chinese cabbage lines tends to be clustered into the same group with a small number of exceptions (Pino Del Carpio et al., 2011; Takuno et al., 2007; Zhao et al., 2005), consistent with our results.

Understanding the genetic information of breeding resource collections is important for breeding programs, and the genetic relationship among genotypes can help breeders in selecting parental lines, and predict the level of heterozygosity of hybrids. As any single DNA marker type may show bias in identifying the genetic distance (Frascaroli and Landi, 2013; Hamblin et al., 2007), we assessed the genetic distance by three types of DNA markers; SSR are multi-allelic markers, CAPS generated based on SNP information of RNA-seq are bi-allelic markers in exon regions, and SNPs detected by RAD-seq are bi-allelic markers.

Table 1
Relationships among rMPVs and rBPVs of 12 traits.

	6 DAS			21 DAS					Final product			
	A	B	C	D	E	F	G	H	I	J	K	L
6 DAS	A	-0.04	-0.05	-0.06	-0.12	-0.01	-0.06	-0.07	-0.05	-0.09	-0.10	-0.02
	B	-0.18	0.93**	0.97**	0.92**	0.87**	0.90**	0.96**	0.39*	0.47**	0.38*	0.07
	C	-0.03	0.80**	0.98**	0.89**	0.90**	0.91**	0.97**	0.45**	0.58**	0.47**	0.21
	D	-0.12	0.92**	0.96**	0.89**	0.88**	0.91**	0.98**	0.40*	0.53**	0.42*	0.16
21 DAS	E	-0.25	0.92**	0.75**	0.85**	0.95*	0.97**	0.95**	0.38*	0.46**	0.33	0.18
	F	0.00	0.79**	0.92**	0.89**	0.82**	0.99**	0.95**	0.40*	0.55**	0.41*	0.30
	G	-0.12	0.88**	0.91**	0.94**	0.91**	0.96**	0.97**	0.37*	0.51**	0.36*	0.27
	H	-0.13	0.92**	0.95**	0.99**	0.89**	0.94**	0.98**	0.40*	0.55**	0.40*	0.22
	I	-0.11	0.05	-0.04	0.00	0.07	-0.01	0.01	0.01	0.90**	0.88**	0.51**
Final	J	-0.15	0.19	0.14	0.19	0.15	0.14	0.16	0.18	0.91**	0.88**	0.59**
product	K	-0.16	0.13	0.26	0.25	0.05	0.17	0.16	0.22	0.69**	0.76**	0.40*
	L	-0.13	0.09	0.18	0.14	0.26	0.42	0.33	0.23	0.21	0.22	0.10

rMPV, relative ratio between F₁ and mid parent values (MPV) (The upper right).rBPV, relative ratio between F₁ and best parent values (BPV) (The lower left).

A, Cotyledon area at 6 DAS.

B, leaf length of the first largest leaf at 21 DAS; C, leaf width of the first largest leaf at 21 DAS; D, B × C.

E, leaf length of the second largest leaf at 21 DAS; F, leaf width of the second largest leaf at 21 DAS; G, E × F.

H, D + G.

I, Total biomass; J, harvested biomass; K, Circumference of harvested plant; L, Plant height of harvested plant.

*, p < 0.05; **, p < 0.01

DAS, days after sowing.

The genetic distance measured using the three types of DNA markers showed a high correlation (correlation coefficient is greater than 0.65). Though the ascertainment bias of SNP markers is controversial (Hamblin et al., 2007), the three types of DNA markers together may predict the genetic distances among inbred lines.

4.3. Estimation of yield heterosis from genetic distance of parental lines or the early developmental growth

Methods for predicting hybrid performance from the parental generations would benefit breeding programs. Heterosis has been suggested to be associated with the genetic distance between parental lines of F₁ hybrids (Biton et al., 2012; Cheres et al., 2000; Flint-Garcia

et al., 2009; Godshalk et al., 1990; Hua et al., 2002; Jagosz, 2011). In this study, we examined the relationship between GD calculated by three types of DNA markers and the level of mid parent and better parent heterosis such as cotyledon area at 6 DAS, leaf length x width of largest leaf at 21 DAS, and harvested biomass. Correlations were not observed between GD and rMPV/rBPV in any parameter examined, indicating that it is difficult to predict the hybrid performance from the genetic distance of parental lines.

In our previous study using in *A. thaliana*, the heterosis phenotype was obvious in early developmental stages and heterosis was maintained and enhanced at the later stages. We suggested that an early developmental growth advantage is important for displaying later heterosis (Fujimoto et al., 2012). Early developmental heterosis was

Table 2
Relationships between GD and rMPV or rBPV.

	6 DAS			21 DAS					Final product			
	A	B	C	D	E	F	G	H	I	J	K	L
GD vs. rMPV	0.08	-0.09	-0.12	-0.10	-0.33	-0.30	-0.32	-0.20	0.24	0.14	0.22	-0.15
GD vs. rBPV	0.18	-0.17	-0.12	-0.12	-0.39*	-0.25	-0.29	-0.20	0.20	0.21	0.24	-0.36*

rMPV, relative ratio between F₁ and mid parent values (MPV); rBPV, relative ratio between F₁ and best parent values (BPV).

A, Cotyledon area at 6 DAS.

B, leaf length of the first largest leaf at 21 DAS; C, leaf width of the first largest leaf at 21 DAS; D, B × C.

E, leaf length of the second largest leaf at 21 DAS; F, leaf width of the second largest leaf at 21 DAS; G, E × F.

H, D + G.

I, Total biomass; J, harvested biomass; K, Circumference of harvested plant; L, Plant height of harvested plant.

*, p < 0.05

DAS, days after sowing.

GD, genetic distance.

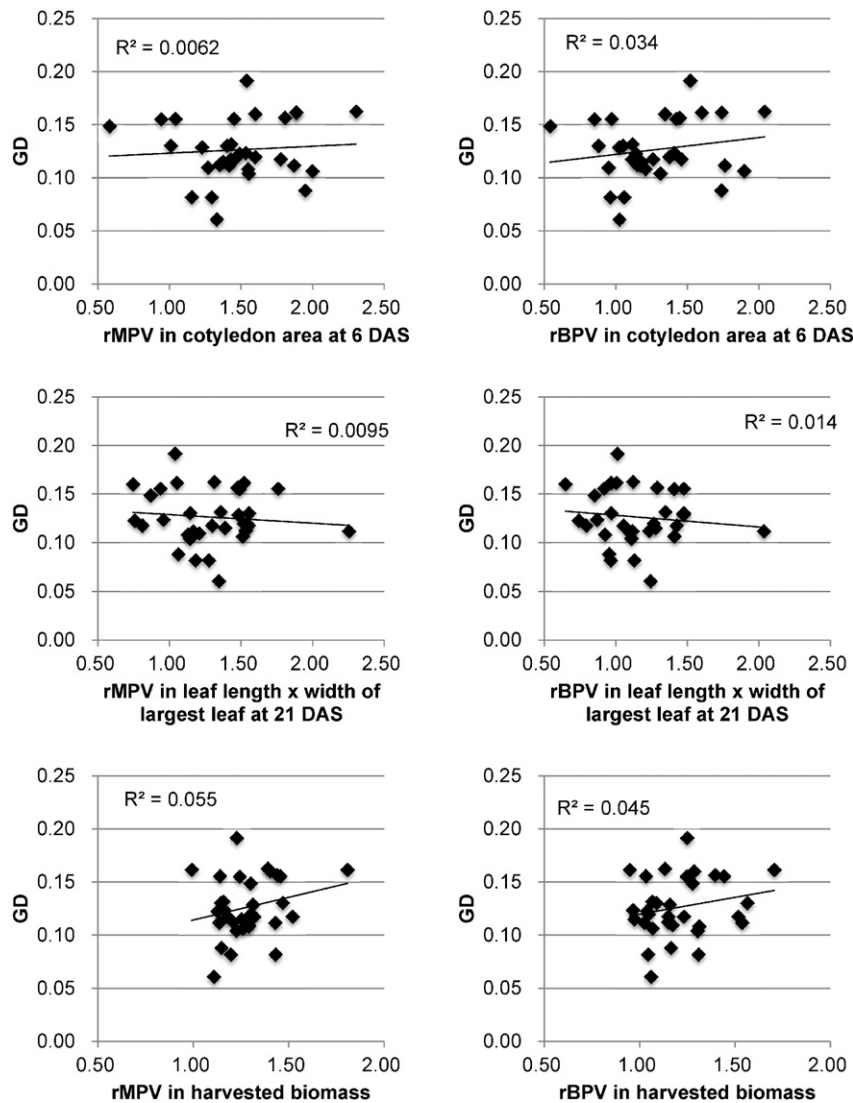


Fig. 2. Relationship between genetic distance (GD) and relative ratio of plant size or biomass between F_1 and mid parent values (rMPV) or between F_1 and best parent values (rBPV). DAS, days after sowing.

observed in crops such as rice, maize, and wheat, and the level of heterosis is trait-dependent (Flint-García et al., 2009; He et al., 2010; Li et al., 2014; Springer and Stupar, 2007; Xing et al., 2014). We considered the prediction of yield heterosis from the early developmental stages could be useful to save time and labor. We examined the heterosis level in three developmental time points, cotyledon (6 DAS), seedling (21 DAS), and harvesting stages. In rMPV, we found moderate correlation between leaf size at 21 days seedling and harvested biomass. In crops, seed production is important for yield but early seedling development does not directly correlate to yield-related traits. However early seedling development is more closely related to yield in leafy vegetables such as Chinese cabbage, suggesting that the assessment of heterosis level in seedling stages will be useful for the first screening of parental combinations of F_1 hybrid cultivars during the breeding step. As heterosis of yield is a difficult trait to measure, multiple trials to examine the relationship between early and late stage heterosis will be required.

Conflict of interest statement

The authors declare that they have no conflict of interest.

Acknowledgments

We thank Dr. Yoshihito Takahata and Dr. Yu Takahashi for providing the materials and Dr. Masanori Yamasaki for much advice on statistical analysis. We also thank Mr. Mahiro Yoshida, Mr. Shuhei Konno, Dr. Masaki Yasugi, Dr. Mie Hojyo, Ms. Satoko Matsuoka, Ms. Chiori Kume, Ms. Yuki Ishizaki, Mr. Hirofumi Abe, and Mr. Yusuke Ebe for their technical assistance. This work was supported by a grant-in-aid for Scientific Research on Innovative Areas (24113509) (The Japan Society for the Promotion of Sciences), by a Program for Revitalization Promotion, Japan Science and Technology Agency (JST) (241FT0125), and by PRESTO (JST) to RF.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.plgene.2015.10.003>.

References

Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., Selker, E.U., Cresko, W.A., Johnson, E.A., 2008. Rapid SNP discovery and genetic mapping using

- sequenced RAD markers. *PLoS One* 3, e3376. <http://dx.doi.org/10.1371/journal.pone.0003376>.
- Barth, S., Busimi, A.K., Friedrich, Utz.H., Melchinger, A.E., 2003. Heterosis for biomass yield and related traits in five hybrids of *Arabidopsis thaliana* L. Heynh. *Heredity* 91, 36–42. <http://dx.doi.org/10.1038/sj.hdy.6800276>.
- Biton, I., Shevtsov, S., Ostersetzer, O., Mani, Y., Lavee, S., Avidan, B., Ben-Ari, G., 2012. Genetic relationships and hybrid vigour in olive (*Olea europaea* L.) by microsatellites. *Plant Breed.* 131, 767–774. <http://dx.doi.org/10.1111/pbr.12001>.
- Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. <http://dx.doi.org/10.1093/bioinformatics/btu170>.
- Catchen, J., Hohenlohe, P.A., Bassham, S., Amores, A., Cresko, W.A., 2013. Stacks: an analysis tool set for population genomics. *Mol. Ecol.* 22, 3124–3140. <http://dx.doi.org/10.1111/mec.12354>.
- Cheres, M.T., Miller, J.F., Crane, J.M., Knapp, S.J., 2000. Genetic distance as a predictor of heterosis and hybrid performance within and between heterotic groups in sunflower. *Theor. Appl. Genet.* 100, 889–894. <http://dx.doi.org/10.1007/s001220051366>.
- Crow, J.F., 1998. 90 years ago: the beginning of hybrid maize. *Genetics* 148, 923–928.
- Dreisigacker, S., Melchinger, A.E., Zhang, P., Ammar, K., Flachenecker, C., Hoisington, D., Warburton, M.L., 2005. Hybrid performance and heterosis in spring bread wheat, and their relations to SSR-based genetic distances and coefficients of parentage. *Euphytica* 144, 51–59. <http://dx.doi.org/10.1007/s10681-005-4053-2>.
- Flint-Garcia, S.A., Buckler, E.S., Tiffin, P., Ersoz, E., Springer, N.M., 2009. Heterosis is prevalent for multiple traits in diverse maize germplasm. *PLoS One* 4, e7433. <http://dx.doi.org/10.1371/journal.pone.0007433>.
- Frascaroli, E., Landi, P., 2013. Divergent selection in a maize population for germination at low temperature in controlled environment: study of the direct response, of the trait inheritance and of correlated responses in the field. *Theor. Appl. Genet.* 126, 733–746. <http://dx.doi.org/10.1007/s00122-012-2014-4>.
- Fujimoto, R., Taylor, J.M., Shirasawa, S., Peacock, W.J., Dennis, E.S., 2012. Heterosis of *Arabidopsis* hybrids between C24 and Col is associated with increased photosynthesis capacity. *Proc. Natl. Acad. Sci. U. S. A.* 109, 7109–7114. <http://dx.doi.org/10.1073/pnas.1204464109>.
- Geleta, L.F., Labuschagne, M.T., Viljoen, C.D., 2004. Relationship between heterosis and genetic distance based on morphological traits and AFLP markers in pepper. *Plant Breed.* 123, 467–473. <http://dx.doi.org/10.1111/j.1439-0523.2004.01017.x>.
- Godshalk, E.B., Lee, M., Lamkey, K.R., 1990. Relationship of restriction fragment length polymorphisms to single-cross hybrid performance of maize. *Theor. Appl. Genet.* 80, 273–280. <http://dx.doi.org/10.1007/BF00224398>.
- Guo, Y., Chen, S., Li, Z., Cowling, W.A., 2014. Center of origin and centers of diversity in an ancient crop, *Brassica rapa* (turnip rape). *J. Hered.* 105, 555–565. <http://dx.doi.org/10.1093/jhered/esu021>.
- Hamblin, M.T., Warburton, M.L., Buckler, E.S., 2007. Empirical comparison of simple sequence repeats and single nucleotide polymorphisms in assessment of maize diversity and relatedness. *PLoS One* 2, e1367. <http://dx.doi.org/10.1371/journal.pone.0001367>.
- Hatakeyama, K., Horisaki, A., Niikura, S., Narusaka, Y., Abe, H., Yoshiaki, H., Ishida, M., Fukuoka, H., Matsumoto, S., 2010. Mapping of quantitative trait loci for high level of self-incompatibility in *Brassica rapa* L. *Genome* 53, 257–265. <http://dx.doi.org/10.1139/g10-001>.
- He, G., Zhu, X., Elling, A.A., Chen, L., Wang, X., Guo, L., Liang, M., He, H., Zhang, H., Chen, F., Qi, Y., Chen, R., Deng, X.W., 2010. Global epigenetic and transcriptional trends among two rice subspecies and their reciprocal hybrids. *Plant Cell* 22, 17–33. <http://dx.doi.org/10.1105/tpc.109>.
- Hua, J.P., Xing, Y.Z., Xu, C.G., Sun, X.L., Yu, S.B., Zhang, Q., 2002. Genetic dissection of an elite rice hybrid revealed that heterozygotes are not always advantageous for performance. *Genetics* 162, 1885–1895.
- Jagosz, B., 2011. The relationship between heterosis and genetic distances based on RAPD and AFLP markers in carrot. *Plant Breed.* 130, 574–579. <http://dx.doi.org/10.1111/j.1439-0523.2011.01877.x>.
- Li, A., Liu, D., Wu, J., Zhao, X., Hao, M., Geng, S., Yan, J., Jiang, X., Zhang, L., Wu, J., Yin, L., Zhang, R., Wu, L., Zheng, Y., Mao, L., 2014. mRNA and small RNA transcriptomes reveal insights into dynamic homeolog regulation of allopolyploid heterosis in nascent hexaploid wheat. *Plant Cell* 26, 1878–1900. <http://dx.doi.org/10.1105/tpc.114.124388>.
- Lowe, A.J., Moule, C., Trick, M., Edwards, K.J., 2004. Efficient large-scale development of microsatellites for marker and mapping applications in *Brassica* crop species. *Theor. Appl. Genet.* 108, 1103–1112. <http://dx.doi.org/10.1007/s00122-003-1522-7>.
- Márquez-Lema, A., Velasco, L., Pérez-Vich, B., 2010. Transferability, amplification quality, and genome specificity of microsatellites in *Brassica carinata* and related species. *J. Appl. Genet.* 51, 123–131. <http://dx.doi.org/10.1007/BF03195720>.
- Metzker, M.L., 2010. Sequencing technologies – the next generation. *Nat. Rev. Genet.* 11, 31–46. <http://dx.doi.org/10.1038/nrg2626>.
- Meyer, R.C., Törjék, O., Becher, M., Altmann, T., 2004. Heterosis of biomass production in *Arabidopsis*. Establishment during early development. *Plant Physiol.* 134, 1813–1823. <http://dx.doi.org/10.1104/pp.103.033001>.
- Murray, M.G., Thompson, W.F., 1980. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* 8, 4321–4325. <http://dx.doi.org/10.1093/nar/8.19.4321>.
- Nei, M., Tajima, F., Tateno, Y., 1983. Accuracy of estimated phylogenetic trees from molecular data II. Gene frequency data. *J. Mol. Evol.* 19, 153–170. <http://dx.doi.org/10.1007/BF02300753>.
- Paritosh, K., Yadava, S.K., Gupta, V., Panjabi-Massand, P., Sodhi, Y.S., Pradhan, A.K., Pental, D., 2013. RNA-seq based SNPs in some agronomically important oleiferous lines of *Brassica rapa* and their use for genome-wide linkage mapping and specific-region fine mapping. *BMC Genomics* 14, 463. <http://dx.doi.org/10.1186/1471-2164-14-463>.
- Pino Del Carpio, D., Basnet, R.K., De Vos, R.C.H., Maliepaard, C., Visser, R., Bonnema, G., 2011. The patterns of population differentiation in a *Brassica rapa* core collection. *Theor. Appl. Genet.* 122, 1105–1118. <http://dx.doi.org/10.1007/s00122-010-1516-1>.
- Rafalski, A., 2002. Applications of single nucleotide polymorphisms in crop genetics. *Curr. Opin. Plant Biol.* 5, 94–100. [http://dx.doi.org/10.1016/S1369-5266\(02\)00240-6](http://dx.doi.org/10.1016/S1369-5266(02)00240-6).
- Ramchiary, N., Nguyen, V.D., Li, X., Hong, C.P., Dhandapani, V., Choi, S.R., Yu, G., Piao, Z.Y., Lim, Y.P., 2011. Genic microsatellite markers in *Brassica rapa*: development, characterization, mapping, and their utility in other cultivated and wild *Brassica* relatives. *DNA Res.* 18, 305–320. <http://dx.doi.org/10.1093/dnares/dsr017>.
- Shimizu, M., Fujimoto, R., Ying, H., Pu, Z.J., Ebe, Y., Kawanabe, T., Saeki, N., Taylor, J.M., Kaji, M., Dennis, E.S., Okazaki, K., 2014. Identification of candidate genes for fusarium yellow resistance in Chinese cabbage by differential expression analysis. *Plant Mol. Biol.* 85, 247–257. <http://dx.doi.org/10.1007/s11103-014-0182-0>.
- Springer, N.M., Stupar, R.M., 2007. Allele-specific expression patterns reveal biases and embryo-specific parent-of-origin effects in hybrid maize. *Plant Cell* 19, 2391–2402. <http://dx.doi.org/10.1105/tpc.107.052258>.
- Suwabe, K., Iketani, H., Nunome, T., Kage, T., Hirai, M., 2002. Isolation and characterization of microsatellites in *Brassica rapa* L. *Theor. Appl. Genet.* 104, 1092–1098. <http://dx.doi.org/10.1007/s00122-002-0875-7>.
- Suwabe, K., Tsukazaki, H., Iketani, H., Hatakeyama, K., Kondo, M., Fujimura, M., Nunome, T., Fukuoka, H., Hirai, M., Matsumoto, S., 2006. Simple sequence repeat-based comparative genomics between *Brassica rapa* and *Arabidopsis thaliana*: the genetic origin of clubroot resistance. *Genetics* 173, 309–319.
- Takuno, S., Kawahara, T., Ohnishi, O., 2007. Phylogenetic relationships among cultivated types of *Brassica rapa* L. em. Metzger, as revealed by AFLP analysis. *Genet. Resour. Crop. Evol.* 54, 279–285. <http://dx.doi.org/10.1007/s10722-005-4260-7>.
- The Brassica rapa Genome Sequencing Project Consortium., 2011. The genome of the mesopolyploid crop species *Brassica rapa*. *Nat. Genet.* 43, 1035–1039. <http://dx.doi.org/10.1038/ng.919>.
- Xing, N., Fan, C., Zhou, Y., 2014. Parental selection of hybrid breeding based on maternal and paternal inheritance of traits in rapeseed (*Brassica napus* L.). *PLoS One* 9, e103165. <http://dx.doi.org/10.1371/journal.pone.0103165>.
- Yu, C.Y., Hu, S.W., Zhao, H.X., Guo, A.G., Sun, G.L., 2005. Genetic distances revealed by morphological characters, isozymes, proteins and RAPD markers and their relationships with hybrid performance in oilseed rape (*Brassica napus* L.). *Theor. Appl. Genet.* 110, 511–518. <http://dx.doi.org/10.1007/s00122-004-1858-7>.
- Zhao, J., Wang, X., Deng, B., Lou, P., Wu, J., Sun, R., Xu, Z., Vromans, J., Koornneef, M., Bonnema, G., 2005. Genetic relationships within *Brassica rapa* as inferred from AFLP fingerprints. *Theor. Appl. Genet.* 110, 1301–1314. <http://dx.doi.org/10.1007/s00122-005-1967-y>.