Theor Appl Genet (2005) 110: 1027–1037
DOI 10.1007/s00122-005-1921-z

ORIGINAL PAPER

R. M. Gaafar · U. Hohmann · C. Jung

Bacterial artificial chromosome-derived molecular markers for early bolting in sugar beet

Received: 5 January 2005 / Accepted: 5 January 2005 / Published online: 16 February 2005

Abstract Early bolting in sugar beet (*Beta vulgaris* L.) is controlled by the dominant gene *B*. From an incomplete physical map around the *B* gene, 18 bacterial artificial chromosomes (BACs) were selected for marker development. Three BACs were shotgun-sequenced, and 61 open reading frames (ORFs) were identified. Together with 104 BAC ends from 54 BACs, a total number of 55,464 nucleotides were sequenced. Of these, 37 BAC ends and 12 ORFs were selected for marker development. Thirty-one percent of the sequences were found to be single copy and 24%, low copy. From these sequences, 15 markers from ten different BACs were developed. Ten polymorphisms were determined by simple agarose gel electrophoresis of either restricted or non-restricted PCR products. Another five markers were determined by tetra-primer amplification refractory mutation system-PCR. In order to select candidate BACs for cloning the gene, genetic linkage between seven markers and the bolting gene was calculated using 1,617 plants from an F_2 population segregating for early bolting. The recombination values ranged between 0.0033 and 0.0201. In addition, a set of 41 wild and cultivated *Beta* accessions differing in their early bolting character was genotyped with seven markers. A common haplotype encompassing two marker loci and the *b* allele was found in all sugar beet varieties, indicating complete linkage disequilibrium between these loci. This suggests that the bolting gene is located in close vicinity to these markers, and the corresponding BACs can be used for cloning the gene.

Communicated by F. Salamini

R. M. Gaafar · U. Hohmann · C. Jung (✉)
Plant Breeding Institute,
Christian-Albrechts-University of Kiel,
Olshausenstr. 40, 24098 Kiel, Germany
E-mail: c.jung@plantbreeding.uni-kiel.de
Tel.: +49-431-8807364
Fax: +49-431-8802566

Introduction

Sugar beet (*Beta vulgaris* L. ssp. *vulgaris*) is a diploid species with 18 chromosomes and a relatively small genome size of 758 Mbp (Arumuganathan and Earle 1991). It is closely related to the wild beet *B. vulgaris* L. ssp. *maritima*. Natural populations are highly heterozygous due to a gametophytic self-incompatibility system with a series of *S* alleles. In the first year, cultivated beets form a leaf rosette and a storage root. Shoot elongation is initiated after a period of low temperatures (6–10 weeks at $<4°C$) under long-day conditions (Vince-Prue 1975). Occasionally, bolting plants appear under field conditions, with drastically reduced root yield and sugar content. An early bolting gene, *B*, was identified from wild beets (Munerati 1931; Abegg 1936), which was mapped with RFLP markers to chromosome 2 of sugar beet (Boudry et al. 1994). The dominant *B* allele enables bolting short after germination without vernalization; however, incomplete penetrance of this gene has been observed (Owen 1954; Owen and McFarlane 1958; Shimamoto et al. 1990; Sadeghian 1993; Boudry et al. 1994). The gene is widely distributed in wild beet populations but completely absent from cultivated beets. Temperature and day length are the two main environmental factors that affect the bolting behaviour in sugar beet. A second gene for long-day requirement has been proposed in close vicinity to the *B* locus (Abe et al. 1997). The detection of the *B* allele in elite breeding material and in seed lots is of major importance. Molecular markers that are closely linked to the bolting gene could be useful for marker-assisted selection of seed impurities. A genetic map around the *B* locus has been constructed (El-Mezawy et al. 2002). Two AFLP markers have been found to be closely linked to *B* gene, mapping in distances 0.14 cM and 0.23 cM from the *B* gene, respectively. Using these two markers and other AFLP markers from a genetic window spanning 7.8 cM around the *B* locus as probes, bacterial artificial

chromosome (BAC) clones have been selected, and a first physical map has been constructed from this region of the genome with a purpose to clone the gene from its position in the genome (Hohmann et al. 2003).

Selected BACs from the initial BAC contigs had to be remapped to determine their order in the physical map. Therefore, BAC-based marker assays are needed which enable genotyping of a large number of plants at low costs. For genotyping, different methods can be used relying on SNPs present among different alleles of the *B*-linked markers. Cleaved amplified polymorphic sequences (CAPS) are polymorphic restriction fragments visible after restriction digest of PCR fragments, followed by agarose gel electrophoresis (Konieczny and Ausubel 1993). SNPs can also be detected using allele-specific PCR primers designed in such a way that the 3′ nucleotide of a primer corresponds to the site of the SNP (Ugozzoli and Wallace 1991). A simple SNP genotyping method involving a single PCR reaction, followed by gel electrophoresis, was reported (Ye et al. 2001). This technique, named tetra-primer amplification refractory mutation system (ARMS)-PCR, adopts the principles of the tetra-primer PCR method (Ye et al. 1992) and ARMS.

The development of CAPS and tetra-primer ARMS markers relies on sequence information around the gene of interest. In the course of a project to clone the bolting gene (Hohmann et al. 2003), BACs have been partially sequenced, giving a wealth of sequence information for further marker development. In this paper, we report on CAPS and tetra-primer ARMS markers developed from BAC sequences. After genotyping a panel of wild and sugar beet accessions, two of these markers were found to be in complete linkage disequilibrium with the bolting gene, which indicates that the corresponding BACs can be used for cloning the gene.

Materials and methods

Plant materials and DNA extraction

A nonbolting sugar beet line (A906001, *bb*) was crossed with an early-bolting line as a pollinator (93167P, *BB*). An F_2 population was produced by selfing one F_1 plant. Early bolting was determined as shoot elongation during a period between 6 weeks and 15 weeks after sowing. Plants that did not show shoot elongation during this period were scored as nonbolting individuals. In total, 1,617 F_2 plants were investigated. Of these, 300 plants were found to be nonbolting plants and used for calculating recombination values. 24 wild beet (*B. vulgaris* L. ssp. *maritima*) accessions and 17 sugar beet hybrid cultivars kindly provided by Dr. L. Frese from the Federal Centre for Breeding Research on Cultivated Plants (Braunschweig) were selected for marker genotyping. Wild beets were grown in the greenhouse to determine their bolting behaviour. Genomic DNA was extracted using the CTAB method as described by Saghai-Maroof et al. (1984), with slight modifications.

BAC sequencing

A BAC library had been constructed from a doubled haploid sugar beet (KWS2320) and screened with 14 molecular markers that have been shown to be closely linked to the *B* gene (El-Mezawy et al. 2002). In total, 54 candidate BACs from the gene *B* region were identified and assembled into three major contigs (Hohmann et al. 2003). The ends of the BAC clones were sequenced by the company DLMBC, Berlin, using the T7 and SP6 vector primers. Three BACs, GJ01, GJ70 and GJ131, from the central gene *B* region were shotgun sequenced by the company AGOWA, Berlin. Potential open reading frames (ORFs) were detected by GeneQuest software (DNASTAR, GATC Biotech, Konstanz, Germany). The ORFs were chosen for marker development. Sequence homology searches of BAC ends and BAC shotgun sequences to database were done using the BIOMAX server of the KWS SAAT (Einbeck, Germany).

Primer design and PCR analysis

PCR primers were designed using the PrimerSelect software (DNASTAR, GATC Biotech). The length of the primers ranged from 19 bp to 23 bp. PCR reactions were carried out in a total volume of 25 µl 1× reaction buffer, containing 2 mM $MgCl_2$, each dNTP at 0.2 mM, 1 U *Taq* DNA polymerase (Invitrogen Life Technologies), 10 ng genomic template DNA and 10 pmol each primer in a thermocycler (PCR Express GRADIENT thermocycler, Hybaid, Ashford, UK) under the following conditions: 3 min at 94°C, followed by 30 cycles of 1 min at 94°C, 1 min at 60°C for BAC end markers or 58°C for BAC ORF markers (Table 1) and 1 min at 72°C. BAC-PCR was carried out with 1 ng template DNA. Aliquots (5 µl) of PCR products were separated by agarose gel electrophoresis and visualized under UV light. Primer sequences and PCR conditions are available on request from the authors.

Copy-number determination of BAC sequences

BAC-PCR products were purified using a Montage PCR_{96} Cleanup Kit (Millipore, Bedford, Mass., USA), radioactively labelled with α-[^{32}P]-dATP and α-[^{32}P]-dCTP and used as probes for subsequent filter hybridization. The labelled probes were either hybridized to high-density colony filters containing 3,072 BACs from the sugar beet library, which is equal to 0.5 genome equivalents, or to Southern filters with 10 µg *Hin*dIII- and *Eco*RI-restricted sugar beet DNA. Copy number was judged from the number of positive colonies and the number of Southern bands.

Table 1 Characters of bacterial artificial chromosomes (*BAC*)-derived markers

	Scoring type, restriction enzyme	Number of sequenced base pairs	Number of SNPs	SNPs/1,000 bp	Number of indels (size range in base pairs)
BAC end marker					
GJ01T7	Dominant	–	–	–	–
GJ10SP6	–	316	2	6.3	0
GJ10T7	Dominant, *Hin*fI	–	–	–	–
GJ15SP6	–	428	13	30.4	3 (1–2)
GJ17SP6	Codominant, *Hin*fI	–	–	–	–
GJ17T7	Dominant	–	–	–	–
GJ18SP6	–	415	8	19.3	0
GJ18T7	–	564	5	8.9	0
GJ21SP6	–	509	4	7.9	2 (1)
GJ22T7	Dominant	–	–	–	–
GJ29SP6	–	461	0	0	0
GJ29T7	Codominant *Alu*I	–	–	–	–
GJ37T7	Dominant	–	–	–	–
GJ44SP6	–	467	6	2.2	0
GJ44T7	–	583	0	0	0
GJ49SP6	Codominant, *Mbo*I	–	–	–	–
GJ55SP6	–	519	0	0	0
GJ55T7		559	0	0	0
GJ70SP6	–	360	3	8.3	0
GJ70T7	–	233	0	0	0
GJ73SP6	–	498	0	0	0
GJ73T7	–	372	3	8.1	0
GJ131SP6	Codominant, *Alu*I	–	–	–	–
GJ131T7	Dominant	–	–	–	–
Total (average)		6284	44	(7.0)	5
BAC ORF marker					
GJ01co15F2[a]	–	797	0	0	0
GJ01co15F3[a]	–	630	2	3.2	0
GJ01co36F4[a]	–	692	1	1.4	0
GJ70co3F4[a]	–	808	2	2.5	0
GJ70co9F1[b]	–	494	0	0	0
GJ70co9F3[a]	–	874	0	0	0
GJ131co15F2[a]	–	761	3	3.9	0
GJ131co15F6[a]	–	701	4	5.7	1 (3)
Total (average)	–	5757	12	(2.08)	1

The SNPs were determined by comparing sequences from both parents of the mapping population
– Not determined
[a]Exon region was amplified by using intron primers
[b]Intron region was amplified by using exon primers

Marker analysis

For CAPS analysis, candidate sequences were searched for restriction sites with the programme MapDraw from the DNASTAR software package (GATC Biotech). Five microlitres from each PCR product were digested with 1 U restriction enzyme for 3 h or overnight. Then, the digested PCR products were separated on 2% (w/v) agarose gels. The following restriction enzymes were employed in this analysis: *Alu*I, *Hin*cII, *Hin*dIII, *Hin*fI, *Hph*I, *Mbo*I, *Mbo*II, *Mse*I, *Pst*I, *Rsa*I and *Taq*I (Amersham Biosciences, Freiburg, Germany).

Prior to ARMS analysis, allelic sequences of bolting (93167P) and nonbolting (A906001) genotypes were determined by directly sequencing of the purified PCR products, with a MegaBACE 500 sequencer using Dynamic ET terminator cycle sequencing chemistry (Amersham Biosciences). DNA sequence analyses and SNP detection were done using SeqMan II from the DNASTAR software package (GATC Biotech). In a next step, allele-specific primers (inner primers) were designed to contain one-base pair mismatches at position-2 from the 3′ terminus. The procedure for selecting a nucleotide for mismatch in ARMS-PCR has been previously described (Little 1997). The PCR reaction was carried out in a total volume of 25 µl, containing 10 ng of template DNA, 10 pmol of each inner primer (allele-specific primers), 5 pmol of each outer primer (primers used to amplify the PCR products), 2 mM $MgCl_2$, 1× PCR buffer and 1 U *Taq* DNA polymerase. For selecting the optimum annealing temperature, a gradient PCR was performed on a PCR Express GRADIENT thermocycler (Hybaid), with annealing temperatures ranging from 56°C to 67°C. An 8-µl aliquot of the PCR products was mixed with 5 µl loading buffer and separated on a 2% agarose gel containing ethidium bromide.

Linkage analysis

Hühn (1995) described calculation of recombination values, using the recessive class of a segregating population only. Thus, the 300 nonbolting (*bb*) F_2 plants from a population of 1,617 F_2 plants, in total, were genotyped, and recombination values were determined accordingly. All markers utilized were scored codominantly. Linked significance tests for testing linkage between markers and *B* locus were done with the LOD SCORE test. Based on the recombination values between each marker and the *B* locus, the recombination values among markers were calculated. Three-point linkage analysis was done in order to find the most likely order of markers in relation to the *B* locus (Liu 1998). The recombination values were converted into map distances (centiMorgans), using Kosambi's equation (Kosambi 1944).

Results

Sequence analysis

Shotgun sequencing of three BACs representing the central part of the gene *B* region (Hohmann et al. 2003) yielded a total sequence of 151 kbp, which is approximately 42.6 % of the BAC inserts. Sixty-one ORFs with >200 bp were found using the DNASTAR software. Homology search revealed that only two ORFs showed significant homology to a known gene, while 12 ORFs showed homology to hypothetical proteins. Two BAC sequences representing ORFs with potential coding capacity were used as molecular markers. The DNA sequence of marker GJ01co36F4 showed homology to a mitochondrial processing peptidase (*E*-value: 7^{-58}, accession no. AF297643.1). The sequence of marker GJ131co15F6 showed 79% identity to a Myb transcription factor protein (T25B15.30, At3g52250) over a 128-bp region. BLASTP analysis (*E*-value: 7^{-87}) revealed 50% identity over 353 amino acids. This Myb protein shows similarity to a human nuclear receptor corepressor 2 (N-CoR2) (accession no. Q9Y618), which is involved in chromatin binding and a component of a HDAC3 complex.

From 54 BACs, 104 BAC ends were sequenced, summing up to a total of 55.5 kbp. BLASTN and BLASTP analysis of the 104 BAC end-derived sequences revealed no significant homology to any known gene. Twenty-seven BAC ends showed homology to known repetitive elements like repetitive or satellite DNA, rDNA and retrotransposons. Seventy-seven BAC end sequences, which did not show homology to repetitive sequences, were used for marker development. The copy number of the selected BAC ends was determined in two different ways: genomic Southern blots with DNA from the bolting (*BB*) and from the nonbolting (*bb*) beet parents were hybridized with 74 BAC end probes and 12 ORF-derived probes. Autoradiograms were evaluated according to number and quality of hybridization signals. Sequences (31.08%) were classified as single copy (1–2 bands), 24.3% were low copy (3–10 bands), and 25.8% were multicopy (11–20 bands) (Fig. 1). BAC end sequences (18.9%) were repetitive due to strong signals without distinguishable bands. Only one sequence (GJ47T7) gave a ladder pattern typical for tandemly repeated elements, demonstrating that preselection by sequence analysis was successful.

In a second approach the copy number of all 74 BAC ends was determined by colony hybridization with high-density filters containing 3,072 BAC clones spotted in duplicate, which is equivalent to 50% of the beet genome (Fig. 1). A correlation coefficient between genomic Southern and colony hybridization data was calculated for all probes with number of bands < 16, and a positive correlation was found ($r = 0.75$).

Marker development

On the basis of the BAC end sequences, 74 primer combinations were developed. Primer lengths ranged between 19 bp and 23 bp. After PCR with the corresponding BACs followed by agarose gel electrophoresis, all primer combinations gave fragments of the expected size, which ranged between 206 bp and 583 bp. The primer combinations of 37 amplicons, which had been classified as single or low copy (see above), were used for PCR with genomic DNA of the three beet genotypes, KWS2320 (*bb*), 93167P (*BB*) and A906001 (*bb*). Thirty-six primer combinations gave fragments of the expected size. None of the primer combinations gave polymorphisms between *BB* and *bb* genotypes visible after agarose gel electrophoresis, whereas five primer combinations yielded fragments only with the *bb* genotypes (Table 1; Fig. 2). Thus, they were polymorphic

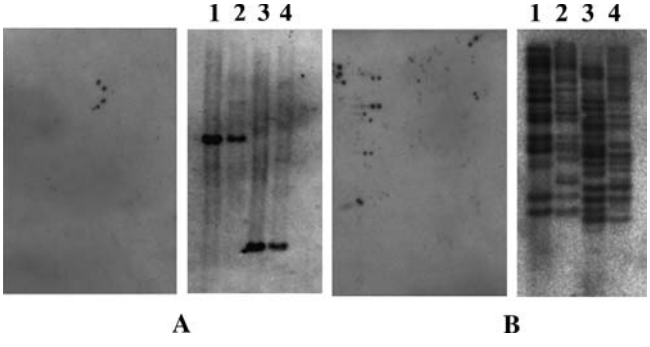

Fig. 1 Copy-number determination of bacterial artificial chromosomes (BAC)-derived sequences by hybridization on genomic DNA filters and colony filters containing 3,072 BACs (half-genome equivalent). Filters were hybridized with radioactively labelled probes GJ21SP6 (**a**) and GJ13T7 (**b**). Probe GJ21SP6 is a single-copy probe, whereas GJ13T7 is multicopy. Sugar beet DNA (*lanes 1, 3* 93167P; *lanes 2, 4* A906001) was restricted with *Hin*dIII (*lanes 1, 2*) and *Eco*RI (*lanes 3, 4*). Exposure time: 72 h

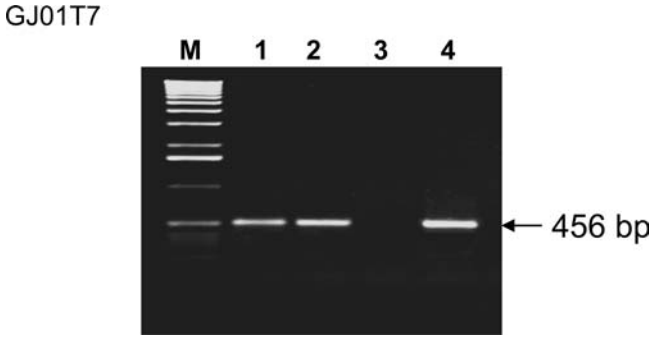

Fig. 2 Phenotype of the dominant BAC marker GJ01T7. Five microlitres of the PCR reaction were separated in a 1% agarose gel. *M* 1-kb ladder, *lane 1* donor BAC GJ01, *lane 2* KWS2320 (*bb*), *lane 3* 93167P (*BB*), *lane 4* A906001 (*bb*)

between both parents of the mapping population and could be scored as dominant markers.

In an attempt to find more polymorphisms, PCR fragments of 36 primer combinations were restricted with enzymes *Alu*I, *Hin*cII, *Hin*dIII, *Hin*fI, *Hph*I, *Mbo*I, *Mbo*II, *Mse*I, *Pst*I, *Rsa*I and *Taq*I prior to gel electrophoresis. As a result, another five polymorphic fragments were detected which could be mapped as CAPS markers (Table 1; Fig. 3). For marker GJ29T7, three fragments (369, 89, and 77 bp) were visible with the *BB* genotype, while only two (369 bp and 166 bp) were visible with the *bb* genotype after restriction with *Alu*I and separation of the restricted DNA on a 2% agarose gel. For 26 BAC ends (72%), no polymorphisms could be found after restriction digest.

To increase the efficiency of marker development, the ARMS technique was applied to uncover and map additional SNPs. Fourteen primer combinations derived from those BAC ends lacking CAPS, and STS polymorphisms were used to amplify DNA from the parents of the mapping population (93167P, *BB* and A906001, *bb*). The PCR products were sequenced directly after purification. For each primer combination, both sequences together with the BAC end sequences were compared. Eight out of 14 BAC ends displayed SNPs between both parents. Two primer combinations, GJ15SP6 and GJ21SP6, uncovered indels (Table 1). The number of SNPs per BAC end ranged from 2 (GJ10SP6) to 13 (GJ15SP6). In total, 44 SNPs were detected among 6,284 bp, corresponding to 1 SNP/143 bp (Table 1), which is higher as compared to ORF-derived markers (see below). Among 44 BAC end-derived SNPs, 22 (50%) were transitions, and 22 (50%) were transversions while, eight SNPs (66.7%) were transitions, and four SNPs (33.3%) were transversion in the case of BAC ORF-derived SNPs.

ORF-related sequences from the gene *B* region are most likely single-copy sequences, and hence, are suitable markers closely linked to the gene *B* or even representing it. ORF-derived markers were developed as follows: first, possible introns were determined; second, from 12 ORFs with more than 450 bp, primers were designed, and the corresponding PCR products from both parents were directly sequenced. According to primer design, ten primer combinations amplified exons, whereas the remaining two amplified introns. One primer combination failed to amplify the expected frag-

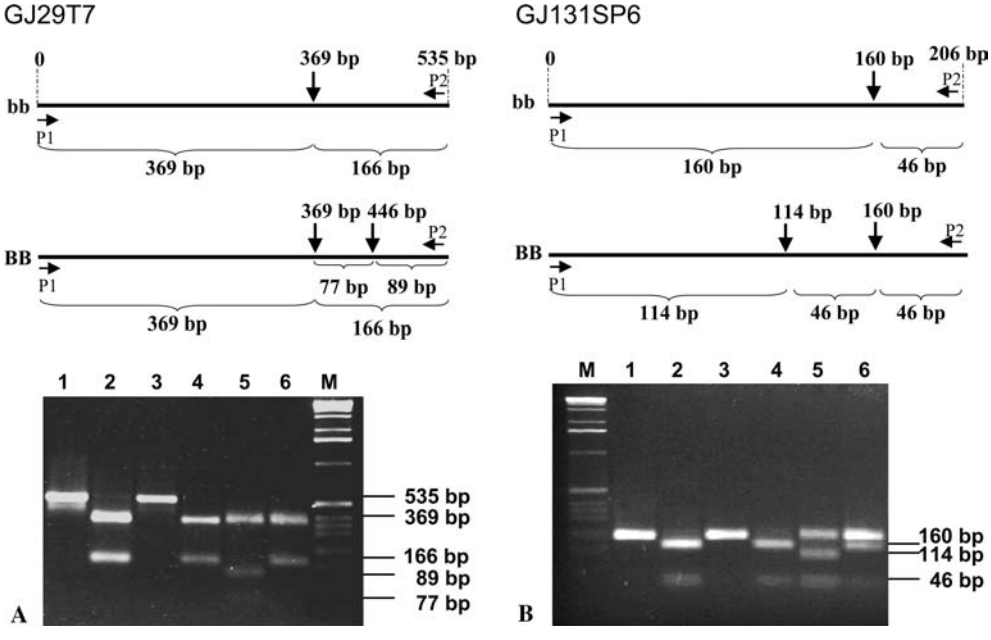

Fig. 3 Phenotype of the two codominant cleaved amplified polymorphic sequences markers GJ29T7 (**a**) and GJ131SP6 (**b**) developed from two BAC end sequences. The PCR products were digested with *Alu*I. Ten microlitres of the restriction volume were separated in a 1% agarose gel. *M* 1-kb ladder, *lane 1* nondigested PCR product from GJ29/GJ131, *lane 2* GJ29/B131, *lane 3* KWS2320 (*bb*) donor for BAC library, *lane 4* KWS2320 (*bb*) donor for BAC library, *lane 5* 93167P (*BB*), *lane 6* A906001 (*bb*). PCR products in *lanes 2*, *4*, *5* and *6* were digested with *Alu*I

ments as determined by agarose gel electrophoresis, whereas three primer combinations failed to give amplicons with the bolting parent and were not further investigated. Eight primer combinations gave fragments of the expected size. By comparing the sequences, five primer combinations showed single nucleotide polymorphisms (Table 1). According to genomic Southern analysis and BAC library screening, they were single- or low-copy sequences. The SNP frequency based on potential ORFs was lower as compared to sequences from BAC ends. In total, 12 SNPs were detected among 5,757 bp, which corresponds to 1 SNP/480 bp (Table 1).

In a next step, a number of ORFs and BAC ends were turned to molecular markers, using the ARMS technique. The functionality and reproducibility of an ARMS assay heavily relies on primer design. ARMS marker assays for three BAC ORFs (GJ01co36F4, GJ70co3F4 and GJ131co15F6) and two BAC end SNPs (GJ18T7 and GJ44SP6) are described in Table 2. The primers were designed to have a mismatch at two nucleotides upstream from the 3′ terminus of the primer-binding site. Then, the tetra-primer PCR reaction (PCR using four primers in one reaction) was performed at the optimized annealing temperature. After PCR amplification, the allele-specific PCR fragments were separated on 2% agarose gels.

In general, a reduced concentration of the outer primers (1:2 = outer:inner primers) enhanced the amplification of the two shorter allele-specific products (Fig. 4). Gradient PCR was used in order to optimize annealing temperatures for each assay as determined by maximum yield of PCR products visible after gel electrophoresis. The allele-specific PCR products were amplified with the lowest possible annealing temperature within the range tested (58–67.8°C). The wide temperature ranges allowed running the different marker assays with the same PCR programme.

Electrophoretic patterns obtained with two ARMS, GJ18T7 and GJ44SP6, are shown in Fig. 4. The PCR with marker GJ18T7 yielded three fragments. The 564-bp fragment was amplified with the two outer primers and served as a template for the amplification of the two allele-specific fragments. The 289-bp fragment was amplified due to a C at position 263 bp. It was linked to the nonbolting allele, while the 322-bp fragment was linked to the bolting allele. Thus, F_2 individuals could be clearly genotyped as homozygous and heterozygous individuals. In the case of marker GJ44SP6, a 467-bp fragment was visible with all genotypes, the nonbolting genotype (*bb*) showed a 322 bp allele-specific fragment, whereas the bolting genotype showed a 193-bp fragment. Heterozygotes (*Bb*) showed both fragments (C allele, 322 bp and G allele, 193 bp).

For markers GJ10SP6 and GJ70SP6, allele-specific assays were developed without mismatches within the primer site, because all attempts to develop tetra-primer PCR assays failed. Polymorphic fragments were obtained using one primer pair for each allele, one outer primer and one allele-specific primer, in a single PCR reaction. Thus, each of the two allele-specific fragments was amplified in two separate PCR reactions.

Linkage analysis

Genotyping was substantially accelerated by using only the nonbolting recessive (*bb*) plants from F_2 populations. Recombination values between the *B* locus and each marker were calculated (Table 3). All markers were found to be in close genetic distance to the bolting gene. BAC ORF marker GJ01co36F4 showed the lowest recombination value (0.0033) with the *B* locus. Four markers GJ17SP6, GJ18T7, GJ29T7 and GJ44SP6, were in a genetic distance of 0.0050 recombination units to the *B* gene. Based on the LOD SCORE test, all markers showed significant linkage to the *B* locus (Table 3).

Screening of *Beta* accessions with BAC-derived markers

A purpose of this analysis was to look for linkage disequilibrium around the *B* gene. Sugar beet specific

Table 2 Amplification refractory mutation system (ARMS)-PCR primers and conditions for allele-specific ARMS markers development

Marker name	SNP position (bp)	SNP (bolting parent /nonbolting parent)	Primer pair[a]	Annealing temperature[b] (°C)	Amplicon size (bp)
GJ18T7	263	(A/C)	FIP (A allele)/ROP	60.0	320
			RIP (C allele)/FOP		289
			FOP/ ROP		564
GJ44SP6	170	(G/C)	FIP (C allele)/ROP	62.0	322
			RIP (G allele)/FOP		193
			FOP/ ROP		467
GJ01co36F4	437	(G/A)	FIP (A allele)/ROP	62.0	279
			RIP (G allele)/FOP		457
			FOP/ ROP		693
GJ70co3F4	368	(A/T)	FIP (A allele)/ROP	57.0	462
			RIP (T allele)/FOP		395
			FOP/ ROP		808
GJ131co15F6	351	(C/T)	FIP (T allele)/ROP	62.0	277
			RIP (C allele)/FOP		373
			FOP/ ROP		608

[a] *FIP* Forward inner primer, *RIP* reverse inner primer, *FOP* forward outer primer, *ROP* reverse outer primer

[b] Optimized by gradient PCR

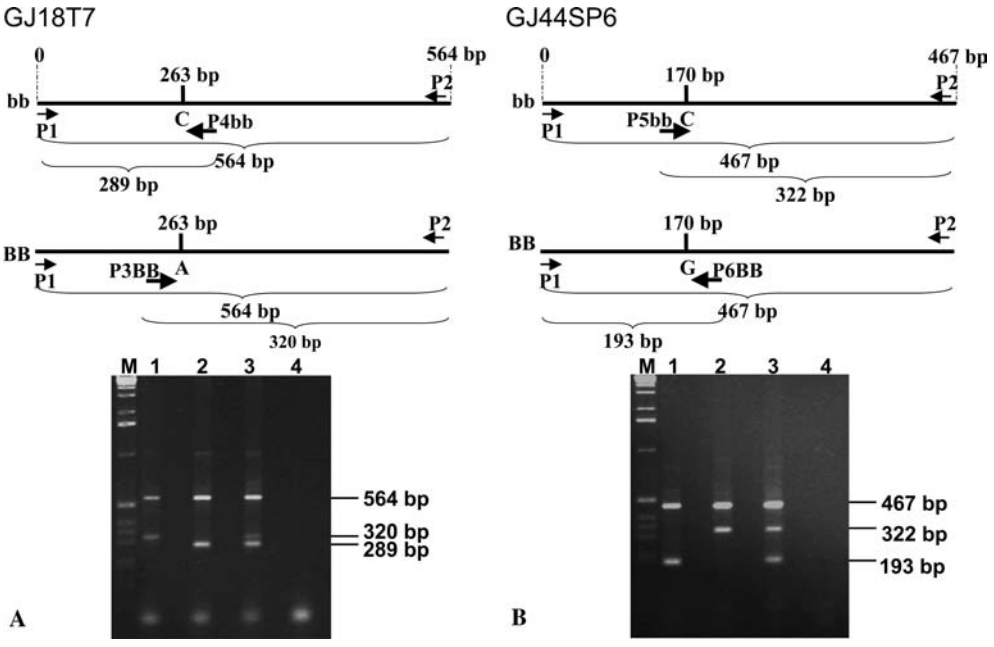

Fig. 4 Phenotype of the tetra-primer amplification refractory mutation system (ARMS)-PCR markers GJ18T7 (**a**) and GJ44SP6 (**b**). *P1* and *P2* represent the outer primer pairs. Eight microlitres of the PCR product were separated on a 2% agarose gel. Allele-specific primers had mismatches at position-2 at their 3′-termini. *GJ 18T7* A/C SNP was present at position 263. Primer P4bb amplifies the nonbolting allele; primer P3BB amplifies the bolting allele. *M* 1-kb ladder, *lane 1* 93167P (*BB*), *lane 2* A906001 (*bb*), *lane 3* F_2 heterozygous plant (*Bb*), *lane 4* water control. *GJ44SP6* G/C SNP was present at position 170. Primer P5bb amplifies the nonbolting allele; primer P6BB amplifies the bolting allele

haplotypes should be indicative for tight associations between markers and bolting gene. Seventeen sugar beet varieties originating from Germany and the US as well as 24 wild beet accessions from different parts of Europe were genotyped with seven BAC-derived markers (Table 4). All sugar beets are biennial, while wild beets can be divided into two classes, biennial and annual. Regarding the *B* locus, all sugar beets are supposed to have the constitution *bb*, while annual wild beets are most likely *BB* or *Bb*.

Since *B. vulgaris* is an outcrossing species, a large degree of heterozygosity was observed at the majority of marker loci investigated. These markers will not be suitable to distinguish between annual and biennial beets. Altogether, ten different sugar beet and 23 different wild beet genotypes were found, demonstrating high recombination activity within this limited region of the beet genome covering 3.8 cM (data not shown). No marker alone was able to differentiate unambiguously between biennial wild beet and annual sugar beet.

However, four markers in combination differentiated between annual wild beet and biennial sugar beet accessions, with one class of material being homozygous at a given locus. Ten out of eleven biennial wild beet accessions were homozygous at marker locus GJ131co15F6. All biennial wild beets were homozygous at marker locus GJ44SP6 (Table 4). The 'annual' allele, however, was also present in 15 of the 17 sugar beet varieties, which were either homozygous or heterozygous. Complete linkage disequilibrium (LD) was found for all sugar beet varieties between loci GJ70co3F4, GJ29T7 and the *B* gene. Hence, a unique sugar beet haplotype including the *b* allele was characteristic for all sugar beet varieties, demonstrating high selection intensity against the *B* allele. However, 11 out of 17 biennial beet accessions were also homozygous at the corresponding marker loci.

Table 3 Two-point linkage analysis of BAC-derived and RFLP-derived markers

Marker	Recombination value (R)	Variance	Distance (cM)	LOD SCORE
GJ01co36F4	0.0033	0.0000056	0.17	142.1
GJ18T7	0.0050	0.0000083	0.25	139.8
GJ17SP6	0.0050	0.0000083	0.25	139.8
GJ29T7	0.0050	0.0000083	0.25	139.8
GJ44SP6	0.0050	0.0000083	0.25	139.8
GJ70co3F4	0.0084	0.0000138	0.42	135.4
GJ131co15F6	0.0084	0.0000138	0.42	135.4
pKP374	0.0201	0.0000328	1.02	122.5

All markers were significantly linked to the *B* locus

Discussion

BACs have been proven to be suitable as resources for marker development from a defined region of the genome. We have used BAC end sequences and putative ORFs from randomly sequenced BACs, which had been assigned to chromosome 2 of sugar beet where the *B* gene is located. CAPS and ARMS marker assays

Table 4 Marker genotypes of 41 accessions from wild and cultivated species of the genus *Beta* (see 'Introduction') including the parents of the mapping population

Accession number	Origin	Type	Marker genotypes						
Parents			GJ17SP6	GJ18T7	GJ01co36F4	GJ131co15F6	GJ70co3F4	GJ29T7	GJ44SP6
93167P	Germany	Annual	381 bp	A	G	C	A	89 bp	G
A906001	Germany	Biennial	316 bp	C	A	T	T	166 bp	C
B. vulgaris L. ssp . *maritima*									
930034	Spain	Annual	381 bp	A	G/A	T	T	89 bp and 166 bp	G
930043	Belgium	Annual	381 bp and 316 bp	A/C	A	C/T	T	166 bp	G
991971	Greece	Annual	381 bp and 316 bp	C	A	T	T	166 bp	G
960066	Israel	Annual	381 bp and 316 bp	C	A	T	T	89 bp and 166 bp	G
991968	Italy	Annual	381 bp and 316 bp	A	G/A	C/T	T	166 bp	G
950407	Italy	Annual	381 bp and 316 bp	A/C	G/A	C/T	A	166 bp	G
991969	Italy	Annual	381 bp and 316 bp	A/C	A	C/T	T	166 bp	G
930029	Spain	Annual	381 bp and 316 bp	A/C	G/A	C/T	A/T	166 bp	G
930031	Portugal	Annual	316 bp	A	G	T	T	166 bp	G
950409	Spain	Annual	381 bp and 316 bp	A/C	G/A	C/T	T	166 bp	G
960059	Tunisia	Annual	381 bp and 316 bp	A/C	G/A	C/T	T	89 bp and 166 bp	G
960078	Tunisia	Annual	381 bp	A	G/A	C	T	89 bp and 166 bp	G
991972	Tunisia	Annual	381 bp and 316 bp	A/C	A	C/T	T	166 bp	G
930028	Denmark	Biennial	316 bp	A/C	G/A	T	T	89 bp and 166 bp	G/C
991957	England	Biennial	381 bp and 316 bp	A	–	T	T	89 bp	G
991947	France	Biennial	381 bp	A	A	T	T	89 bp	G
991948	France	Biennial	381 bp and 316 bp	A/C	A	T	T	89 bp and 166 bp	G/C
991954	France	Biennial	381 bp	A/C	G/A	T	T	89 bp and 166 bp	G/C
991958	France	Biennial	381 bp	A/C	G/A	T	T	166 bp	G
930038	Greece	Biennial	381 bp and 316 bp	C	A	T	T	166 bp	C
960055	Greece	Biennial	381 bp	A/C	A	C	T	166 bp	G
991943	Italy	Biennial	381 bp	A	G	T	T	166 bp	G
930033	Italy	Biennial	381 bp	A	G	T	T	166 bp	C
960037	Sweden	Biennial	381 bp and 316 bp	A	A	T	T	89 bp and 166 bp	G
B. vulgaris L. ssp . *vulgaris*									
940026	Germany	Biennial	316 bp	C	A	C/T	T	166 bp	G/C
940027	Germany	Biennial	316 bp	C	A	T	T	166 bp	G
940028	Germany	Biennial	381 bp and 316 bp	A/C	G/A	C/T	T	166 bp	C
970102	Germany	Biennial	316 bp	C	A	T	T	166 bp	G/C
970103	Germany	Biennial	316 bp	C	A	C/T	T	166 bp	G/C
970104	Germany	Biennial	316 bp	C	A	T	T	166 bp	G/C
980005	Germany	Biennial	316 bp	C	A	T	T	166 bp	G
980008	Germany	Biennial	381 bp	A	G	C/T	T	166 bp	G/C
010238	Germany	Biennial	381 bp and 316 bp	A/C	G/A	T	T	166 bp	G
010240	Germany	Biennial	316 bp	C	A	T	T	166 bp	G
010243	Germany	Biennial	316 bp	C	A	T	T	166 bp	G
010239	Germany	Biennial	381 bp and 316 bp	A/C	G/A	C/T	T	166 bp	G/C
010242	Germany	Biennial	316 bp	C	A	T	T	166 bp	G/C
010241	Germany	Biennial	381 bp and 316 bp	A/C	G/A	T	T	166 bp	G/C
930179	USA	Biennial	316 bp	C	A	T	T	166 bp	C
930182	USA	Biennial	316 bp	–	–	T	T	166 bp	G
930184	USA	Biennial	381 bp	A	G	C	T	166 bp	G

– Missing data

have been established, using PCR primers complementary to these sequences. These markers should be used for two purposes: the first is to determine recombination rates between the respective BACs and the *B* gene, which is important for ordering the BACs within the physical map, and the second is to use these sequences as molecular markers for genotyping a number of wild and sugar beet accessions. Interestingly, marker sequence GJ131co15F6 showed similarity to the Myb protein with unknown function in vernalization in *Arabidopsis thaliana* and to a nuclear receptor corepressor in human which is involved in chromatin binding and a component of an HDAC3 complex. The vernalization (*Vrn1)* gene of *A. thaliana* belongs to this class of transcription factors (Levy et al. 2002). Recently, it has been shown that this gene, together with other histone deacetylation/methylation factors, is involved in chromatin remodelling and transition to flowering in *A. thaliana* (Bastow et al. 2004; Sung and Amasino 2004).

Altogether 49 BAC-derived sequences have been investigated. PCR-based marker assays were developed for ten sequences, which rely on SNPs found between the two parental lines. The SNP frequencies were varying according to the sequence from which they were identified. The average numbers of SNPs/1,000 bp

were 2.08 for BAC ORFs and 7.0 for BAC ends. Thus, BAC ends, which are mostly noncoding, were more suitable for marker development. However, selection of single-/low-copy sequences was found to be essential. Schneider et al. (2001) found 3.5 SNPs/1,000 bp within potential coding regions of 37 genes, using EST-derived primers and two different inbred lines of sugar beet. In a study performed with 36 maize inbred lines and 17 ESTs, 16.4 SNPs/1,000 bp and 7.9 indels/1,000 bp were reported (Ching et al. 2002). The average SNP frequency was 32.2 SNPs/1,000 bp in noncoding regions and 8.06 SNPs/1,000 bp in coding regions. In melon, the average frequencies of SNPs and indels were 1/441 bp and 1/1,666 bp, respectively (Morales et al. 2004). The genome-wide scanning for SNPs in the human genome revealed 0.833 SNPs/1,000 bp (Zhao et al. 2003). For genic regions, the SNP densities in intronic, exonic and adjoining untranslated regions were 0.821, 0.528 and 0.751 SNPs/1,000 bp, respectively (Zhao et al. 2003). Here, we found that 66.7% of the SNPs derived from BAC ORFs were transitions, and the remaining ones were transversions, which is in accordance with other reports. When using sugar beet ESTs, the proportion of transitions was 62% (Schneider et al. 2001). An identical frequency of 68% for transitions was reported for maize (Ching et al. 2002) and melon (Morales et al. 2004).

BAC and YAC sequences have been considered as important sources for marker development for fine mapping of genes. Tartarini et al. (1999) have reported the development of PCR-based markers for the *Vf* scab resistance gene in apple from BAC end sequences. Cevik and King (2002) have described the development of PCR-based markers from BACs within a contig around the *Sd-1* aphid resistance locus of *Malus*. In addition, BAC end and partial BAC sequences were also used to develop PCR-based markers for late blight resistance in potato in order to increase the map resolution in the *RB* region (Bradeen et al. 2003). Jamsari et al. (2004) reported the development of diagnostic markers derived from BAC end sequences, which were used to construct a high-density map around the *M* gene of *Asparagus officinalis*.

One of the simplest marker assays relies only on PCR followed by gel electrophoresis, which requires large indels. As expected, only five BAC sequences could be genotyped in this way. Restriction digest of PCR fragments to create CAPS was successfully applied. Five sequences (11.1%) could be genotyped in this way. A problem became obvious with CAPS marker GJ131SP6. After digestion of PCR products, the intact PCR fragment (206 bp) was still present in addition to the expected restriction fragments. This phenomenon was also reported in a study about CAPS markers for the *Fom-2* fusarium wilt resistance gene in melon (Zheng et al. 1999). This could be explained by multiple copies of the respective sequence in the genome, resulting in amplification of two or more different PCR fragments with different restriction sites. To avoid this problem, only single-copy sequences should be used for marker development.

Currently a variety of techniques is used for genotyping of SNPs (Törjek et al. 2003). Tetra-primer ARMS-PCR was used for SNP genotyping in humans (Ye et al. 2001), *Mycobacterium tuberculosis* (Fan et al. 2003) and barley (Chiapparino et al. 2004). It relies on PCR amplification of two allele-specific amplicons, using four primers in one PCR reaction. Tetra-primer ARMS-PCR uses an extra mismatch at the 2-position of the 3′ terminus of each allele-specific primer, offering the advantage that only one PCR is needed, while ARMS-PCR requires two separate reactions. In our study, we have successfully used ARMS-PCR for converting two BAC ends (GJ10SP6 and GJ70SP6) into allele-specific markers. Another five SNPs (two BAC end and three BAC ORF sequences) were genotyped by tetra-primer ARMS-PCR that uses four primers at the same time to amplify both allele-specific amplicons in one PCR reaction. The insertion of a mismatch at position-2 from the 3′ termini of the allele-specific primers was very effective for ensuring high allele specificity. Ye et al. (2001) reported that a touchdown PCR programme reduced artificial products. In our study, we used gradient PCR in order to find the most suitable annealing temperature. Ye et al. (2001) showed that a 1:10 ratio between outer and inner primers increased efficiency of amplification of the allele-specific products and reduced the amount of artifacts. We found that this ratio rather leads to weak amplification of the outer fragment, which was also observed by Chiapparino et al. (2004). Therefore, we used a 1:2 ratio between outer and inner primers.

Seven BAC-derived markers closely linked to the *B* gene were used for genotyping 24 wild beet and 17 sugar beet accessions in an attempt to detect LD between markers and *B* locus. A preliminary study like this has been presented by Hansen et al. (2001), who detected two AFLP markers associated with the early-bolting character by comparing four different subpopulations of wild beets. Although a large variation became obvious in our study, resulting in as many as 23 different wild beet and 10 different sugar beet genotypes, three markers showed a close association with the annual character. For determining associations, the material was grouped into three classes: annual wild beet, biennial wild beet, and sugar beet (which is always biennial). Since breeders have been strongly selecting against the *B* allele, a marker tightly linked to this locus was expected to differentiate between the different groups of beet accessions. One sugar beet haplotype was found reflecting strong selection pressure towards nonbolting beets. However, this haplotype was also found in a number of annual wild beets. Therefore, the corresponding markers can only be used in combination with other markers to detect wild beet introgressions in sugar beet. Possibly both haplotypes are not equal by descent. If annuality (*BB*) has been the rule for wild beets, mutations within the *B* gene resulted in biennial types still having the same haplotype. This could explain the presence of GJ70co3F4-GJ29T7 haplotypes

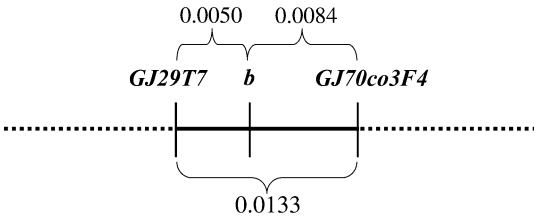

Fig. 5 Graphical presentation of marker order around the *B* gene on chromosome 2 of sugar beet. The flanking markers were found to be in complete linkage disequilibrium with the *b* allele for all sugar beet varieties tested (see Table 4). Recombination values were calculated according to Hühn (1995)

in biennial beets. Alternatively, early bolting in wild beets is due to other loci or to environmental factors which have been shown to induce early bolting in cultivated beets (Owen 1954; Shimamoto et al. 1990; Boudry et al. 1994; Abe et al. 1997).

Taking together the data from recombination mapping and LD analysis, a map around the *B* gene can be drawn. The gene is flanked by markers GJ29T7 and GJ70co3F4, with distances of 0.0050 and 0.0084 recombination units, respectively (Fig. 5). First studies on plants demonstrate that LD is dependent upon the outcrossing rate of a species and can range between a few hundred base pairs in the case of maize and 250 kb in *A. thaliana* (Neale and Savolainen 2004). Since sugar beet is an outcrossing species, the LD found here might suggest that the physical distance between both loci is very small. Unfortunately, both reside on different contigs and the gap between them has not been closed, so far.

Future work will focus on cloning the region between markers GJ29T7 and GJ70co3F4. The gene will be identified after sequencing the respective BAC and comparing candidate sequences with sequences from an EMS mutant collection that has been recently established in our institute, and that encompasses several mutants with altered vernalization requirement. The cloning of the *B* gene will enable breeding of winter beets, which require full control of the early-bolting character. In contrast to traditional spring beet cultivation, winter beets will be sown before winter. Their yielding potential is supposed to be much higher than that of spring beets, because plant development takes place early in spring under central European conditions. However, only beets, which are completely resistant to bolting, even after long periods of cold temperatures, can be grown as winter beets. The manipulation of the *B* gene by transgenic technology is the initial step towards breeding winter beets which flower early during the breeding phase, but are completely bolting resistant during cultivation in the field.

Acknowledgements The authors thank Verena Kowalewski, Birgit Defant and Monika Bruisch for excellent technical assistance. R.M. Gaafar is grateful to the Egyptian government for their financial support. This project was funded by the seed companies Strube-Dieckmann (Nienstädt, Germany) and KWS SAAT (Einbeck, Germany) and the German Federal Ministry for Education, Research and Technology (GABI grant no. 0312284).

References

Abe J, Guan G-P, Shimamoto Y (1997) A gene complex for annual habit in sugar beet (*Beta vulgaris* L.). Euphytica 94:129–135

Abegg FA (1936) A genetic factor for the annual habit in beets and linkage relationship. J Agric Res 53:493–511

Arumuganathan K, Earle ED (1991) Nuclear DNA content of some important plant species. Plant Mol Biol Rep 9:208–218

Bastow R, Mylne JS, Lister C, Lippman Z, Martienssen RA, Dean C (2004) Vernalization requires epigenetic silencing of FLC by histone methylation. Nature 427:164–167

Boudry P, Wieber R, Saumitou-Laprade P, Pillen K, Van Dijk H, Jung C (1994) Identification of RFLP markers closely linked to the bolting gene *B* and their significance for the study of the annual habit in beet (*Beta vulgaris* L.). Theor Appl Genet 88:852–858

Bradeen JM, Naess SK, Song J, Haberlach GT, Wielgus SM, Buell CR, Jiang J, Helgeson JP (2003) Concomitant reiterative BAC walking and fine genetic mapping enable physical map development for broad-spectrum late blight resistance region, *RB*. Mol Genet Genomics 269:603–611

Cevik V, King GJ (2002) Resolving the aphid resistance locus *Sd-1* on a BAC contig within a sub-telomeric region of *Malus* linkage group 7. Genome 45:939–945

Chiapparino E, Lee D, Donini P (2004) Genotyping single nucleotide polymorphisms in barley by tetra-primer ARMS-PCR. Genome 47:414–420

Ching A, Caldwell KS, Jung M, Dolan M, Smith OS, Tingey S, Morgante M, Rafalski AJ (2002) SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. BMC Genet 3:19–32

El-Mezawy A, Dreyer F, Jacobs G, Jung C (2002) High resolution mapping of the bolting gene *B* of sugar beet. Theor Appl Genet 105:100–105

Fan X-Y, Hu Z-Y, Xu F-H, Yan Z-Q, Guo S-Q, Li Z-M (2003) Rapid detection of *rpoB* gene mutations in rifampin-resistant *Mycobacterium tuberculosis* isolates in Shanghai by using the amplification refractory mutation system. J Clin Microbiol 41:993–997

Hansen M, Kraft T, Ganestam S, Säll T, Nilsson N-O (2001) Linkage disequilibrium mapping of the bolting gene in sea beet using AFLP markers. Genet Res 77:61–66

Hohmann U, Jacobs G, Telgmann A, Gaafar RM, Alam S, Jung C (2003) A bacterial artificial chromosome (BAC) library of sugar beet and a physical map encompassing the bolting gene *B*. Mol Genet Genomics 269:126–136

Hühn M (1995) Determining the linkage of disease-resistance genes to molecular markers: The LOD-SCORE method revisited with regard to necessary sample sizes. Theor Appl Genet 90:841–846

Jamsari A, Nitz I, Reamon-Büttner SM, Jung C (2004) BAC-derived diagnostic markers for sex determination in asparagus. Theor Appl Genet 108:1140–1146

Konieczny A, Ausubel FM (1993) A procedure for mapping *Arabidopsis* mutations using codominant ecotype-specific PCR-based markers. Plant J 4:403–410

Kosambi DD (1944) The estimation of map distances from recombination values. Ann Eugen 12:172–175

Levy YY, Mesnage S, Mylne JS, Gendall AR, Dean C (2002) Multiple roles of *Arabidopsis VRN1* in vernalization and flowering time control. Science 297:243–246

Little S (1997) ARMS analysis of point mutations. In: Taylor GR (ed) Laboratory methods for the detection of mutations and polymorphisms in DNA. CRC, Boca Raton, pp 45–51

Liu B-H (1998) Statistical genomics: linkage, mapping and QTL analysis. CRC, New York

Morales M, Roig E, Monforte AJ, Arús P, Garcia-Mas J (2004) Single-nucleotide polymorphisms detected in expressed sequence tags of melon (*Cucumis melo* L.). Genome 47:352–360

Munerati O (1931) L'eredità della tendenza alla annualità nella commune barbabietola coltivata. Z Zuecht, Reihe A, Pflanzenzuecht 17:84–89

Neale DB, Savolainen O (2004) Association genetics of complex traits in conifers. Trends Plant Sci 9:325–330

Owen FV (1954) The significance of single gene reactions in sugar beets. Proc Am Soc Sugar Beet Technol 8:392–398

Owen FV, McFarlane JS (1958) Successive annual backcrosses to a nonbolting inbred line of sugar beets. J Am Soc Sugar Beet Technol 10:124–132

Sadeghian YW (1993) Bolting sugar beet: genetics and physiological aspects. The Swedish University of Agricultural Sciences, Department of Plant Breeding Research, Svalöv

Saghai-Maroof MA, Soliman KM, Jorgensen RA, Allard RW (1984) Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location and population dynamics. Proc Natl Acad Sci USA 81:8014–8018

Schneider K, Weisshaar B, Borchardt DC, Salamini F (2001) SNP frequency and allelic haplotype structure of *Beta vulgaris* expressed genes. Mol Breed 8:63–74

Shimamoto Y, Tanada T, Abe J (1990) Analysis for bolting of sugar beet by means of test crosses of biennial lines with the annual line. Proc Jpn Soc Sugar Beet Technol 32:134–137

Sung S, Amasino RM (2004) Vernalization and epigenetics: how plants remember winter. Curr Opin Plant Biol 7:4–10

Tartarini S, Gianfranceschi L, Sansavini S, Gessler C (1999) Development of reliable PCR markers for the selection of the *Vf* gene conferring scab resistance in apple. Plant Breed 118:183–186

Törjek O, Berger D, Meyer RC, Müssig C, Schmid KJ, Rosleff Sörensen T, Weisshaar B, Mitchell-Olds T, Altmann T (2003). Establishment of a high-efficiency SNP-based framework marker set for *Arabidopsis*. Plant J 36:122–140

Ugozzoli L, Wallace RB (1991) Allele-specific polymerase chain reaction. Methods Enzymol 2:42–48

Vince-Prue D (1975) Photoperiodism in plants. Mcgraw-Hill, London

Ye S, Humphries S, Green F (1992) Allele specific amplification by tetra-primer PCR. Nucleic Acids Res 20:1152

Ye S, Dhillon S, Ke X, Collins AR, Day INM (2001) An efficient procedure for genotyping single nucleotide polymorphisms. Nucleic Acids Res 29:17e88

Zhao ZM, Fu YX, Hewett-Emmett D, Boerwinkle E (2003) Investigating single nucleotide polymorphism (SNP) density in the human genome and its implications for molecular evolution. Gene 312:207–213

Zheng XY, Wolff DW, Baudracco-Arnas S, Pitrat M (1999) Development and utility of cleaved amplified polymorphic sequences (CAPS) and restriction fragment length polymorphisms (RFLPs) linked to the *Fom-2* fusarium wilt resistance gene in melon (*Cucumis melo* L.). Theor Appl Genet 99:453–463