

# **Fine mapping of the bolting gene of sugar beet (*Beta vulgaris* L.) using BAC-derived sequences**

Dissertation

zur Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Christian-Albrechts-Universität zu Kiel



Vorgelegt von

M.Sc. Reda Mohamed Gaafar

aus Tanta, Ägypten

Kiel, 2005

Referent:  
Koreferent:  
Tag der mündlichen Prüfung:  
Zum Druck genehmigt:  
Kiel, den 03.02.05

Prof. Dr. Christian Jung  
Prof. Dr. Frank Kempken  
03.02.05

gez. Prof. Dr. Jürgen Grotemeyer

---

Der Dekan

**Table of Contents**

Table of Contents .....	I
List of figures .....	III
List of tables .....	V
Abbreviations .....	VII
<b>1. Introduction .....</b>	<b>1</b>
1.1. Biology of <i>Beta vulgaris</i> .....	1
1.2. Centers of origin of the species .....	2
1.3. Reproduction of <i>B. vulgaris</i> .....	2
1.4. Bolting gene and annual habit in sugar beet .....	3
1.5. DNA markers for genome analysis .....	3
1.6. Genetic linkage maps in sugar beet .....	5
1.6.1. Association mapping .....	5
1.7. Flowering time genes .....	6
1.7.1. Autonomous promotion pathway .....	7
1.7.2. Photoperiodic promotion pathway .....	7
1.7.3. Gibberellin pathway .....	7
1.7.4. Vernalization promotion pathway .....	7
1.8. Candidate gene mapping .....	11
1.9. Physical mapping around the <i>B</i> gene of sugar beet .....	12
1.10. Objectives of the present study .....	13
<b>2. Materials and methods .....</b>	<b>14</b>
2.1. Chemicals and materials .....	14
2.1.1. Enzymes .....	16
2.1.2. Primers .....	17
2.1.3. Organisms .....	21
2.2. Methods .....	23
2.2.1. DNA analysis .....	23
2.2.2. Screening of BAC libraries .....	25
2.2.3. Sequencing .....	25
2.2.4. Primer design .....	26
2.2.5. Polymerase chain reaction (PCR) .....	27
2.2.6. Southern analysis .....	28
2.2.7. Molecular marker analysis .....	30
2.2.8. Linkage analysis .....	32
2.2.9. Isolation and mapping of sugar beet flowering and vernalization orthologous genes .....	33
2.2.10. Sequence homology analysis .....	34
<b>3. Results .....</b>	<b>35</b>
3.1. Mapping populations .....	35
3.1.1. Verification of phenotypic scoring of F <sub>2</sub> population .....	35
3.1.2. Restructuring of the subpopulation 4 (960701) based on F <sub>3</sub> phenotypic data .....	37
3.1.3. Increasing number of non-bolting plants used for two-point analysis .....	37
3.2. Environmental impact on bolting character .....	38
3.3. BAC sequencing .....	39
3.3.1. BAC end sequencing and homology search .....	39
3.3.2. Shotgun sequencing of BAC clones and sequence analysis .....	40
3.4. Copy number determination of BAC sequences .....	41
3.4.1. Copy number determination of BAC ends .....	41
3.4.2. Copy number determination of BAC-derived ORF sequences .....	43
3.5. Molecular marker analysis .....	43
3.5.1. BAC-derived markers .....	43

## Table of Contents

	II
3.5.2. CAPS markers .....	45
3.5.3. SNAP and ARMS markers.....	51
3.5.4. Conversion of RFLPs into PCR markers .....	61
3.6. Linkage analysis .....	64
3.6.1. Marker genotyping in populations segregating for early bolting .....	64
3.6.2. Construction of a linkage map around the <i>B</i> gene .....	66
3.6.3. Two-point linkage analysis of non-bolting plants.....	67
3.6.4. Recombination mapping with graphical genotypes .....	68
3.7. Genotyping of wild and cultivated beets.....	69
3.8. Mapping of flowering and vernalization gene sequences .....	73
3.8.1. Marker development .....	74
3.9. Linkage analysis with the markers derived from flowering and vernalization gene sequences of <i>Arabidopsis</i> .....	76
3.10. Integrated physical and genetic map .....	77
4. <b>Discussion</b> .....	80
4.1. Segregation distortion of the <i>B</i> locus .....	80
4.2. Bolting mechanism in sugar beet and its significance .....	83
4.3. Molecular markers for the <i>B</i> gene .....	83
4.3.1. SNP characterization .....	85
4.3.2. Comparison between different markers .....	86
4.4. Copy number of BAC sequences .....	88
4.5. Fine scale mapping.....	88
4.6. Evaluation of <i>B</i> linked markers in wild <i>Beta</i> accessions .....	89
4.7. PCR mapping of sugar beet flowering time and vernalization orthologous genes .....	91
4.8. Relationship between recombination rate and physical distance .....	91
4.9. Outlook.....	92
5. <b>Summary</b> .....	94
6. <b>Zusammenfassung</b> .....	96
7. <b>References</b> .....	98
8. <b>Appendix</b> .....	114
8.1. Copy number determination of BAC-derived sequences.....	114
8.2. Markers and populations that have been mapped in this study.....	116
8.3. Sequences of BAC ends, PCR products (non-bolting, A906001 and bolting, 93167P) that have been used in the development of markers.....	117
9. <b>Acknowledgments</b> .....	118
10. <b>Curriculum vitae</b> .....	119
11. <b>Erklärung</b> .....	120

## List of figures

<b>Figure 1:</b> Genetic and physical map around <i>B</i> locus.....	12
<b>Figure 2:</b> pGEM-T vector map and sequence reference points.....	21
<b>Figure 3:</b> Schematic representation of the pedigree of the mapping populations used in the study. ....	22
<b>Figure 4:</b> Diagrammatic presentation showing the design of ORF primers.....	27
<b>Figure 5:</b> Principle of CAPS markers.....	30
<b>Figure 6:</b> Schematic representation of the SNAP and tetra-primer ARMS-PCR methods .....	32
<b>Figure 7:</b> A possible explanation for the occurrence of deviating segregation ratios of $F_3$ families due to imperfect isolation of $F_2$ plants.....	36
<b>Figure 8:</b> The synthetic $F_2$ population selected from 960701 with 299 plants (subpopulation 7). .....	37
<b>Figure 9:</b> Distribution of bolting time points in three different populations, subpop 5, subpop 6 (96701), and population 950619 sown in early May 2003.....	39
<b>Figure 10:</b> Frequency distribution of BAC ends with respect to copy numbers as determined by genomic Southern hybridization and BAC filter screening.....	42
<b>Figure 11:</b> PASA BAC-derived markers of the bolting gene <i>B</i> of sugar beet. ....	45
<b>Figure 12:</b> BAC-derived CAPS markers GJ10T7 ( <b>A</b> ) and GJ17SP6 ( <b>B</b> ) of the bolting gene <i>B</i> of sugar beet.....	48
<b>Figure 13:</b> BAC-derived CAPS markers GJ29T7 ( <b>A</b> ) and GJ49SP6 ( <b>B</b> ) of the bolting gene <i>B</i> of sugar beet.....	49
<b>Figure 14:</b> BAC-derived CAPS marker GJ131SP6 of the bolting gene <i>B</i> of sugar beet. ....	51
<b>Figure 15:</b> A part of multiple sequence alignments of PCR products amplified from different genotypes (parents) using GJ44SP6 derived primers .....	54
<b>Figure 16:</b> BAC-derived SNAP markers GJ10SP6 ( <b>A</b> ) and GJ70SP6 ( <b>B</b> ) of the bolting gene <i>B</i> of sugar beet.....	56
<b>Figure 17:</b> BAC-derived ARMS markers GJ18T7 ( <b>A</b> ) and GJ44SP6 ( <b>B</b> ) of the bolting gene <i>B</i> of sugar beet.....	58
<b>Figure 18:</b> BAC-derived ARMS markers GJ01co36F4 ( <b>A</b> ) and GJ70co3F4 ( <b>B</b> ) of the bolting gene <i>B</i> of sugar beet.....	59
<b>Figure 19:</b> BAC-derived ARMS marker GJ131co15F6 of the bolting gene <i>B</i> of sugar beet .....	60
<b>Figure 20:</b> Partial linkage maps of Chromosome 2 covering the gene <i>B</i> region.....	62
<b>Figure 21:</b> RFLP-derived CAPS markers pKP374 ( <b>A</b> ) and pKP730 ( <b>B</b> ) of the bolting gene <i>B</i> of sugar beet.....	64

## List of figures

---

	IV
<b>Figure 22:</b> Genetic map (28.3 cM interval) around the <i>B</i> locus on chromosome 2 of sugar beet .....	66
<b>Figure 23:</b> Graphical genotypes of 17 non-bolting ( <i>bb</i> ) recombinants between marker loci or marker loci and <i>B</i> locus.....	69
<b>Figure 24:</b> PASA VRN1_Bv marker derived using Arabidopsis <i>VRN1</i> sequence .....	75
<b>Figure 25:</b> PASA VRN2_Bv marker derived using Arabidopsis <i>VRN2</i> sequence .....	75
<b>Figure 26:</b> PASA CO_Bv marker derived using Arabidopsis <i>CO</i> sequence .....	76
<b>Figure 27:</b> Genetic map of the gene <i>B</i> region on chromosome 2. Two sugar beet vernalization related markers were integrated.....	77
<b>Figure 28:</b> Genetic and physical maps of the sugar beet chromosomal region encompassing the <i>B</i> gene.....	79
<b>Figure 29:</b> Graphical representation of marker order around the <i>B</i> gene on chromosome 2 of sugar beet.....	90

**List of tables**

<b>Table 1:</b> Taxonomic division of the genus <i>Beta</i> (based on Letschert <i>et al.</i> , 1994).....	1
<b>Table 2:</b> Flowering time and vernalization genes cloned from <i>Arabidopsis</i> .....	9
<b>Table 3:</b> Accessions of wild and cultivated species of the genus <i>Beta</i> . ....	23
<b>Table 4:</b> Summary of the F <sub>3</sub> -progeny analysis (full sister) for determination of bolting phenotype. ...	35
<b>Table 5:</b> Results of F <sub>3</sub> -tests from progenies of the isolated F <sub>2</sub> plants of subpopulation 4 (960701).....	36
<b>Table 6:</b> Phenotyping of F <sub>2</sub> population 960701 (subpopulations 4, 5, and 6) and of population 950619 according to F <sub>2</sub> phenotyping as well as F <sub>3</sub> phenotyping of subpopulation 4. ....	38
<b>Table 7:</b> BAC end sequences with significant homology to known plant genes. ....	40
<b>Table 8:</b> Characterization of the three shotgun libraries.. ....	41
<b>Table 9:</b> ORFs sequence homology to ESTs.....	41
<b>Table 10:</b> Copy number of different ORF sequences determined by genomic Southern hybridization and BAC filter screening.....	43
<b>Table 11:</b> List of the restriction enzymes selected for CAPS development based on restriction map analysis of BAC-end sequences using MapDraw <sup>TM</sup> . ....	46
<b>Table 12:</b> BAC-derived SNP markers, number of SNPs per BAC end, SNP frequency of the BAC ends, and number of Indels detected. ....	53
<b>Table 13:</b> BAC-ORF derived SNP markers, number of SNPs per BAC-ORF, SNP frequency of the BAC-ORFs, and number of Indels detected. ....	55
<b>Table 14:</b> PCR-based markers developed from BAC-ends or ORF-derived sequences. ....	61
<b>Table 15:</b> RFLP-derived SNPs, number of SNPs per RFLP sequence, SNP frequency of the RFLP sequence, and number of Indels detected.....	62
<b>Table 16:</b> Chi-square analysis of the marker loci used for map construction around the <i>B</i> gene .....	65
<b>Table 17:</b> Recombination values between markers and bolting gene <i>B</i> based on analysis of non-bolting plants of subpopulations 5 and 6 (960701) and of population 950619.....	67
<b>Table 18:</b> Comparison of recombination values and distances between marker loci and <i>B</i> locus, which were calculated using two different methods.....	68
<b>Table 19:</b> Genotyping of 41 accessions of the section I of the genus <i>Beta</i> including wild accessions and sugar beet varieties. ....	71
<b>Table 20:</b> Production of specific amplicons with homology to conserved domains of cloned <i>A. thaliana</i> flowering genes.....	74
<b>Table 21:</b> Results of copy number determination of the selected BACs.....	114

**List of tables**

---

**Table 22:** Copy number of different contig sequences from shotgun libraries analyzed by Southern hybridization. .... 115

**Table 23:** List of markers developed from BAC-Ends, ORF-, RFLP-, and cloned-PCR sequences.. 116



**Abbreviations**

A	adenine
AFLP	amplified fragment length polymorphism
APS	ammonium persulfate
ARMS	amplification refractory mutation system
ATP	adenosine-5'-triphosphate
BAC	bacterial artificial chromosome
BLAST	basic local alignment search tool
bp	base pair
BSA	bovine serum albumin
C	cytosine
°C	degree Celsius
CAPS	cleaved amplified polymorphic sequence
cDNA	complementary DNA
cM	centiMorgan
cm	centimeter
CTAB	cetyl trimethyl ammonium bromide
dATP	2'-deoxyadenosine-5'-triphosphate
dCTP	2'-deoxycytidine-5'-triphosphate
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	2'-deoxyribonucleotide-5'-triphosphate
DTT	1,4-dithiothreitol
EDTA	ethylene diamine tetra acetic acid
EST	expressed sequence tag
F <sub>1</sub>	first generation after a cross
F <sub>2</sub>	second generation after a cross
F <sub>3</sub>	third generation after a cross
G	guanine
h	hour
IPTG	isopropyl- β- D- thiogalactoside
ISSR	inter-simple sequence repeat amplified polymorphism
kb	kilobase
LOD	logarithm of odds
M	molar
MAS	marker assisted selection
μg	microgram
mg	milligram
min	minute
μl	microliter
ml	milliliter
mM	millimolar
mRNA	messenger RNA
ng	nanogram
nm	nanometer
OPA <sup>+</sup>	One-Phor-All buffer
PASA	PCR amplification of specific alleles
PCR	polymerase chain reaction
pd(N) <sub>6</sub>	random hexamer
pmol	picomolar
PVP	polyvinylpyrrolidone

QTL	quantitative trait loci
RAPD	random amplified polymorphic DNA
RFLP	restriction fragment length polymorphism
RNA	ribonucleic acid
RNase	ribonuclease
rpm	rounds per minute
RT	room temperature
SCAR	sequence characterized amplified region
SDS	sodium dodecyl sulfate
sec	second
SNAP	single nucleotide amplification polymorphism
SNP	single nucleotide polymorphism
SSC	sodium chloride- sodium citrate-buffer
SSPE	sodium chloride- sodium phosphate-EDTA-Buffer
SSR	simple sequence repeat
STS	sequence tagged site
T	thymine
TEMED	N,N,N',N',-tetramethylethylenediamine
Tris	tris-(hydroxymethyl)-aminomethane
Tween 20	polyoxyethylenesorbitan monolaurate
U	unit for enzyme activity
UV	ultra violet
v/v	volume per volume
vol.	volume
w/v	weight per volume
X-Gal	5- bromo- 4- chloro- 3- indolyl- $\beta$ - D- galactoside
YAC	yeast artificial chromosome

## 1. Introduction

### 1.1. Biology of *Beta vulgaris*

Sugar beet (*Beta vulgaris* L.) belongs to the family *Chenopodiaceae*. This family consists of approximately 1400 species which are divided into 105 genera (Watson and Dallwitz, 1992). Members of this family are dicotyledonous and usually herbaceous in nature. Economically important species in this family include sugar beet, fodder beet/mangolds, red table beet, Swiss chard/leaf beet (all *Beta vulgaris*), and spinach (*Spinacia oleracea*).

The genus *Beta* is divided into four sections according to Letschert *et al.*, 1994 (Table 1). The section *Beta* includes three species *Beta patula* (2n= 18), *Beta macrocarpa* (2n= 18, 2n= 36) and *Beta vulgaris* and its related subspecies (Table 1). Section *Corollinae* contains five species with chromosome numbers ranging from 2n= 18 to 2n= 54. Section *Nanae* has one species called *B. nana* with 2n= 18. The fourth section *Procumbentes* includes three species *B. patellaris* (2n= 18, 2n= 36), *B. procumbens* (2n= 18) and *B. webbiana* (2n= 18).

Sugar beet (*Beta vulgaris* L.) is a diploid species with 18 chromosomes (2n= 2x= 18). The sugar beet haploid genome size is 758 Mbp and the nuclear DNA content per haploid genome is 1.57 pg (Arumuganathan and Earle, 1991).

**Table 1:** Taxonomic division of the genus *Beta* (based on Letschert *et al.*, 1994).

Species	Chromosome number (2n)
<b>Section 1: <i>Beta</i></b>	
<i>B. vulgaris</i> ssp. <i>vulgaris</i> L.	18
<i>B. vulgaris</i> ssp. <i>maritima</i> L.	18
<i>B. adanensis</i>	18
<i>B. patula</i>	18
<i>B. macrocarpa</i>	18, 36
<b>Section 2: <i>Corollinae</i></b>	
<i>B. macrorhiza</i>	36
<i>B. lomatogona</i>	18
<i>B. corolliflora</i>	36
<i>B. trigyna</i>	54
<i>B. intermedia</i>	18
<b>Section 3: <i>Nanae</i></b>	
<i>B. nana</i>	18
<b>Section 4: <i>Procumbentes</i></b>	
<i>B. procumbens</i>	18
<i>B. webbiana</i>	18
<i>B. patellaris</i>	18, 36

In fact, cultivated sugar beet is a biennial species. However, under certain conditions it shows an annual growth habit (Smith, 1987). The biennial sugar beet plant develops a large succulent taproot in the first year and a seed stalk in the second year. Typically sugar beet root crops are planted in the spring and harvested in the autumn of the same year. For seed production, however, an over wintering period of cold temperatures of 4-7°C (vernalization) is required for the plant to bolt in the next growing season and for the reproductive stage to be initiated (Smith, 1987).

During the first growing season, the vegetative stage, the sugar beet plant is described as having glabrous leaves that are ovate to cordate in shape and dark green in color. The leaves form a rosette from an underground stem. A white, fleshy taproot develops, prominently swollen at the junction of the stem (Duke, 1983). During the second growing season, the reproductive stage, a flowering stalk elongates (bolts) from the root. This angular seed stalk forms an inflorescence and grows approximately 1.2-1.8 meters tall. A large petiolate leaf develops at the base of the stem with small leaves, further up the stem there are less petiolate leaves and finally sessile leaves developing. At the leaf axils, secondary shoots form a series of indeterminate racemes (Forster *et al.*, 1997). These flowers are small, sessile and occur singly or in clusters.

Sugar beets produce a perfect flower consisting of a tricarpelate pistil surrounded by five stamens and a perianth of five narrow sepals. Petals are absent and each flower is subtended by a slender green bract (Smith, 1987). The ovary forms a fruit, which is embedded in the base of the perianth of the flower. The fruit contains a single seed whose shape varies from round to kidney-shaped. The ovaries are enclosed by the common receptacle of the flower cluster (Duke, 1983). A monogerm seed is formed when a flower occurs singly, while the multigerm beet seed is formed by an aggregation of two or more flowers (Cooke and Scott, 1993).

## 1.2. Centers of origin of the species

The centre of origin of beet (*Beta*) is believed to be the Middle East, near the Tigris and Euphrates Rivers. It is thought that wild beets spread west into the Mediterranean and north along the Atlantic sea coast. Geographic isolation of wild beets on the Canary Islands led to the creation of several distinct species (*B. patellaris*, *B. webbiana* and *B. procumbens*) that are largely annual (Cooke and Scott, 1993).

The dispersal of wild types north into the mountains of Turkey, Iran, and the Caucasus Mountains of Russia, also led to the establishment of the species *B. trigyna*, *B. lomatogona*, and *B. macrorhiza* (Cooke and Scott, 1993). These species are somewhat perennial in growth habit. Finally, wild beet spread east through most of Eastern Asia. Cultivated sugar beet originated from wild maritima beet (*B. vulgaris* ssp. *maritima*) through breeding selection (Cooke and Scott, 1993).

## 1.3. Reproduction of *B. vulgaris*

Sugar beet seed is only produced by biennial flowers during the second year although certain conditions during the first year may cause premature bolting. Flowers reach anthesis about 5 to 6 weeks after the initiation of reproductive development. Anthesis continues for a period of several weeks. After dehiscence of the mature anthers the globular pollen is transmitted largely by wind and occasionally by insects. Sugar beet pollen is extremely sensitive to moisture. However, under dry conditions its viability is lost within 24 hours.

The primary method of pollination is cross-pollination due to the lack of synchrony between pollen release and receptiveness of the stigma. Since the pollen can be carried by the wind over long distances, breeding stock and commercial seed production fields must ensure the isolation of flowering sugar beet plants. Sugar beet is strongly self-sterile setting few or no seeds under strict isolation. The underlying genetic mechanisms may be explained by two series of multiple sterility alleles (S1 -Sn, Z1 -Zn) (Stander, 1995). The setting of some seeds after selfing, so-called pseudo-compatibility, is due to a breakdown of the incompatibility-mechanism (Stander, 1995). Pseudo-compatibility is pronounced in varying degrees in different genotypes and is highly influenced by environmental conditions, especially temperature (Stander, 1995). There is a self-fertility gene which, when introduced, can create plants which are self-fertile (Smith, 1987).

#### 1.4. Bolting gene and annual habit in sugar beet

In sugar beet, bolting (stem elongation) is the first clearly visible step of transition from vegetative to reproductive growth. Annual habit in wild beet is controlled by a dominant gene *B*, which induces bolting under long days without the cold requirement usually essential for biennial cultivars (Abe *et al.*, 1997). Variation for vernalization requirement is determined by the polymorphism of a major gene (Munerati, 1931). The dominant *B* allele enables bolting within the year the plant was sown without any vernalization and is suspected to cause the invasive character of weed beet. If young beet plants are grown at low temperatures (0°C to +10°C) for several days, this results in vernalization of the biennial cultivars (Van Dijk *et al.*, 1997). The vernalization requirements for wild and cultivated forms differ among populations. Boudry *et al.* (2002) found evidence for an adaptive significance of this trait according to geographical climate conditions. The wild *B. vulgaris* ssp. *maritima* beets often show an annual habit, since they start flowering in their first year of growth, and usually die after seed set due to unfavorable environmental conditions during hot and dry summers in southern Europe. Flowering in the first season without vernalization is mediated by genetic adaptation to different environmental factors. First year flowering is genetically related to a single locus called *B* (bolting), and the inheritance of this trait is only partly Mendelian (Boudry *et al.*, 2002). In sugar beet, genes for cold sensitivity have been selected against so that the plants exhibit low levels of bolting. This is based primarily, but not exclusively, on homozygosity of “*bb*” in diploid or “*bbb*” in triploid sugar beet varieties. The *B* allele removes vernalization requirement, but it is not unambiguously dominant.

The *B* allele is highly selected in weeds because of the fast crop turnover. While this allele is absent in the cultivated beets, it is extremely frequent among both the ruderal populations in the seed-production area (Boudry *et al.*, 1993) and the natural coastal populations of Biscay and the Mediterranean (Van Dijk *et al.*, 1997). The penetrance of *B* was found to be influenced by both environmental and genetic factors (Abegg, 1936; Owen, 1954; Owen and McFarlane, 1958; Shimamoto *et al.*, 1990; Boudry *et al.*, 1994a). Long day length is essential for the homozygous (*BB*) plants to induce bolting where they are no longer annuals with well-developed rosettes when planted in early August (Owen, 1954). However, more complicated behavior was recognized in heterozygous (*Bb*) plants (Abegg, 1936; Owen, 1954; Owen and McFarlane, 1958; Shimamoto *et al.*, 1990; Boudry *et al.*, 1994b). The plants usually behave as annuals in favorable environments for bolting, but when planted in late June and July, some plants are annuals while others behave as biennials. Owen (1954) discussed that much of the observed variation in heterozygous *B* plants probably developed from genes introduced with *b* (bienniality) that influenced the expression of the *B* gene. Boudry *et al.* (1994b) found genes, modifying the penetrance of annual habit in wild coastal beets. Abe *et al.* (1997) suggested that the annual habit of sugar beet is partly controlled by a gene complex consisted of two closely linked loci, *B* and *Lr* (Long day requirement), and controlled singly by *B* for cold non-requirement under long days, a favorable environment for bolting. Therefore, the detection efficiency of *B* allele is reduced due to its incomplete penetrance (Owen, 1954; Boudry *et al.*, 1994a).

Bolting of sugar beet is undesirable, as it results in reduction of yield and gives rise to problems during harvesting and sugar extraction. Flowering and bolting can be induced in many plant species by exogenous treatment of gibberellins, which indicates that GAs play a role in the vernalization-induced transition to flowering (Thomas *et al.*, 1998).

#### 1.5. DNA markers for genome analysis

In the past two decades, molecular marker techniques have been developed as a result of the needs of genome analysis. These techniques range from molecular assays for genetic mapping, gene cloning, and marker assisted plant breeding to genome fingerprinting and for the investigation of genetic relationships. Genetic markers are based on DNA polymorphisms in the nucleotide sequences of genomic regions which either detected by restriction enzymes, or two priming sites. Not all types of markers are the same; the information content depends on the method that was used to obtain the

marker data and the population in which the markers were scored. For example, it is not always possible to distinguish genome fragments that are present in homozygous condition from heterozygous fragments. In a heterogeneous population like an F<sub>2</sub>, co-dominant markers like restriction fragment length polymorphisms (RFLPs, Botstein *et al.*, 1980) and co-dominantly scored amplified fragment length polymorphisms (AFLPs, Vos *et al.*, 1995) yield more information than dominant markers like random amplified polymorphic DNAs (RAPDs, Welsh and McClelland, 1990) and dominantly scored AFLPs.

RFLPs are co-dominant and are able to identify a unique locus. Therefore, they are very informative. RFLP mapping together with molecular cloning of genes provide the stage for establishing syntenic relationships for a number of plant and animal species. In contrast to RFLPs, PCR-based techniques developed during the last fifteen years require only approximately 10% of DNA amount as template for PCR amplification of large quantities of the target sequence. However, the unpredictable behavior of short primers which is affected by numerous reaction conditions, inheritance in a dominant manner, and population specificity are the main disadvantages of RAPDs.

Microsatellites, or simple sequence repeats (SSRs), simple sequence length polymorphisms (SSLPs), short tandem repeats (STRs), simple sequence motifs (SSMs), and sequence target microsatellites (STMs) represent a class of repetitive sequences which are widely distributed in all eukaryotic genomes. The variation in number and length of the repeats is a source of polymorphism even between closely related individuals. Such microsatellite sequences can be easily amplified by PCR using a pair of flanking locus-specific oligonucleotides as primers and detect DNA length polymorphisms (Litt and Luty, 1989; Weber and May, 1989).

Polymorphisms corresponding to differences at a single nucleotide position (substitution, deletion or insertion) occur approximately every 1.3 kb in human (Cooper *et al.*, 1985; Kwok *et al.*, 1996) and are referred to single nucleotide polymorphisms or SNPs. Most polymorphisms of this type have only two alleles and are also called biallelic loci. Positional cloning based on SNPs may accelerate the identification of disease traits and a range of biologically informative mutations (Wang *et al.*, 1998). The increasing production of genomic sequence data in conjunction with improved methods for SNP analysis are leading to the systematic generation of today's genetic maps consisting of more and more SNPs. Many methods have been developed for SNP genotyping (Landegren *et al.*, 1998; Bhattaramakki and Rafalski, 2001). These methods detect single-nucleotide variations and usually rely on an expensive DNA sequencing equipment for genotyping, as for example in performing single-base extension or pyrosequencing (Ronaghi *et al.*, 1998). Others are based on hybridization assays that require radioactivity or reporter molecules such as the TaqMan system (PE Biosystem, Warrington, U.K.). These methods are developed for high throughput, with cost being a secondary issue. So far, the expense has limited the uptake of this class of DNA markers.

Therefore, there is a need for simple and accurate genotyping assays that can be used in laboratories that do not have access to sophisticated equipment. A solution to this problem is the detection of a SNP site by an appropriate restriction endonuclease whose recognition sequence has been altered or introduced by the SNP. Three widely used types of SNP detection using PCR method are cleaved amplified polymorphic sequences (CAPS; Konieczny and Ausubel, 1993; Thiel *et al.*, 2004), derived CAPS (dCAPS; Michaels and Amasino, 1998; Neff *et al.*, 1998), and single strand conformation polymorphism (SSCP; Orita *et al.*, 1989). CAPS markers detect polymorphisms that occur in restriction sites and dCAPS markers are created during PCR amplification by introducing a restriction site at the position of a SNP using specially designed primers. While, SSCP technique separates denatured double stranded DNA on a non-denaturing gel, and thus allows the secondary structure, as well as the molecular weight, of single stranded DNA to determine gel mobility.

A simple and economical method involving a single PCR was reported for SNP genotyping (Fan *et al.*, 2003; Chiapparino *et al.*, 2004), in which the tetra-primer ARMS (amplification refractory mutation

system)-PCR procedure (Ye *et al.*, 2001), employing two primer pairs, was used to amplify two different alleles of a SNP in a single PCR reaction. This method has been shown to be reliable, robust, inexpensive, and easy to use. The informativeness and ease of genotyping are the most important criteria determining choice of marker assay.

SNPs and insertion-deletions, which are the basis of most differences between alleles, are considered as an essentially inexhaustible source of polymorphic markers used in the high-resolution genetic mapping of traits, and for association studies, which are based on candidate genes or possibly whole genomes (Rafalski, 2002). Schneider *et al.* (2001) found 3.5 SNPs/1000 bp within potential coding regions of 37 genes using EST-derived primers and two different inbred lines of sugar beet.

In sugar beet many agronomical useful traits have been mapped using RFLP, AFLP and RAPD markers, e.g. *Hs1<sup>pro-1</sup>* (Jung *et al.*, 1992) and *Hs2<sup>pro-7</sup>* (Löptien, 1984; Lange *et al.*, 1993) for nematode resistance, *Rr1* (Büttner *et al.*, 1995) for rhizomania resistance, as well as QTLs (Setiawan *et al.*, 2000) for resistance to *Cercospora beticola* leaf spot disease and for the bolting gene (Boudry *et al.*, 1994b; El-Mezawy *et al.*, 2002).

## 1.6. Genetic linkage maps in sugar beet

Segregation analysis can be applied to a segregating population that is derived from common ancestors. Markers that co-segregate (always present or absent together) must be linked, i.e. they must be located in vicinity to each others in the genome. However, in some cases due to recombination events, the linkage between the markers may be lost. The frequency with which the linkage between co-segregating markers is broken is an indication of the genetic distance between the markers. An extensive analysis of the linkage between large numbers of molecular markers yields information on their arrangement in the genome. Such analysis can finally result in the construction of a genetic linkage map, on which all markers are arranged in separate linkage groups or chromosomes. On such a map, the distances between markers reflect the degree of observed recombination events. Linkage maps and physical maps are related, but this relation is usually not linear (Schmidt *et al.*, 1995).

Various genetic maps of sugar beet have been published. The first genetic map of sugar beet was established using isozyme markers, where Wagner and Wircke (1991) studied the genetic controls of five isozyme systems. Then using morphological, isozyme, and RFLP markers, the first linkage map of sugar beet was constructed (Pillen *et al.*, 1992). Barzen *et al.* (1995) developed an extended map which includes RAPD and RFLP marker loci. Schondelmaier *et al.* (1995) have developed a linkage map based on RFLP markers. This map was later extended by 120 AFLP markers (Schondelmaier *et al.*, 1996). Combining of different sugar beet linkage maps was described by Schumacher *et al.* (1997). Taken into account localization and mapping of agronomically useful traits in sugar beet, major genes of hypocotyl color, monogerm, and rhizomania resistance were localized using RFLP markers (Barzen *et al.*, 1992). RFLP markers closely linked to the *B* gene have been identified (Boudry *et al.*, 1994b). RFLP markers were used to localize four genes of nematode resistance in different sugar beet lines (Heller *et al.*, 1996). These RFLP markers were successfully used for positional cloning of the *Hs1<sup>pro1</sup>* from sugar beet (Cai *et al.*, 1997). Recently a high resolution genetic map of the bolting gene *B* of sugar beet based on AFLP markers has been published (EL-Mezawy *et al.*, 2002).

### 1.6.1. Association mapping

An alternative approach to the common method for mapping of plant genes is association mapping. Linkage disequilibrium (LD) is the nonrandom allelic association of genetic markers and phenotypes in a population. Disequilibrium can also be produced and maintained by selection favoring one combination of alleles over another. Significant LD is unlikely to remain for many generations, except if there are tightly linked markers to a defined gene. Association mapping allows a more finer mapping with high resolution than the normal bi-parental cross approach. LD mapping approach proves the

associations between phenotypic variation and haplotypes in a genomic region (Olsen *et al.*, 2004). LD is affected by biological factors such as recombination and other factors like selection, inbreeding, and fluxes in population size (Gaut and Long, 2003). The LD has been successfully applied in studies of *Drosophila* (Long *et al.*, 1998), maize (Thornsberry *et al.*, 2001), human (Fullerton *et al.*, 2000), and *Arabidopsis thaliana* (Olsen *et al.*, 2004) to identify genes and specific polymorphisms within genes underlying natural phenotypic variation.

Linkage disequilibrium has been originally developed for mapping human disease genes (Risch, 2000). In plants, since many of the important plant traits are difficult to assess phenotypically, or may require many years of testing, therefore the use of collections of well-characterized inbred lines, such as commercial cultivars, could result in better phenotypic data and thus better mapping (Rafalski, 2002). The applicability of whole genome scan-type association mapping depends on the extent of linkage disequilibrium in the population. If the level of LD is low, whole genome scan may be impractical because of the excessive number of markers required (Kruglyak, 1999). In European human populations, LD extends at least to 60 kb (Reich *et al.*, 2001), while in plants, little is known about the extent of LD, even though many laboratories are working on this issue (Thornsberry *et al.*, 2001). Additional complication is that the extent of LD is population-dependent as well as expected to vary across the genome as a function of recombination frequency and history. Therefore, LD measured in a wild population may be completely different from that is observed in a breeding population, which has been subjected to bottlenecks. LD mapping seems, therefore, to be more useful for precise estimation of QTL positions, while classical mapping is more useful for a genome-wide scans for QTL.

Thornsberry *et al.* (2001) and Pritchard (2001) found association of *Dwarf8* gene with flowering time variation in 92 maize inbred lines by candidate gene approach. Such association genetic methodologies for mapping complex traits in existing populations will be further developed and will allow rapid and sensitive mapping of simple and complex traits. Simko (2004) described the haplotype association mapping method in autotetraploid potato, which allows exploitation of biodiversity without developing new populations. In sugar beet, Kraft *et al.* (2000) have investigated the relationship between linkage and linkage disequilibrium among mapped AFLP markers in a selection of sugar beet breeding lines. Another study on linkage disequilibrium mapping of the bolting gene has been reported by Hansen *et al.* (2001). The authors described linkage disequilibrium among 137 AFLP markers in four natural populations of sea beet. Two AFLP markers were found to be in linkage disequilibrium with the bolting (*B*) gene.

### 1.7. Flowering time genes

Flowering is a result of interaction between internal and external factors. Flowering time is influenced by both environmental conditions (which include day-length, temperature, light quality, and nutrient deprivation) and developmental factors associated with the age of the plant (Koornneef *et al.*, 1998). A large number of flowering mutants are available in *Arabidopsis thaliana* and have been classified into early mutants, which advance flowering in comparison to the wild type, and late mutants, that delay flowering time. Genetic analyses of late-flowering mutants identified more than 20 genes. The mutants involved in the control of this process were physiologically classified into four genetic pathways on the basis of their responsiveness to environmental factors (Simpson *et al.*, 1999; Devlin and Kay, 2000; Samach and Coupland, 2000). Recently, molecular studies have shown that flowering time in plants is under complex genetic control and is regulated by several pathways involving vernalization, perception of day-length, and response to phytohormones such as gibberellin, as well as an autonomous pathway (Simpson and Dean, 2002).



### 1.7.1. Autonomous promotion pathway

The first class of flowering time genes belongs to the autonomous pathway, which promotes the transition from vegetative to reproductive development under both long-day and short-day conditions. This pathway was identified via a group of mutations that are late-flowering under all photoperiods, and are highly responsive to vernalization (Martinez-Zapater and Somerville, 1990). These mutants (e.g., *fca*, *fy*, *fpa*, *ld*, *fve*) contain higher levels of *FLC* mRNA than do wild type plants or late-flowering mutants affected in the long-day or GA pathways (Koornneef *et al.*, 1991; Koornneef *et al.*, 1998). These mutants delay flowering by increasing expression of *FLC*. Therefore, this pathway negatively regulates *FLC* expression in wild-type plants. Several genes in the autonomous pathway have been cloned. *FCA* encodes a novel, nuclear RNA binding protein with two N-terminal RRM (RNA recognition motif)-type RNA binding domains (Macknight *et al.*, 1997; Quesada *et al.*, 2003) and a C-terminal WW protein-protein interaction domain (Mcknight *et al.*, 1997).

Interestingly, *FPA* also encodes an RNA binding protein, suggesting a general role for post-transcriptional regulation in the autonomous pathway. *LUMINIDEPENDENS (LD)* may encode a transcription factor. How these proteins regulate *FLOWERING LOCUS C (FLC)* expression is not known. MADS-box genes like *FLOWERING LOCUS C (FLC)* and *SHORT VEGETATIVE PHASE (SVP)* belong to this pathway. Both of them negatively regulate the transition, but *FLC* is the most central regulator of this pathway (Michaels and Amasino, 1999; Hartmann *et al.*, 2000). Compared to late flowering genes, less is known about early flowering genes. *TERMINAL FLOWER (TFL)* controls both flowering time and the identity of the shoot meristem (Shannon and Meekss-Wagner, 1991; Alvarez *et al.*, 1992). Therefore, *TFL* provides a link between the control of flowering time and flowering initiation.

### 1.7.2. Photoperiodic promotion pathway

The second pathway is the photoperiodic pathway (also called the long day pathway) which promotes flowering only under long-day conditions but has no effect under short days. Three genes are involved in this pathway *CONSTANS (CO)*, *GIGANTEA (GI)*, and *FLOWERING LOCUS T (FT)*. *CONSTANS (CO)* is a zinc-finger protein (Putterill *et al.*, 1995). *GI* controls the expression of *CO* which in turn activates the floral pathway integrators, *FT* and *SOC1 (SUPPRESSOR OF OVEREXPRESSION OF CO1)*. This results in the activation of floral identity genes *APETALA1 (API)* and *LEAVY (LFY)*, which promotes the floral transition (Simpson, 2003). Recent studies have demonstrated that regulation of *CO* protein stability is central to the photoperiodic regulation of flowering (Valverde *et al.*, 2004). Flowering is induced when *CO* mRNA expression coincides with the exposure of plants to light, which only occurs in long day photoperiods. Photoreceptors regulate *CO* stability so that *CO* protein abundance is high enough to induce flowering only in long days.

### 1.7.3. Gibberellin pathway

The day-length independent pathway (also called the gibberellin pathway) stimulates flowering via the plant hormone gibberellin (GA). The growth regulator GA has been thought to induce flowering in many plants species (Bernier, 1988). Flowering in Arabidopsis is also promoted by GA. Exogenous treatment of GA accelerates flowering in Arabidopsis particularly under short days (Chandler and Dean, 1994) and strong GA biosynthesis mutants fail to flower under short days and show a slight delay in flowering under long days (Wilson *et al.*, 1992). Therefore, GA is absolutely necessary for flowering under short days in Arabidopsis.

### 1.7.4. Vernalization promotion pathway

In this pathway, flowering is promoted by exposure of imbibed seeds or vegetative plants to a period of cold temperature. Vernalization prepares the plant to flower but does not itself induce flowering.

The *vrn* mutants may be defective either in the perception of cold temperature or in the transduction of the cold signal by the vernalization promotion pathway. There are four genes, *VRN1*, *VRN2*, *VRN3*, and *VRN4* in *Arabidopsis*, which are related to vernalization based on the data of genetic mapping (Chandler *et al.*, 1996). *VRN1* and *VRN2* are located on the chromosomes III and IV, respectively. The winter annual habit in *Arabidopsis* is conferred by a single dominant gene, which is named *FRIGIDA* (*FRI*). Studies of natural variation have also shown that a dominant allele of another gene, *FLOWERING LOCUS C* (*FLC*), is necessary for *FRI* to confer a winter-annual habit (Lee *et al.*, 1994a; Koornneef *et al.*, 1994). Vernalization overcomes the effect of *FRI* by repressing *FLC* expression and this repression is stably maintained after a return to warm growth conditions. Thus, the epigenetic repression of *FLC* is a key feature of vernalization (Sung and Amasino, 2004).

Genes involved in the stable maintenance of *FLC* repression have recently been isolated (Gendall *et al.*, 2001). *VRN1* encodes a DNA binding protein and *VRN2* encodes a homologue to Polycomb group proteins indicating that the repression of *FLC* by vernalization requires chromatin-remodeling factors. Vernalization causes changes in histone methylation within the *FLC* locus. These changes result in silenced chromatin states in other organisms. Analysis of methylation patterns of *vrn* mutants revealed a "histone code" that specify a silenced chromatin state mediates the memory of winter (Bastow *et al.*, 2004).

Increasing numbers of genes that affect flowering time in *Arabidopsis* are being cloned, and many of them encode transcription factors (Ratcliffe and Riechmann, 2002). Table 2 shows the most of the cloned flowering genes of *Arabidopsis thaliana*.

**Table 2:** Flowering time and vernalization genes cloned from Arabidopsis.

Genes	Symbol	Gene Product/Functional Type	Pathway/Role	Reference
ANTHOCYANINLESS 2	<i>ANL2</i>	Transcription factor (HB family)	Floral repressor	(Kubo <i>et al.</i> , 1999; Weigel <i>et al.</i> , 2000)
APETALA1	<i>AP1</i>	MADS domain	Floral promoter, floral homeotic genes	(Mandel <i>et al.</i> , 1992)
APETALA3	<i>AP3</i>	MADS domain	Floral promoter, floral homeotic genes	(Jack <i>et al.</i> , 1992)
CIRCADIAN CLOCK ASSOCIATED 1	<i>CCA1</i>	Transcription factor (Myb-related group)	Circadian clock, floral repressor	(Wang <i>et al.</i> , 1997; Wang and Tobin, 1998)
CKB3	<i>CKB3</i>	Regulatory subunit of Ser/Thr protein kinase	Circadian clock	(Sugano <i>et al.</i> , 1998; Sugano <i>et al.</i> , 1999)
CONSTANS	<i>CO</i>	Transcription factor (Zinc coordinating type)	Floral promoter, photoperiod pathway	(Putterill <i>et al.</i> , 1995)
CRYPTOCHROME 1/ELONGATED HYPOCOTYL 4	<i>CRY1/HY4</i>	Photoreceptor	Light perception	(Ahmad and Cashmore, 1993)
CRYPTOCHROME 2/FHA	<i>CRY2/FHA</i>	Photoreceptor	Light perception, photoperiod pathway	(Guo <i>et al.</i> , 1998)
EARLY FLOWERING 3	<i>ELF3</i>	Novel nuclear protein	Circadian clock	(Covington <i>et al.</i> , 2001; Hicks <i>et al.</i> , 2001; Liu <i>et al.</i> , 2001)
EMBRYONIC FLOWER 1	<i>EMF1</i>	Putative novel type of transcription factor	Floral repressor, multiple pathways	(Aubert <i>et al.</i> , 2001)
EMBRYONIC FLOWER 2	<i>EMF2</i>	Chromatin related protein	Floral repressor, multiple pathways	(Yoshida <i>et al.</i> , 2001)
FCA	<i>FCA</i>	RNA binding protein	Floral promoter, autonomous pathway	(Macknight <i>et al.</i> , 1997)
FERTILIZATION INDEPENDENT ENDOSPERM	<i>FIE</i>	Chromatin related protein	Floral repressor	(Kinoshita <i>et al.</i> , 2001)
FLAVIN BINDING, KELCH REPEAT, F BOX	<i>FKF1</i>	PAS protein	Circadian clock	(Nelson <i>et al.</i> , 2000)
FLORAL PROMOTING FACTOR 1	<i>FPF1</i>	Novel class of protein	Floral promoter	(Kania <i>et al.</i> , 1997)
FLOWERING LOCUS C	<i>FLC</i>	Transcription factor (MADS family)	Floral repressor, multiple pathways	(Michaels and Amasino, 1999; Sheldon <i>et al.</i> , 1999)
FLOWERING LOCUS T	<i>FT</i>	Phosphatidylethanolamine binding protein	Floral promoter, multiple pathways	(Kardailsky <i>et al.</i> , 1999; Kobayashi <i>et al.</i> , 1999)
FPA	<i>FPA</i>	RNA binding protein	Floral promoter, autonomous pathway	(Schomburg <i>et al.</i> , 2001)
FRIGIDA	<i>FRI</i>	Novel class of protein	Floral repressor	(Johanson <i>et al.</i> , 2000)
FWA	<i>FWA</i>	Transcription factor (HB family)	Floral repressor	(Soppe <i>et al.</i> , 2000)
GA REQUIRING 1	<i>GAI</i>	Enzyme	GA regulation	(Sun and Kamiya, 1994)
GIBBERELLIC ACID INSENSITIVE	<i>GAI</i>	Transcription factor (GRAS family)	GA regulation	(Peng <i>et al.</i> , 1997)
GIGANTEA	<i>GI</i>	Novel nuclear localized protein	Floral promoter, photoperiod pathway	(Fowler <i>et al.</i> , 1999; Park <i>et al.</i> , 1999)
LATE ELONGATED HYPOCOTYL	<i>LHY</i>	Transcription factor (Myb-related group)	Circadian clock related, floral repressor	(Schaffer <i>et al.</i> , 1998)

Table 2: continued.

LIKE HETEROCHROMATIN PROTEIN 1	<i>LHP1</i>	Chromatin related protein	Floral repressor	(Gaudin <i>et al.</i> , 2001)
LOV KELCH PROTEIN 2	<i>LKP2</i>	PAS protein	Circadian clock	(Schultz <i>et al.</i> , 2001)
LUMINIDEPENDENS	<i>LD</i>	Novel protein	Floral promoter, autonomous pathway	(Lee <i>et al.</i> , 1994b)
MADS AFFECTING FLOWERING 1/FLOWERING LOCUS M	<i>MAF1/FLM</i>	Transcription factor (MADS family)	Floral repressor, multiple pathways	(Ratcliffe <i>et al.</i> , 2001; Scortecci <i>et al.</i> , 2001)
PHYTOCHROME A/ELONGATED HYPOCOTYL 8	<i>PHYA/HY8</i>	Photoreceptor	Light perception	(Sharrock and Quail, 1989)
PHYTOCHROME B/ELONGATED HYPOCOTYL 3	<i>PHYB/HY3</i>	Photoreceptor	Light perception	(Sharrock and Quail, 1989)
PHYTOCHROME INTERACTING FACTOR 3	<i>PIF3</i>	Transcription factor (bHLH family)	Photoreceptor signal transduction	(Ni <i>et al.</i> , 1998)
REPRESSOR OF GA1-3	<i>RGA</i>	Transcription factor (GRAS family)	GA regulation	(Silverstone <i>et al.</i> , 1998)
SHORT INTERNODES	<i>SHI</i>	Transcription factor (Zinc coordinating type)	GA regulation	(Fridborg <i>et al.</i> , 1999)
SHORT VEGETATIVE PHASE	<i>SVP</i>	Transcription factor (MADS family)	Floral repressor	(Hartmann <i>et al.</i> , 2000)
SQUAMOSA PROMOTER BINDING PROTEIN-LIKE 3	<i>SPL3</i>	Transcription factor (SBP family)	Floral promoter, multiple pathways	(Cardon <i>et al.</i> , 1997)
SUPPRESSOR OVEREXPRESSION CONSTANS 1	<i>SOC1/AGL20</i>	Transcription factor (MADS family)	Floral promoter, multiple pathways	(Borner <i>et al.</i> , 2000; Lee <i>et al.</i> , 2000; AGAMOUS-LIKE 20 Samach <i>et al.</i> , 2000)
TERMINAL FLOWER 1	<i>TFL1</i>	Phosphatidylethanolamine binding protein	Floral repressor, multiple pathways	(Bradley <i>et al.</i> , 1997)
TIMING OF CAB EXPRESSION 1	<i>TOC1</i>	Putative transcriptional regulator	Circadian clock	(Strayer <i>et al.</i> , 2000)
VERNALIZATION 1	<i>VRN1</i>	Transcriptional factor (B3 family)	Floral promoter, vernalization pathway	(Chandler <i>et al.</i> , 1996)
VERNALIZATION 2	<i>VRN2</i>	Chromatin related protein	Floral promoter, vernalization pathway	(Gendall <i>et al.</i> , 2001)
ZEITLUPE	<i>ZTL</i>	PAS protein	Circadian clock	(Somers <i>et al.</i> , 2000)

MADS, it is in reference to the four proteins (MCM1, AG, DEF A and SRS), which have a conserved domain in common.

## 1.8. Candidate gene mapping

Candidate gene mapping is based on the hypothesis that a gene with known function that may affect genetic control of a trait can be considered as a candidate gene for this trait. For QTL, candidate genes analysis uses known genes as markers to identify their map position on chromosomes. For this purpose, the following approaches can be used: Southern hybridization, with a cloned gene as probe, is one approach that localizes the candidate gene sequence on a molecular linkage map (Gebhardt *et al.*, 1994). Instead, linkage maps based on expressed sequence tags (ESTs) were considered as have been done in rice (Kurata *et al.*, 1994) and maize (Chao *et al.*, 1994). However, this depends on the availability of cloned genes and a hybridization-based marker is not suitable for marker-assisted selection (MAS) screening of plants on a large scale. An alternative approach is to use the public databases as sequence sources for mapping functional genes that are of agronomic interest and to develop PCR-based and gene-specific markers. This approach takes the advantage of high conservation of genes between related species and uses consensus sequence motifs as primers to amplify analogous sequences (Schneider *et al.*, 1999). Conservation of genic regions is well documented in closely related species within the same family: for example, *Arabidopsis* and *Brassica oleracea*; rice and barley; and tomato, potato and capsicum (Schmidt, 2002).

Using available sequences information of various species (e.g. *Arabidopsis* and rice) as RFLP probes in closely related species has led to the extensive genome comparison and showed conservation of gene order within homoeologous chromosomal segments. Such studies demonstrated that tomato and potato maps were very similar (Bonierbale *et al.*, 1988), a close genomic relationship between cereals (Gale and Devos, 1998a), and between the mustard *Brassica* and *Arabidopsis* (Teutonico and Osborn, 1994). Wheat, rye, barley, maize, and sorghum seem to have genetic view similar to rice; i.e., the genes are in much the same order but the larger genomes have more uncoding DNA between genes (Gale and Devos, 1998b).

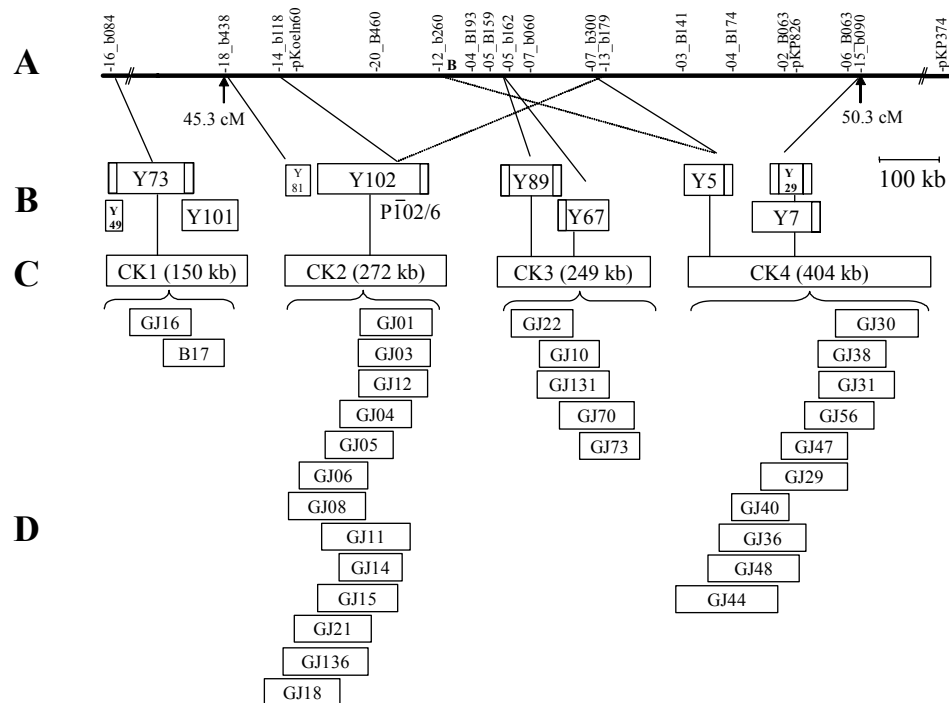
Map position of homologous genes is partly conserved across species, providing a way by which results from one species can be used to identify genes affecting a QTL in a related species. For example, humans and mice show rather extensive conservation of gene order. Therefore, a marker tightly linked to a QTL in mouse has a good chance of also being tightly linked to the homologous gene in human (Nadeau, 1989). Many commercial crops (cotton and sugar beet) are relying on the co-linearity information of these model crops (Cook, 1998; Adam, 2000). Schneider *et al.* (1999) described the PCR-based cloning of functional gene homologues in *Beta vulgaris*. Using primers derived from 18 homologous ESTs and conserved regions of 32 heterologous ESTs encoding gene products related to carbohydrate and nitrogen metabolism, 42 functional markers have been developed and assigned to the nine linkage groups of sugar beet.

Recently, the PCR-based cloning approach was used for cloning of resistance gene analogous conferring resistance to nematodes in sugar beet. The PCR amplification of resistance domains was based on a set of degenerated primers (Tian *et al.*, 2004).

In *Arabidopsis thaliana*, the flowering time and vernalization genes have been identified and characterized. These gene sequences are important as candidate genes for *B* gene in sugar beet, in identification and isolation of their sugar homologues. On the other hand, these sequences provide a source for development of sugar beet specific markers for flowering time and vernalization. This at the end could enable the determination of their map positions and the relation between them and bolting gene *B* in sugar beet.

### 1.9. Physical mapping around the *B* gene of sugar beet

In a former project a genetic linkage map of *B* gene locus has been constructed by El-Mezawy *et al.* (2002) using 224 AFLP and 2 RFLP marker loci. This linkage map was based on 775 F<sub>2</sub> plants. From these marker loci, 17 loci were linked to *B* within a distance of 5 cM (45.3-50.3 cM) (Figure 1A).



**Figure 1:** Genetic and physical map around *B* locus. (A) The genetic map of the gene *B* region based on 775 plants of the population 960701 (El-Mezawy *et al.*, 2002). (B) The YAC-contigs and YAC clones used for screening of the BAC library. (C) The first BAC-contigs (minimal tiling paths) obtained by RFLP-fingerprinting. The CK-contigs of the *b*-genotype BAC library (D) The BAC-contigs consist of 2 to 13 different BACs. The size bar is valid for YAC contigs only (modified after Hohmann *et al.*, 2003).

Some of these AFLP marker loci were used for selecting YAC clones from a YAC library, which were employed in construction of YAC-contigs (Jacobs, 2001) (Figure 1B). Four AFLP marker loci (P04\_B174, P04\_B193, P05\_B159 and P05\_b162) were used for fine mapping of *B* locus using 1359 plants F<sub>2</sub> population 960701 (subpopulation 4) (El-Mezawy *et al.*, 2002). By increasing map resolution, the marker locus P04\_B193 showed the lowest recombination rate and mapped 0.14 cM from the *B* gene. Using these four AFLP marker loci and 10 single copy YAC-end sequences as probes, 54 BAC clones were selected (Hohmann *et al.*, 2003).

These selected BAC clones were used to construct a physical map around the *B* gene (Figure 1C&D). This physical map consists of four BAC contigs (Hohmann *et al.*, 2003). One small contig (CK1, 2 BAC clones) and three large contigs (CK2, 13 BAC clones; CK3, 5 BAC clones and CK4, 10 BAC clones) could be constructed.

**1.10. Objectives of the present study**

The present study focuses mainly on:

- 1- Genetic fine mapping of the bolting *B* gene region using BAC sequences derived CAPS and ARMS markers.
- 2- Use of PCR-based BAC-derived markers for genotyping a number of wild and sugar beet accessions in order to detect the presence of any linkage disequilibrium (LD) between the markers and *B* locus.
- 3- Development of candidate gene markers from Arabidopsis derived sequences and mapping of these markers in a segregating F<sub>2</sub> population in order to specify their distances to the *B* gene.

## 2. Materials and methods

### 2.1. Chemicals and materials

Chemicals	Cat. No.	Suppliers
$\alpha$ [ <sup>32</sup> P]-dATP	AA0004	Amersham Buchler, Braunschweig
$\alpha$ [ <sup>32</sup> P]-dCTP	AA0005	Amersham Buchler, Braunschweig
$\beta$ -Mercaptoethanol	50407	Biomol Feinchem. GmbH., Hamburg
Agar	5210.2	Roth, Karlsruhe
Agarose, Metaphore	50180	BMA, Rockland, USA
Ammonium acetate	1116	Merck, Darmstadt
Ammonium persulfate	15523-012	Life Technologies, USA
Ampicillin	A0104	Duchefa, Haarlem, The Netherlands
ATP	A-2383	Sigma, Steinheim
Bacto <sup>®</sup> -Trypton	0123-17-3	Difco Laboratories, Detroit, USA
Bacto <sup>®</sup> -Yeast extract	0118-17-0	Difco Laboratories, Detroit, USA
Boric acid	1.12015	Merck, Darmstadt
Bromophenol blue	50206	Biomol, Hamburg
BSA	8076.2	Roth, Karlsruhe
Chloramphenicol	16785	Serva, Feinbiochemica, Heidelberg
Chloroform	3313.2	Roth, Karlsruhe
CTAB	16530	Serva, Feinbiochemica, Heidelberg
DMSO	4720.1	Roth, Karlsruhe
dNTPs	R0192	MBI-Fermentas
DTT	6908.1	Roth, Karlsruhe
DYEnamic ET Terminator Cycle Sequencing Kit	US81070	Amersham Biosciences, Freiburg
EDTA	8043.2	Roth, Karlsruhe
Ethanol p.a.	9065.2	Roth, Karlsruhe
Ethidium bromide	21252	Serva, Feinbiochemica, Heidelberg
Ficoll <sup>™</sup> 400	F-4375	Sigma Chemical Co., St. Louis, USA
Fugi imaging plates	IP-MA2340	raytest, Straubenhardt, Germany
Glycerin	3783.2	Roth, Karlsruhe
Herring sperm DNA	18580	Serva, Feinbiochemica, Heidelberg
IPTG	05684	Biomol Feinchem.GmbH, Hamburg
Isoamylalcohol	1.00979	Merck, Darmstadt
Isopropanol p.a.	9866.2	Roth, Karlsruhe



Chemicals	Cat. No.	Suppliers
Lithium chloride	1.05679	Merck, Darmstadt
Magnesium acetate	5819	Merck, Darmstadt
Magnesium chloride	5833	Merck, Darmstadt
Magnesium sulfate X 7 H <sub>2</sub> O	5886	Merck, Darmstadt
Montage PCR <sub>96</sub> Cleanup Kit	MANU 03010	Millipore Corporation, Bedford, MA 01730 U.S.A.
N, N-Dimethylformamide	1.009684	Merck, Darmstadt
NucleoFast 96 PCR clean-up plates	743100.10	MACHEREY-NAGEL GmbH & Co. KG, Germany
NucleoSpin® Extract Kit	740590.50	MACHEREY-NAGEL GmbH & Co. KG, Germany
Nylon membrane, Hybond-N+	RPN203B	Amersham Biosciences, Freiburg
Oligonucleotide	-	MWG-Biotech, Ebersberg
One-Phor-All buffer	27-0901	Pharmacia, Freiburg
pGEM®-vector	A3600	Promega, Madison, USA
Phenol	51292	Biomol, Hamburg
Plasmid Large Construct-Kit	12462	Qiagen, Hilden
Plasmid Maxi Kit	12262	Qiagen, Hilden
Plasmid Midi Kit	12243	Qiagen, Hilden
Potassium chloride	3957.1	Roth, Karlsruhe
Potassium acetate	6779.1	Roth, Karlsruhe
PVP	7443	Merck, Darmstadt
SDS, p.a.	20763	Sigma, St. Louis, USA
Seakem Agarose ( <i>low melt</i> )	50082	FMC Bioproducts, Rockland, ME, USA
Seakem LE Agarose	5005	FMC Bioproducts, Rockland, ME, USA
Sephadex G50	17-0042-01	Amersham Biosciences, Freiburg
Sodium chloride	3957.2	Roth, Karlsruhe
Sodium citrate X 2 H <sub>2</sub> O	3580.2	Roth, Karlsruhe
Sodium hydroxide	6771.2	Roth, Karlsruhe
TEMED	3787	Biomol, Hamburg
Tris	08003	Biomol, Hamburg
Triton X-100	6223	Roth, Karlsruhe
Tween 20	9127.1	Roth, Karlsruhe
Urea	A120.1	Roth, Karlsruhe
X-Gal	2315.4	Roth, Karlsruhe
Xylene cyanol FF	95600	Fluka Chemie, Buchs, Switzerland

**2.1.1. Enzymes****2.1.1.1. Enzymes**

Enzyme	Cat. No.	Supplier
<i>Taq</i> DNA-Polymerase	10342-020	Invitrogen Life technologies
T4-Ligase	E70005	Amersham Biosciences, Freiburg
RNAse IA	27-0323-01	USB, Cleveland, USA
Klenow Fragment	27-0929	Amersham Biosciences, Freiburg

**2.1.1.2. Restriction endonucleases**

Enzyme	Recognition site*	Cat. No.	Supplier
<i>AccI</i>	5'-GT↓ATAC-3'	E1001Y	Amersham Biosciences, Freiburg
<i>AluI</i>	5'-AG↓CT-3'	ER0011	MBI-Fermentas
<i>BclI</i>	5'-T↓GATCA-3'	ER0721	Amersham Biosciences, Freiburg
<i>BglII</i>	5'-A↓GATCT-3'	E1021Y	Amersham Biosciences, Freiburg
<i>BcuI</i>	5'-GG↓CC-3'	ER0151	MBI-Fermentas
<i>DpnI</i>	5'-GATC-3'	ER1701	MBI-Fermentas
<i>DraI</i>	5'-TTT↓AAA-3'	E1037Y	Amersham Biosciences, Freiburg
<i>EcoRI</i>	5'-G↓AATTC-3'	E1040Y	Amersham Biosciences, Freiburg
<i>EcoRV</i>	5'-GAT↓ATC-3'	E1042Y	Amersham Biosciences, Freiburg
<i>HaeIII</i>	5'-GGCC-3'	E1051Y	Amersham Biosciences, Freiburg
<i>HincII</i>	5'-GTYRAC-3'	ER0491	MBI-Fermentas
<i>HindIII</i>	5'-A↓AGCTT-3'	E1060Y	Amersham Biosciences, Freiburg
<i>HinfI</i>	5'-G↓ANTC-3'	ER0801	MBI-Fermentas
<i>HphI</i>	5'-GGTGA (N) <sub>8/7</sub> ↓-3'	ER1101	MBI-Fermentas
<i>MboI</i>	5'-↓GATC-3'	ER0811	MBI-Fermentas
<i>MboII</i>	5'-↓GAAGA-3'	ER0821	MBI-Fermentas
<i>MseI</i>	5'-T↓TAA-3'	R0525M	Biolabs, Schwalbach
<i>NdeI</i>	5'-CA↓TATG-3'	ER0581	MBI-Fermentas
<i>RsaI</i>	5'-G↓TAC-3'	ER0211	MBI-Fermentas
<i>PstI</i>	5'-CTGCA↓G-3'	E1073Y	Amersham Biosciences, Freiburg
<i>Sau3AI</i>	5'-GATC↓-3'	E1082Y	Amersham Biosciences, Freiburg
<i>TaqI</i>	5'-T↓CGA-3'	ER0671	Amersham Biosciences, Freiburg
<i>TruI</i>	5'-T↓TAA-3'	ER0982	MBI-Fermentas
<i>XhoI</i>	5'C↓TCGAG-3'	E1094Y	Amersham Biosciences, Freiburg

\* Restriction site is marked with an arrow

## 2.1.2. Primers

All primer sequences are available from Prof. C. Jung, Plant Breeding Institute, Christian-Albrechts University of Kiel, Olshausenstrasse 40, 24098 Kiel (c.jung@plantbreeding.uni-kiel.de).

### 2.1.2.1. Primers derived from the BAC ends used for markers

BAC	Primers	Length (bp)	BAC	Primers	Length (bp)
GJ01	GJ01SP6P1	21-mer	GJ21	GJ21SP6P1	21-mer
	GJ01SP6P2	20-mer		GJ21SP6P2	21-mer
	GJ01T7P1	20-mer		GJ21T7P1	22-mer
	GJ01T7P2	21-mer		GJ21T7P2	23-mer
GJ08	GJ08SP6P1	19-mer	GJ22	GJ22SP6P1	21-mer
	GJ08SP6P2	19-mer		GJ22SP6P2	22-mer
	GJ08T7P1	19-mer		GJ22T7P1	22-mer
	GJ08T7P2	19-mer		GJ22T7P2	22-mer
GJ10	GJ10SP6P1	21-mer	GJ23	GJ23SP6P1	20-mer
	GJ10SP6P2	21-mer		GJ23SP6P2	22-mer
	GJ10T7P1	21-mer		GJ23T7P1	20-mer
	GJ10T7P2	22-mer		GJ23T7P2	22-mer
GJ12	GJ12SP6P1	21-mer	GJ26	GJ26SP6P1	22-mer
	GJ12SP6P2	21-mer		GJ26SP6P2	20-mer
	GJ12T7P1	22-mer		GJ26T7P1	19-mer
	GJ12T7P2	23-mer		GJ26T7P2	19-mer
GJ13	GJ13SP6P1	21-mer	GJ29	GJ29SP6P1	21-mer
	GJ13SP6P2	21-mer		GJ29SP6P2	23-mer
	GJ13T7P1	21-mer		GJ29T7P1	21-mer
	GJ13T7P2	21-mer		GJ29T7P2	22-mer
GJ15	GJ15SP6P1	24-mer	GJ33	GJ33SP6P1	21-mer
	GJ15SP6P2	21-mer		GJ33SP6P1	21-mer
GJ17	GJ17SP6P1	22-mer		GJ33T7P1	21-mer
	GJ17SP6P2	20-mer		GJ33T7P2	21-mer
	GJ17T7P1	23-mer	GJ37	GJ37SP6P1	20-mer
	GJ17T7P2	22-mer		GJ37SP6P2	20-mer
GJ18	GJ18SP6P1	20-mer		GJ37T7P1	20-mer
	GJ18SP6P2	20-mer		GJ37T7P2	22-mer
	GJ18T7P1	21-mer	GJ44	GJ44SP6P1	23-mer
	GJ18T7P2	20-mer		GJ44SP6P2	20-mer

<b>BAC</b>	<b>Primers</b>	<b>Length (bp)</b>	<b>BAC</b>	<b>Primers</b>	<b>Length (bp)</b>
GJ44	GJ44T7P1 GJ44T7P2	24-mer 21-mer			
GJ49	GJ49SP6P1 GJ49SP6P2 GJ49T7P1 GJ49T7P2	22-mer 19-mer 22-mer 24-mer	GJ73	GJ73SP6P1 GJ73SP6P2 GJ73T7P2 GJ73T7P2	21-mer 21-mer 23-mer 19-mer
GJ55	GJ55SP6P1 GJ55SP6P2 GJ55T7P1 GJ55T7P2	20-mer 22-mer 22-mer 21-mer	GJ76	GJ76SP6P1 GJ76SP6P2 GJ76T7P1 GJ76T7P2	21-mer 21-mer 21-mer 23-mer
GJ65	GJ66SP6P1 GJ66SP6P2 GJ68T7P1 GJ65T7P2	20-mer 20-mer 24-mer 21-mer	GJ80	GJ80SP6P1 GJ80SP6P2 GJ80T7P1 GJ80T7P2	19-mer 21-mer 21-mer 22-mer
GJ67	GJ67SP6P1 GJ67SP6P2 GJ67T7P1 GJ67T7P2	21-mer 22-mer 21-mer 21-mer	GJ131	GJ131SP6P1 GJ131SP6P2 GJ131T7P1 GJ131T7P2	20-mer 20-mer 22-mer 19-mer
GJ70	GJ70SP6P1 GJ70SP6P2 GJ70T7P1 GJ70T7P2	20-mer 20-mer 22-mer 23-mer			

**2.1.2.2. Primers derived from the RFLP sequences and ORFs-derived primers**

<b>Primers</b>	<b>Length (bp)</b>	<b>Primers</b>	<b>Length (bp)</b>
pKP374: P1	22-mer	GJ01co015_F2P1	20-mer
P2	20-mer	GJ01co015_F2P2	20-mer
pKP490: P1	20-mer	GJ01co015_F3P1	20-mer
P2	21-mer	GJ01co015_F3P2	20-mer
pKP591: P1	19-mer	GJ01co036_F4P1	19-mer
P2	22-mer	GJ01co036_F4P2	21-mer
pKP826: P1	21-mer	GJ01co036_F4P3	20-mer
P2	21-mer	GJ01co038_F2P1	22-mer
P3	21-mer	GJ01co038_F2P2	20-mer
pKP730: P1	20-mer	GJ01co038_F2P3	20-mer
P2	20-mer	GJ70co003_F4P1	20-mer
pKP814: P1	20-mer	GJ70co003_F4P2	22-mer
P2	21-mer	GJ70co003_F4P3	20-mer
pKP884: P1	22-mer	GJ70co003_F6P1	20-mer
P2	23-mer	GJ70co003_F6P2	21-mer
pKP907: P1	20-mer	GJ70co003_F6P3	23-mer
P2	23-mer	GJ70co009_F1P1	21-mer
pKP943: P1	20-mer	GJ70co009_F1P2	21-mer
P2	20-mer	GJ70co009_F3P1	22-mer
pKP967: P1	20-mer	GJ70co009_F3P2	22-mer
P2	22-mer	GJ131co008_F2P1	19-mer
pKP1159:P1	21-mer	GJ131co008_F2P2	19-mer
P2	21-mer	GJ131co008_F2P3	21-mer
Koeln60: P1	20-mer	GJ131co008_F3P1	21-mer
P2	19-mer	GJ131co008_F3P2	23-mer
P3	21-mer	GJ131co008_F3P3	21-mer
P4	21-mer	GJ131co015_F2P1	23-mer
P5	23-mer	GJ131co015_F2P2	22-mer
P6	22-mer	GJ131co015_F6P1	20-mer
		GJ131co015_F6P2	20-mer

**2.1.2.3. Allele specific and ARMS primers**

Primers	Length (bp)	Primers	Length (bp)
GJ10SP6P3A	22-mer	GJ70SP6P3A	24-mer
GJ10SP6P3B	22-mer	GJ70SP6P3B	24-mer
GJ10SP6P4A	21-mer	GJ70SP6P4A	23-mer
GJ10SP6P4B	21-mer	GJ70SP6P4B	23-mer
GJ10SP6P5A	20-mer	GJ70SP6P5A	23-mer
GJ10SP6P5B	20-mer	GJ70SP6P5B	23-mer
GJ10SP6P6A	23-mer	GJ70SP6P6A	19-mer
GJ10SP6P6B	23-mer	GJ70SP6P6B	19-mer
GJ10SP6P7A	24-mer	GJ70SP6P8A	22-mer
GJ10SP6P7B	24-mer	GJ70SP6P8B	22-mer
GJ10SP6P8A	22-mer	GJ70T7P3A	21-mer
GJ10SP6P8B	22-mer	GJ70T7P3B	21-mer
GJ18T7P3B	23-mer	GJ01co036F4_P4B	21-mer
GJ18T7P4A	23-mer	GJ01co036F4_P5A	23-mer
GJ44SP6P3B	24-mer	GJ70co003F4_P4A	26-mer
GJ44SP6P4A	22-mer	GJ70co3F4_P5B	24-mer
GJ44SP6P5A	25-mer	GJ131co015F6_P3A	21-mer
GJ44SP6P5B	25-mer	GJ131co015F6_P4	21-mer
GJ44SP6P6A	24-mer	GJ131co015F6_P5A	21-mer
GJ44SP6P6B	24-mer	GJ131co015F6_P6B	22-mer
GJ70SP6P10A	24-mer	GJ131co015F6_P8B	23-mer
GJ70SP6P10B	24-mer		

**2.1.2.4. Flowering and vernalization genes-derived primers**

Primers	Length (bp)	Primers	Length (bp)
CO_Ara_P1	21-mer	Vrn1_Ara_P3	19-mer
CO_Ara_P2	20-mer	Vrn1_Ara_P4	21-mer
CO_Ara_P3	21-mer	Vrn2_Ara_P1	22-mer
CO_Ara_P4	21-mer	Vrn2_Ara_P2	22-mer
CO_Bv_P1A	21-mer	Vrn2_Ara_P4	20-mer
CO_Bv_P2A	22-mer	Vrn2_Bv_P1A	21-mer
FT_Ara_P1	23-mer	Vrn2_Bv_P1B	21-mer
FT_Ara_P2	22-mer	Vrn2_Bv_P2A	21-mer
Vrn1_Ara_P1	21-mer	Vrn2_Bv_P2B	21-mer
Vrn1_Ara_P2	21-mer		

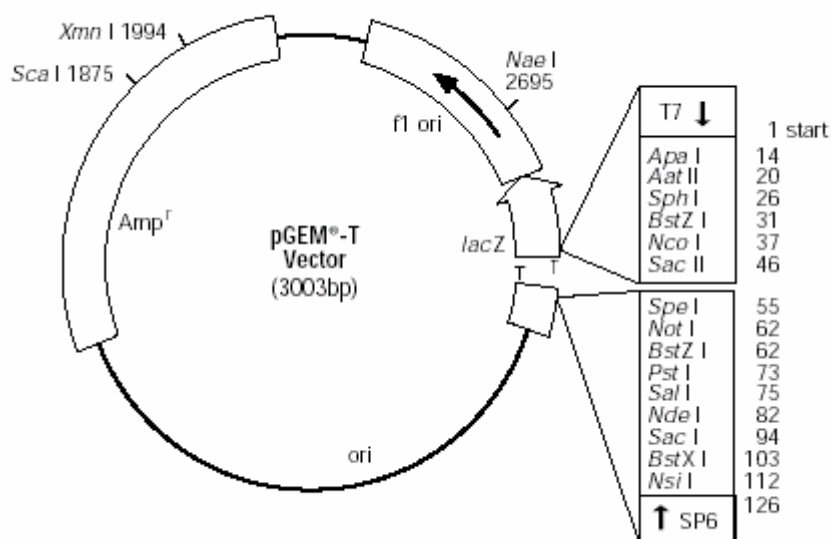
### 2.1.2.5. Primers used for inserts sequencing

Primers	Sequence	Length (bp)
M13_univer_for	5'-GTTGTAAAACGACGGCCAGT-3'	20-mer
M13_univer_rev	5'-CACAGGAAACAGCTATGACC-3'	20-mer
pCR-Blunt_for	5'-GCCGCCAGTGTGATGGATATCTGC-3'	24-mer
pCR-Blunt_rev	5'-CGAGCTCGGATCCACTAGTAACG-3'	23-mer

### 2.1.3. Organisms

#### 2.1.3.1. Bacterial strains and vectors

*E. coli* strain Top-10 (Invitrogen, Germany) was used for cloning of plasmids with PCR product inserts. The pGEM-T (Promega, Cat. No. A3600) was used for transformation of PCR fragments up to 2kb.

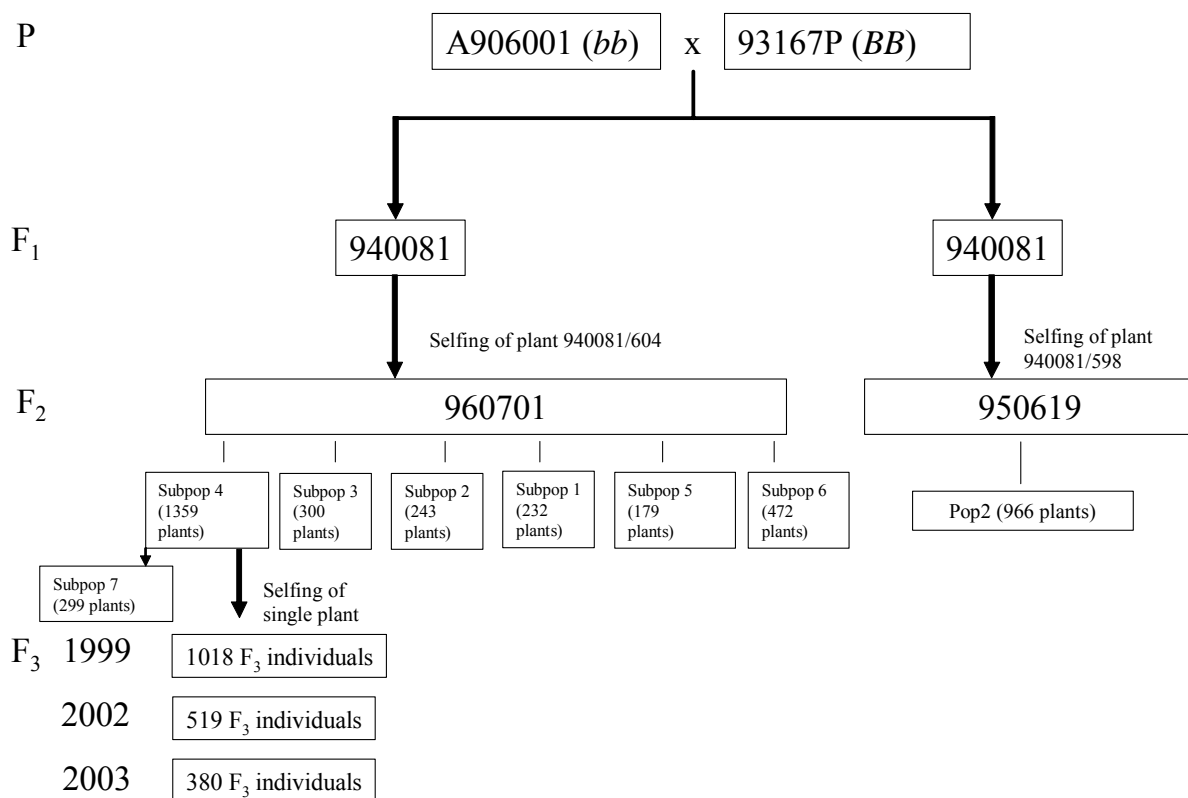


**Figure 2:** pGEM-T vector map and sequence reference points.

#### 2.1.3.2. Plant material

The plant materials used for genetic mapping were F<sub>2</sub> plants from a large population (El-Mezawy *et al.*, 2002). This population consisted of 2134 individuals, which included four subpopulations 1 (subpop1), 2 (subpop 2), 3 (subpop 3), and 4 (subpop 4). For genetic mapping of BAC-derived markers, 299 F<sub>2</sub> plants from subpopulation 4 were selected after the reconstruction of this subpopulation. For the verification of the bolting/non-bolting phenotype of F<sub>2</sub> plants from subpopulation 4 (1359 plants), F<sub>3</sub> populations were produced by selfing individual F<sub>2</sub> plants. This verification was done in years 2002 and 2003. After phenotyping, the subpopulation 4 was

reconstructed by selecting three classes: (a) non-bolting (*bb*) plants, (b) heterozygous (*Bb*)  $F_2$  plants with 3:1 segregation in  $F_3$ , and (c) bolting (*BB*) plants which had the same bolting phenotypes in  $F_2$  and  $F_3$ . After that, these three classes were grouped to form subpopulation 7 (subpop 7, Figure 3). In order to map, using only the non-bolting plants, the mapping population was increased. Two new subpopulations, subpop 5 (960701/95) and subpop 6 (960701/96) were produced from the same  $F_1$  plant, which was used to produce the four subpopulations (subpop 1, subpop 2, subpop 3, and subpop 4). In addition, a new population (Pop 2, 950619) from different  $F_1$  plant of the same crossing was produced (Figure 3). Only the 300 non-bolting  $F_2$  plants from subpopulations 5 and 6 and from population 2 were genotyped with the BAC markers.



**Figure 3:** Schematic representation of the pedigree of the mapping populations used in the study.

For marker genotyping of wild and cultivated beets, 24 wild beet (*B. vulgaris* L. ssp. *maritima*) accessions and 17 sugar beet hybrid cultivars were selected (Table 3). A single seed from each accession was grown in the greenhouse at 20°C during the day and 18°C at night with 16 h of daylight. The bolting behavior of wild beets was determined. Plants were bolting after 15 weeks considered being annuals (no vernalization requirement). The other plants, which did not bolt after 15 weeks, were considered biennial (vernalization is needed).



**Table 3:** Accessions of wild and cultivated species of the genus *Beta*.

No.	Accession number	Origin	No.	Accession number	Origin
<i>B. vulgaris</i> L. ssp. <i>maritima</i>			<i>B. vulgaris</i> L. ssp. <i>vulgaris</i>		
	Kiel	Original			
1	930034	-	25	940026	Rizor Germany
2	930043	-	26	940027	Steffi Germany
3	991971	960126	27	940028	Hilma Germany
4	960066	-	28	970102	Evasion Germany
5	991968	960116	29	970103	HM1370 Germany
6	950407	930037	30	970104	HM1371 Germany
7	991969	960122	31	980005	Meta Germany
8	930029	-	32	980008	Tellus Germany
9	930031	-	33	010238	Tatjana Germany
10	950409	930052	34	010240	Arivan Germany
11	960059	-	35	010243	Achat Germany
12	960078	-	36	010239	Coriana Germany
13	991972	960130	37	010242	Mars Germany
14	930028	-	38	010241	Ribella Germany
15	991957	960069	39	930179	FC902 USA
16	991947	980327	40	930182	FC708 USA
17	991948	980325	41	930184	FC607 USA
18	991954	980324			
19	991958	960013			
20	930038	-			
21	960055	-			
22	991943	980326			
23	930033	-			
24	960037	-			

## 2.2. Methods

### 2.2.1. DNA analysis

#### 2.2.1.1. Isolation of genomic DNA

Genomic DNA was extracted using the CTAB-method as described by Saghai-Marouf *et al.* (1984) with slight modifications. 2.5 to 5 g fresh leaf material or liquid nitrogen-frozen material were ground into fine powder under liquid nitrogen conditions. The fine powder was subsequently added to pre-warmed 50 ml Falcon tubes containing 21.5 ml 1× CTAB buffer and 260 µl β-mercaptoethanol. After short vortexing, the tubes were incubated at 65°C for 60-90 min. Subsequently the suspension was cooled down for 10 min. at room temperature (RT). After adding 0.6 vol. chloroform/isoamylalcohol (24:1), the suspension was gently mixed for 10 min. using an overhead shaker (Reax2, Heidolph). Centrifugation was performed at 7500 rpm (Beckmann J2-21 ME) for 20 min. at RT. The supernatant was transferred to a new tube. After 30 min. at 37°C RNase-IA (10 µl, 10 mg/ml) treatment, genomic DNA was precipitated with 13.5 ml cold isopropanol (0.6 vol.) and centrifuged at 7500 rpm for 10 min. at RT. The sediment was washed with 1 ml washing buffer I for 20 min., followed by centrifugation for 10 min. at 7500 rpm. A second washing step was performed with 0.5 ml washing buffer II for 10 min. followed by centrifugation at 7500 rpm. The sediment was dried for 20 min. at RT and then resuspended in 1× TE buffer (100-300 µl according to the pellet size). This method was

applied for DNA extraction from parent plants, because a large amount of DNA was required for Southern analysis.

Genomic DNA extraction from the new populations (subpop 5, subpop 6, and Pop2) was done following the same procedure but with some modifications. About 0.5 g fresh healthy leaves were freeze dried in 2 ml micro-centrifuge tubes for 3 days using a Zirbus-Lypholizer VaCo I. The plant samples were ground to get fine powder by using a swing mill (Retsch type MM2, Eydam). 1.25 ml 65°C pre-warmed 1× CTAB buffer was added to the powder, vortexed, and subsequently incubated in a shaking water bath at 65°C for 30 min. After cooling for 2 min., 500 µl chloroform/isoamylalcohol (24:1) were added and mixed by inverting on a rotating wheel (Reax2, Heidolph) for 10 min. After centrifugation at 14,000 rpm for 15 min. at 4°C (Beckman GS-15R-centrifuge), the DNA containing supernatant was transferred to 2 ml sterile micro-centrifuge tube and treated with 3 µl RNase-IA (10 mg/ml) and incubated for 30 min. at 37°C. DNA was precipitated with 700 µl of isopropanol (RT) and mixed gently. The precipitated DNA was transferred to a new sterile 2 ml micro-centrifuge tube and washed with 1 ml washing buffer I, followed by two washing steps with 0.5 ml washing buffer II. After that the tubes were centrifuged at 14,000 rpm for 5 min at RT. The dried DNA was resuspended in 100-300 µl 1× TE.

#### **1× CTAB (DNA extraction buffer)**

Tris/HCl (pH 7.5)	200 mM
EDTA (pH 8.0)	20 mM
NaCl	1.4 M
CTAB	1% (w/v)
β-Mercaptoethanol*	1% (v/v)

\*Add fresh directly before use

#### **Washing buffer I**

Na-acetate	200 mM
Ethanol	76% (v/v)

#### **Washing buffer II**

NH <sub>4</sub> acetate	10 mM
Ethanol	76% (v/v)

#### **1× TE buffer**

Tris/HCl (pH 8.0)	10 mM
EDTA	1.0 mM

### **2.2.1.2. Determination of DNA concentration**

DNA concentration was determined by using gel electrophoresis. For each sample, 10 µl (1 µl genomic DNA and 9 µl 1× bromophenol blue buffer) were loaded and separated on 0.75% agarose gel containing ethidium bromide (0.05%). Electrophoresis was done at constant voltage (36 V) for 30 min. (electrophoresis equipment type Mini-sub cell with power supply type Pac 3000, Bio-Rad). λ-DNA samples of 10, 50, and 100 ng were used as standard to determine DNA concentrations. DNA was visualized using UV light (302 nm) with Gel Documentation system (Bio-Rad). The DNA concentration was determined by visual comparison with standard λ-DNA. After that, the DNA concentration of each sample was adjusted with 1× TE buffer to 10 ng/µl.

#### **10× TAE buffer**

Tris/HCl	400 mM
Acetic acid	200 mM
EDTA	10 mM

pH 8.0, adjusted with concentrated HCl

#### **10× Bromophenol blue buffer**

TAE-buffer	1×
Glycerin	50% (v/v)
Bromophenol blue	0.1% (w/v)

### **0.75% Agarose gel**

Agarose gel	0.75% (w/v)
TAE buffer	1.0×
Ethidium bromide	0.05% (w/v)

## **2.2.2. Screening of BAC libraries**

### **2.2.2.1. Selected BACs from gene *B* region**

For screening of KWS2320 (*bb*)-derived BAC library (GJ library), probes generated from cloned YAC-ends and closely linked AFLP marker fragments were used to hybridize high-density BAC filters. In total, ten YAC-ends and four AFLP probes were used. Pooled probes (3-4 probes) were first used to hybridize the high-density BAC filters, then, the positive BAC clones were spotted on nylon membranes and rehybridized using individual probes. The BAC ends of the positively identified clones were sequenced using SP6 and T7 sequencing primers as described in 2.2.3.

In case of 000520 (*BB*)-derived BAC library (DS library), gene *B* flanked probes from KWS2320 (*bb*)-derived BAC library were used. Fifty different probes from gene *B* region were used. As described above, the pools of 3-4 probes were used to hybridize the high-density filters of this library. After that individual probes were used to identify the individual positive BAC clones. The positive candidate BAC clones were sequenced as described in 2.2.3. The screening of both libraries has been done during the GABI-BOLT project.

### **2.2.2.2. Shotgun libraries**

Three BACs (GJ01, GJ70 and GJ131) KWS2320 (*bb*)-derived BAC library were selected for construction of the shotgun libraries. These BACs were selected because they are from the central gene *B* region. Shotgun sequencing of the three BACs was done by the company AGOWA, Berlin. First, the BAC DNA was mechanically sheared, then the 1.5 to 2 kb fragments were shotgun cloned into pCR-Blunt vector (Invitrogen) and at last, from 400 plasmid clones/BAC, 96 plasmids from each shotgun libraries were sequenced using pCR-Blunt-forward primer. Sequence analysis and identification of potential ORFs (open reading frames) for marker development were done using the GeneQuest<sup>TM</sup> software (DNASTAR, GATC Biotech, Konstanz, Germany).

## **2.2.3. Sequencing**

Sequencing of all BAC ends from the positively identified BAC clones after chromosome walking was performed by a service company (DLMBC, Berlin) on an ABI 373 upgrade automated sequencer using the Big Dye terminator cycle sequencing chemistry (ABI). The BAC clones were grown overnight on solid LB-agar medium in 2 ml eppendorf tubes and then sent for sequencing. In addition, for SNP determination in some BAC ends and RFLP sequence-derived PCR products from the plants of the mapping population were done by DLMBC. The sequencing of the most of PCR products for marker development was done in Kiel on a MegaBACE 500 DNA Sequencing System using Dynamic ET terminator cycle sequencing chemistry (Amersham Biosciences, Freiburg, Germany). The two sequencing systems are based on the dideoxy termination method (Sanger *et al.*, 1977). PCR products were generated for the single/low copy BAC end-derived sequences. The PCR products were purified using the Montage PCR<sub>96</sub> Cleanup Kit (Millipore Corporation, Bedford, MA 01730 USA) or with NucleoFast 96 PCR clean-up plates (MACHEREY-NAGEL GmbH & Co. KG, Germany). The PCR products were directly transferred onto the membranes; the contaminants (primers and nucleotides) were removed by ultrafiltration (400-600 mbar and 10-15 min.). Then, the PCR products were washed

two times for 5-10 min. with bidist. H<sub>2</sub>O. The purified PCR products were recovered by adding 50 µl H<sub>2</sub>O or 10 mM Tris and gentle shaking (220 rpm) for 20 min. on the Heidolph shaker (Heidolph Instruments, UniMax1010). The concentration of the purified PCR products was determined (2.2.1.2.). Aliquots of the PCR products were used for setting up sequencing reactions (2.2.3.1.) or labeling reactions (2.2.6.2.).

### **2.2.3.1. Preparation of sequencing reactions**

For each sequencing reaction, 20-50 ng purified PCR product were used. One µl of 5 pmol/µl sequencing primer was added to 5 µl PCR product (20-50 ng) followed by the addition of 4 µl sequencing reagent premix (1:7 diluted premix) and thus the total sequencing reaction volume was 10 µl. After that the probes were thoroughly mixed, shortly centrifuged (30 sec), and kept on ice in the dark until cycle sequencing.

### **2.2.3.2. Cycling conditions**

The following sequencing PCR program was used; the cycling program was repeated for 25 cycles.

95°C	20 sec
50-54°C	15 sec
60°C	60 sec

The sequencing program was performed using a Perkin Elmer thermocycler (Gene Amp PCR system 9600, Perkin Elmer, Applied Biosystem, Weiterstadt).

### **2.2.3.3. Purification of sequencing reactions**

After the sequencing cycling program was completed, 1 µl of ammonium acetate (7.5 M) was added to each tube. Then, 30 µl of 100% ethanol were added to each reaction and mixed by vortexing. The tubes were centrifuged at room temperature (RT) for 60 min. at 4000 rpm using Multifuge3<sub>S-R</sub> centrifuge (Heraeus). The supernatant was removed by inverting the plate on a paper towel for 1 min. The DNA pellets were washed with 100 µl 70% ethanol, and after that the plate was centrifuged at 4000 rpm at room temperature (RT) for 20 min. The supernatant was removed and the DNA pellets were dried for 30 min. at room temperature (RT), then the pellets were resuspended in 10 µl of 0.1% Tween 20 and vortexed vigorously for 10-20 sec. The samples were sequenced directly or stored at -20°C in the dark, after being transferred into Hard-Shell<sup>TM</sup> Thin-Wall microplates (Biozym, Oldendorf) sealed with rubber PCR sealing mat (SEALMAT FORMAT 96, Biozym, Oldendorf).

### **2.2.3.4. Samples run on MegaBace**

The purified sequencing samples were briefly centrifuged to collect the samples at the bottom of the tubes and to remove air bubbles. The samples were injected at 2KV for 40-50 sec injection time. After sample injection, the running buffer plate was placed. The run of samples was carried out at 8 KV for 120 min. Sequence data were analyzed and edited with the Laser Gene software package (DNASTAR; GATC Biotech, Konstanz, Germany).

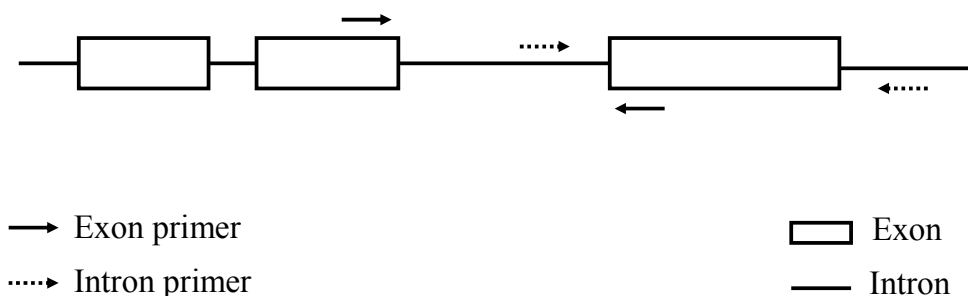
## **2.2.4. Primer design**

After editing of BAC-end sequences using SeqMan<sup>TM</sup>II from the DNASTAR software package (GATC Biotech, Konstanz, Germany), the primers were designed with the PrimerSelect<sup>TM</sup> software (DNASTAR, GATC Biotech, Konstanz, Germany). Generally, the primers have been designed in a way that, they should not form primer dimers and have no internal complementarity. Primers should have 50-60% GC content and length of 19-23 bp. Ideally, one or two bases on the 3'-end are either a G or a C that results in more stable primer binding and ensures recognition by the polymerase. The

primers for the BAC-end sequences were designed to have 60°C-annealing temperature. For allele specific primers, the primers were designed manually. The 3'-end of the primers terminates on the SNP position. For ARMS primers, a second mismatch was introduced at the second last position of the 3'-end. The rules for selecting a nucleotide for additional mismatch in the ARMS-PCR have been previously described by Little (1997). The allele specific primers were checked with the PrimerSelect™ for the primer dimer formation and internal complementarity.

From shotgun sequencing of three BACs (GJ01, GJ70 and GJ131), larger contigs with potential open reading frames (ORFs) were identified using the software GeneQuest™ (DNASTAR; GATC Biotech, Konstanz, Germany). Contig information was kindly provided by Dr. Uwe Hohmann (Kiel). So-called ORF primers were designed to amplify large Exon or Intron sequence regions. First the ORFs have been identified using the GeneQuest™ program (DNASTAR; GATC Biotech, Konstanz, Germany), then exon primers were designed, which were placed approx. 80 bp within exon regions and used to amplify the intron sequences. Intron primers were also placed approx. 80 bp within intron regions and amplified the exon sequences (Figure 4). The ORF primers usually amplified 600-800 bp amplicons.

For amplification of identified gene sequences of *Arabidopsis thaliana* *CONSTANS* (*CO*), *Vernalization 1* (*VRN1*) and 2 (*VRN2*) from sugar beet, heterologous primers complementary to the gene sequences were designed. Since the gene regions are supposed to be conserved, normal primers were used. The primer pairs have been placed on different positions in order to amplify the coding sequences (Exons) or domains (motifs).



**Figure 4:** Diagrammatic presentation showing the design of ORF primers.

### 2.2.5. Polymerase chain reaction (PCR)

The protocol for polymerase chain reaction was described by Mullis and Falona (1987). Thermo-stable DNA polymerase (Taq polymerase) was used to catalyze the extension of the annealed oligonucleotide primers. Briefly, the template DNA is first denatured by heating in the presence of large molar excess of each the two oligonucleotide primers and 4 dNTPs. The reaction mixture is then cooled to a melting temperature ( $T_m$ ) that allows the oligonucleotide primers to anneal specifically to their target sequences. The amplicon is synthesized by 3' extension of the annealed primers. The cycle of denaturation, annealing, and DNA synthesis is then repeated, thus the DNA is amplified. PCR reaction was performed in 25  $\mu$ l containing 10 ng template DNA, 10 pmol of forward and reverse primers, 1 unit of Taq-polymerase, 0.2 mM dNTPs, and 2 mM  $MgCl_2$  in 1 $\times$  PCR buffer (pH 8.3). The PCR was carried out in PCR Express GRADIENT thermocycler (Hybaid, Ashford, U.K.) for 30 cycles with 1 min. denaturation step at 94°C, 1 min. annealing at  $T_m-3^\circ C$  and 1 min. at 72°C.

Amplification mixture		PCR conditions	
10× PCR-buffer	2.50 $\mu$ l	94°C, 3 min.	1×
2 mM dNTPs	2.50 $\mu$ l	94°C, 1 min.	} 30×
25 mM MgCl <sub>2</sub>	2.00 $\mu$ l	T <sub>m</sub> -3°C, 1 min.	
Sterile water	14.85 $\mu$ l	72°C, 1 min.	
10 pmol primer-forward	1.00 $\mu$ l	72°C, 3 min.	1×
10 pmol primer-reverse	1.00 $\mu$ l		
Taq polymerase (1U/ $\mu$ l)	0.15 $\mu$ l		
Template DNA (10 ng)	1.00 $\mu$ l		
Total	25 $\mu$ l		

For gradient PCR, the conditions were as previously described except for a range of annealing temperatures (50°C-67°C) tested to determine the optimal temperature for specific binding to the target DNA.

## 2.2.6. Southern analysis

### 2.2.6.1. Southern transfer

The standard alkaline method described by Southern (1975) was used. 5  $\mu$ g of genomic DNA from bolting (930190) and non-bolting (A906001) parents of the mapping population were digested with two different enzymes, *EcoRI* and *HindIII*. The digested DNA was separated by electrophoresis on 0.8% agarose gel with 2 V/cm for 16-18 h. The agarose gel was photographed, exposed to UV light (302 nm) for 2 min., and then transferred with the transfer buffer onto Hybond<sup>®</sup> Nylon+ membrane (Amersham Biosciences, Freiburg, Germany) by capillary diffusion blotting for overnight, using 0.5 M NaOH/1.5 M NaCl as blotting solution. The DNA was fixed to the filter at 80°C for 20 min.

For preparing of BAC filters with restricted DNA from candidate BACs, the DNA was extracted using the Large Construct kit (Qiagen, Hilden, Germany). 1  $\mu$ g of BAC DNA was digested with *EcoRI* and *HindIII*. The digested DNA was separated on a 0.8% agarose gel for 20 h at 25 V. The membranes were prepared as described above.

For KWS2320 derived-BAC library, high-density colony filters were prepared by spotting the bacteria in duplicate (4 × 4 double offset spotting pattern) on one Biodyne B transfer membrane (Pall, Portsmouth, U.K.) using a BioGrid spotting roboter (Biorobotics, Cambridge, U.K.). Each filter (22 × 22 cm) contained 18,432 clones, which were grown overnight on LB/X-Gal/IPTG plates. The filters were processed according to Sambrook *et al.* (1989). For the 000520 derived-BAC library, the spotting pattern was 3 × 3 and Q-Pix was used (Genetix, Christchurch, UK). The high-density BAC filters were provided by Dr. Uwe Hohmann (Kiel).

### 2.2.6.2. Probe labeling

PCR products, originating from the corresponding BAC-ends or inserts from the shotgun library, were used for labeling. PCR products were purified with Montage PCR<sub>96</sub> Cleanup Kit (Millipore Corporation, Bedford, MA 01730 USA) and stored at -20°C till use. DNA probes were radioactively labeled with  $\alpha$ - [<sup>32</sup>P]-dATP and  $\alpha$ - [<sup>32</sup>P]-dCTP according to the random priming method (Feinberg and Vogelstein, 1983). 40-60 ng DNA were used to generate the probe. The DNA was denatured at 100°C in a boiling water bath for 10 min. and immediately placed on ice. The labeling mixture was added and the solution was mixed with a pipette. The labeling reaction was performed for 45 min. at 37°C in a water bath. The non-incorporated radioactive nucleotides were trapped by filtration through a Sephadex G-50 column.

**Labeling mixture**

10× Klenow buffer	10.0 µl
Pd(N) <sub>6</sub> (50U/ml)	10.0 µl
GT-Mix (0.5 mM)	4.0 µl
α[ <sup>32</sup> P]-dATP	1.0 µl
α[ <sup>32</sup> P]-dCTP	1.0 µl
Klenow (5U/µl)	2.0 µl
Probe DNA (40-60 ng)	<u>70.0 µl</u>
H <sub>2</sub> O added up to	100.0 µl

**2.2.6.3. Hybridization and signal detection**

Hybridization of blots, washing steps, and signal detection were performed as described by Sambrook and Russel (2001) with small modifications. Prehybridization buffer was warmed at 60°C in a shaking water bath. Sonified Herring sperm DNA solution (100 µg/µl; 2 ml for 100 ml prehybridization buffer) was denatured at 100°C for 10 min. and immediately added to the prehybridization buffer at a final concentration of 20 mg/100 ml per prehybridization. The membranes were placed in a plastic box and incubated in the buffer for 2-3 h at 60°C with slow motion shaking in a water bath. The prehybridized filters were then transferred to a 250 ml hybridization tube containing 50 µl prehybridization buffer/cm<sup>2</sup> filter. The tube was placed at 60°C in the hybridization oven for 30 min. The double-stranded labeled DNA was denatured in a water bath at 100°C for 10 min. Subsequently, the labeled probe was added to the hybridization tube and the hybridization was performed by placing the tube on a rotating wheel in a hybridization oven at 60°C for overnight.

The membranes were washed twice with washing buffers I and II (washing buffer I: 2× SSC/1% SDS and washing buffer II: 0.5× SSC/1% SDS) at 60°C in a shaking water bath for 10 min. The filters were wrapped with plastic film, placed on phosphoimaging plate, and then placed into an exposing cassette. After exposure of the filters for 48 to 72 h to Fuji imaging plates (Fuji film BAS-IP-MA2340) (23 × 25, 23 × 40 or 35 × 43 cm<sup>2</sup>), the image was read with Fuji Bio Imaging Analyzer FLA 5000 v 1.0 software at the 200 µm pixel. In addition, further analysis and modification of image file were performed with Aida Image Analyzer v. 3.0 software.

**2.2.6.4. Filter stripping**

The hybridized filters were stripped and rehybridized subsequently to other probes. Filters were treated twice with 200 ml of stripping buffer (0.05× SSC/0.001M EDTA pH= 8.0), containing 0.1% SDS, for 20 min. at 60°C, then the filters were rinsed with current water. Finally, the filters were air dried and stored between Whatman papers.

**20× SSPE**

NaH <sub>2</sub> PO <sub>4</sub>	200 mM
NaCl	3 M
EDTA	20 mM

**100× Denhardt's solution**

PVP	2% (w/v)
BSA	2% (w/v)
Ficoll <sup>TM</sup> 400	2% (w/v)

**Washing solution**

SSC	0.5 ×
SDS	0.2 %

**GT-Mix**

dGTP	0.5 mM
dTTP	0.5 mM

**Prehybridization solution**

Denhardt's solution	5 ×
SSPE	5 ×
SDS	0.2 %
hering sperm-DNA	0.02 % (w/v)

**Transfer buffer**

NaOH	0.5 M
NaCl	1.5 M

**10× Klenow buffer**

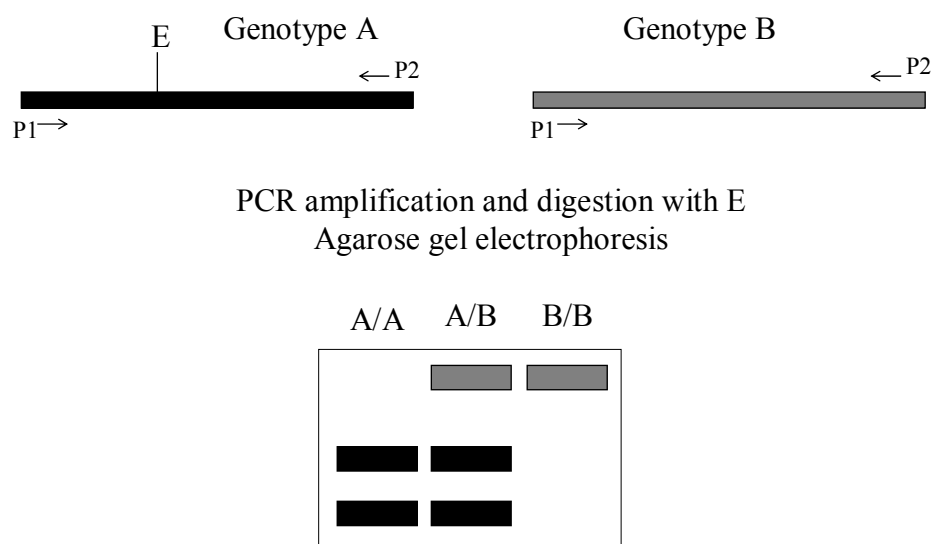
Tris/HCl pH 7.2	0.5 M
MgCl <sub>2</sub>	0.1 M
DTT	1.0 mM
BSA	0.2 %

**20× SSC**

Na <sub>3</sub> -citrate	0.3 M
NaCl	3.0 M

**2.2.7. Molecular marker analysis****2.2.7.1. PASA/CAPS markers**

In the CAPS technique, gene-specific primers are used to amplify template DNA (Figure 5), and polymorphic nucleotides are detected by the loss or gain of recognition sites for a restriction endonuclease (Konieczny and Ausubel, 1993).



**Figure 5:** Principle of CAPS markers. The presence of a recognition site (E) for a restriction endonuclease allows the cleavage of the amplified fragment at one site in genotype A and not in genotype B.

For developing of PASA/CAPS markers from the BAC sequences, two BAC-end specific PCR primers were used to amplify products from bolting and non-bolting genotypes that are parents of the mapping populations. If the PCR product was successfully obtained only from one of the genotypes, this is considered to be a PASA (PCR amplification of specific allele) marker. This means that, PASA is a dominant marker; the target sequence could be amplified by PCR from one genotype and not from the other. In case that, two specific PCR products of the same size were amplified from bolting and non-bolting genotypes, the PCR products were digested with suitable restriction enzymes. The restriction enzymes were selected after restriction site analysis of BAC-end sequences, which was done using MapDraw<sup>TM</sup> program (DNASTAR; GATC Biotech, Konstanz, Germany). Amplification was carried out in a PCR Express GRADIENT thermocycler (Hyaid, Ashford, U.K.). PCR conditions were as in 2.2.5.



### Restriction enzyme digestion and analysis of amplified products

Volumes of 2.5 µl from each PCR product were digested with 1U of appropriate enzyme. The total reaction volume was 10 µl. The restriction digestion was carried out at the suitable temperature and incubated for 2-3 hours. The CAPS markers were co-dominantly scored after separation on 2% agarose gels. Electrophoresis was done at constant voltage (120 V) for 80 min. (Maxi Gel System M, peQLab Biotechnology Ltd.).

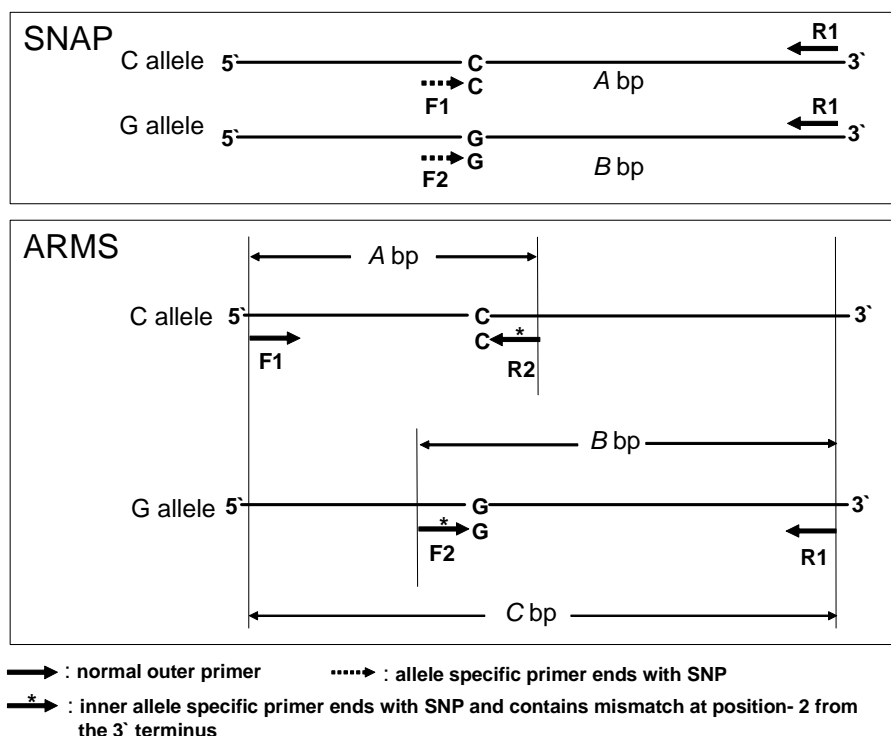
#### **2.2.7.2. Allele specific markers**

PCR products from bolting and non-bolting plants were amplified using specific primers. The PCR products were purified with a Millipore Kit. The purified PCR products were sequenced (2.2.3.). SNPs (single nucleotide polymorphisms) were detected after sequence comparison and alignment of the sequences using MegAlign™ program (DNASTAR; GATC Biotech, Konstanz, Germany). The original BAC sequences were used as reference in the alignment of the sequences obtained from the PCR products of the two genotypes. The presence of SNPs depends on two types of base pair substitutions: 1) Transition (C↔T or G↔A) and/or 2) Transversion (C↔G, T↔A, C↔A or T↔G).

SNPs can be genotyped using allele-specific PCR primers. The 3'-end of the primer corresponds to the site of the SNP (Ugozzoli and Wallace, 1991). The allele specific primer preferentially amplifies one of the specific alleles (Cha *et al.*, 1992). Detection of SNPs by allele specific PCR shares many of the advantages of CAPS and dCAPS, including co-dominance (as long as an allele specific primer is used for each allele) and the convenience of detecting the polymorphism on standard agarose gels.

Moreover, allele-specific PCR procedures have the additional benefit that endonuclease restriction of the amplicon is not necessary after PCR amplification. For SNP detection, a modification of tetra-primer ARMS-PCR method was used in this study (Ye *et al.*, 2001).

In a single PCR reaction two primer pairs were used to amplify two allele specific fragments (Figure 6). Two allele specific primers (inner primers) were designed in such a way that, they encompass a mismatch at position -2 from the 3'-terminus (see 2.2.4. primer design).



**Figure 6:** Schematic representation of the SNAP and tetra-primer ARMS-PCR methods. The SNP is shown as G→C substitution. SNAP: allele specific forward primer (F1 or F2) ends with C or G. With normal reverse primer (R1) amplified PCR products for C allele or G allele are obtained in two different PCR reactions. ARMS: two allele-specific amplicons are generated using two pairs of primers, one pair (F1R2) producing an amplicon specific for the C allele and the other pair (F2R1) producing an amplicon representing the G allele. To enhance allelic specificity a mismatch at position-2 from 3'-terminus was incorporated in the inner primers. By positioning the two outer primers at different distances to the polymorphic nucleotide, two allele-specific amplicons are created differing in length and allowing them to be discriminated by gel electrophoresis.

## 2.2.8. Linkage analysis

### 2.2.8.1. Statistical analysis

The Chi-square ( $\chi^2$ ) analysis was used to test the hypothesis of 1:2:1 or 3:1 expected segregation ratio. The  $\chi^2$  was calculated using the following equation:

$$\chi^2 = \sum_{i=1}^a (O_i - E_i)^2 / E_i$$

where  $O_i$  and  $E_i$  are the observed and expected counts for class  $i$ , respectively, with the degree of freedom ( $a-1$ ) where  $a$  is the number of classes. The rejection area for significance level  $\alpha$  is  $\chi^2 \geq \chi^2_{a-1}$ , where  $\chi^2_{a-1}$  is the critical value for significance level  $\alpha$  from a chi-square distribution with degree of freedom ( $a-1$ ).

### 2.2.8.2. Mapping procedure

Linkage analysis of the markers was performed using MAPMAKER/EXP 3.0b (Lander *et al.*, 1987; Lincoln *et al.*, 1992). The calculation based on a minimum LOD score of 10 and a maximum

recombination value (REC-value) of 0.30. DRAWMAP v2.1 (Voorrips, 2002) was used to plot the genetic linkage map. The map distances were calculated using Kosambi's mapping function (Kosambi, 1944):

$$X = 0.25 \ln \left( \frac{1 + 2\theta}{1 - 2\theta} \right) \quad \text{for } 0 \leq \theta < 0.50$$

where  $X$  is the Kosambi distance and  $\theta$  is the recombination value.

Linkage analysis using the MAPMAKER/EXP 3.0b was done with 299 plants from  $F_2$  plants of subpopulation 7 (2.1.3.2.) with the verified phenotype in the  $F_3$  population.

To facilitate the linkage analysis, 300 non-bolting  $F_2$  recessive plants from the  $F_2$  populations (subpopulation 5 and 6, 960701 and population 2, 960619) were used. Where, according to Hühn (1995), two-point linkage analysis could be done based only on the recessive class of the mapping population using the following equation:  $\hat{R} = \frac{2Z_1 + Z_2}{2(Z_1 + Z_2 + Z_3)}$

Where  $\hat{R}$  is the maximum likelihood estimate of the recombination fraction  $R$  and  $Z_1, Z_2, Z_3$  are the observed absolute frequencies of different classes of recessive non-bolting individuals. Significance tests for testing linkage between markers and  $B$  locus were done with the LOD SCORE test (Hühn, 1995). The variances were calculated using the following formula:  $V(\hat{R}) = \frac{R(1-R)}{2N}$  where  $V(\hat{R})$  refers to the variance of  $\hat{R}$  which depends on  $R$ = the recombination fraction and  $N$ = total number of non-bolting  $F_2$  plants analyzed (Hühn, 1995).

For two-point analysis, the recombination value between each two markers can be calculated using recombination between each marker and the  $B$  locus with the following equation:

$$\hat{R}_{M1M2} = \frac{\hat{R}_{M1B} + \hat{R}_{M2B}}{1 + 4(\hat{R}_{M1B} * \hat{R}_{M2B})} \quad \text{where, } \hat{R}_{M1M2} \text{ is the recombination value between marker 1 (M1)$$

and marker 2 ( $M_2$ ),  $\hat{R}_{M1B}$  and  $\hat{R}_{M2B}$  are the recombination values between the marker 1, marker 2 and  $B$  locus, respectively. Finally, three-point linkage analysis was done in order to find the most likely order of markers in relation to the  $B$  locus (Liu, 1998). This is an "addition formula" for the Kosambi mapping function for a given order  $M1-B-M2$ . For the order  $M1-M2-B$  one attains another formula.

## 2.2.9. Isolation and mapping of sugar beet flowering and vernalization orthologous genes

### 2.2.9.1. PCR fragments excision and purification

The polymorphic PCR bands were positioned, cut from the gel (2% low melting agarose) and placed in 1.5 ml Eppendorf tubes. The DNA extraction from gel was done using the NucleoSpin® Extract Kit. Gel lysis was done using 300  $\mu$ l of lysis buffer (NT1) and incubation at 50°C for 5-10 min. until the gel slices were dissolved. The samples were then centrifuged for 1 min. at 8,000 rpm after loading into the NucleoSpin® Extract columns. The flow-through was discarded and the silica membranes were washed with 600  $\mu$ l of the washing buffer (NT3). The columns were centrifuged for 1 min at 11,000 rpm. A second washing step was done for 2 min at 11,000 rpm. The columns were then dried. DNA was eluted using 25-50  $\mu$ l of elution buffer (NE) and centrifuged for 1 min. at 11,000 rpm after incubation at room temperature for 1 min.

### 2.2.9.2. Ligation and plasmid cloning

The pGEM-T vector of Promega (Madison, USA) was used for cloning PCR-products. pGEM-T vector contain T7 and SP6 RNA polymerase promoters flanking a multiple cloning region within the  $\alpha$ -peptide cloning region of  $\beta$ -galactosidase enzyme. Insertional inactivation of the peptide allows recombinant clones to be identified by color screening on indicator plates.

The PCR products were ligated to the pGEM-T vector. Calculation of the amount of DNA needed for ligation was determined according to pGEM-T and pGEM-T Easy Vector System protocol (Promega, Cat. No. A3600):

$$\frac{\text{ng of vector} \times \text{kb size of insert}}{\text{kb size of vector}} \times \text{insert : vector molar ratio} = \text{ng of insert}$$

Where the amount of vector was 50 ng, its 3 kb and the insert to vector ratio is 3:1. The fragments to be cloned ranged from 200 to 1700 bp. The ligation was performed with 3 U of T4 ligase in a total volume of 50  $\mu$ l at 16°C overnight.

### 2.2.9.3. Heat shock transformation of *E. coli*

*E. coli* cells were transformed using the heat shock method. A 100  $\mu$ l cell suspension was mixed with 100 ng of DNA in a chilled tube and incubated on ice for 30 min. The mixture was heated at 42°C for 60-90 sec and then placed on ice immediately. After heat shock, 0.8 ml SOC medium was added to the cells and incubated at 37°C, with shaking, for 1 h. The white colonies were selected after plating on LB/ampicillin/IPTG/X-Gal plates overnight at 37°C.

### 2.2.9.4. Screening colonies of transformed *E. coli*

White colonies (containing insert) were picked and resuspended in 10  $\mu$ l of sterile H<sub>2</sub>O. Inserts were amplified using plasmid primers (M13F and M13R) and 1  $\mu$ l of the bacterial suspension in a 25  $\mu$ l PCR reaction. The rest of bacterial suspension was plated on LB/ampicillin indicator plates and grown for overnight at 37°C. The PCR products were amplified with M13 forward and reverse primers using white bacterial suspensions as template. The amplicons were purified and sequenced as described in 2.2.3.

### 2.2.9.5. Sequence analysis

The sequences from the cloned PCR fragments were edited and analyzed using DNASTAR software (GATC Biotech, Konstanz, Germany).

### 2.2.10. Sequence homology analysis

BAC-end sequences were analyzed for their copy number and for homology to ESTs available from BIOMAX server of Kleinwanzlebener SAAT AG, Einbeck. The edited cloned sequences from flowering and vernalization genes were used to screen the databases. First the database screening was performed with the BLAST 2.2.2 and then WU-BLAST 2.0 software package. The following databases were screened:

<http://www.arabidopsis.org/blast/>  
<http://www.ncbi.nlm.nih.gov/blast/>

Also, the GABI sugar beet EST database, Sputnik, dbv1.0.3, GABI (<http://www.gabi.de>) and the sugar beet database (<http://www.genomics.msu.edu/sugarbeet/>) were screened.

### 3. Results

#### 3.1. Mapping populations

##### 3.1.1. Verification of phenotypic scoring of F<sub>2</sub> population

In this study, the plant materials used for mapping were derived from two different populations. The first population 960701 consists of six subpopulations (subpopulations 1, 2, 3, 4, 5, and 6). The second population is 950619. The production of plant materials is described in 2.1.3.2. The population 950619 and subpopulations 5 and 6 were produced in this study, while the subpopulations 1-4 (960701) were produced by El-Mezawy *et al.* (2002). The segregating F<sub>2</sub> subpopulations 1-4 (960701) were used by El-Mezawy *et al.* (2002) for fine mapping of the bolting gene (*B*) of sugar beet. From the selfed F<sub>2</sub> plants in the years 1999/2000, 1018 F<sub>3</sub> progenies (924 bolting and 94 non-bolting F<sub>3</sub> full sister families) were phenotypically scored regarding to bolting behavior. However, sample sizes of the F<sub>3</sub> families were relatively small.

In this study, the same mapping subpopulation 4 (960701) was used for genetic fine mapping of the gene *B* using BAC-derived markers. For fine mapping of the bolting gene, phenotypes of the F<sub>2</sub> plants used for mapping must be determined, unambiguously. Therefore, bolting and non-bolting F<sub>2</sub> plants (960701, subpopulation 4) were selfed. In the years 2002 and 2003, F<sub>3</sub> progenies were phenotyped again after growing of 519 and 380 families, respectively (see Figure 3). In addition, to improve the reliability of the F<sub>3</sub> phenotypic data, the numbers of F<sub>3</sub> plants were gradually increased from 4 up to 32 plants per F<sub>3</sub> family. For the F<sub>3</sub> progeny analysis of 916 F<sub>2</sub> plants (subpopulation 4, 960701), at least four F<sub>2</sub> plants were grown for each F<sub>3</sub> family and for 811 F<sub>2</sub> plants, at least 8 F<sub>3</sub> plants were grown (Table 4). The classification into bolting and non-bolting plants was determined 105 days after sowing (Table 6).

**Table 4:** Summary of the F<sub>3</sub>-progeny analysis (full sister) for determination of bolting phenotype. The F<sub>3</sub>-families were produced by selfing of single F<sub>2</sub> plants of population 960701/subpopulation 4. Only when the number of F<sub>3</sub>-families was >7 plants, the F<sub>3</sub>-phenotypic data were considered to be safe.

No. of examined F <sub>3</sub> individuals/F <sub>3</sub> -family	No. of F <sub>3</sub> -families	% F <sub>3</sub> -families
0-3	443	32.6
4-7	105	7.7
>7	811	59.7
<b>Total</b>	<b>1359</b>	<b>100</b>

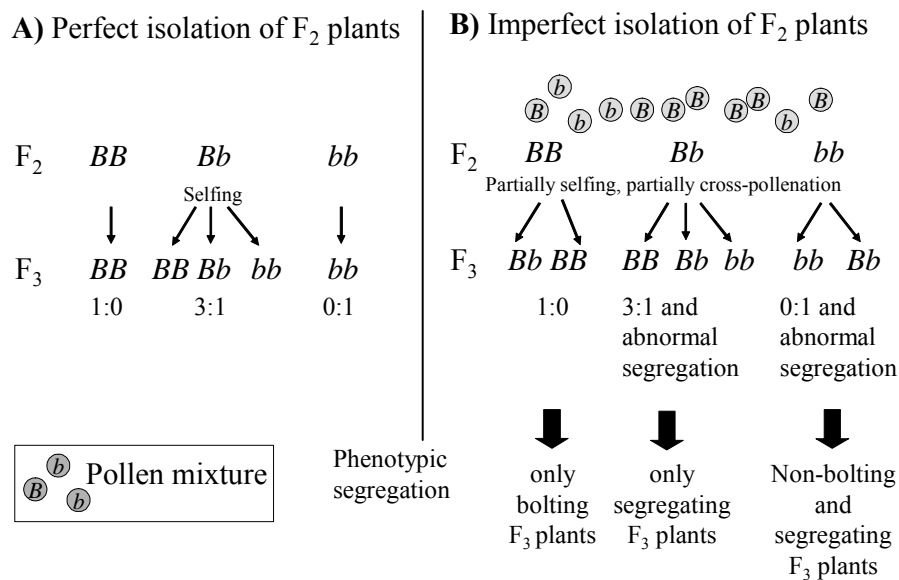
The  $\chi^2$  analysis was performed to test the expected 3:1 segregation ratio in the subpopulation 4 (960701) F<sub>3</sub> families. Significant  $\chi^2$  values have been observed (Table 6). Because of this deviation from expected segregation and for more reliable phenotypic data of subpopulation 4 (960701), the F<sub>3</sub> full sister progenies (produced by selfing of F<sub>2</sub> plants) with n>7 plants were examined (Table 5). In addition, the corresponding F<sub>2</sub> plants (with deviated segregation ratios) were not included in the further linkage analysis. Table 5 contains the different groups of F<sub>2</sub> plants, which were used for the linkage analysis. However, on the basis of segregation of F<sub>3</sub> families of 746 F<sub>2</sub> plants (n>7 plants), the  $\chi^2$  value (277.1) was highly significant.

**Table 5:** Results of F<sub>3</sub>-tests from progenies of the isolated F<sub>2</sub> plants of subpopulation 4 (960701). Only F<sub>3</sub> families with >7 plants were used for this analysis. The  $\chi^2$  values were compared with table values (df= 2,  $\alpha$ = 0.05). Only the populations, which did not deviate significantly from the expected  $\chi^2$  value, are specified in the table.

Only non-bolting plants	Bolting: non-bolting plants were not significant with $\chi^2$ test	Only bolting plants	$\chi^2$
F <sub>2</sub> parent: <i>bb</i>	F <sub>2</sub> parent: <i>Bb</i>	F <sub>2</sub> parent: <i>BB</i>	
73	297	376	277.1*

$\alpha_{0.05} = 5.99$ , df= 2. \*:  $\chi^2$  is significant at P= 0.05.

In fact, the extreme deviation from the expected 3:1 segregation ratio that has been found in F<sub>3</sub> populations might be explained by inappropriate isolation during selfing of individuals from the F<sub>2</sub> populations (Figure 7).



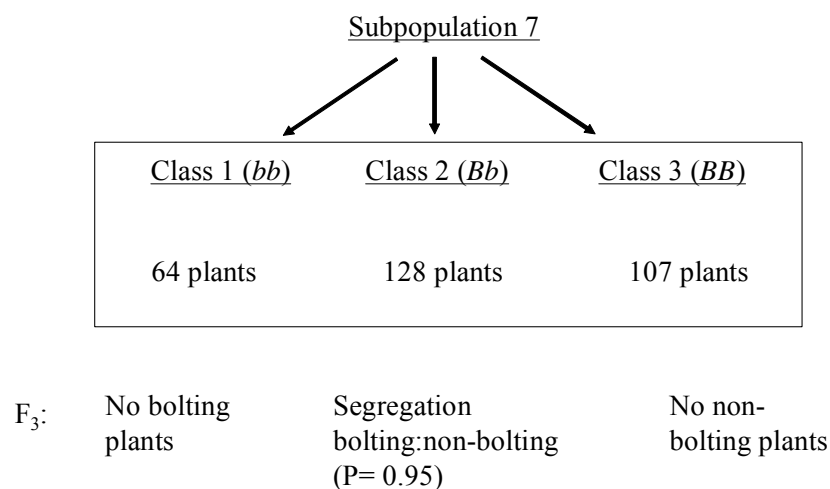
**Figure 7:** A possible explanation for the occurrence of deviating segregation ratios of F<sub>3</sub> families due to imperfect isolation of F<sub>2</sub> plants. **A)** Perfect isolation of F<sub>2</sub> plants. Three different segregation ratios 1:0, 3:1 and 0:1 (bolting: non-bolting) are expected in F<sub>3</sub> families. **B)** Inappropriate isolation of F<sub>2</sub> plants. In addition to the normal segregation ratios of F<sub>2</sub> plants in F<sub>3</sub> families, abnormal segregation might be expected in case of the two last classes.

By selfing of F<sub>2</sub> plants, different segregation ratios (1:0, 3:1, and 0:1 bolting: non-bolting) are expected for F<sub>3</sub> families, if genotypes of *BB*, *Bb*, and *bb* for their parents were considered. This holds true for complete isolation of the plants (Figure 7A). The three phenotypic classes (bolting F<sub>3</sub> plants, 3:1 segregating F<sub>3</sub> plants, and non-bolting F<sub>3</sub> plants) from the F<sub>3</sub> families that are in agreement with their genotypic classes (*BB*, *Bb*, and *bb*) are supposed to be reliable group of plants without cross-pollination (Figure 7B). In the first class (bolting F<sub>3</sub> plants, where its F<sub>2</sub> genotype was determined on the basis of its phenotypic segregation in F<sub>3</sub>), cross-pollination will have no effect on the phenotype in the following generation. While, in the other two classes (3:1 segregating F<sub>3</sub> plants and non-bolting F<sub>3</sub> plants), a pollen cloud of *B*-pollens, (falling on the plants of the other two classes) could theoretically

lead to wrong classification of the groups (Figure 7B). Since cross-pollination can take place in the segregating  $F_3$  families and in case of bolting  $F_3$  family, therefore, only the group of  $F_3$  families derived from  $bb$   $F_2$  plants was used for the estimation of the recombination rates (see 3.6.3.). This group consisted only of non-bolting plants, and their  $F_2$  parents must be of  $bb$  genotype. Only  $F_3$  data from this group can be regarded as absolutely reliable because no cross-pollination has occurred. Non-bolting  $F_3$  families are never be found in the progenies of  $BB$  or  $Bb$  plants.

### 3.1.2. Restructuring of the subpopulation 4 (960701) based on $F_3$ phenotypic data

The first solution for the deviated segregation of the subpopulation 4 (960701) was to reconstruct the subpopulation 4. The examination of a larger number of  $F_3$  phenotypes of the  $F_3$ -families provided a more reliable determination of the genotype of  $F_2$  plants (subpopulation 4, 960701). Consequently, the subpopulation 4 was reconstructed in a way that, from 746 plants with verified genotype in  $F_3$  (Table 5), 299  $F_2$  plants (subpopulation 7) were selected for linkage analysis using MAPMAKER/EXP 3.0b program. This subpopulation 7 (299  $F_2$  plants) included three classes of plants: The first class included all the non-bolting (64)  $F_2$  plants verified in  $F_3$ . The second class contained the 128  $F_2$  plants showing an expected segregating ratio of 3:1 in  $F_3$  (the  $\chi^2$  values were non-significant with  $P=0.95$ , Figure 8). The third class included the 107 bolting  $F_2$  plants verified in  $F_3$  (including 42 recombinant plants scored with the four previously mapped AFLP markers). Eight BAC-derived markers, one YAC marker, and one RFLP-derived CAPS marker were used to map the  $B$  gene using these 299  $F_2$  plants (verified phenotype in  $F_3$ ).



**Figure 8:** The synthetic  $F_2$  population selected from 960701 with 299 plants (subpopulation 7). The genotypes of these plants have been verified by phenotypic data in  $F_3$  families.

According to Hühn (1995) the linkage analysis could be performed only on the recessive individuals. Therefore, the second solution to overcome the deviated segregation of the subpopulation 4 (960701) was to use only the non-bolting plants (64) from subpopulation 4 for fine mapping. This class of plants was considered to be the most phenotypically reliable class of plants (see 3.1.1.).

### 3.1.3. Increasing number of non-bolting plants used for two-point analysis

In order to increase the resolution of the genetic map based on the non-bolting plants, two more subpopulations 5 and 6 were produced from population 960701. In addition, a new  $F_2$  mapping

population 950519 was produced to prove the deviation from 3:1 segregation as has been observed in subpopulation 4 (960701). The new population 950519 originated from a different F<sub>1</sub> plant of the same cross. Plants were sown in May 2003. In this way, plants were not exposed to cold night temperatures and thus any effect of vernalization on the bolting behavior was eliminated.

The first bolting plants were observed 6 weeks after sowing. The phenotyping of plants was done every three days within the period of 9 to 12 weeks, thus a differentiation with respect to early and late bolting could be observed (Table 6).

**Table 6:** Phenotyping of F<sub>2</sub> population 960701 (subpopulations 4, 5, and 6) and of population 950619 according to F<sub>2</sub> phenotyping as well as F<sub>3</sub> phenotyping of subpopulation 4.

F <sub>2</sub> -Population	Total number of plants	No. of bolting plants	No. of non-bolting plants	Ratio of bolting/non-bolting	$\chi^2$
Phenotyping 9 weeks after sowing					
960701/95 (subpop 5)	179	134	45	2.98	0.002 ns
960701/96 (subpop 6)	472	365	107	3.41	1.35 ns
950619	965	730	235	3.11	0.215 ns
Phenotyping 12 weeks after sowing					
960701/95 (subpop 5)	179	144	35	4.11	2.78 ns
960701/96 (subpop 6)	472	386	86	4.49	11.33*
950619	965	786	179	4.39	20.99*
Phenotyping 15 weeks after sowing					
960701 (subpop 4) <sup>1</sup>	1359	1053	306	3.44	4.47*
960701 (subpop 4)	811 <sup>2</sup>	738	73	10.1	110.7*

$\alpha_{0.05} = 3.84$ ,  $df = 1$ . \*:  $\chi^2$  is significant at  $P = 0.05$ . ns: not significant at  $P = 0.05$ .

<sup>1</sup>: 1359 F<sub>2</sub> plants of subpopulation 4 phenotyped by El-Mezawy (2001).

<sup>2</sup>: from 811 F<sub>3</sub>-full sister progenies (with  $n > 7$  plants), observed values for the bolting behavior were available.

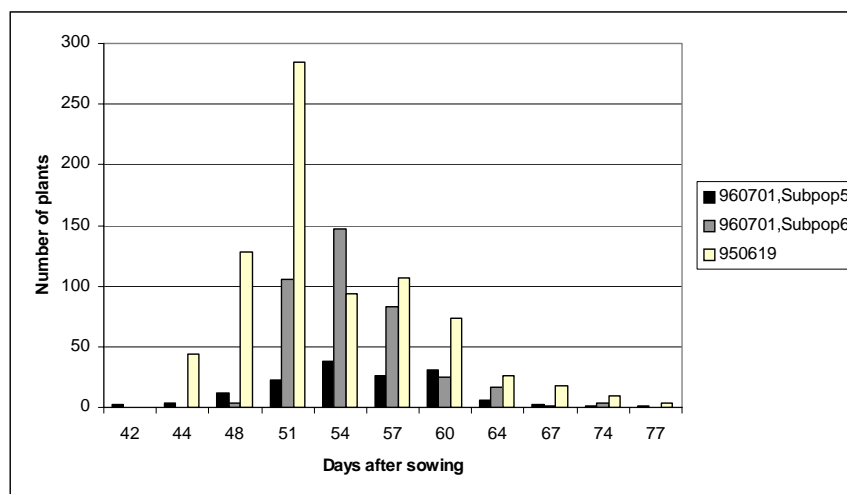
If scoring the members of bolting and non-bolting plants was done 9 weeks after sowing, segregation ratios of 143:45 (subpopulation 5), 365:107 (subpopulation 6), and 730:235 (population 950619) were found. These ratios were not significantly deviating from the expected values ( $P = 0.05$ ,  $df = 1$ ). With the determination of the bolting character after 12 weeks, the segregation ratios have been changed (386 bolting and 86 non-bolting plants in subpopulation 6). For population 950619, 786 bolting and 179 non-bolting F<sub>2</sub> individuals were found (Table 6). These segregation ratios significantly deviated from the expected values ( $P = 0.05$ ,  $df = 1$ ). From the extended subpopulations 5 and 6 (960701) and the new population (950619), 35, 86, and 179 non-bolting plants, respectively, were used for two-point linkage analysis (Table 6). In total, 300 non-bolting plants were genotyped with 8 molecular markers (7 BAC-derived markers and 1 RFLP-derived CAPS marker).

### 3.2. Environmental impact on bolting character

Boudry *et al.* (1994b) have reported that there are effects of different environmental conditions, e.g. temperature and day length, on the expression and penetrance of the bolting behavior. They found that



growing season affected significantly the segregation ratio of bolting. They described that complete penetrance of the *B* gene is only achieved under spring growing conditions.



**Figure 9:** Distribution of bolting time points in three different populations, subpop 5, subpop 6 (96701), and population 950619 sown in early May 2003. Bolting time points were determined as the day at which first stem elongation (time point) occurred.

By sowing in late spring and growing under long days, in comparison to sowing in March/April, the plants developed well and bolted relatively fast. Figure 9 shows the distribution and frequency of bolting plants during the first 12 weeks after sowing. If the bolting point time was 6 weeks after sowing (as found in subpopulation 5), then approx. 70% of the bolting plants have bolted in the first 9 weeks. In addition, bolting plants have also been observed after 12 weeks from sowing.

### 3.3. BAC sequencing

#### 3.3.1. BAC end sequencing and homology search

##### 3.3.1.1. Criteria for selection of BACs

A BAC library was constructed from a double haploid non-bolting line (KWS2320). This BAC library consisted of 57,600 clones and approximately represents 8.7 genome equivalent (Hohmann *et al.*, 2003). For screening of the BAC library during the GABI-BOLT project, probes from 10 YAC ends and four AFLP markers closely linked to *B* gene were used. 54 different candidate BAC clones could be identified from the *B* gene region. The number of 54 candidate BACs has been reduced to 52 BACs, because two clones, which contained more than one insert were eliminated. 104 BAC ends were sequenced (DLMBC, Berlin). For sequencing of the different inserts, ends of the BAC clones, the SP6 and T7 sequencing primers were used. It is possible to differentiate between the insert ends in the SP6 end and T7 end of an insert.

### 3.3.1.2. Primer design and sequence analysis

Sequence data were edited and analyzed with SeqMan<sup>TM</sup>II from DNASTAR software (GATC Biotech, Konstanz, Germany). The average read length ranged from 523 bp (SP6 ends) to 534 bp (T7 ends). The total sequence length was 55,464 bp. Homology search on nucleotide and amino acid levels (BLASTN and BLASTX analyses) was done with the purpose to screen the databases for any homology of the BAC-end sequences to bolting or flowering related genes. Specific primers for each BAC end sequence were designed using PrimerSelect<sup>TM</sup> software (DNASTAR, GATC Biotech, Konstanz, Germany).

For 70.8% of the BAC ends no clear sequence homology was found. 22.4% had homology to repetitive DNAs (retrotransposons, satellite DNA from *Beta vulgaris* and *B. vulgaris* repetitive DNA (clone pDRV2, and 18S rDNA). 6.7% of the BAC-end sequences showed homology to plant genes (Table 7).

**Table 7:** BAC end sequences with significant homology to known plant genes.

BAC No.	BAC end	Homology/function (Accession No.)	E-value
GJ18	SP6	<i>Beta vulgaris</i> cell wall invertase (AJ277458)	3 <sup>-27</sup>
GJ44	SP6	Tissue specific promoter (AX499164)	1 <sup>-33</sup>
GJ48	SP6	<i>Beta vulgaris</i> root-cDNA (BI698340)	2 <sup>-45</sup>
GJ83	SP6	<i>Beta vulgaris</i> root-cDNA (BI698297)	3 <sup>-54</sup>
GJ65	T7	Glyphosate resistant transgenic plant (BD008402)	3 <sup>-59</sup>
GJ67	T7	Auxin transport protein (AF507080)	3 <sup>-21</sup>
GJ68	T7	Glyphosate resistant transgenic plant (BD008402)	1 <sup>-94</sup>

BAC-end sequences with homology to flower formation, development, vernalization, genes involved in gibberellins synthesis, circadian rhythm, and/or signal transduction are of particular interest for cloning of the *B* gene. With one exception, no clear homology to candidate genes was found. The T7-end sequence of BAC GJ67 showed homology to an auxin transporter protein from Arabidopsis. Such protein enables the directed movement of substances (macromolecules, small molecules or ions) into, out of, within, or between cells.

### 3.3.2. Shotgun sequencing of BAC clones and sequence analysis

Three BACs (GJ01, 140 kb; GJ70, 150 kb; and GJ131, 180 kb) were selected for shotgun sequencing. These three BACs represented the closest two contigs to the bolting *B* gene. Sequencing of the three BACs was performed by the company AGOWA (Berlin) in the following way: first, the BACs were sheared mechanically and size-fractionated. Second, the fragments within the range of 1.5 and 2 kb were shotgun cloned into the vector PCR Blunt (Invitrogen). Third, from each BAC, 96 plasmids were sequenced.

After sequence analysis with DNASTAR 43-50, different sequence contigs for each BAC were assembled and 15-23 different ORFs (>200 bp) could be identified (Table 8). Approximately, for each BAC, 50 kb non-redundant sequence information resulted from the assembly of sequences of 96 plasmid clones (1× equivalent).

Sequence comparisons of the contigs by BLASTN and BLASTX as well as Pedant analysis for identification of ORFs with functional domains were done. A BLASTN with GJ01 contig 36 showed a significant homology to a mitochondrial processing peptidase with  $2^{-38}$  and GJ131 contig 15 revealed 50% identity over 353 amino acids of the *Myb* transcription factor protein (T25B15.30, At3g52250) over a 128 bp region with E-value  $7^{-87}$ . No more significant hits to known Arabidopsis genes were found.

**Table 8:** Characterization of the three shotgun libraries. Data was supplied by Dr. Uwe Hohmann.

BAC number	GJ01	GJ70	GJ131
Size (kb)	140	148	182
Total sequence length (bp)	154,704	142,663	141,895
Total sequence length, non-redundant (bp)	52,859	49,709	48,225
Number of contigs	43	50	50
Contigs with ORFs (>200 bp)	23	23	15

A BLAST search was conducted at the nucleotide level with the sequence contigs (with identified ORFs) from the three partially sequenced BACs against the EST database of the BIOMAX server of KWS. For similarity search, the standard threshold (for identification of the homology regions) as well as the default gap penalties were chosen. From all the ORFs that have been identified, 17 ORFs with homology to unknown and/or hypothetical proteins and 16 ORFs with homology to ESTs were found (Table 9). ESTs with the highest E-values are summarized in Table 9.

**Table 9:** ORFs sequence homology to ESTs. Data was supplied by Dr. Uwe Hohmann.

BAC	Sequence contig No.	E-value (BLASTN)	Homology to ESTs	Accession No.
GJ01	15	$1^{-57}$	Storage root	BQ590275
GJ01	33	$9^{-76}$	Partial cell wall invertase gene	AJ277458
GJ01	38	$1^{-70}$	Chitinase 1 gene	X79301
GJ01	43	$2^{-75}$	Root	BI698347
GJ70	03	$1^{-57}$	Storage root, resistance gene	BQ591697
GJ70	28	$1^{-98}$	Leaf	BQ588292
GJ131	07	$1^{-70}$	Root	BI643405
GJ131	28	$1^{-66}$	Inflorescence	BQ582527

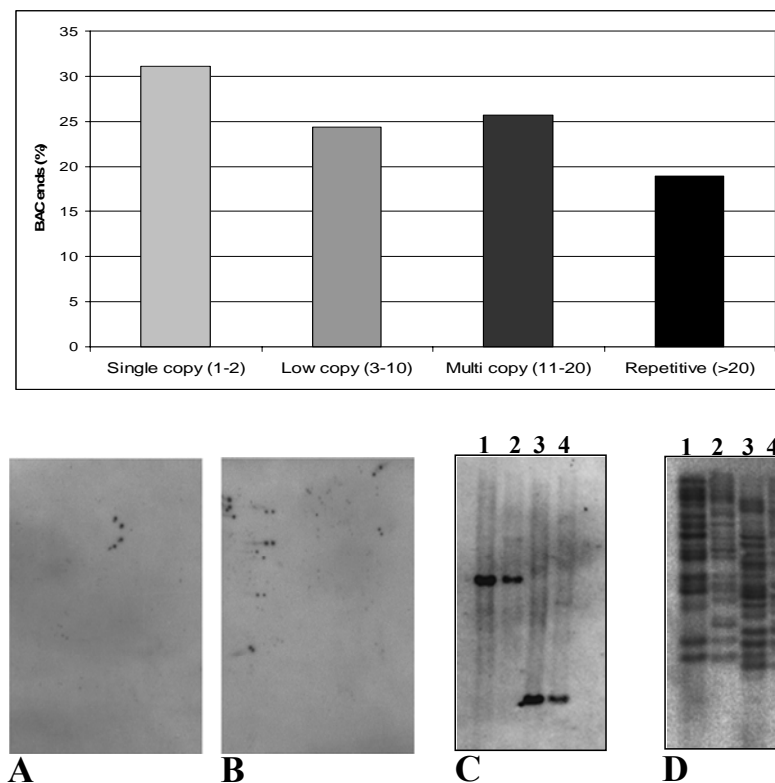
The potential ORF contigs with the longest ORFs were selected for marker development (see 3.5.3.).

### 3.4. Copy number determination of BAC sequences

#### 3.4.1. Copy number determination of BAC ends

The copy number of the BAC ends was determined using two methods: genomic Southern hybridization and genomic library screening. In the first method, genomic Southern blots were prepared from plant 93167P, homozygous for the bolting gene (*BB*), and A906001, homozygous for the recessive allele (*bb*), by restriction of the genomic DNA with two different enzymes (*EcoRI* and

*HindIII*). Then the filters were hybridized with 74 BAC end probes. Autoradiograms were evaluated according to number and quality of hybridization signals. 31% of the sequences (1-2 bands) could be regarded as single copy, while 24% were considered to be low copy (3-10 bands). 25% of the BAC ends were considered as multi-copy (11-20 bands), while 19% were repetitive due to strong signals without distinguishable bands. One of the sequences gave a ladder pattern typical for tandem repeated elements, demonstrating that preselection by sequence analysis was successful. The copy numbers of 23 BAC ends were determined by A. Telgmann (Telgmann, 2002).



**Figure 10:** Frequency distribution of BAC ends with respect to copy numbers as determined by genomic Southern hybridization and BAC filter screening. The probes GJ21SP6 (A, C) and GJ13T7 (B, D) were used. A and B are high-density filters (3072 BAC clones were twice spotted), C and D are genomic filters prepared by restriction of the bolting 93167P (lane 1 and 3) and non-bolting A906001 (lane 2 and 4). DNA digested with *Hind* III (lane 1 and 2) and *Eco*RI (Lane 3 and 4). Hybridization with BAC end probes GJ21SP6 (A, C) and GJ13T7 (B, D). The filters were exposed for 72 h.

In a second approach, copy number of all 74 BAC ends was determined by colony hybridization with high-density filters containing 3072 BAC clones spotted in duplicate, which is equivalent to 50% of the beet genome. An example is shown in Figure 10. Correlation coefficient ( $r$ ) between genomic Southern and colony hybridization data was calculated for all probes with number of bands  $<15$ . A correlation was calculated in the following way: first, the number of signals on the BAC filters was counted for each BAC end probe. In the next step, the number of bands on the genomic Southern membranes (in case of *Eco*RI and *Hind*III) was determined. Finally a correlation coefficient ( $r$ ) between the number of signals on BAC filters and number of bands found on genomic filters ( $<15$  bands) was estimated. A positive correlation coefficient ( $r=0.75$ ) has been found. This calculated correlation coefficient ( $r$ ) indicated strong positive correlation between the number of bands on the genomic Southern and the number of positive signals that could be detected on the high-density filters.

### 3.4.2. Copy number determination of BAC-derived ORF sequences

The copy number of BAC-derived ORF sequences was determined in the same way as described for BAC ends (3.4.1.). Probes from ORF sequences were hybridized to filters with genomic sugar beet DNA and high-density filters (half genome equivalent of sugar beet). The ORF probes were purified PCR products (450-800 bp), produced by amplification of intron and exon sequence regions. Intron regions were identified by using the GeneQuest<sup>TM</sup> software (DNASTAR, GATC Biotech, Konstanz, Germany). Introns and exons were identified and long ORFs were chosen for further marker analysis. The intron and exon probes were prepared by PCR amplification from the corresponding BAC clones (Table 10).

**Table 10:** Copy number of different ORF sequences determined by genomic Southern hybridization and BAC filter screening. Single copy: 1-2 bands, low copy: 3-10 bands, multi-copy: 11-20 bands, and repetitive: smear (no countable bands).

GJ01		GJ70		GJ131	
Sequence contig	Copy number	Sequence contig	Copy number	Sequence contig	Copy number
15	Multi-copy	3	Single copy	8	Single copy
36	Single copy	9	Repetitive	15	Single copy

Only one repetitive sequence and one multi-copy sequence were found. The remaining 4 ORF probes were single copies, in particular whenever the probes were generated from exon regions. Detailed results of determination of the copy number from BAC-derived sequences could be seen in appendix 8.1.

## 3.5. Molecular marker analysis

### 3.5.1. BAC-derived markers

The reasons for selecting molecular markers from BACs were as follows:

First, the BACs selected for marker development were identified with probes derived from previously mapped AFLP markers in close vicinity of *B* gene. To confirm the origin of the BACs from *B* gene region, re-mapping of BAC ends was essential. Second, BAC sequences were useful resources for development of molecular markers of known physical distances. BAC sequences were turned into PCR-based markers, in order to increase the marker density within the *B* gene region as well as to use them as diagnostic markers in detection of early bolting gene *B* in wild beets and breeding lines. Third, the markers were also helpful for the construction of physical maps and for the orientation of BACs around the *B* gene.

#### 3.5.1.1. Criteria for selection of BAC-derived sequences for marker development

Single- and low-copy BAC-end sequences were selected for marker development. In addition, single- and low-copy ORF-derived sequences were selected from partial sequences of shotgun libraries of three BACs (GJ01, GJ70 and GJ131). The BAC-derived sequences were selected from all four BAC contigs encompassing the gene *B*. In total, 37 BAC-end sequences from BACs of the four different contigs were selected. The sequences originated from genotype KWS2320 (the donor of the BAC

library), which is a non-bolting genotype (*bb*). For marker development, amplicons from both parents of the mapping population were produced and analyzed for polymorphisms. Potential CAPS markers were identified by computer analysis using MapDraw<sup>TM</sup> from the DNASTAR software package (GATC Biotech, Konstanz, Germany). With this program, recognition sites for restriction endonucleases were determined (see 3.5.2.).

### 3.5.1.2. PASA markers

Thirty-seven BAC-end sequences were analyzed for polymorphisms. The sequence specific primers were designed within the GABI-BOLT project. Using these primers, PCR products were amplified from the two parents 93167P (*BB*) and A906001 (*bb*) of the mapping populations. After agarose gel electrophoresis of the PCR products, polymorphisms with respect to the presence or absence of five PCR products were found. From the T7 ends of BAC GJ01, GJ17, GJ22, GJ37, and GJ131, only amplicons from non-bolting genotype A906001 were obtained. The primers, derived from a non-bolting genotype KWS2320, failed to generate amplicons from the bolting genotype 93167P of the mapping population, except for GJ131T7. The markers were designated as PASA (PCR amplification of specific alleles) marker GJ01T7, GJ17T7, GJ22T7, GJ37T7, and GJ131T7 (Figure 11).

#### Marker GJ01T7:

The GJ01T7-specific primer combination was used in a PCR experiment with DNA from the non-bolting genotypes KWS2320 (*bb*) and A906001 (*bb*), the bolting genotype 93167P (*BB*), and DNA from BAC GJ01 as a control. An amplicon with the expected size of 456 bp was observed for the non-bolting genotypes KWS2320 and A906001 as well as for control. A PCR fragment (bolting allele) for the bolting genotype 93167P was not detected. According to the PCR amplification of specific alleles (PASA), the type of marker was designated as a dominant PASA marker GJ01T7 (Figure 11).

#### Marker GJ17T7:

The GJ17T7-specific primer combination was used in a PCR experiment with DNA from the non-bolting genotypes KWS2320 (*bb*) and A906001 (*bb*), the bolting genotype 93167P (*BB*), and DNA from BAC GJ17 as a control. An amplicon with the expected size of 521 bp was observed for the non-bolting genotypes KWS2320 and A906001 as well as for the control. A PCR fragment (bolting allele) for the bolting genotype 93167P was not detected. According to the PCR amplification of specific alleles (PASA), the type of marker was designated as a dominant PASA marker GJ17T7 (Figure 11).

#### Marker GJ22T7:

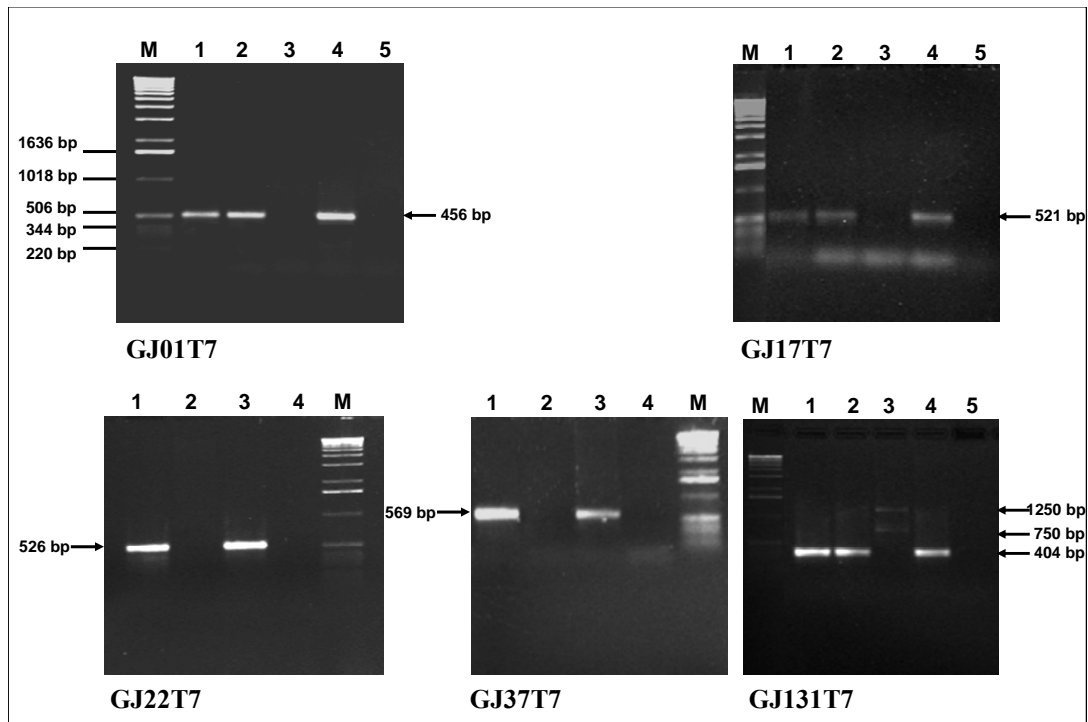
The GJ22T7-specific primer combination was used in a PCR experiment with DNA from the non-bolting genotypes KWS2320 (*bb*) and A906001 (*bb*), and the bolting genotype 93167P (*BB*). An amplicon with the expected size of 526 bp was observed for the non-bolting genotypes KWS2320 and A906001. A PCR fragment (bolting allele) for the bolting genotype 93167P was not detected. According to the PCR amplification of specific alleles (PASA), the type of marker was designated as a dominant PASA marker GJ22T7 (Figure 11).

#### Marker GJ37T7:

The GJ37T7-specific primer combination was used in a PCR experiment with DNA from the non-bolting genotypes KWS2320 (*bb*) and A906001 (*bb*), and the bolting genotype 93167P (*BB*). An amplicon with the expected size of 569 bp was observed for the non-bolting genotypes KWS2320 and A906001. A PCR fragment (bolting allele) for the bolting genotype 93167P was not detected. According to the PCR amplification of specific alleles (PASA), the type of marker was designated as a dominant PASA marker GJ37T7 (Figure 11).

### Marker GJ131T7:

The GJ131T7-specific primer combination was used in a PCR experiment with DNA from the non-bolting genotypes KWS2320 (*bb*) and A906001 (*bb*), the bolting genotype 93167P (*BB*), and DNA from BAC GJ01 as a control. An amplicon with the expected size of 404 bp was observed for the non-bolting genotypes KWS2320 and A906001 as well as for control. Two PCR fragments of 750 and 1250 bp for the bolting genotype 93167P were detected. According to the PCR amplification of specific alleles (PASA), the type of marker was designated as a co-dominant PASA marker GJ131T7 (Figure 11).



**Figure 11:** PASA BAC-derived markers of the bolting gene *B* of sugar beet. Agarose gel electrophoresis of PCR products produced with primer combinations from the T7 ends of the following BACs: GJ01, GJ17, GJ22, GJ37 and GJ131. 5  $\mu$ l of the PCR reactions were loaded in each slot and separated on a 1% agarose gel for 60 min. at 60 V. **GJ01T7**, **GJ17T7**, **GJ131T7**) the size of the expected PCR fragments were 456, 521, and 404 bp, respectively. 1: BAC GJ01, GJ17 or GJ131; 2: KWS2320; 3: bolting parent 93167P; 4: non-bolting parent A906001, 5: control without DNA template. **GJ22T7**, **GJ37T7**) the size of the expected PCR fragments were 526, 569 bp, respectively. 1: KWS2320; 2: bolting parent 93167P; 3: non-bolting parent A906001, 4: control without template DNA. M: 1kb DNA size marker (Gibco-BRL).

### 3.5.2. CAPS markers

For those amplicons which were not polymorphic between both parents 93167P (*BB*) and A906001 (*bb*), restriction analysis of the BAC-end sequences were accomplished. The sequences representing the selected BAC ends which are the sequences from the KWS2320 non-bolting genotype, the donor of the BAC library, were searched for restriction sites using MapDraw<sup>TM</sup> from the DNASTAR software package (GATC Biotech, Konstanz, Germany). Restriction enzymes that were available,

cheap, and supposed to cut the BAC-end sequence were chosen. In this way, recognition sites for 23 different restriction enzymes could be detected. The following table lists the names of BAC-end sequences and the respective enzymes, which were used for cleavage (Table 11).

**Table 11:** List of the restriction enzymes selected for CAPS development based on restriction map analysis of BAC-end sequences using MapDraw™.

BAC end	Restriction enzymes tested
GJ01SP6	<i>DpnI, HaeIII, HinfI, NdeI, MboII, PstI &amp; RsaI (AfaI)</i>
GJ08SP6	<i>AluI, EcoRI, HindIII, HinfI &amp; MboI (Sau3AI)</i>
GJ10SP6	<i>AluI, BglII, HindIII, HinfI* &amp; MboI (Sau3AI)</i>
GJ10T7	<i>AluI, DraI, HindIII, HinfI &amp; MboI (Sau3AI)</i>
GJ13T7	<i>AluI, HaeIII, MseI, RsaI (AfaI) &amp; MboI (Sau3AI)</i>
GJ17SP6	<i>AluI, HinfI* &amp; MboI (Sau3AI)</i>
GJ18SP6	<i>DpnI, HphI, MboI (Sau3AI), MboII, MseI, PstI, &amp; TaqI</i>
GJ18T7	<i>AluI, BclI, DpnI, DraI, HinfI, HphI, MboI (Sau3AI), MseI &amp; TaqI</i>
GJ22SP6	<i>AluI, HincII, MboI (Sau3AI)</i>
GJ29SP6	<i>AluI, DpnI, HaeIII, HinfI, HphI, MboI (Sau3AI), MboII, MseI, PstI, TaqI &amp; VspI</i>
GJ29T7	<i>AluI*, HaeIII, HinfI, MboI (Sau3AI)</i>
GJ33SP6	<i>AluI, MboII &amp; RsaI (AfaI)</i>
GJ37SP6	<i>AluI, HindIII, HphI, MboI (Sau3AI) &amp; MseI</i>
GJ44SP6	<i>AccI, AluI, BcuI, HindIII, HincII, HinfI, HphI, MseI, PstI, RsaI (AfaI)</i>
GJ44T7	<i>AluI, HindIII, HincII, HinfI, HphI, MboII, MseI, PstI, RsaI (AfaI)</i>
GJ49SP6	<i>AluI, HindIII, HphI, MboI (Sau3AI)* &amp; MseI</i>
GJ49T7	<i>AluI, HindIII, HphI, MboI (Sau3AI) &amp; MseI</i>
GJ55SP6	<i>AluI, EcoRI, EcoRV, MboI (Sau3AI) &amp; PstI</i>
GJ55T7	<i>AluI, EcoRI, EcoRV, MboI (Sau3AI) &amp; PstI</i>
GJ65SP6	<i>AluI, EcoRI, HphI, MboI (Sau3AI) &amp; MseI</i>
GJ65T7	<i>AfaI, AluI, HindIII, MseI, RsaI (AfaI)</i>
GJ67SP6	<i>AluI, EcoRI, EcoRV, MboI (Sau3AI) &amp; PstI</i>
GJ67T7	<i>AluI, EcoRI, EcoRV, MboI (Sau3AI) &amp; PstI</i>
GJ70SP6	<i>AluI, HinfI</i>
GJ73SP6	<i>AluI, HphI, HinfI &amp; MboI (Sau3AI)</i>
GJ73T7	<i>AluI, HphI, HinfI &amp; MboI (Sau3AI)</i>
GJ76SP6	<i>AluI, EcoRI, EcoRV, MboI (Sau3AI) &amp; PstI</i>
GJ76T7	<i>AluI, EcoRI, EcoRV, MboI (Sau3AI) &amp; PstI</i>
GJ131SP6	<i>AluI*, HinfI, HphI &amp; MboI (Sau3AI)</i>

\*: Informative restriction enzyme.

PCR was done with DNA from 93167P (*BB*), A906001 (*bb*), KWS2320, and the corresponding BAC. The amplicons were digested with up to 23 different restriction endonucleases and separated by agarose gel electrophoresis on 1.5-2% agarose gels. In this way, length polymorphisms for another five BAC-end sequences (GJ10T7, GJ17SP6, GJ29T7, GJ49SP6, and GJ131SP6) were detected. They were designated as CAPS (Cleaved Amplified Polymorphic Sequence) markers. Four of these CAPS markers (GJ017SP6, GJ029T7, GJ49SP6, and GJ131SP6) were co-dominant, while the marker GJ10T7 was scored as dominant.

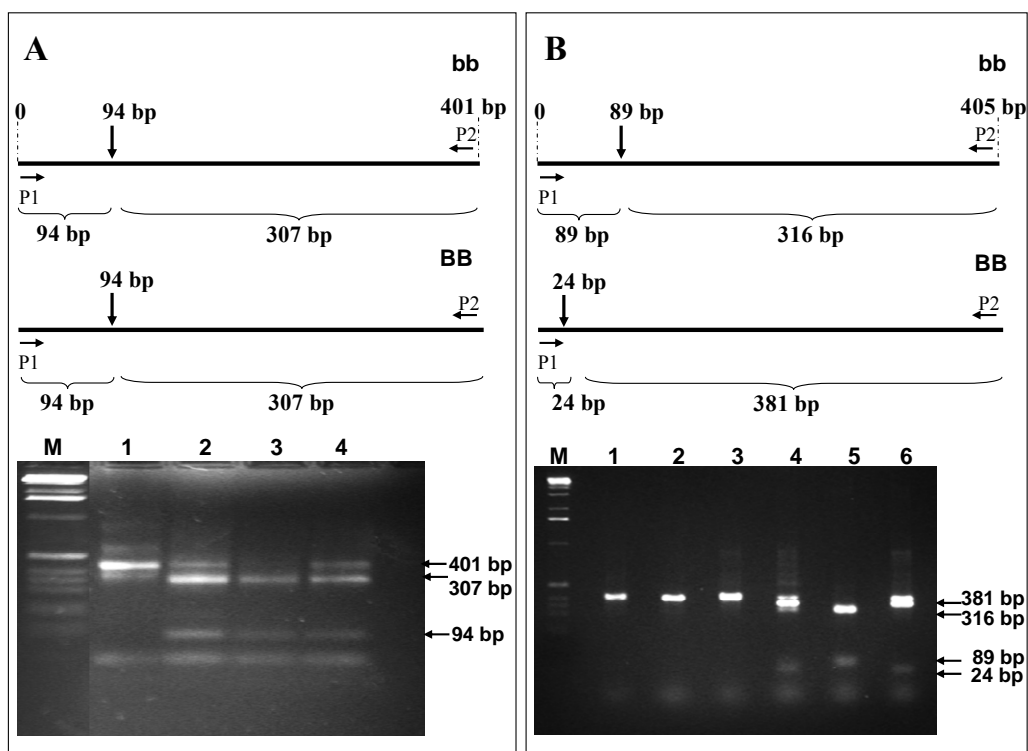


Marker GJ10T7:

A single PCR fragment of the expected size (401 bp) was amplified with the primers derived from GJ10T7 end sequence in 93167P (*BB*), A906001 (*bb*), and KWS2320. The PCR products were digested with *Hinf*I at 37°C for three hours. The digested PCR products were separated on a 1.5% agarose gel. Two fragments (94 and 307 bp) were observed in the bolting 93167P parent after gel electrophoresis, while an extra fragment 401 bp could be seen in A906001 (*bb*) and KWS2320, that is considered to be a non-restricted PCR copy. This type of polymorphism was described as a dominant CAPS marker (Figure 12A).

Marker GJ17SP6:

Using the specific primers derived from GJ17SP6 end sequence, the expected 405 bp PCR product was generated from 93167P (*BB*), A906001 (*bb*), and a heterozygous plant. A restriction digestion of the PCR products was performed at 37°C for three hours with *Hinf*I. After separation of digested PCR products on a 2% agarose gel, two fragments (89 and 316 bp) were visible in non-bolting A906001. Two different fragments (24 and 381 bp) were observed in bolting parent 93167P. In the heterozygous plant, all four fragments were observed. This marker was considered to be a co-dominant CAPS marker. Thus, it has differentiated between homozygous (*BB*, *bb*) and heterozygous (*Bb*) plants (Figure 12B).

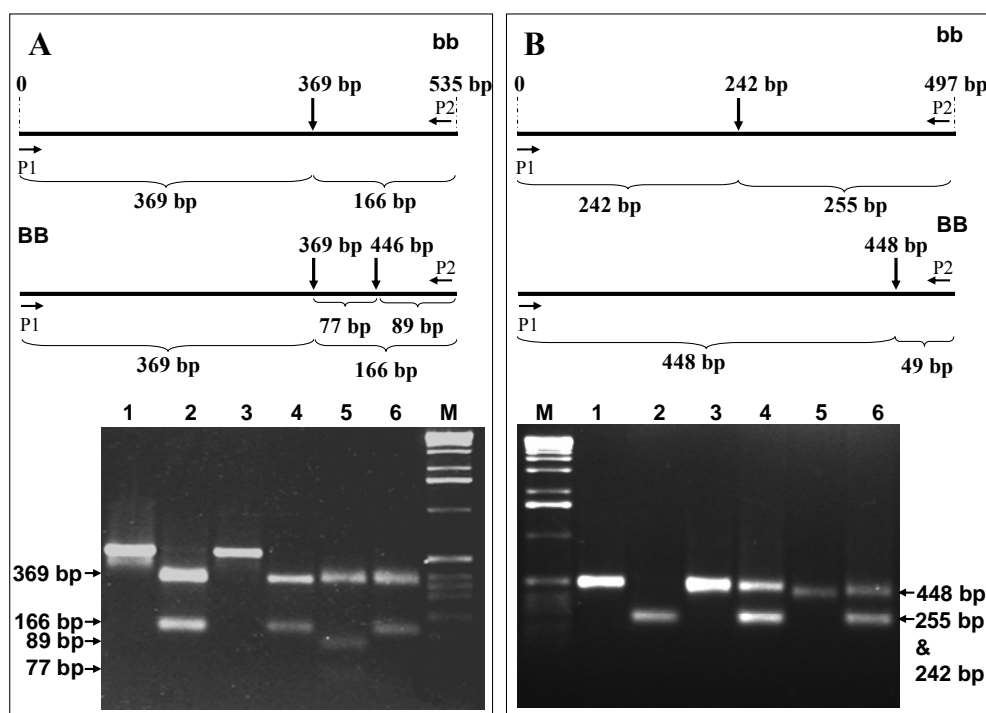


**Figure 12:** BAC-derived CAPS markers GJ10T7 (**A**) and GJ17SP6 (**B**) of the bolting gene *B* of sugar beet. Bottom is agarose gel electrophoresis of restricted PCR products amplified with primer combinations from the T7 end of BAC GJ10 and SP6 end of BAC GJ17. 5  $\mu$ l of the PCR reactions were digested with *Hinf*I in a 10  $\mu$ l total volume and separated on a 1.5-2% agarose gel for 60 min. at 60 V. **A**) GJ10T7; 1: undigested PCR product amplified from KWS2320 genotype; 2-4: *Hinf*I restricted PCR products from non-bolting and bolting genotypes, three fragments 401, 307, and 94 bp were observed in KWS2320 and non-bolting A906001; 3: only two fragments 307 and 94 bp were visible in bolting parent 93167P. Top, is graphical representation of *Hinf*I restriction sites within the 401 bp fragment amplified from the T7 end of BAC GJ10. Arrows indicate primer positions. P1 and P2 are forward and reverse primers, respectively. Vertical arrows indicate the positions of *Hinf*I restriction sites. **B**) GJ17SP6, 1, 2, and 3: undigested PCR products of heterozygous plant, non-bolting, and bolting parents. *Hinf*I digested PCR products from heterozygous plant, non-bolting, and bolting parent; 5: two fragments 89 and 316 bp were visible in non-bolting parent A906001; 6: other *Hinf*I restriction site observed at 24 bp and resulted in two different fragments of 24 and 381 bp in bolting parent 93167P; 4: heterozygous plant. M: 1kb DNA size marker (Gibco-BRL). Top, is a graphical representation of *Hinf*I restriction sites within the 405 bp fragment amplified from the SP6 end of BAC GJ17. Arrows indicate primer positions. P1 and P2 are forward and reverse primers, respectively. Vertical arrows indicate the positions of *Hinf*I restriction sites.

#### Marker GJ29T7:

After restriction site analysis of GJ29T7 end sequence, *Alu*I was found to cut the non-bolting A906001 (*bb*) PCR fragment at 369 bp. *Alu*I restriction of the PCR product amplified from non-bolting A906001 (*bb*) genotype, using the GJ29T7P1P2 primer combination, resulted in expected electrophoretic bands pattern. In the non-bolting parent (A906001), two bands were observed, 369 bp and 166 bp in length. As expected, two bands were observed for genotype KWS2320 as well. In contrast, in the bolting parent 93167P (*BB*), an additional potential restriction site was present at base pair position 446. Here,

the 166 bp fragment was cut into two shorter fragments of 89 bp and 77 bp (Figure 13A). The GJ29T7 CAPS marker is regarded to be a co-dominant CAPS marker.



**Figure 13:** BAC-derived CAPS markers GJ29T7 (A) and GJ49SP6 (B) of the bolting gene *B* of sugar beet. Bottom, is agarose gel electrophoresis of restricted PCR products amplified with primer combinations from the T7 end of BAC GJ29 and SP6 end of BAC GJ49. 5  $\mu$ l of the PCR reactions were digested with *AluI* and *Sau3AI* in a 10  $\mu$ l total volume and separated on a 1.5-2% agarose gel for 60 min. at 60 V. **A)** GJ29T7; 1 and 3: undigested PCR products amplified from GJ29 and KWS2320; 2, 4, and 6: two fragments of 369 and 166 bp were observed in GJ29, KWS2320, and non-bolting A906001 after *AluI* digestion; 5: three fragments 369, 89, and 77 bp were visible in bolting parent 93167P. Top, is a graphical representation of *AluI* restriction sites within the 535 bp fragment amplified from the T7 end of BAC GJ29 in non-bolting (*bb*) and bolting (*BB*) genotypes. Arrows indicate primer positions. P1 and P2 are forward and reverse primers, respectively. Vertical arrows indicate the positions of *AluI* restriction sites. **B)** GJ49SP6; 1 and 3: undigested PCR products amplified from GJ49 and KWS2320; 2, 4, 5, and 6: *Sau3AI* digested PCR products; 2: two fragments of 242 and 255 bp were visible as one band in GJ49; 4: two bands of 497 and 255/242 bp were visible for non-bolting genotypes KWS2320 and A906001, respectively; 5: *Sau3AI* restriction site observed at 448 bp and resulted in two different fragments of 448 and 49 bp in the bolting parent 93167P. M: 1kb DNA size marker (Gibco-BRL). Top, is a graphical representation of *Sau3AI* restriction sites within the 497 bp fragment amplified from the SP6 end of BAC GJ49. Arrows indicate primer positions. P1 and P2 are forward and reverse primers, respectively. Vertical arrows indicate the positions of *Sau3AI* restriction sites.

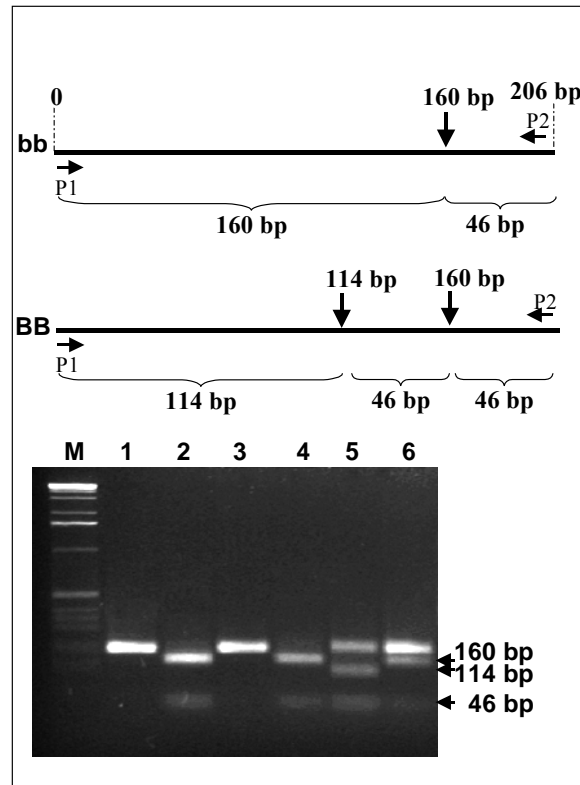
#### Marker GJ49SP6:

A potential *Sau3AI* restriction site was detected at base pair position 242 of the BAC-derived GJ49SP6 end sequence. The *Sau3AI* enzyme cut the PCR fragments amplified from GJ49, KWS2320, and non-bolting A906001 parent into two fragments 242 and 255 bp (Figure 13B). The size difference of the

two fragments could not be resolved under conditions of electrophoresis used. Thus, the two fragments were visible as one band. Interestingly, the 497 bp PCR fragment was present in the amplicons of both KWS2320 and non-bolting genotypes. This may indicate that the digestion was not complete or the presence of another copy in the genome but without the *Sau3AI* restriction site. The bolting parent (93167P) has different cutting site for *Sau3AI* at 448 bp which resulted in fragments of 448 bp and 49 bp. The GJ49SP6 marker was referred to be a co-dominant CAPS marker.

#### Marker GJ131SP6:

With the GJ131SP6 CAPS marker, the expected electrophoretic bands pattern was observed for the digested PCR products of GJ131 and KWS2320, a 206 bp PCR product was digested with *AluI*. Two fragments of 160 bp and 46 bp, as expected from the restriction site analysis on the BAC end sequence, were observed for the non-bolting genotype A906001. In the bolting parent (93167P), another *AluI* restriction site was detected at base pair position 114. This way, a different banding pattern was observed with a bolting (*BB*) specific fragment of 114 bp. However, there was another copy of the PCR product (206 bp) amplified from bolting and non-bolting genotypes that could not be digested with *AluI*. The non-digested copy of the PCR product may result from a duplication of the sequence in the sugar beet genome. Thus, the GJ131SP6 end is low copy in the sugar beet genome.



**Figure 14:** BAC-derived CAPS marker GJ131SP6 of the bolting gene *B* of sugar beet. Top, is a graphical representation of *AluI* restriction sites within the 206 bp fragment amplified from the SP6 end of BAC GJ131. Arrows indicate primer positions. P1 and P2 are forward and reverse primers, respectively. Vertical arrows indicate the positions of *AluI* restriction sites. Bottom, is agarose gel electrophoresis of restricted PCR products amplified with primer combinations from the SP6 end of BAC GJ131. 5  $\mu$ l of the PCR reactions were digested with *AluI* in 10  $\mu$ l total volume and separated on a 1.5% agarose gel for 60 min. at 60 V. 1 and 3: undigested PCR products amplified from GJ131 and KWS2320; 2, 4, and 6: two fragments 160 and 46 bp were observed in GJ131, KWS2320, and non-bolting A906001 after *AluI* digestion; 5: two fragments 160 and 46 bp were visible in bolting parent 93167P. M: 1kb DNA size marker (Gibco-BRL).

### 3.5.3. SNAP and ARMS markers

For the following BAC ends GJ10SP6, GJ15SP6, GJ18SP6, GJ18T7, GJ21SP6, GJ29SP6, GJ44SP6, GJ44T7, GJ55SP6, GJ55T7, GJ70SP6, GJ70T7, GJ73SP6, and GJ73T7, no polymorphisms could be obtained with the techniques described above. However, these BAC-end sequences were important for genetic mapping because they are from BACs of the central contigs of the gene *B* region. The genetic mapping with these BACs facilitated the integration of the genetic and physical maps, and the determination of the correlation of genetic and physical distances in the gene *B* region on chromosome 2. Conventional gel electrophoresis is not sensitive enough to discriminate between single base pair mutations in the PCR products.

Therefore, an alternative strategy was chosen for mapping these BAC ends: detection of SNPs present between the two parental lines by sequencing and genotyping using allele specific primers and ARMS (amplification refractory mutation system, Ye *et al.*, 2001).

### 3.5.3.1. Principle of SNAP and ARMS

Different steps are required for SNP detection based on a simple PCR assay using amplification of specific alleles. They are listed as follows: 1) PCR products of the two parents were amplified with primers derived from BAC sequences and were sequenced. 2) The SNPs were detected by comparison of sequences of the two parents. Then, allele specific primers were designed, with their 3'-end terminating of the SNP position. 3) Allele specific fragments were amplified with PCR using allele specific primers. After that, the allele specific fragments were separated on 2% agarose gels. Figure 6 demonstrates the conversion of SNPs into SNAP or tetra-primer-ARMS markers.

In fact, two methods were used for converting SNPs into allele specific markers. The first method was SNAP (single nucleotide amplified polymorphisms), in which allele specific primers were designed. The primers were designed in such a way that the 3'-ends are located at the SNP position. For the amplification of each of the two alleles, one regular and one allele specific primer were used, which allowed only the amplification of one of the two alleles. The second allele was not amplified with this primer, since the primer extension at the second allele was not allowed due to the mismatch at the 3'-end.

The second method was ARMS (amplification refractory mutation system), which is based on the following principle: the allele specific primers were designed with their 3'-end located at the SNP. The deliberate introduction of an additional second mismatch at -2 bp from the 3'-termini. This permitted a better reproducible amplification by enhancing the selectivity and eliminating the false-positive extension of the double-mismatched primers. Two different allele specific (nested) primers were used at the same time; both allele specific amplicons were synthesized at the same time in one PCR reaction. The SNP position in the amplicon should be asymmetric left and right from the SNP base. In this way, the allele-specific PCR products, that are synthesized, are of different sizes.

### 3.5.3.2. SNP discovery within BAC sequences

Fourteen primer combinations (Table 12), derived from BAC-end sequences of nine different BACs, were used to amplify DNA from the following beets as template: (1)- Non-bolting parent (A906001), (2)- Bolting parent (93167P), (3)- KWS\_P1 non-bolting (ZR15\_77) (4)- KWS\_P2 bolting (ZR15\_76), and (5)- KWS\_F1 (ZR15\_78). The KWS lines were chosen in order to find more polymorphisms and to determine alleles characteristic for bolting and non-bolting genotypes. The DNA from KWS F1 was used as a control. After sequence comparison of amplicons from KWS lines and our parents (A906001 and 93167P), the common SNPs were converted into ARMS. In this way, the developed markers should be useful, not only to distinguish between bolting and non-bolting plants from the mapping population used in this study, but also in other populations with different genetic background.

The PCR products were directly sequenced after purification. SNPs or Indels were detected by sequence comparison using SeqMan<sup>TM</sup> from the DNASTAR software package (GATC Biotech, Konstanz, Germany). For nine of the 14 BAC-end sequences, SNPs could be detected (Table 12).

Nucleotide diversity was calculated as follows: The total number of SNPs per BAC sequence was divided by the length of the BAC sequence (bp) and then multiplied by 1000. For example, the nucleotide diversity on average  $\frac{2}{316} \times 1000 = 6.3$ ; this means that 6.3 SNPs occur within 1000 bp. The SNP nucleotide diversity between bolting and non-bolting genotypes varied between 2.2 and 30.4.

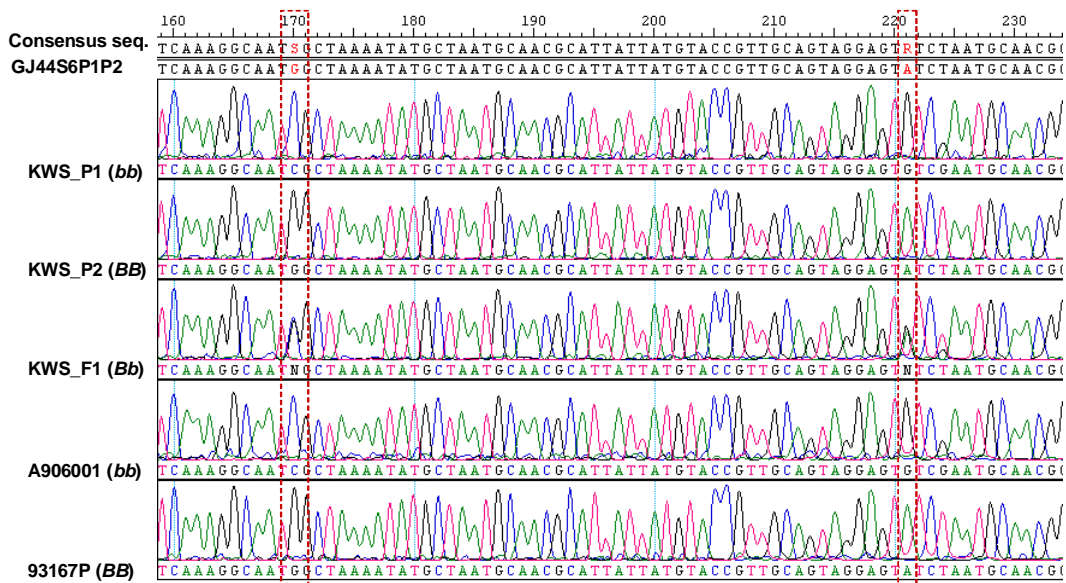
**Table 12:** BAC-derived SNP markers, number of SNPs per BAC end, SNP frequency of the BAC ends, and number of Indels detected.

BAC end	Amplicon size (bp)	Number of SNPs	No. of SNPs/1000 bp	Number of Indels (size range in bp)
GJ10SP6	316	2	6.3	-
GJ15SP6	428	13	30.4	3 (1-2)
GJ18SP6	415	8	19.3	0
GJ18T7	564	5	8.9	0
GJ21SP6	509	4	7.9	2 (1)
GJ29SP6	461	1	12.8	0
GJ44SP6*	467	6	2.2	0
GJ44T7	583	0	0	0
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	4	10.8	0
<b>Total</b>	<b>6284 bp</b>	<b>46</b>	<b>7.3</b>	<b>5</b>

\*: SNP present in all five genotypes.

A total of 24 SNPs (52.2%) were found to be transitions, while 22 (47.8%) were transversions. Transitions are purine-purine (A ↔ G) or pyrimidine-pyrimidine (C ↔ T) substitutions, whereas transversions are purine-pyrimidine or pyrimidine-purine (A ↔ C, A ↔ T, G ↔ C, G ↔ T) exchanges. Figure 15 shows the detection of SNPs after aligning the sequences from different lines using SeqMan™ from DNASTAR software package (GATC Biotech, Konstanz, Germany).

In addition, five Indels were detected. Three of these Indels were found in GJ15SP6 (one nucleotide deletion at positions 111 and 167 bp and 2 bp deletion at position 365-366 bp in non-bolting parent A906001). A one nucleotide deletion at position 325 in the bolting parent 93167P and a one bp deletion at position 469 in the non-bolting parent A906001 were found in GJ21SP6. Five BAC ends showed SNPs between the two KWS parents. The other five BAC ends did not show any SNP and were not further evaluated. Using all sequence information from sequencing the four beet parents together, 46 SNPs were found among 6,284 bp, which corresponds to an average SNP frequency of 7.3 SNPs/1000 bp (Table 12).



**Figure 15:** A part of multiple sequence alignments of PCR products amplified from different genotypes (parents) using GJ44SP6 derived primers. The forward primer (P1) was used as sequencing primer. Two SNPs are visible at positions 170 and 221 bp. Genotyping for the beet lines was as follows: KWS\_P1: CC and GG; KWS\_P2: GG and AA; KWS\_F1: C/G and G/A; bolting parent 93167P: GG and AA, and non-bolting parent A906001: CC and GG.

In addition to these BAC-end sequences, partial sequences from three BACs (GJ01, GJ70, and GJ131) from the central gene *B* region were also used for ARMS marker development. First, possible introns were determined in the following way: the standard genetic code was used and the ORF prediction was adjusted to predict ORFs with a minimum of 15 amino acids in length. The GeneQuest<sup>TM</sup> software (GATC Biotech, Konstanz, Germany) was used for ORF analysis. Second, from 12 ORFs, with more than 450 bp sequence length, primers were designed. According to the primer positions, 10 primer combinations should amplify exons, whereas the remaining two should amplify introns. For amplification of exons, the primers were placed at the flanking intron regions. For amplification of introns, the primers were placed within the flanking exon regions. The distance between primer binding sites and the intron-exon borders was approximately 80 bp. PCR products were separated on 1% agarose gels. According to gel electrophoresis, one primer combination (GJ70co3F6) completely failed to amplify the expected fragments, three primer combinations (GJ01co38F2, GJ131co8F2, and GJ131co8F3) failed to give amplification with the bolting parent and were not further investigated, while the other eight primer combinations gave fragments with the expected size (Table 13).

After multiple sequence alignment, five sequences showed single nucleotide polymorphisms between all genotypes (Table 13). According to genomic Southern analysis and BAC high-density filters, they were single or low copy sequences. For primer combination GJ131co15F6, a deletion of three nucleotides was found in the bolting parents (93167P and KWS P2 ZR15\_76) at position from 229 to 231 bp. Sixteen SNPs were found between all bolting and non-bolting parents in all five polymorphic ORFs. Except for GJ01co36F4, which was polymorphic only between 93167P and A906001, no SNP was found between the KWS genotypes. In summary, in sequences of all four parental lines, 16 SNPs were detected among 5,757 bp which corresponded to 1 SNP/359.8 bp (Table 13). Ten SNPs (62.5%)



were found to be transitions and six (37.5%) were transversions. The SNP nucleotide diversity based on potential ORF-derived sequences ranged from 0 to 6.3, which was lower as compared to sequences derived from BAC ends. Only one Indel was found for the GJ131co15F6 ORF sequence.

**Table 13:** BAC-ORF derived SNP markers, number of SNPs per BAC-ORF, SNP frequency of the BAC-ORFs, and number of Indels detected. SNP frequency was determined by sequencing amplicons from beet lines No. KWS\_P1 (ZR15\_77), KWS\_P2 (ZR15\_76), KWS\_F1 (ZR15\_78), 93167P (*BB*), and A906001 (*bb*) parents.

BAC ORF	Amplicon size (bp)	Number of SNPs	No. of SNPs/1000 bp	No. of Indels (size range in bp)
GJ01co15F2*	797	0	0	0
GJ01co15F3*	630	4	6.3	0
GJ01co36F4*	692	1	1.4	0
GJ70co3F4*	808	3	3.7	0
GJ70co9F1§	494	0	0	0
GJ70co9F3*	874	0	0	0
GJ131co15F2*	761	4	5.2	0
GJ131co15F6*	701	4	5.7	1 (3)
<b>Total</b>	<b>5757</b>	<b>16</b>	<b>2.8</b>	<b>1</b>

\*: Exon region was amplified by using intron primers.

§: Intron region was amplified by using exon primers.

### 3.5.3.3. SNAP and ARMS assays

For genotyping, all plants of the populations could have been sequenced at their respective SNP positions. However, sequencing of all plants is laborious and expensive compared to the PCR-based methods. Therefore, the SNPs were converted into PCR-based SNAP or ARMS markers. The SNAP method was successfully used to convert two SNPs from BAC-end derived sequences (GJ10SP6 and GJ70SP6) into allele specific markers. This type of marker is dominant, where only one allele can be detected in one PCR reaction.

#### Marker GJ10SP6:

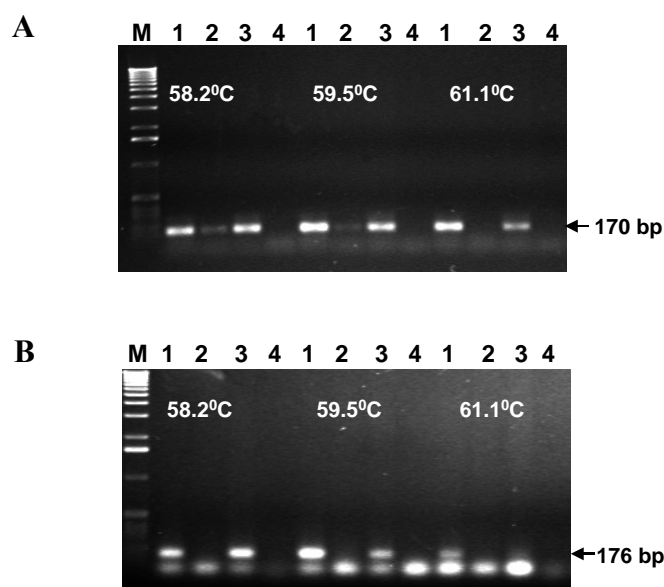
After comparison of sequenced PCR products of parental lines (93167P, *BB*; A906001, *bb*) two SNPs were detected. Allele specific primers for bolting and non-bolting alleles were designed (see 3.5.3.1.). One SNP (A in bolting 93167P and G in non-bolting A906001) at position 149 bp was converted into a dominant SNAP marker as described above. For amplification of specific alleles, two primers were used. With an allele specific primer (P6A for G allele) and a normal forward primer (P1) a 170 bp allele specific fragment was amplified in non-bolting parent A906001 and KWS2320 genotype. The annealing temperature at 61.1°C was found to be the optimal temperature for specific amplification of the 170 bp fragment (Figure 16A).

The GJ10SP6 SNAP marker was considered to be dominant. In order to detect both alleles in one PCR reaction, multiplexing of the allele specific primers of different alleles (bolting and non-bolting) was done. Unfortunately, this was not successful although the primers were expected to amplify differently sized fragments. The BAC GJ10 is a part of the central contig from the gene *B* region. Therefore, the marker GJ10SP6 was genotyped by sequencing of the F<sub>2</sub> plants.

### Marker GJ70SP6:

For GJ70SP6, three SNPs were detected between bolting and non-bolting genotypes. The SNP at position 222 bp was chosen to design the allele specific primers P10A and P10B with 3'-end of G (for bolting genotype 93167P) and T (in non-bolting A906001 genotype). The primer P10A and a normal primer P1 were used to amplify the allele specific fragment of 176 bp from non-bolting parent A906001 and KWS2320 genotype (Figure 16B). The gradient PCR (ranged from 58.2°C to 67.8°C) was used to optimize the annealing temperature. The non-bolting allele was amplified at 58.2°C and 59.5°C (Figure 16B). No amplicon was obtained in the bolting parent (93167P). This marker was dominantly scored.

The amplification of the bolting allele fragment with the following allele specific primers P3B, P4B, P5B, P6B, and P8B of GJ70SP6 failed. Consequently, a second mismatch was introduced to the allele specific primers (P10A and P10B) at -2 position from 3'-ends in order to increase the amplification specificity. Despite of adding the second mismatch, it was difficult to specifically amplify the bolting fragment with the annealing temperature optimized before.



**Figure 16:** BAC-derived SNAP markers GJ10SP6 (A) and GJ70SP6 (B) of the bolting gene *B* of sugar beet. Agarose gel electrophoresis of allele-specific PCR products amplified using allele specific primers. 5  $\mu$ l of the PCR reactions were loaded in each slot and separated on a 1% agarose gel for 60 min. at 60 V. **A)** GJ10SP6, the G allele specific fragment (170 bp) was amplified from non-bolting genotypes only (1: KWS2320; 3: A906001) while no amplification occurred in 2: bolting genotype (93167P). **B)** GJ70SP6, the G allele specific fragment (176 bp) was amplified from non-bolting genotypes only (1: KWS2320; 3: A906001) while no amplification occurred in 2: bolting genotype (93167P). 4: control without template DNA. M: 1kb DNA size marker (Gibco-BRL).

#### 3.5.3.4. ARMS markers

Due to the dominant nature of SNAP markers, two PCR reactions, one for bolting allele and the other for non-bolting allele, have to be done. Therefore, a second method (ARMS) was used for SNP genotyping. The co-dominant markers (e.g. ARMS) are useful for genotyping of sugar beet

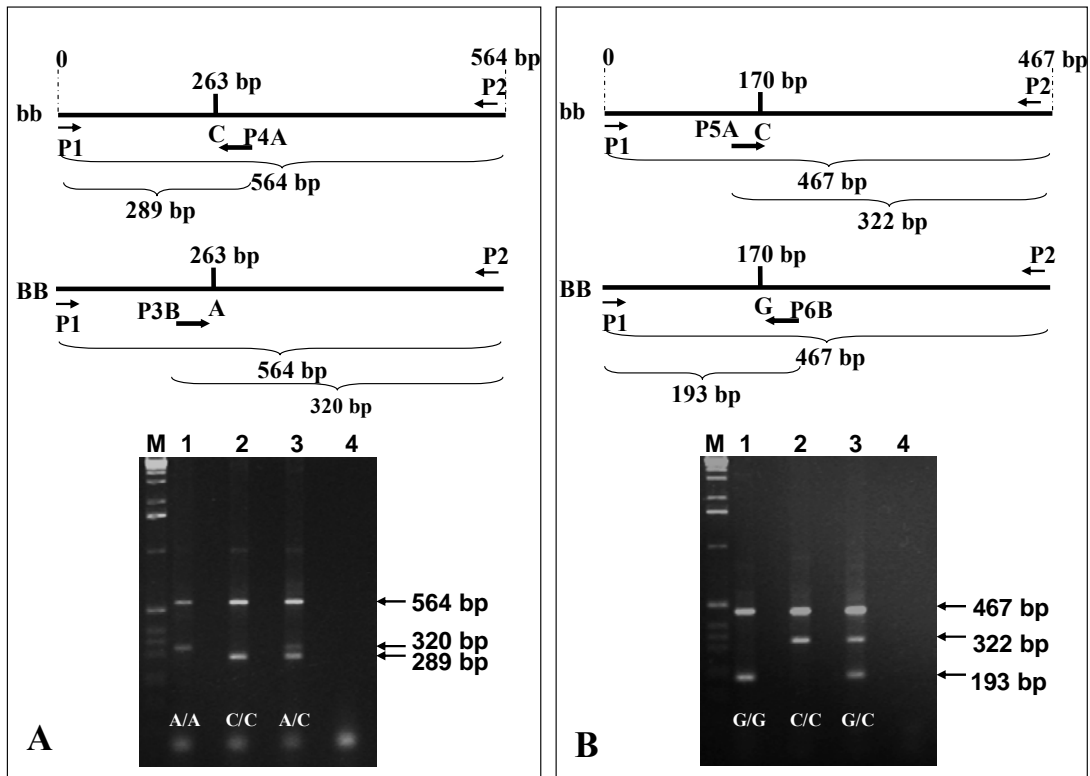
populations as it is an outbreeding species. The SNPs of five BAC-derived sequences were successfully converted into PCR-based ARMS markers using tetra-primer ARMS-PCR method (see 3.5.3.1.). Out of these five sequences, two were derived from BAC ends (GJ18T7 and GJ44SP6), and the other three were from ORF sequences (GJ01co36F4, GJ70co3F4 and, GJ131co15F6). The three single copy ORF sequences (GJ01co36F4, GJ70co3F4 and, GJ131co15F6) were selected from BACs GJ01, GJ70 and GJ131, respectively. These BACs were physically mapped to the two central contigs (CK2 and CK3) of the *B* gene region (see Figure 1, Introduction). The genetic distances of these BACs to the *B* gene were determined by genotyping 299 plants (960701, subpopulation 7), and 300 non-bolting plants (subpopulations 5, 6 (960701), and population 950619). The ARMS technique was chosen because it is easier and cheaper as compared to sequencing.

#### Marker GJ18T7:

As described above (see 3.5.3.1.), in an ARMS assay four primers were used in one PCR reaction (tetra-primer ARMS) in order to amplify two allele specific fragments. The PCR with outer primers (P1 and P2) and two inner (nested) allele specific primers (P4A and P3B) yielded three fragments. The allele specific primers bind specifically to both the C (P4A) and the A (P3B) at position 263 bp. 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, corresponding to *B* and *b* alleles. The 320 bp fragment was the product obtained with primer combination P3B and P2 and is linked to bolting allele (*B*). The 289 bp amplicon, obtained with primer pairs P1 and P4A and is linked to the non-bolting allele (Figure 17A). Thus, F<sub>2</sub> individuals could be clearly genotyped into homozygous and heterozygous individuals.

#### Marker GJ44SP:

In the tetra-primer ARMS marker assay of GJ44SP6 (Figure 17B), all genotypes produced a major band of 467 bp. The bolting homozygous plants (*BB*) had an additional band of the *B*-linked allele (193 bp). The non-bolting (*bb*) homozygous plants had an amplicon of the *b*-linked marker allele of different size (322 bp). In heterozygous plants (*Bb*), both allele specific bands (G allele, 193 bp and C allele, 322 bp) were detected. This marker was considered to be co-dominant and it differentiated between homozygous and heterozygous plants.



**Figure 17:** BAC-derived ARMS markers GJ18T7 (A) and GJ44SP6 (B) of the bolting gene *B* of sugar beet. Top, is a graphical representation of SNP position of the T7 and SP6 ends of BAC GJ18 and GJ44, respectively. Arrows indicate the positions of outer forward and reverse primers (P1 and P2), respectively as well as the inner allele specific primers. Vertical bars indicate the SNP positions. Bottom, is agarose gel electrophoresis of the PCR products amplified with the four primers. 8  $\mu$ l of the PCR reactions were separated on a 2% agarose gel for 90 min. at 90 V. **A)** GJ18T7 SNP (C $\rightarrow$ A polymorphism) **B)** GJ44SP6 SNP (C $\rightarrow$ G polymorphism). 1: bolting (*BB*) genotype that showed A/A or G/G, respectively. 2: non-bolting (*bb*) that showed C/C or C/C, respectively. 3: bolting heterozygous (*Bb*) that showed A/C or G/C, respectively. 4: control without template DNA. M: 1kb DNA size marker (Gibco-BRL). The optimal annealing temperatures were 60°C (A) and 62°C (B).

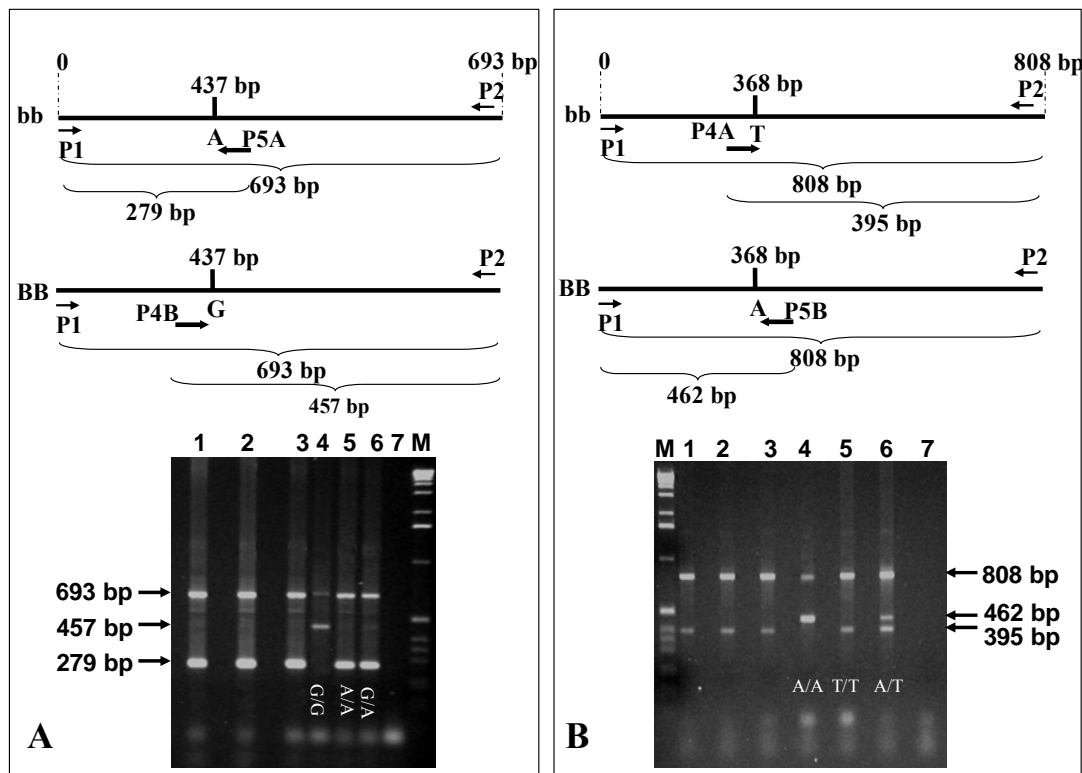
#### Marker GJ01co36F4:

The marker assay of the sequence GJ01co36F4 relied on four primers, two outer (P1 and P2) and two inner allele specific primers (P5A and P4B) placed at SNP position (437 bp). In one PCR reaction, three PCR fragments were generated (Figure 18A). One major fragment of 693 bp was present in all genotypes (bolting, *BB*; non-bolting, *bb*; and; heterozygous, *Bb* plants). The fragment of 457 bp (G allele), which is the *B*-linked marker allele, could be seen in the bolting parent (93167P). Specific fragment of 279 bp (A allele), linked to the non-bolting marker allele, was amplified with the A inner allele specific primer. Both fragments 457 and 279 bp were visible in the heterozygous plant (G/A).

#### Marker GJ70co3F4:

At base pair position 368 of the ORF-derived sequence GJ70co3F4, a SNP was found between the bolting parent (93167P) and non-bolting parent (A906001). Using the ARMS method (described in 3.5.3.1.), the allele specific primers amplified two allele specific fragments, 462 bp in the bolting (A/A) and 395 bp in the non-bolting (T/T) parents. These two fragments were visible in the

heterozygous (A/T) plant. The larger fragment of 808 bp that was amplified by primer combination P1P2, could be seen in all genotypes (Figure 18B).

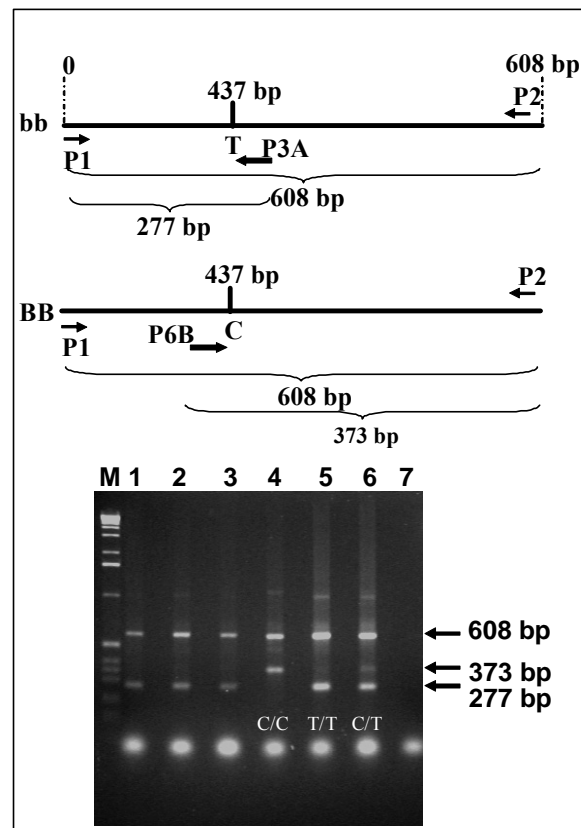


**Figure 18:** BAC-derived ARMS markers GJ01co36F4 (A) and GJ70co3F4 (B) of the bolting gene *B* of sugar beet. Top is a graphical representation of SNP position of the BAC-derived ORFs GJ01co36F4 and GJ70co3F4. Arrows indicate the positions of outer forward and reverse primers (P1 and P2), respectively as well as the inner allele specific primers. Vertical bars indicate the SNP positions. Bottom, is agarose gel electrophoresis of the PCR products amplified with the four primers. 8  $\mu$ l of the PCR reactions were separated on a 2% agarose gel for 90 min. at 90 V. **A)** GJ01co36F4 SNP (A $\rightarrow$ G polymorphism) **B)** GJ70co3F4 SNP (T $\rightarrow$ A polymorphism). 1: bolting (*BB*) genotype that showed G/G or A/A, respectively. 2: non-bolting (*bb*) that showed A/A or T/T, respectively. 3: bolting heterozygous (*Bb*) that showed G/A or A/T, respectively. 4: control without template DNA. M: 1kb DNA size marker (Gibco-BRL). The optimal annealing temperatures were 62°C and 57°C, respectively.

#### Marker GJ131co15F6:

For the ORF-derived sequence GJ131co15F6, the SNP that was detected between bolting and non-bolting parent at position 437 bp, was chosen for the establishment of a PCR-based assay for genotyping (Figure 19). In the bolting parent, a C was sequenced while in non-bolting a T was found. After the tetra-primer ARMS-PCR, three fragments were visible. The major fragment of 608 bp was found in the amplicons of all genotypes. In the bolting parent (93167P), the 373 bp fragment (C/C) of the *B*-linked marker allele was visible. The 277 bp fragment (T/T) was present in the non-bolting parent (A906001). These two allele specific fragments (373 and 277 bp) were clearly seen in the heterozygous plant (C/T).

In this way, inexpensive and simple PCR-based assays for 14 BAC ends and three ORF-derived markers were established. Allele-specific amplicons could be developed by the synthesis of allele-specific primers for 4 BAC ends (GJ10SP6, GJ70SP6, GJ18T7, and GJ44SP6) and three ORF sequences.



**Figure 19:** BAC-derived ARMS marker GJ131co15F6 of the bolting gene *B* of sugar beet. Top, is a graphical representation of SNP position of the BAC-derived ORF sequence GJ131co15F6. Arrows indicate the positions of outer forward and reverse primers (P1 and P2), respectively as well as the inner allele specific primers. Vertical bars indicate the SNP positions. Bottom, is agarose gel electrophoresis of the PCR products amplified with the four primers. 8  $\mu$ l of the PCR reactions were separated on a 2% agarose gel for 90 min. at 90 V. GJ131co15F6 SNP (C $\rightarrow$ T polymorphism). 1, 2, and 3: non-bolting F<sub>2</sub> plants; 4: bolting (*BB*) genotype that showed C/C; 5: non-bolting (*bb*) that showed T/T; 6: bolting heterozygous (*Bb*) that showed C/T; 7: control without DNA template. M: 1kb DNA size marker (Gibco-BRL). The optimal annealing temperature was 62°C.

Altogether PCR-based markers were developed from 14 BAC ends and three ORF-derived sequences (Table 14). Nine markers were scored co-dominantly, while for the remaining eight markers, only PCR products of the *b* linked marker allele were visible on the agarose gel. Thus the marker allele and *B* gene are coupled in repulsion. Although the presence/absence scoring of the dominant markers is more suitable for screening of the breeding materials, for fine mapping the use of co-dominant markers is more informative.

The co-dominant markers were very suitable for fine mapping of the *B* gene (see 3.6.4.). Therefore, the co-dominant markers were used to genotype the F<sub>2</sub> populations and the wild beet plants. The markers derived from sequences of GJ37T7 and GJ49SP6 were not further investigated for mapping because the two BACs, GJ37 and GJ49, could not be integrated into any of the contigs (see Figure 1).

**Table 14:** PCR-based markers developed from BAC-ends or ORF-derived sequences.

Marker type	Number of markers	Marker name
PASA marker	5	GJ01T7, GJ17T7, GJ22T7, GJ37T7, GJ131T7
CAPS marker	5	GJ10T7, GJ17SP6, GJ29T7, GJ49SP6, GJ131SP6
SNAP or ARMS	7	GJ10SP6, GJ70SP6 or GJ01co36F4, GJ18T7, GJ44SP6, GJ70co3F4, GJ131co15F6
Scoring type		
Dominant marker	8	GJ01T7, GJ10SP6, GJ10T7, GJ17T7, GJ22T7, GJ37T7, GJ70SP6, GJ131T7
Co-dominant marker	9	GJ01co36F4, GJ17SP6, GJ18T7, GJ29T7, GJ44SP6, GJ49SP6, GJ70co3F4, GJ131co15F6; GJ131SP6

### 3.5.4. Conversion of RFLPs into PCR markers

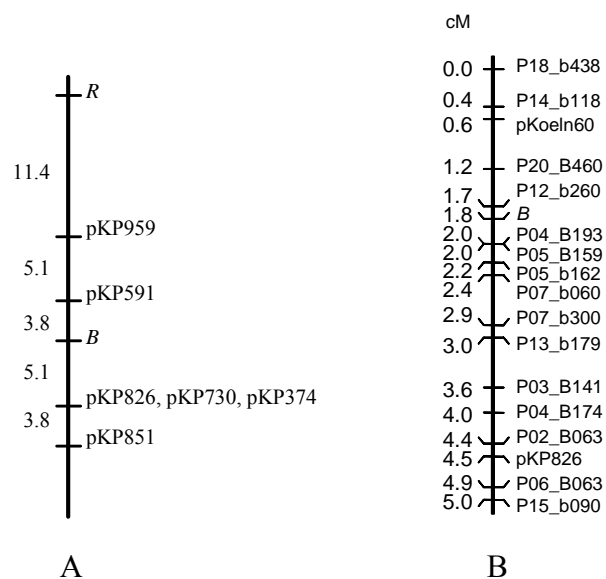
The following RFLP markers had been previously mapped to chromosome 2 of sugar beet in close distance to the *B* gene: pKP374, pKP591, pKP730, pKP826 and Koeln60 (Boudry *et al.*, 1994b; El-Mezawy *et al.*, 2002). Partial linkage maps of the gene *B* are shown in Figure 20.

The aim was to genotype markers pKP374 or pKP730, and pKP591 in the mapping subpopulation 7, in order to compare their location with the BAC-derived markers. The markers pKP374 and GJ17SP6 were used to screen the mapping population for recombinants within a 5 cM region around the *B* gene. Mapping of RFLP markers is very laborious and time consuming. Therefore, the RFLP markers were converted into PCR-based markers. As mentioned above, the PCR-based markers are easy and simple.

The RFLPs pKP374 and pKP730 were turned into PCR markers in the following way: the primers were designed based on DNA sequences of inserts used for RFLP mapping. After PCR amplification, the expected genomic fragments from bolting and non-bolting parents, the amplicons, were investigated for size polymorphism by agarose gel electrophoresis. If no polymorphism was seen, the PCR products were cleaved with different restriction endonucleases. The restriction endonucleases used, were selected after restriction site analysis using MapDraw<sup>TM</sup> from the DNASTAR software package (GATC Biotech, Konstanz, Germany). The restricted PCR products were resolved by agarose gel electrophoresis. In this way, CAPS markers were generated from the RFLP sequences. The two RFLP sequences pKP374 and pKP730 were converted into CAPS markers.

To get more polymorphisms between bolting and non-bolting genotypes, all 12 RFLP sequences were analyzed for SNPs. PCR products were amplified with 12 primer combinations (derived from the RFLP sequences) and DNA from bolting parent 93167P, non-bolting parent A906001, P2 bolting parent (KWS P2, ZR15\_76), P1 non-bolting parent (KWS P1, ZR15\_77) and KWS F1 (ZR15\_78) genotypes. The PCR products were sequenced and aligned. In total, 47 SNPs were detected within 6595 bp. The number of SNPs/amplicon ranged from 1 SNP (pKP730) to 20 SNPs that were detected in pKP814 (Table 15). 26 of the SNPs (54.2%) were classified as transitions and 22 (46.8%) were transversions. A high number (6) of Indels was detected in pKP814. The nucleotide diversity ranged from 0 to 39.4 with an average of 7.28.

SNPs derived from the three RFLP sequences Koeln 60, pKP591, and pKP814 were used as molecular markers for genotyping the recombinant plants by sequencing (subpopulation 7).



**Figure 20:** Partial linkage maps of Chromosome 2 covering the gene *B* region. **A)** Linkage map of the gene *B* based on a segregating backcross progeny (Boudry *et al.*, 1994b). **B)** Linkage map of the *B* gene based on 775 plants of F<sub>2</sub> population (El-Mezawy *et al.*, 2002).

**Table 15:** RFLP-derived SNPs, number of SNPs per RFLP sequence, SNP frequency of the RFLP sequence, and number of Indels detected.

RFLP	Amplicon size (bp)	SNP position used for genotyping	SNP marker name	Number of SNPs	No. of SNPs/1000 bp	Number of Indels (size range in bp)
Koeln60	639	1019	Koeln60-1019	2	3.12	0
pKP374	542	-	-	14	25.8	0
pKP490	500	-	-	2	4.0	1
pKP591	507	100	pKP591-100	2	3.94	1 (2)
pKP730	452	-	-	1	2.21	0
pKP814	507	427	pKP814-427	20	39.44	6 (4, 7, 2, 1, 2, 8)
pKP826	1130	0	-	0	0	0
pKP884	440	-	-	5	11.36	0
pKP907	410	0	-	0	0	0
pKP943	388	0	-	0	0	0
pKP967	422	0	-	0	0	0
pKP1159	658	-	-	2	3.04	0
<b>Total</b>	<b>6595</b>	-	-	<b>48</b>	<b>7.28</b>	<b>8</b>

-: not determined.



### 3.5.4.1. Features of the RFLP-derived PCR-based markers

#### Marker pKP374:

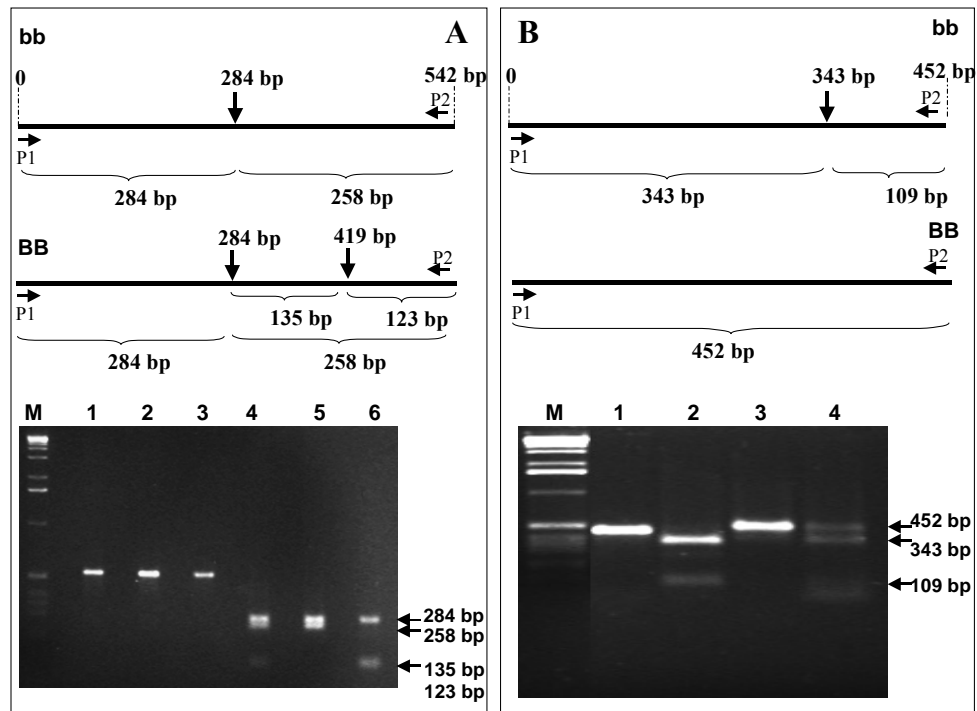
With the primer pair derived from pKP374, a 542 bp PCR-fragment was amplified from bolting (93167P) and non-bolting (A906001) parents. After *Sau3AI* (*MboI*) enzyme digestion at 37°C and resolving on a 2% agarose gel, one major fragment 284 bp was found in both bolting and non-bolting parents (Figure 21). Two additional polymorphic fragments (135 and 123 bp) of the *B*-linked marker allele were observed. Only an additional fragment of 258 bp could be detected in non-bolting parent. All of the four fragments were visible in the heterozygous genotype. Marker pKP374 was a co-dominant marker, which can distinguish between the homozygous and heterozygous plants. The CAPS-marker pKP374 was mapped in the subpopulation 4 (1359 individuals).

#### Marker pKP730:

The 452 bp PCR fragment was amplified from (bolting 93167P & non-bolting A906001 parents), and KWS2320 genotype with primer combination of pKP730. The PCR product was digested with *Sau3AI* at 37°C for three hours. After gel electrophoresis, two fragments (343 bp and 109 bp) were visible in KWS2320 and A906001 parent (Figure 21B). However, the 452 bp fragment was still visible in A906001 parent after restriction digest. This means that there is a copy of this fragment in the genome, which does not have the *Sau3AI* restriction site. Although this marker is co-dominant, it was not used for genotyping of F<sub>2</sub> plants. The pKP374 and pKP730 markers were located at the same map position according to the *B* map from Boudry *et al.* (1994b).

A complete list of the molecular markers, derived from BAC ends, BAC ORFs, and RFLP sequences, that have been developed in this study as well as the number of F<sub>2</sub> plants from mapping populations can be found in appendix 8.2.

The sequences of BAC ends, RFLP sequences used for PASA & CAPS marker development as well as the sequences of bolting and non-bolting alleles (amplified with BAC-end, ORF-, RFLP sequences) are shown in appendix 8.3.



**Figure 21:** RFLP-derived CAPS markers pKP374 (A) and pKP730 (B) of the bolting gene *B* of sugar beet. Bottom, is agarose gel electrophoresis of restricted PCR products amplified with primer combinations from pKP374 and pKP730 sequences. 5  $\mu$ l of the PCR reactions were digested with *Sau3AI* in a 10  $\mu$ l total volume and separated on a 1.5-2% agarose gel for 60 min. at 60 V. **A)** pKP374; 1, 2 and 3: undigested PCR products of heterozygous plant, non-bolting and bolting parents, respectively; 4-6: *Sau3AI* digested PCR products of heterozygous plant, non-bolting and bolting parents, respectively; 5: two fragments 284 and 258 bp were visible in non-bolting parent A906001; 6: *Sau3AI* restriction at 419 bp resulted in two different fragments 135 and 123 bp in bolting parent 93167P; 4: heterozygous plant. M: 1kb DNA size marker (Gibco-BRL). Top, is a graphical representation of *Sau3AI* restriction sites within the 542 bp fragment amplified from the sequence of RFLP pKP374. Arrows indicate primer positions. P1 and P2 are forward and reverse primers, respectively. Vertical arrows indicate the positions of *Sau3AI* restriction sites. **B)** pKP730; 1: undigested PCR product of KWS2320; 2-4: *Sau3AI* digested PCR products of non-bolting genotypes (KWS2320 and A906001), and bolting parent, respectively; 2: two fragments 343 and 109 bp were observed in KWS2320 and non-bolting A906001; 3: intact PCR product, no *Sau3AI* restriction site in bolting parent 93167P; 4: heterozygous plant. Top, is a graphical representation of *Sau3AI* restriction sites within the 452 bp fragment amplified from the sequence of RFLP pKP730. Arrows indicate primer positions. P1 and P2 are forward and reverse primers, respectively. Vertical arrows indicate the positions of *Sau3AI* restriction sites.

### 3.6. Linkage analysis

#### 3.6.1. Marker genotyping in populations segregating for early bolting

For genotyping, the following BAC-derived markers were selected: GJ01co36F4, GJ10SP6, GJ17SP6, GJ18T7, GJ29T7, GJ44SP6, GJ70co3F4, and GJ131co15F6. These eight BAC-derived markers were

selected in such a way to represent the four contigs of the *B* gene region and to anchor the physical map (see Figure 1). Two additional, pKP374 and Y102H12R (Hohmann *et al.*, 2003), markers were included. The pKP374 was used as a flanking marker. The YAC-derived Y102H12R marker was selected because it showed very low recombination to gene *B* (Hohmann *et al.*, 2003). It was also used as a probe for screening the BAC library and selecting BACs from the gene *B* region. The selected BACs were from contig CK2 (see Figure 1). The following plants/populations were selected for genotyping: 299 F<sub>2</sub> plants with verified phenotype in F<sub>3</sub> (subpopulation 7) (see 3.1.2.) and 300 F<sub>2</sub> non-bolting plants (subpopulations 5 & 6, and population 950619). They were used later for 1) determination of the recombination rates between the markers and *B* locus and 2) construction of the linkage map around the *B* gene. The genotypic data from the 300 F<sub>2</sub> non-bolting plants were used for the two-point linkage analysis (Hühn, 1995; see 3.6.3.) and fine mapping (3.6.4.).

To demonstrate that the marker loci (co-dominant) are segregating according to 1:2:1 segregation ratio, the  $\chi^2$  value for each marker locus were calculated. The observed phenotypic frequency of each marker locus is shown in Table 16. The 14 markers used for calculation were co-dominantly scored. The  $\chi^2$  test was performed with degrees of freedom  $df=2$  and at level of significance  $\alpha=0.05$ . Two markers (GJ01co36F6 and GJ17SP6) had the lowest  $\chi^2$  value of 0.8. The four AFLP markers p04\_B174, p04\_B193, p05\_B159 and p05\_b162 had the highest  $\chi^2$  values 19.2, 15.6, 9.4, and 10.1, respectively.  $\chi^2$  values of the AFLP markers showed significant deviation from the expected segregation ratio of 1:2:1, thus indicating the presence of segregation distortion. The calculated  $\chi^2$  values of the BAC-derived markers, pKP374, and Y102H12R markers were not significant, this means that these markers were segregating like Mendelian loci (Table 16).

**Table 16:** Chi-square analysis of the marker loci used for map construction around the *B* gene. A: homozygous non-bolting, B: homozygous bolting, and H: heterozygous

Marker	Scoring type	Population genotyped with this marker	Number of plants genotyped	Observed no. of plants in three phenotypic classes			$\chi^2$
				A	B	H	
GJ01co36F4	co-dominant	subpop 7	255	63	141	73	0.8 ns
GJ10SP6	co-dominant	subpop 7	269	62	154	68	2.3 ns
GJ131co15F6	co-dominant	subpop 7	245	62	144	66	1.1 ns
GJ17SP6	co-dominant	subpop 7	299	60	159	80	3.9 ns
GJ18T7	co-dominant	subpop 7	281	63	152	75	1.7 ns
GJ29T7	co-dominant	subpop 7	289	65	154	75	1.4 ns
GJ44SP6	co-dominant	subpop 7	293	64	157	75	1.9 ns
GJ70co3F4	co-dominant	subpop 7	243	62	132	77	1.8 ns
p04_B174	co-dominant	subpop 7	299	60	187	52	19.2*
p04_B193	co-dominant	subpop 7	299	66	182	51	15.6*
p05_B159	co-dominant	subpop 7	299	62	176	61	9.4*
p05_b162	co-dominant	subpop 7	299	62	177	60	10.1*
pKP374	co-dominant	subpop 7	297	68	164	66	3.1 ns
Y102H12R	co-dominant	subpop 7	293	68	155	73	0.8 ns

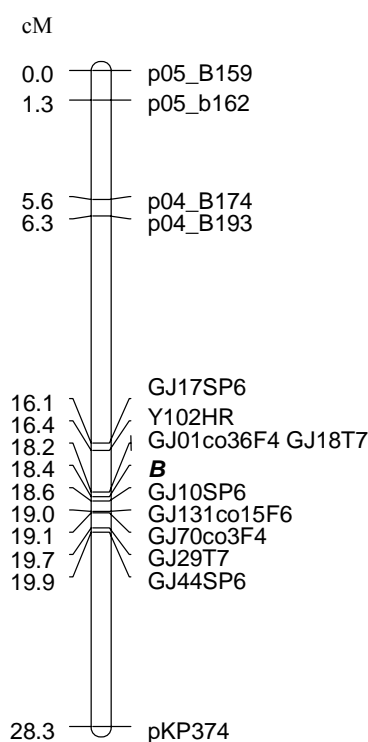
ns:  $\chi^2$  value is not significant.

\*:  $\chi^2$  value is significant,  $\alpha=0.05$ ,  $df=2$ .

### 3.6.2. Construction of a linkage map around the *B* gene

For map construction, the genetic information of 14 marker loci was used. Out of these loci, eight were BAC-derived markers, one RFLP- and one YAC-derived marker. In addition, genetic data of four AFLP markers, previously mapped to the *B* locus by El-Mezawy *et al.* (2002), were integrated. Based on the marker data of the 299  $F_2$  plants (subpopulation 7, with verified phenotype in  $F_3$ ) and using MAPMAKER/EXP 3.0b (LOD= 10 and REC= 0.30), a genetic map around the *B* gene was constructed (Figure 22). LOD score of 10 and REC-value of 0.30 were used to construct the map, in order to get high stringent conditions.

The mapping procedure is summarized as follows: first the raw data was stored in an Excel sheet. Then the file containing the marker data was converted into an ASCII file with extension “raw”. The framework map was constructed as follows: a minimum LOD score of 10.0 for statistical acceptance of linkage and a maximum recombination fraction of 0.30 were chosen for grouping the markers; the marker order within the linkage group was optimized using “order” and “ripple” commands to compare all three-loci permutations of the framework order. Finally, the linkage map was constructed (Figure 22).



**Figure 22:** Genetic map (28.3 cM interval) around the *B* locus on chromosome 2 of sugar beet, including eight BAC-derived markers, four AFLP markers, one YAC- and one RFLP-derived marker, and the bolting gene *B*. The map is based on 299 plants (verified phenotype in  $F_3$ ) of  $F_2$  subpopulation 7. cM: centiMorgan.

The genetic map spanned 28.3 cM with the flanking markers, p05\_B159 and pKP374 of *B* locus. Two BAC-derived markers (GJ01co36F6 and GJ10SP6) were mapped tightly linked to the *B* locus with a map distance of 0.2 cM. These two closest BAC-derived markers (GJ10SP6, BAC-end derived marker and GJ01co36F6, ORF-derived marker) to the *B* locus were derived from BACs from two non-overlapping contigs (see Figure 1). Construction of this linkage map has assisted in the integration of

the eight BAC-derived markers, YAC- and pKP-derived markers, and the four AFLP markers (previously mapped) into one genetic map.

### 3.6.3. Two-point linkage analysis of non-bolting plants

Hühn (1995) has presented formulas (2.2.8.2.) for calculating recombination value (R) between a marker and a gene using only one of the three genotypic classes of a F<sub>2</sub> population. This selection reduced the number of plants to be analyzed for fine mapping. Here, the non-bolting plants of the mapping population 960701 (subpopulations 5 and 6) and population 950619 were chosen. They should all have the genotype *bb*. As described in 3.1.1., the non-bolting plants belong to the most reliable class of plants. Altogether, 300 F<sub>2</sub> non-bolting plants were analyzed. Seven BAC-derived markers and one RFLP-derived marker were used to genotype these 300 non-bolting (*bb*) F<sub>2</sub> plants. After calculation of recombination values, they were converted into map distances using Kosambi's equation (2.2.8.2.). The number of individuals to be analyzed in this linkage analysis was significantly reduced. A small number of individuals was used for calculating the recombination values (R) between the loci and the *B* locus. Therefore, variances were calculated to prove the confidence of the calculated recombination values (Table 17).

In total, eight markers were employed for calculating the recombination values between the markers and the *B* locus. They ranged between 0.0033 and 0.0201 (Table 17). Marker GJ01co36F4 showed the lowest recombination value to the *B* locus. Two markers (GJ70co3F4 and GJ131co15F6) showed recombination values of 0.0084 to the *B* locus. Four BAC-end derived markers GJ17SP6, GJ18T7, GJ29T7, and GJ44SP6 have recombination values of 0.0050 to the *B* locus. As expected, the RFLP-derived pKP374 marker showed the highest recombination value with R= 0.0201 (Table 17).

**Table 17:** Recombination values between markers and bolting gene *B* based on analysis of non-bolting plants of subpopulations 5 and 6 (960701) and of population 950619.

Marker	Recombination value (R)	Standard deviation	Distance (cM)	LOD SCORE
GJ01co36F4	0.0033	0.00236	0.17	142.1*
GJ18T7	0.0050	0.00289	0.25	139.8*
GJ17SP6	0.0050	0.00289	0.25	139.8*
GJ29T7	0.0050	0.00289	0.25	139.8*
GJ44SP6	0.0050	0.00289	0.25	139.8*
GJ70co3F4	0.0084	0.00372	0.42	135.4*
GJ131co15F6	0.0084	0.00372	0.42	135.4*
pKP374	0.0201	0.00573	1.02	122.5*

\*: Marker is significantly linked to *B* locus.

The significance of the genetic linkage between marker loci and *B* locus was determined with the LOD score significance test. The LOD values for all of the markers were >3, which indicated significant linkage between the marker loci and the *B* locus. LOD for each marker locus was calculated according to Hühn (1995), where the LOD is defined as:  $LOD\ SCORE = \log \left( \frac{L(R)}{L(0.50)} \right)$  where L denotes likelihood function, R= recombination value and log= logarithm to the base 10.

Table 18 shows the recombination values and distances between eight BAC-derived markers, one RFLP-derived as well as one YAC-derived marker and the *B* locus. The recombination values were calculated based on two groups of F<sub>2</sub> plants. First, the recombination value for each marker was calculated according to Hühn (1995) based on the 300 non-bolting plants from subpopulations 5 and 6

(960701), and population 950619. The estimation of R-values was done as described above. Second, the R-values were calculated by MAPMAKER/EXP 3.0b after performing multi-point analysis, using genotypic data obtained by genotyping the 299 F<sub>2</sub> plants of subpopulation 7 with the same set of markers. To compare the R-values, being calculated with the two different methods, they were converted manually into map distances using Kosambi's equation (Kosambi, 1944).

**Table 18:** Comparison of recombination values and distances between marker loci and *B* locus, which were calculated using two different methods. Hühn's (1995) method was used to calculate the recombination of markers based on 300 non-bolting plants from subpopulations 5 and 6 (960701) and population 950619. The MAPMAKER/EXP 3.0b calculation was based on 299 F<sub>2</sub> plants of subpopulation 7 (960701).

Marker	Recombination value (Hühn 1995)	Recombination value (MAPMAKER)	cM (Hühn 1995)	cM (MAPMAKER)
GJ01co36	0.0033	0.0010	0.17	0.01
GJ18T7	0.0050	0.0010	0.25	0.01
GJ10SP6	-	0.0040	-	0.20
GJ131co15F4	0.0084	0.0050	0.42	0.25
GJ70co3F4	0.0084	0.0050	0.42	0.25
GJ44SP6	0.0050	0.0110	0.25	0.61
GJ17SP6	0.0050	0.0220	0.25	1.12
GJ29T7	0.0050	0.0250	0.25	1.28
Y102H12R	-	0.0280	-	1.44
pKP374	0.0201	0.0590	1.02	3.16

-: not determined.

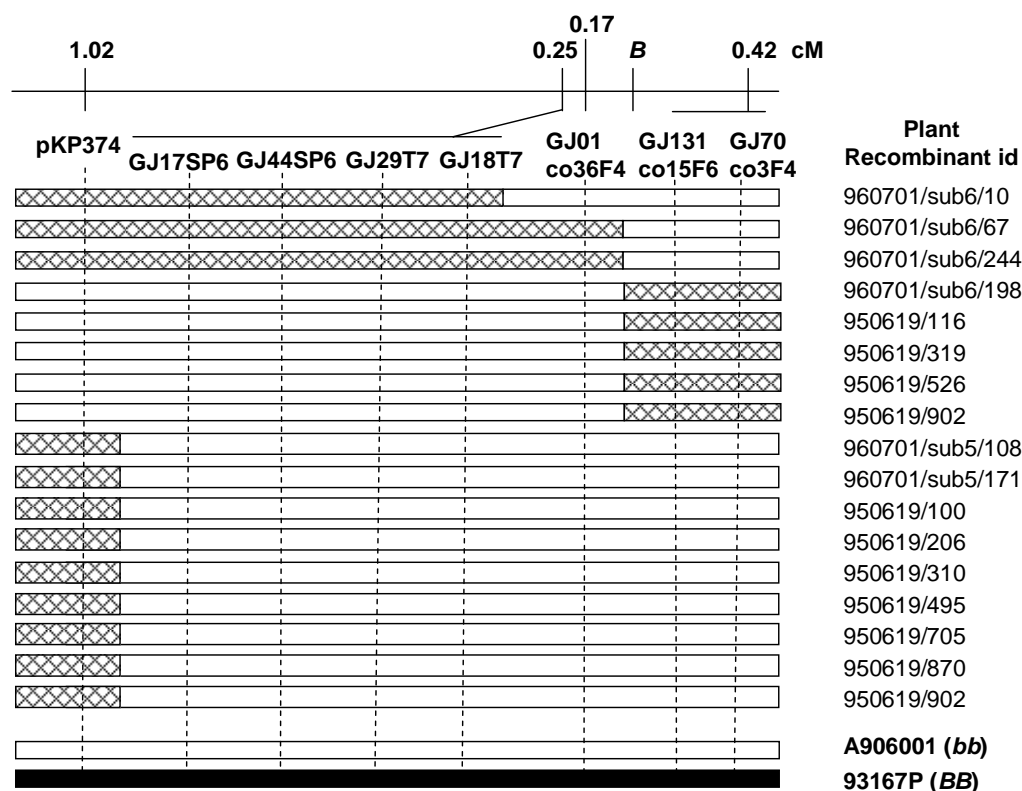
In general, the recombination values calculated using MAPMAKER/EXP 3.0b were slightly lower than those obtained using Hühn's method (1995). However, the R-values (MAPMAKER based) for markers GJ17SP6, GJ29T7, GJ44SP6, and pKP374 were higher than those calculated using Hühn's formula (1995). Interestingly, the markers GJ70co3F4 and GJ131co15F4 showed higher recombination value (0.0084) using Hühn's method (1995) based on one class of plants (non-bolting) than MAPMAKER/EXP 3.0b based on the three plant classes of the subpopulation 7 (960701). Here, one has to consider that R values were calculated based on two different populations and with two different methods. For example based on the two different methods (Hühn's method, 1995 & MAPMAKER/EXP 3.0b), marker pKP374 showed a distance of 1.02 cM and 3.16 cM to the *B* locus, respectively (Table 18). There was a considerable difference between the distance of pKP374 to the *B* locus calculated manually by Kosambi's equation (Kosambi, 1944), based on R value calculated with MAPMAKER/EXP 3.0b, and its distance calculated directly using MAPMAKER/EXP 3.0b.

### 3.6.4. Recombination mapping with graphical genotypes

F<sub>2</sub> plants were defined as recombinants if they showed new allele combinations between linked loci, which were not found in the parents. Seventeen new combinations (recombinants) F<sub>2</sub> plants between the *B* and marker loci, which were not found in the parents, have been identified in the following way: the plants used for this analysis were non-bolting plants and the marker were co-dominantly scored. The plants with the bolting-linked marker band (allele) only or both bands (alleles) (bolting-linked marker band and non-bolting-linked marker band) were considered to be recombinants. Eight markers (seven BAC-derived markers and one RFLP-derived marker) were used for genotyping the 300 non-bolting plants. The most informative individuals (recombinants) were the ones that were homozygous for the marker allele linked in coupling to the recessive allele at the *B* locus of one of the flanking markers and were heterozygous for the other. Using these recombinants, the markers were ordered as

shown in Figure 23. The recombinant plants were arranged in such a way that no double recombination between any two markers was found. From the two-point analysis of the 300 F<sub>2</sub> plants, the absolute genetic distance between each marker and the *B* locus has been estimated (Table 17). Based on these recombinants and the genetic distances between the marker loci and the *B* gene, the linear order of markers is as shown in Figure 23.

The marker GJ10SP6 was used to genotype only the 17 recombinant F<sub>2</sub> plants. Therefore, the distance between GJ10SP6 and *B* locus could not be calculated. However, based on the recombinants arrangement (Figure 23), the marker GJ10SP6 should be placed next to the GJ131co15F6 and GJ70co3F4.



**Figure 23:** Graphical genotypes of 17 non-bolting (*bb*) recombinants between marker loci or marker loci and *B* locus identified from non-bolting F<sub>2</sub> plants of subpopulations 5 and 6 (960701) and population 950619. The black and white boxes indicate homozygous regions for bolting (93167P) and non-bolting alleles (A906001), respectively. The patterned boxes indicate that the marker was heterozygous and crossovers have taken place.

### 3.7. Genotyping of wild and cultivated beets

Recently, natural genetic variation has been exploited for mapping of different traits. In order to test the feasibility of the BAC-derived markers for association mapping and identifying the *B* allele in the wild beet (annual and biennial) accessions, 41 beet accessions (*B. vulgaris* ssp. *vulgaris* and *B. vulgaris* ssp. *maritima*) were genotyped with seven BAC-derived markers (GJ17SP6, GJ18T7, GJ29T7, GJ44SP6, GJ01co36F4, GJ70co3F4, and GJ131co15F6) as well as one RFLP-derived marker (pKP374) mapping in close vicinity of the *B* gene (Figure 22). In addition, these markers were used to

distinguish between the annual and biennial *B. vulgaris* ssp. *maritima* accessions and to determine the haplotype structure around the gene *B* region.

The plant materials were selected according to the following criteria: first, the accessions included different types with respect to their bolting behavior (annual and biennial) as well as cultivated varieties. Second, they were collected from different geographic locations (Table 19). All sugar beet varieties 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 allele constitution *bb*, while annual wild beets are most likely *BB* or *Bb*. The bolting behavior of the wild beets was tested in the greenhouse. The plants which were bolting after 15 weeks considered as annuals, while the other plants, that did not bolt, were considered to be biennials. From 24 *B. vulgaris* ssp. *maritima* accessions, only 13 accessions were annuals and 11 were biennials. All seventeen accessions of cultivated sugar beet varieties were not bolting. Seven BAC-derived markers and one RFLP-derived marker were used to screen those accessions in order to test for association between the marker loci and the *B* gene.

Seventeen biennial *B. vulgaris* ssp. *vulgaris* accessions originated from German and American breeding companies were analyzed (Table 19). Surprisingly, only two alleles were detected for each marker as they have been found in the bolting and non-bolting parents. Two marker loci (GJ29T7 and GJ70co3F4) were homogeneous across cultivated accessions. However, the rest of loci showed a high degree of heterozygosity among the cultivated accessions. The markers (GJ70co3F4 and GJ29T7) are particularly suitable for the detection of the non-bolting alleles in breeding materials because they were homozygous among the cultivated accessions.

Twenty-four different types (annual and biennial) of *B. vulgaris* ssp. *maritima* accessions were selected from different locations (Table 19). Ten out of 11 biennial wild beet accessions were homozygous at marker locus GJ131co15F6. All annual wild beets were homozygous at marker locus GJ44SP6. Marker GJ131co15F6 detected the annual-linked marker allele in nine out of thirteen of annual *B. vulgaris* ssp. *maritima* accessions. Two other markers (GJ17SP6 and GJ18T7) failed to detect the annual *B* allele in one or two accessions, respectively. In summary, annuality could not be assessed by genotyping with a single marker. Notably, all biennial wild beets were homozygous at marker locus GJ44SP6.

The haplotype is defined as a set of closely linked alleles of different genes which are inherited together. This also refers to allele frequencies in the gamete. When analyzing the haplostructure of the 17 *B. vulgaris* ssp. *vulgaris* accessions with 7 BAC-derived markers and one RFLP-derived marker, 10 different haplotypes were observed. In the 24 different *B. vulgaris* ssp. *maritima* accessions, 23 different haplotypes were found. No characteristic annual or biennial haplotype could be detected due to the fact that high number of recombinations occurs quite often within the gene *B* region.

The “annual” allele, however, was also present in 15 out of the 17 sugar beet varieties, which were either homozygous or heterozygous. In contrast, all sugar beet varieties were homozygous at two loci, GJ70co3F4 and GJ29T7. This reflected a high degree of linkage disequilibrium as expected for the *B* allele against which breeders have been selecting with high intensity. One of these markers, GJ70co3F4 had been mapped in a distance of 0.7 cM to the *B* gene (see Figure 22). All biennial wild beet accessions were also homozygous at these marker loci with four exceptions, which were heterozygous at marker locus GJ29T7. The fixed bolting and non-bolting alleles in the cultivated gene pool, indicated complete linkage disequilibrium (LD) in all sugar beet varieties between loci (GJ70co3F4 and GJ29T7 and the *B* gene).



**Table 19:** Genotyping of 41 accessions of the section I of the genus *Beta* including wild accessions and sugar beet varieties. Seven BAC-derived markers and one RFLP-derived marker were used for genotyping.

Accession number	Origin	Type	Markers							
			GJ17SP6	GJ18T7	GJ01co36F4	GJ131co15F6	GJ70co3F4	GJ29T7	GJ44SP6	pKP374
<b>Parents</b>										
93167P	Germany	annual	381 bp	A	G	C	A	89 bp	G	135 bp
A906001	Germany	biennial	316 bp	C	A	T	T	166 bp	C	258 bp
<b><i>B. vulgaris</i> ssp. <i>maritima</i></b>										
930034	Spain	annual	381 bp	A	G/A	T	T	89 & 166 bp	G	258 bp
930043	Belgium	annual	381 & 316 bp	A/C	A	C/T	T	166 bp	G	135 bp
991971	Greece	annual	381 & 316 bp	C	A	T	T	166 bp	G	135 bp
960066	Israel	annual	381 & 316 bp	C	A	T	T	89 & 166 bp	G	135 bp
991968	Italy	annual	381 & 316 bp	A	G/A	C/T	T	166 bp	G	135 bp
950407	Italy	annual	381 & 316 bp	A/C	G/A	C/T	A	166 bp	G	135 & 258 bp
991969	Italy	annual	381 & 316 bp	A/C	A	C/T	T	166 bp	G	135 bp
930029	Spain	annual	381 & 316 bp	A/C	G/A	C/T	A/T	166 bp	G	135 & 258 bp
930031	Portugal	annual	316 bp	A	G	T	T	166 bp	G	258 bp
950409	Spain	annual	381 & 316 bp	A/C	G/A	C/T	T	166 bp	G	258 bp
960059	Tunisia	annual	381 & 316 bp	A/C	G/A	C/T	T	89 & 166 bp	G	258 bp
960078	Tunisia	annual	381 bp	A	G/A	C	T	89 & 166 bp	G	258 bp
991972	Tunisia	annual	381 & 316 bp	A/C	A	C/T	T	166 bp	G	258 bp
930028	Denmark	biennial	316 bp	A/C	G/A	T	T	89 & 166 bp	G/C	135 & 258 bp
991957	England	biennial	381 & 316 bp	A	-	T	T	89 bp	G	135 bp
991947	France	biennial	381 bp	A	A	T	T	89 bp	G	135 & 258 bp
991948	France	biennial	381 & 316 bp	A/C	A	T	T	89 & 166 bp	G/C	258 bp
991954	France	biennial	381 bp	A/C	G/A	T	T	89 & 166 bp	G/C	135 bp
991958	France	biennial	381 bp	A/C	G/A	T	T	166 bp	G	135 bp
930038	Greece	biennial	381 & 316 bp	C	A	T	T	166 bp	C	258 bp
960055	Greece	biennial	381 bp	A/C	A	C	T	166 bp	G	258 bp
991943	Italy	biennial	381 bp	A	G	T	T	166 bp	G	135 bp
930033	Italy	biennial	381 bp	A	G	T	T	166 bp	C	135 bp
960037	Sweden	biennial	381 & 316 bp	A	A	T	T	89 & 166 bp	G	135 bp
<b><i>B. vulgaris</i> ssp. <i>vulgaris</i></b>										
940026	Germany	biennial	316 bp	C	A	C/T	T	166 bp	G/C	135 & 258 bp

Table 19: continued.

940027	Germany	biennial	316 bp	C	A	T	T	166 bp	G	258 bp
940028	Germany	biennial	381& 316 bp	A/C	G/A	C/T	T	166 bp	C	135 & 258 bp
970102	Germany	biennial	316 bp	C	A	T	T	166 bp	G/C	258 bp
970103	Germany	biennial	316 bp	C	A	C/T	T	166 bp	G/C	258 bp
970104	Germany	biennial	316 bp	C	A	T	T	166 bp	G/C	258 bp
980005	Germany	biennial	316 bp	C	A	T	T	166 bp	G	258 bp
980008	Germany	biennial	381 bp	A	G	C/T	T	166 bp	G/C	135 & 258 bp
010238	Germany	biennial	381& 316 bp	A/C	G/A	T	T	166 bp	G	135 & 258 bp
010240	Germany	biennial	316 bp	C	A	T	T	166 bp	G	135 bp
010243	Germany	biennial	316 bp	C	A	T	T	166 bp	G	258 bp
010239	Germany	biennial	381& 316 bp	A/C	G/A	C/T	T	166 bp	G/C	135 bp
010242	Germany	biennial	316 bp	C	A	T	T	166 bp	G/C	258 bp
010241	Germany	biennial	381& 316 bp	A/C	G/A	T	T	166 bp	G/C	258 bp
930179	USA	biennial	316 bp	C	A	T	T	166 bp	C	258 bp
930182	USA	biennial	316 bp	-	-	T	T	166 bp	G	135 bp
930184	USA	biennial	381 bp	A	G	C	T	166 bp	G	258 bp

- missing data.

### 3.8. Mapping of flowering and vernalization gene sequences

A candidate gene approach is considered to be a successful approach to isolate different plant genes. This approach was used also to develop candidate gene markers for plant traits. Using the sequence information of *A. thaliana* flowering time and vernalization genes, specific markers for such genes from sugar beet could be generated. Four genes from *A. thaliana* were used for PCR amplification from an annual sugar beet described in Table 20. They were selected to represent two different flowering pathways (day length and vernalization pathways) in *A. thaliana*. The first two genes were *CONSTANS* (*CO*) and *FLOWERING LOCUS T* (*FT*) that promote flowering specifically in response to long days. *CO* induces early flowering and loss of photoperiod sensitivity. *CO* also activates the *FT* transcription and leads ultimately to flowering. The other two genes were vernalization response genes (*VRN1* and *VRN2*) that accelerate flowering by reduction of *FLC* expression. Since, the environmental factors, like day length and cold temperature, affect the bolting of sugar beet therefore these four gene sequences were selected.

Amplification and cloning of PCR products were done according to the following steps:

Each gene sequence was identified from *A. thaliana*. Then, the primers were designed in such a way to amplify conserved domains or long exons. Since gene regions are supposed to be conserved, normal primers were designed. The bolting (93167P) and non-bolting (A906001) parents were used for amplification to look for polymorphisms between them. The Arabidopsis DNA was used as a positive control and to see if the expected fragment was amplified. Gradient PCR was used in order to determine the optimal annealing temperatures for amplification of PCR products from sugar beet bolting (93167P) and non-bolting (A906001) parents, as well as to look for possible polymorphisms between bolting and non-bolting genotypes. If a simple banding pattern was obtained, then the Arabidopsis-derived primers were used for genotyping the F<sub>2</sub> plants of the mapping population. In case of complicated bands pattern, the polymorphic PCR fragments were excised from the agarose gel, extracted, and purified with the Nucleospin Kit. The fragments were cloned into pGEM-T vector. Using blue-white selection, white colonies were randomly picked. Then, the inserts were amplified with PCR using M13 forward and reverse primers. The purified PCR fragments were then sequenced using MegaBACE 500. The sequences of amplicons were used for marker development.

Using primers (based on conserved domains or designed to amplify long exons) derived from Arabidopsis flowering gene sequences, PCR amplification of flowering gene orthologs from sugar beet was done, which may be a candidate for the bolting gene. Here, the objectives were:

- 1) Development PCR based markers in order to map these flowering gene sequences in a F<sub>2</sub> population segregating for bolting character.
- 2) Determination of map position of these sequences to see if they map close to the *B* gene.

**Table 20:** Production of specific amplicons with homology to conserved domains of cloned *A. thaliana* flowering genes, which are important for the transition from vegetative to reproductive phase.

Gene	Function	Functional domain	Annealing temperature range (°C)	Optimal annealing temperature (°C)	Primer name	PCR fragment size with <i>B. vulgaris</i> DNA (bp)	Polymorphism between 93167P and A906001
<i>CO</i>	Flower integrator, positive regulator	B-Box Zinc-finger, CCT-domain (3'-End)	57.1-67.2	57.1	CO_P2 CO_P3	<b>650 (bb)</b>	Yes
<i>FT</i>	Negative regulator of flower development	Phosphatidyl - ethanolamine binding Motif	57.1-67.2	-	FT_P1 FT_P2	-	-
<i>VRN1</i>	Flower activator, repressor of FLC	<i>Myb</i> -related DNA binding protein	57.1-67.2	62.6	Vrn1_P3 Vrn1_P4	<b>220 (bb)</b> <b>190 (BB)</b>	Yes
<i>VRN2</i>	Maintenance of FLC repression	B-Box Zinc-finger, CCT-domain (3'-End)	50.1-60.2	55.0	Vrn2_P1 Vrn2_P4	<b>800 (bb)</b> <b>800, 1200 (BB)</b>	Yes

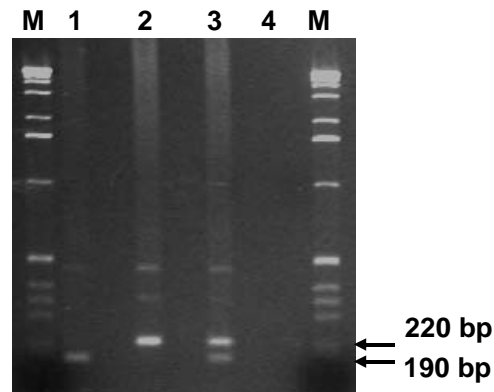
Cloned fragments are in bold.

### 3.8.1. Marker development

For mapping of sugar beet sequences, derived from some of the flowering and vernalization genes of Arabidopsis, simple PCR based marker assays were developed.

#### VRN1 Bv:

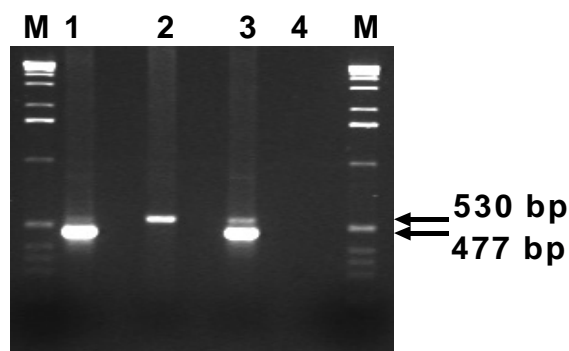
For *VRN1*, it was possible to use Arabidopsis-derived primers (Vrn1\_P3 and P4) directly to map this sequence. After PCR amplification of beet DNA at 62°C, a polymorphism between bolting and non-bolting genotypes was detected using primer combination Vrn1\_P3 and Vrn1\_P4. A 220 bp fragment was present in the *bb* (A906001) as well as in the heterozygous (*Bb*) plant (Figure 24), while a 190 bp fragment was present in the *BB* (93167P) and in heterozygous (*Bb*) plant. This marker is a co-dominant PASA marker that differentiated between homozygous and heterozygous plants. This marker was used to genotype 299 F<sub>2</sub> plants of subpopulation 7 (960701).



**Figure 24:** PASA VRN1\_Bv marker derived using Arabidopsis *VRN1* sequence. Agarose gel electrophoresis of PCR products produced with Vrn\_P3P4 primer combination. 8  $\mu$ l of the PCR reactions were loaded in each slot and separated on a 2% agarose gel for 80 min. at 90 V. 1: bolting parent 93167P; 2: non-bolting parent A906001; 3: DNA of a heterozygous plant; 4: control without template DNA. M: 1kb DNA size marker (Gibco-BRL).

#### VRN2 Bv:

Sugar beet specific primers have been designed based on the consensus sequence (608 bp) of the VRN2\_BB clones. Two primers VRN2\_Bv\_P1 and VRN2\_Bv\_P2B were designed to amplify a fragment of 477 bp. The PCR reaction was performed at 61°C. With the VRN2\_Bv\_P1P2B primer combination, the expected 477 bp fragment could be amplified from bolting (93167P) parent (Figure 25).

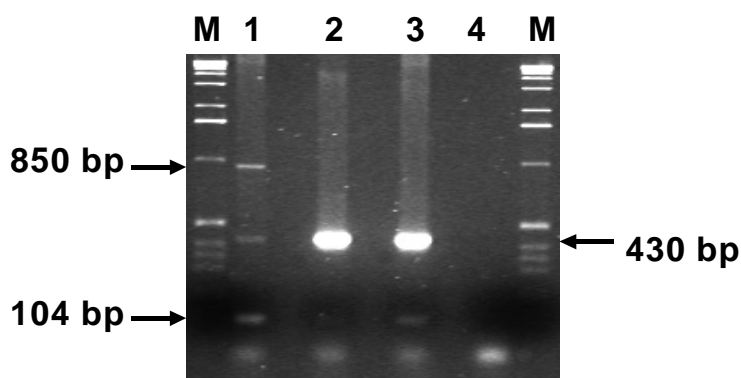


**Figure 25:** PASA VRN2\_Bv marker derived using Arabidopsis *VRN2* sequence. Agarose gel electrophoresis of PCR products produced with VRN2\_Bv\_P1P2B primer combination. 8  $\mu$ l of the PCR reactions were loaded in each slot and separated on a 2% agarose gel for 80 min. at 90 V. 1: bolting parent 93167P; 2: non-bolting parent A906001, 3: DNA of heterozygous plant; 4: control without template DNA. M: 1kb DNA size marker (Gibco-BRL).

An additional 530 bp fragment was obtained with DNA from the non-bolting genotype (A906001). This marker was co-dominant PASA marker, which differentiated between homozygous and heterozygous plants. This marker was used to genotype 299 F<sub>2</sub> plants of subpopulation 7 (960701).

#### CO Bv:

Specific sugar beet-derived primers were designed using the CO\_ *bb* consensus sequence (530 bp). The two primers CO\_Bv\_P1 and CO\_Bv\_P2A were used to amplify specific PCR fragments from bolting and non-bolting parents. With the CO\_Bv\_P1P2A primer combination, the expected PCR fragment (430 bp) was successfully amplified at 61°C in bolting and non-bolting (Figure 26). The 430 bp PCR fragment was visible in the bolting parent as well. In addition, two fragments of 850 bp and 104 bp were amplified from the bolting parent (93167P). With this type of polymorphism, a co-dominant PASA marker was developed. Heterozygous bolting (*Bb*) and homozygous (*BB* or *bb*) plants could be differentiated. 299 F<sub>2</sub> plants (960701, subpopulation 7) were genotyped.



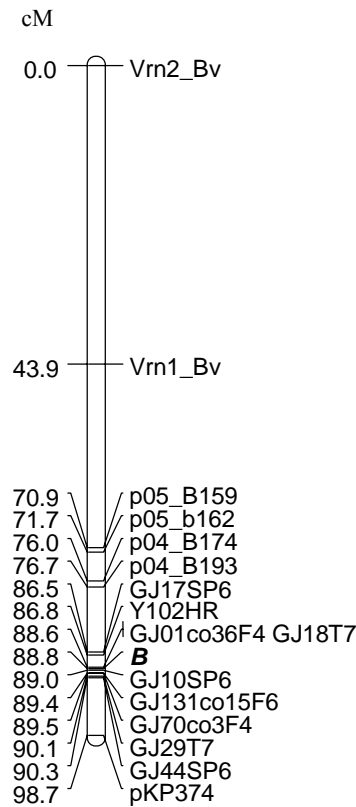
**Figure 26:** PASA CO\_Bv marker derived using Arabidopsis *CO* sequence. Agarose gel electrophoresis of PCR products produced with the CO\_Bv\_P1P2A primer combination. 8  $\mu$ l of the PCR reactions were loaded in each slot and separated on a 2% agarose gel for 80 min. at 90 V. 1: bolting parent 93167P; 2: non-bolting parent A906001; 3: DNA of heterozygous plant; 4: control without template DNA. M: 1kb DNA size marker (Gibco-BRL).

### 3.9. Linkage analysis with the markers derived from flowering and vernalization gene sequences of Arabidopsis

In order to determine the genetic distances between markers derived from flowering and vernalization gene sequences of Arabidopsis and the previously mapped BAC-derived markers from the gene *B* region, the 299 F<sub>2</sub> plants (960701, subpopulation 7) were used. The three molecular markers CO\_Bv, VRN1\_Bv, and VRN2\_Bv were used to genotype the 299 F<sub>2</sub> plants (with verified phenotype in F<sub>3</sub>) of subpopulation 7 (960701). In order to integrate these three markers, a genetic linkage map was constructed based on marker data of the three markers (CO\_Bv, VRN1\_Bv, and VRN2\_Bv), one RFLP-derived marker, eight BAC-sequences derived markers, one YAC end-derived marker, as well as 4 AFLP markers (Figure 27). LOD value of 10 and REC value of 0.30 were used.

It was found that the linear order of the AFLP and BAC-sequences derived markers remained unchanged between the maps. However, the genetic distances were quite different. The total map

length of 98.7 cM was greater than the first linkage map of 28.3 cM of the BAC-sequences derived markers (Figure 22). From this map, most of the markers were clustered together at one side on chromosome 2 while the VRN1\_Bv and VRN2\_Bv were mapped on the other side of the chromosome 2 at 43.9 and 0.0 cM, respectively. Marker CO\_Bv could not be assigned to the same linkage group under the same LOD and REC values.



**Figure 27:** Genetic map of the gene *B* region on chromosome 2. Two sugar beet vernalization related markers were integrated. The map is based on 299 F<sub>2</sub> plants (960701, subpopulation 7). Seventeen markers were mapped with a total map length of 98.7 cM.

### 3.10. Integrated physical and genetic map

The correlation between the physical and genetic maps is important for isolating and cloning of the *B* gene. Therefore, the physical to genetic ratio around the *B* gene was determined. As described in 3.6.2., the two BAC-derived markers GJ10SP6 and GJ01co36F4 flanked the *B* locus with the same map distance of 0.2 cM (Figure 28). Unfortunately, these two BAC-derived markers originated from two BACs GJ10 and GJ01 of two non-overlapping physical contigs, Contig 2 (CK3) and Contig 3 (CK2), respectively (Figure 28). Another, two BAC-derived ORF markers (GJ70co3F4 and GJ131co15F6) developed from two overlapping BACs (GJ70 and GJ131) in Contig 2 (CK3) were mapped at distances of 0.7 cM and 0.6 to the *B* locus, respectively.

Based on PCR analysis with BAC-derived markers from BACs (GJ70 and GJ131), two ORF sequences (GJ70co9 and GJ70co46) could be amplified from BACs (GJ10 and GJ131). Two other ORF-derived sequences from GJ131 (GJ131co8 and GJ131co28) gave amplicons with BAC GJ70. The amplicons generated from all three BACs (GJ10, GJ70, and GJ131) indicated that these BACs

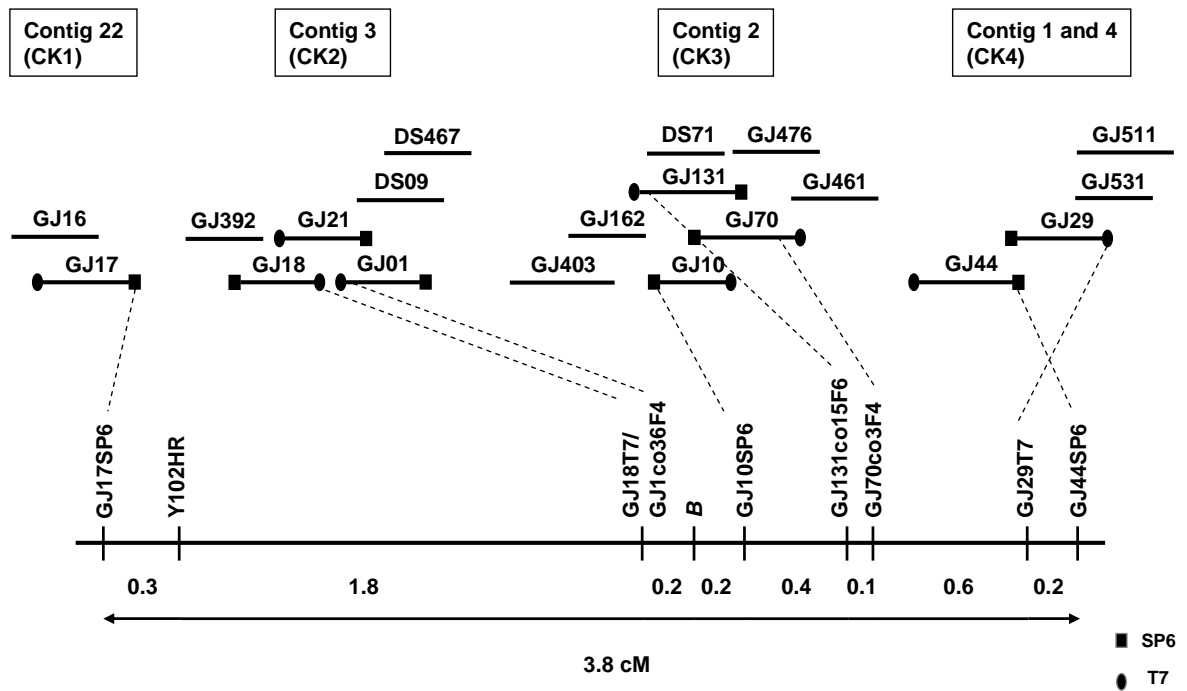
were overlapping. It has been also found that the GJ131co15F6 ORF sequence could be amplified with the BAC GJ10, indicating that this ORF sequence was present on the overlap between the two BACs.

The contigs Contig 22 (CK1), Contig 3 (CK2), Contig 2 (CK3), and Contig 1 and 4 (CK4) spanned approximately 144 kb, 421 kb, 798 kb, and 434 kb, respectively (estimated with FPC software). The total length of all contigs was roughly 1,767 kb. The determination of the ratio between physical and genetic distances for Contig 2 (CK3) could be calculated by using two markers GJ10SP6 and GJ70co3F4 derived from BAC GJ10 and BAC GJ70, respectively. The physical distance between markers GJ10SP6 and GJ70co3F4 was approximately 120 kb and the genetic distance is 0.5 cM. Then the physical to genetic distances for Contig 2 (CK3) was approximately 240 kb/cM. The physical distance between markers GJ29T7 and GJ44SP6 was about 110.8 kb (Telgmann, 2002) and the genetic distance was 0.2 cM, therefore the physical to genetic ratio for Contig 1 and 4 (CK4) was approximately 554 kb/cM.

BACs GJ01 and GJ18 were derived from the Contig 3 (CK2). Its length was approximately 421 kb (estimated with FPC software). The physical distance between markers GJ01co36F4 and GJ18T7 was estimated to be 27 kb. There was no recombination found between these two markers, therefore they mapped at the same position with a 0.2 cM to the *B* gene. Consequently, the physical to genetic ratio could not be calculated for Contig 3 (CK2).

There are three gaps of physically unknown sizes for the *B* gene region. These included a gap between Contig 22 (CK1) and Contig 3 (CK2), Contig 3 (CK2) and Contig 2 (CK3), and Contig 2 (CK3) & Contig 1 and 4 (CK4). The genetic distances were 2.1, 0.4, and 0.6 cM, respectively. From the physical to genetic ratio within the contigs Contig 2 (CK3) and Contig 1 and 4 (CK4) (240 and 554 kb/cM) respectively, one could estimate the size of the physical gap between Contig 2 (CK3) and Contig 3 (CK2) to be 100-200 kb.





**Figure 28:** Genetic and physical maps of the sugar beet chromosomal region encompassing the *B* gene. The genetic map spans 3.8 cM, while the physical map spans 1,767 kb. The physical map consisted of four non-overlapping contigs (Contig 22, Contig 3, Contig 2, and Contig 1 and 4). Dashed lines connect the BAC sequence (end or ORF) to its marker position on the genetic map. The marker GJ17SP6 and GJ44SP6 physically delimited the *B* gene region. cM: centiMorgan. Contig: BAC contig. The information of the physical map was provided by Dr. Uwe Hohmann (Kiel).

## 4. Discussion

Single and low copy BAC sequences from the selected BAC clones were used to develop molecular markers. These molecular markers were used to re-map the BAC sequences. To be sure that the selected BAC clones are from the *B* gene region, re-mapping of BAC-derived sequences is considered to be a very essential step in the procedure of cloning the *B* gene. The hypothesis is that the selected BAC clones are supposed to be physically close to the *B* gene. Therefore, some of the BAC-derived markers are likely to be closer to the *B* gene than any other previously mapped AFLP or RFLP markers. Another important reason to develop BAC-derived markers is that, the AFLP markers are scored dominantly and for fine mapping of such genes, co-dominant markers are needed.

A high-resolution genetic map is a pre-requisite for map-based cloning of the *B* gene. Therefore, in this study, a high-resolution genetic map has been constructed in the vicinity of the *B* gene which contained eight BAC-, one RFLP-, and one YAC-derived markers.

### 4.1. Segregation distortion of the *B* locus

The segregation distortion has been reported to be due to different reasons: statistical bias, genotyping and scoring errors (Plomion *et al.*, 1995), and biological reasons like chromosome loss, viability or lethal genes, genetic isolation mechanisms and genetic load (Bradshaw and Stettler, 1994).

The reliability in evaluation of the phenotype (phenotyping) affects both the order and distances of the markers to the gene locus on the genetic map. In this study, verification of the phenotype and genotype of F<sub>2</sub> plants has been done by scoring the bolting behavior of F<sub>3</sub> families. Although, bolting is an easy to score, environmental factors such as day length and temperature have to be taken into consideration (El-Mezawy *et al.*, 2002). Interactions between bolting phenotype and environment have been observed (Abe *et al.*, 1997; El-Mezawy *et al.*, 2002).

In this study, the  $\chi^2$  analysis of the F<sub>3</sub> families of the subpopulation 4 (960701) as well as F<sub>2</sub> progenies from subpopulations 4 & 6 and population 950619, indicated an extreme deviating segregation from the expected 3:1 segregation. Different segregation ratios have been found among the F<sub>3</sub> families of subpopulation 4. Boudry *et al.* (1994b) found that the growing season (spring or autumn) significantly affected the segregation ratio of annual versus biennial plants in the F<sub>1</sub> progenies. In another study, suppression of bolting has been found when the plants were sown in late May and under restricted day lengths (Abe *et al.*, 1997). They have found an excess of non-bolting plants that led to a segregation that was significantly deviated from its expected 3:1 ratio. Also, Abe *et al.* (1997) have proposed that presence of modifying genes, which alter the effect of the *B* gene, might be involved in the suppression of bolting in the F<sub>2</sub> plants. The genes involved in the suppression were not responsible for thermal-induction but for photo-induction of bolting (Abe *et al.*, 1997). As discussed by Owen (1954), the gene *B* is basically responsible for thermal-induction of bolting.

In this study, late spring sowing has been chosen. Nevertheless, deviation from the expected segregation ratio was observed. In contrast to what have been found by Abe *et al.* (1997), the deviation was due to an excess of bolting plants. The observed segregation ratios of the F<sub>3</sub> families and F<sub>2</sub> plants might be due to: first, imperfect isolation of the plants that resulted in cross-pollination with foreign *B* pollens. Boudry *et al.* (1994b) reported that distortion segregation is likely to be greater in crosses involving wild and cultivated forms of beet than in crosses within each form. Since there was no hybridization barrier between cultivated beets and *B. vulgaris* ssp. *maritima* (Abe *et al.*, 1987), therefore, cross-pollination had likely occurred between them.

Second, the distorted segregation of bolting might be due to the vernalization of plants with a low quantitative vernalization requirement, which allowed floral induction following cold days after early

spring sowing (Desprez, 1980; Evans and Weir, 1981). Third, seed vernalization could probably lead to increased production of bolters in the root crops. Longden *et al.* (1975) reported that rainfall during ripening is considered to be undesirable and low temperatures at this time can affect seed vernalization. Abe *et al.* (1997) reported that, under short day lengths, the bolting gene *B* was suppressed by a number of interacting genes for long day length requirement. One of them is closely linked to *B* and is a part of a gene complex for annuality. In addition, genes acting in the post-vernalization phase of bolting are playing an essential role in determining the bolting tendency of individual plants.

In this study, the segregation ratios varied between the F<sub>3</sub> families of the same cross. This is in agreement with the findings of Abe *et al.* (1997) as different segregations between the families derived from bolting F<sub>1</sub> plants and those from their non-bolting siblings were observed. Due to significant deviation for annual habit, Boudry *et al.* (1994b) has used only the bolting plants to perform mapping of gene *B* with RFLP markers.

Skewed segregation ratios have been reported frequently in interspecific crosses for all types of markers (morphological, isozyme, RFLP, or AFLP). The explanations of these distortions are mostly gametic, zygotic and/or post-zygotic selection, but these are not likely explaining a symmetric and polymodal distribution, especially a 3:1 and 1:3 ratios. In contrast, gene conversion involving DNA heteroduplex formation is known to give such ratios (Nicolas and Rossignol 1983; Nag *et al.*, 1989). Post-meiotic segregation is another process leading to distortion segregation, but with 5:3 and 3:5 ratios. A high rate of conversion and low rate of post-meiotic segregation reflect a high level of heteroduplex formation and efficient repair of the resulting mismatch (Nag *et al.*, 1989). The high rate of heteroduplex formation and mismatch are expected during meiosis pairing in interspecific hybrids (Ky *et al.*, 2000). Another possible reason for segregation distortion is the presence of lethal genes (Nikaido *et al.*, 2000). In this study, there was a decrease in recessive homozygous plants. Therefore, the more probably cause of segregation distortion is the lethal or sublethal genes linked to the recessive allele (non-bolting).

In this study, to solve the problems two approaches were applied:

#### 1. Reconstruction of subpopulation 4

The subpopulation 7 (reconstructed subpopulation 4) was used for map construction with MAPMAKER/EXP 3.0b (Lander *et al.*, 1987). This subpopulation comprised three classes: The first class consisted of the non-bolting recessive plants with verified phenotype in F<sub>3</sub>, the second class covered the heterozygous plants (plants which were heterozygous in F<sub>2</sub> and F<sub>3</sub>), and the third class contained the dominant bolting plants (plants which are bolting in F<sub>2</sub> and F<sub>3</sub>). 299 F<sub>2</sub> plants with verified phenotype in F<sub>3</sub> (high confidence of phenotypic data) were used for construction of the fine map.

For construction of the genetic map, MAPMAKER software was used. Jansen and Nap (2001) reported a novel combination of techniques which demonstrated posterior intervals to the location of markers. But no software which providing such tools is available so far. There are two factors, REC-value and LOD score, that affect the marker-order and therefore, reduce the reliability of the genetic map. In this study, construction of the genetic map was performed under high stringent conditions, where REC-value and LOD score thresholds were 0.30 and 10, respectively. The low REC-value of 0.30 was chosen because the marker loci are thought to be linked and indeed, they could be assigned to the region surrounding the *B* locus in one linkage group on chromosome 2. As pointed out by Remington *et al.* (1999) the marker ordering within each linkage group is a persistent problem in genetic mapping, especially when the number of loci increases. Moreover, genotyping errors can influence both locus ordering and the total genetic length of the map (Lincoln and Lander, 1992).

Therefore, in this study in order to obtain an accurate marker order, the framework map was first optimized using the “ripple” command of MAPMAKER.

Some of the differences between the maps from the two commonly used linkage programs (JOINMAP and MAPMAKER) arise from different algorithms for estimating distances between markers. MAPMAKER estimates distance between markers from the recombination fractions of those markers without weighting (Lander *et al.*, 1987). JOINMAP, on the other hand, uses local weighting of the two flanking markers on either side of the interval to be estimated, with weights based on LOD (Stam and Van Ooijen, 1995).

## 2. Two-point linkage analysis based on non-bolting plants

Only the non-bolting plants have been used for the two-point linkage analysis. The use of markers, which are easily scored and co-dominant, guaranteed high quality of the genotypic data. In addition, to avoid miss scoring of the genotypic data, marker genotypes were scored two times and very carefully. In addition, the marker test was repeated if some plants had an unclear genotype and for recombinants. Hühn (1995) described a method for calculation of recombination values using only the recessive class of a segregating F<sub>2</sub> population. For this analysis, 300 non-bolting (*bb*) F<sub>2</sub> plants from a population of 1617 F<sub>2</sub> plants (i.e. 3234 gametes) were genotyped. The observed absolute frequencies of different classes of marker phenotype of recessive non-bolting individuals were used to estimate the recombination values between marker loci and *B* locus. All BAC-derived markers were found to be in close genetic distance to the bolting gene. The marker GJ01co36F4 showed the lowest recombination value (0.0033) with the *B* locus. Four markers GJ17SP6, GJ18T7, GJ29T7, and GJ44SP6 were at the same genetic distance of 0.0050 recombination units to the *B* gene. Based on the LOD SCORE test (Hühn, 1995), all markers showed significant linkage to the *B* locus.

In this study, comparison of recombination values, calculated with the Hühn’s formula (Hühn, 1995) (based on the 300 F<sub>2</sub> non-bolting plants) with the recombination values calculated using MAPMAKER software (based on 299 F<sub>2</sub> plants of subpopulation 7 with verified phenotype in F<sub>3</sub>), showed no clear difference between the recombination values for each marker. However, the genetic distances calculated with MAPMAKER were quite higher than those which have been calculated manually using Kosambi’s equation based on MAPMAKER recombination values and recombination values produced with Hühn’s formula (Hühn, 1995).

In this study,  $\chi^2$  analysis of the markers, used in the linkage analysis, revealed that the previously found four AFLP markers (El-Mezawy *et al.*, 2002) showed high significant segregation distortion. A similar phenomenon has been described in other studies with AFLP markers by Hansen *et al.* (1999) and Nikaido *et al.* (2000). This is explained by the non-specific amplification of multiple fragments with the same molecular weight (Nikaido *et al.*, 1999), which would distort the segregation ratios and make it difficult to map fragments to unique positions. It has been also reported that, in experimental populations of sugar beet several regions in the genome showed severely distorted segregation (Oleo *et al.*, 1993; Pillen *et al.*, 1993; Schondelmaier *et al.*, 1995).

Interestingly, the PCR-based markers GJ10SP6, GJ131co15F6, and GJ70co3F4, derived from BAC clones that were selected with AFLP markers (GJ10 and GJ131 with p05\_B159 and p05\_b162; GJ70 with p04\_B193), were mapped at different positions compared to the map position of the AFLP markers. This disagreement may be due to the fact that AFLPs may hold more than one specific fragment, which originate from duplicated sequences. The same observation was found by Jamsari (2003). It has been reported that marker systems such as AFLPs, which are based on polymorphisms in short sequence stretches, are more likely to be duplicated and resulted in incorrect locus order on the linkage maps (Quint *et al.*, 2003; Frisch *et al.*, 2004).

## 4.2. Bolting mechanism in sugar beet and its significance

Sugar beet is biennial and its transition from vegetative to generative phase (bolting) must be photo-thermally induced (Smith, 1987). This transition requires exposure to temperatures around 5°C to 8°C for 8 to 14 weeks, followed by exposure to increasing day length (12 hours or more). A photoperiodic requirement after vernalization is essential to initiate bolting (Cooke and Scott, 1993). One of the crucial factors in the sugar beet breeding relates to the susceptibility to cold temperatures (vernalization response) and subsequent bolting. Varieties vary in their sensitivity to bolting and there is a minimum time of exposure to cold temperature required by all varieties. Easy bolting lines require only 6 weeks of exposure to low temperatures, while bolting "resistant" lines may require 12 weeks or more. Wild beet populations (*Beta vulgaris* ssp. *maritima*) showed a wide range of genetic variation for vernalization (Van Dijk and Boudry, 1992).

Firstly, the *B* gene was observed in commercial cultivars by Munerati (1931). Basically, the *B* gene is responsible for thermal induction of bolting, which eliminates the requirement for cold treatment in biennial sugar beets (Owen, 1954). However, long days and moderate temperatures are both necessary for homozygous *B* plants to induce bolting. When planted in early August, they are no more annuals with well-developed rosettes (Abe *et al.*, 1997). Heterozygous bolting plants show a much more complicated behavior. Usually they behave as annuals in May and early June plantings, but when planted in late June or July some plants are annuals while others behave as biennials (Owen, 1954; Shimamoto *et al.*, 1990). Much of the observed variability is probably due to genes introduced with the *b* gene, which influence the post-vernalization phase of bolting, such as response to day length and de-vernalization (Abe *et al.*, 1997).

In wild beet, bolting occurred without any requirement for vernalization and this was found to be controlled by a single dominant gene, called *B* gene (Boudry *et al.*, 1994b). Early flowering is due to the interaction of genetic factors and environmental conditions. Temperature is considered to be the main environmental factor where it is involved in vernalization, growth rate, and photoperiod (Lexander, 1980; Smit, 1983). Beets can be de-vernalized when exposed to high temperatures. During warm weather, when the days are long, sugar beet and other plants tend to develop seed stalks. Stout (1946), Margara (1960), and Curtis (1964) demonstrated the critical action of the biochemical substances by grafting biennial beets to annual beets. Pharis and King (1985) provided a detailed review about the possible role of gibberellins (GA) in vernalization. Changes in GA occur soon after vernalization and during flower initiation in bolting sugar beets but not in non-bolting ones (Pocock and Lenton, 1980). GA applications on non-vernalized non-bolting (*bb*) genotypes induced bolting but not flowering (Margara, 1960).

Although, bolting has been considered as an essential factor in sugar beet breeding for a long time, it is still a problem for sugar beet cultivators. Regardless its importance as a breeding tool, it can cause considerable damage if allowed to contaminate breeding stocks or commercial seed crops. Therefore, isolation and cloning of bolting gene will aid our understanding of the control of flowering in sugar beet and may have a general relevance to biennials.

## 4.3. Molecular markers for the *B* gene

A fine genetic map was essential for precisely localizing the bolting *B* gene in sugar beet. Since DNA libraries are important tools for map-based cloning of genes, a BAC library from the sugar beet has been constructed (Hohmann *et al.*, 2003). In the last years, BAC libraries became an important tool for analysis of whole genomes. Here, the aim was to precisely localize the gene *B* to one of the physical contigs from the gene *B* region and to identify BAC (s) that should contain the *B* gene. BAC sequences of selected BAC clones from four physically non-overlapping contigs (Hohmann *et al.*, 2003) were re-mapped. Re-mapping of 37 BAC end-derived sequences was essential because of two reasons: first, to

confirm the origin of BAC clones from the gene *B* region and second, to develop PCR-based markers (closely *B*-linked), which can be used as markers for the detection of the early bolting marker allele *B* in wild and breeding lines.

In this study, thirty-seven BAC end sequences, eight ORF sequences from three partially sequenced BACs, and twelve RFLP clone sequences were analyzed for polymorphisms. Nineteen PCR-based markers were developed. From the selected BAC sequences five PASA, five CAPS, two SNAP, and five ARMS markers were developed. In addition, two RFLP clone sequences were converted into PCR-based CAPS markers. Eight BAC-derived markers and one RFLP-derived CAPS were used for genetic mapping of the *B* gene.

BAC and YAC sequences have been considered as important sources for marker development and for fine mapping of genes. Tartarini *et al.* (1999) have reported the development of PCR-based markers for *Vf* scab resistance gene in apple from BAC end sequences. BAC sequences have been used to develop STS markers co-segregating with the *MXC3* gene which confers resistance to leaf rust in poplar (Stirling *et al.*, 2001). In other studies, CAPS markers were developed for fine mapping of different plant resistance genes, e.g. aphid resistance locus *Sd-1* in *Malus* (Cevik and King, 2002) and late blight resistance region (*RB*) in potato (Bradeen *et al.*, 2003). Recently, four diagnostic PCR-based CAPS markers have been developed for sex (*M*) gene in *Asparagus officinalis* using BAC end sequences. These markers were later used for genetic fine mapping and assisted in construction of a physical map of the *M* gene (Jamsari *et al.*, 2004). In another interesting study, BAC sequences (sequences from BAC subcloning and BAC ends) from a large-insert DNA library, were used to develop SSR markers which were useful for fine mapping the region surrounding the root-knot nematode resistance (*Ma*) locus of Myrobalan plum (Claverie *et al.*, 2004). Alpert and Tanksley (1996) reported the utilization of converted YAC-end sequences as CAPS markers for high-resolution mapping of fruit weight quantitative trait locus in tomato.

In this study, as a first step in re-mapping of BAC end sequences, 37 BAC-derived primer combinations were used. Five dominant length polymorphisms (PASA) were found between bolting (93167P) and non-bolting (A906001) parents. No PCR products from the bolting (93167P) parent were obtained. On the other hand, amplicons of expected sizes, from the non-bolting genotype, were obtained. These five PASA markers are linked in coupling with the non-bolting allele (*bb*). Although dominant PASA markers are less informative than co-dominant markers, they are useful for the genetic mapping (Reamon-Büttner and Jung, 2000; Meksem *et al.*, 2001). In spite of the usefulness of dominant markers in genetic mapping, only the co-dominant markers were used for genetic mapping of the *B* gene. Sugar beet is an outbreeding crop. It has been reported that the co-dominant markers are particularly well suited for outbreeding species (Rae *et al.*, 2000).

Although CAPS markers require additional work after PCR amplification, they are more accurate, easier to score, consistent, and co-dominant in contrast to PASA. CAPS markers have been shown to be more reliable and convenient for genomic mapping and population genetics (Konieczny and Ausubel, 1993; Drenkard *et al.*, 1997). Seven CAPS markers were successfully developed after cleavage of monomorphic amplicons with restriction endonucleases. Out of the seven markers, three (GJ17SP6, GJ29T7, and pKP374) have been chosen for mapping. These three CAPS markers were co-dominant markers, which distinguished between the bolting (*B*) and non-bolting (*b*) alleles in the heterozygous genotype. In practice, first the BAC end sequences were analyzed for the presence of sites for restriction endonucleases using DNASTAR software. The BAC library has been constructed using a non-bolting genotype (KWS2320). Therefore, only restriction endonucleases that should cut the sequences from non-bolting alleles have been identified with the recognition site analysis. This way was laborious and inefficient to develop CAPS markers from the BAC-end sequences. In addition, a large collection of restriction enzymes was needed and this technique was ultimately limited by the fact that a CAPS must be present (Neff *et al.*, 1998). An alternative is the sequencing of both alleles (*B*

and *b*) and a comparative analysis of recognition sites, which would increase the applicability of the CAPS. Presuming that there is at least one polymorphic nucleotide within the fragment, options for CAPS and dCAPS marker design can be readily identified (Michaels and Amasino, 1998; Neff *et al.*, 1998). Those BAC-end derived CAPS markers (GJ17SP6, GJ29T7, and GJ131SP6) and one RFLP-derived CAPS (pKP374) discriminated successfully between bolting (ZR15\_76) and non-bolting (ZR15\_77) as well as F<sub>1</sub> (ZR15\_78) genotypes. These CAPS markers are applicable in different genetic backgrounds.

The CAPS markers were co-dominant and increased the reliability of scoring of marker data than the PASA markers. However, one has to consider the requirement for cleavage with restriction endonucleases. For GJ29T7 and GJ131SP6, *AluI* restriction endonuclease has been used. This enzyme is not expensive (0.068 €/U). *HinfI* (0.011 €/U) that was applied for GJ10T7 and GJ17SP6 is even cheaper than *AluI*. *MboI* (0.136 €/U) utilized in marker pKP374 was relatively expensive in contrast to those used in case of BAC end-derived CAPS markers.

In this study, intact non-cleaved PCR fragments were found in addition to the cleaved PCR products (e.g. marker GJ131SP6, 206 bp). Other authors have described this phenomenon (Zheng *et al.*, 1999) particularly when the PCR products were cleaved without prior purification and even with purified PCR fragments. This might be due to the multiple-copy nature of the fragment in the genome, which was confirmed by Southern analyses (Zheng *et al.*, 1999). In this study, the multi-copy nature of those BAC ends (e.g. GJ131SP6) was proven by Southern analysis and this might be the cause of the presence of the intact PCR fragments after restriction analysis. To avoid this problem, only single copy sequences were used for marker assay development. For this reason and because of the significance of GJ131 as a BAC from the central *B* gene region, a shotgun library was constructed from this BAC clone to identify single copy sequences for marker development.

#### 4.3.1. SNP characterization

##### SNP density:

It has been suggested from DNA sequencing data, of several individuals of plant or animal species, that DNA markers can be developed for every gene in the genome (Schuler, 1998; Weissenbach, 1998; Lin *et al.*, 1999; Collins, 2000), particularly single nucleotide polymorphism (SNP) markers. In this study, for uncovering of SNPs, the bolting (93167P) and non-bolting (A906001) parents as well as two other bolting (KWS, ZR15\_76) and non-bolting (KWS, ZR15\_77) genotypes were used. Fourteen BAC ends and eight BAC ORFs primer combinations were used for PCR amplification from the four genotypes; the obtained PCR products were directly sequenced. The average number of SNPs/1000 bp was 7.3 for BAC end-derived sequences and 2.8 for BAC ORF-derived sequences. This result is in agreement with Schneider *et al.* (2001), who used EST-derived primers to determine SNPs in sugar beet. They found 3.5 SNPs/1000 bp within possible coding regions of 37 genes, using two different inbred lines of sugar beet. In a study performed with 36 maize inbred lines and 17 ESTs, 16.4 SNPs/1000 bp and 7.9 indels/1000 bp were reported (Ching *et al.*, 2002). Zhao *et al.* (2003) have investigated the single nucleotide polymorphism (SNP) density across the human genome and in different genic categories using two SNP databases. The average numbers of SNPs per 10 kb were 8.33, 8.44, and 8.09 in the human genome, intergenic regions, and genic regions, respectively. In genic regions, the SNP density in intronic, exonic, and adjoining untranslated regions was 8.21, 5.28, and 7.51 SNPs per 10 kb, respectively (Zhao *et al.*, 2003).

In agreement with other published data for human, sugar beet, and melon (Wang *et al.*, 1998; Schneider *et al.*, 2001; Morales *et al.*, 2004), 62.5% of the SNPs derived from BAC ORFs were transitions and the remaining 37.5% were transversions.

In conclusion, the relatively high frequency of SNPs in the sugar beet genome provided an excellent resource for marker development. Different marker assays used to be developed in order to use their polymorphisms for genotyping and mapping.

#### ARMS as a simple technique for SNP assay and mapping:

SNP genetic markers derived from genome sequencing have been used for the development of saturated maps (Froelich *et al.*, 1999), candidate gene identification (Cargill *et al.*, 1999), and high-throughput screening (Germer *et al.*, 2000). Because of their availability, stability, simplicity compared to microsatellite markers, and suitability for high-throughput genotyping, they are considered to be useful markers for association mapping studies or linkage analysis (Kruglyak, 1999; Lander and Schork, 1994), positional cloning (Collins, 1995), and in population genetics (Brumfield *et al.*, 2003). In comparison with commonly used SSRs, biallelic SNPs are less polymorphic (expected heterozygosity is lower) but match the informativeness of RFLPs, when haplotypes, instead of individual SNPs are considered (Rafalski, 2002).

For SNP genotyping, several techniques are available which have been extensively reviewed (Gut, 2001; Shi, 2001; Vignal *et al.*, 2002). One key feature of most techniques, except those based on hybridization, is the two step separation: 1) generation of allele-specific molecular reaction products, 2) separation and detection of the allele-specific products for their identification. From these procedures single-strand conformational polymorphisms (SSCP, Orita *et al.*, 1989), denaturing gradient gel electrophoresis (DGGE), enzymatic mutation detection (Youil *et al.*, 1995), microarray or variant detector arrays (Wang *et al.*, 1998; Hacia *et al.*, 1999; Dong *et al.*, 2001; Qi *et al.*, 2001), and heteroduplex analysis (Lichten and Fox, 1983) have been discussed. Recently, Törjék *et al.* (2003) have described a high-throughput SNaPshot<sup>TM</sup>-microsequencing for SNPs mapping in Arabidopsis. The choice of a method for a particular assay depends on many factors including, cost, throughput, equipment needed, difficulty of assay development, and potential for multiplexing.

In this study, the previously mentioned procedures are labor-intensive, time-consuming, and SNP mapping of large numbers needs significant investment in DNA sequencing. Therefore, a simple, rapid, low cost, and reliable method was chosen, the amplification refractory mutation system (ARMS) for SNP mapping (Ye *et al.*, 2001; Fan *et al.*, 2003).

In plant breeding programs, an immediate, direct, efficient and practical use of linked markers is the marker-assisted selection (MAS). From the practical point of view, markers to be used in MAS need to be simple, fast, and cost effective.

#### **4.3.2. Comparison between different markers**

In this study, for 26 BAC end sequences (72%), no polymorphisms could be found after restriction digest. Because the marker development efficiency for both PASA and CAPS was low, the tetra-primer ARMS technique was employed to map additional SNPs. Five ARMS markers were successfully developed in this study which means that this approach is much better than CAPS technique. The lower efficiency of CAPS markers can be explained by the fact that the majority of single base changes do not generate restriction site differences.

However, to uncover SNPs between bolting and non-bolting genotypes, sequencing of PCR products would need some investment. In this study, to reduce the cost of the SNPs discovery, the sequencing chemistry was diluted 1:7 (undiluted premix: dilution buffer). Therefore, the costs for each sequencing reaction could be reduced to 0.99 €.

Allelic variation can be detected directly on agarose gels without additional manipulation of PCR products. This can be achieved by developing of PCR based markers. In this study, the SNPs were



converted into SNAP or ARMS markers. Drenkard *et al.* (2000) have demonstrated SNAP (single nucleotide amplified polymorphisms) assays which were used to map SNPs in *Arabidopsis*. In this study, the SNAP allele specific primers contained a single base pair mismatch at position -2 from the 3'-end of the two alleles. For the SNAP marker GJ70SP6, introduction of a mismatch lead to higher and improved specificity. Drenkard *et al.* (2000) have also introduced an additional mismatch within the last four bases of the primer to increase primer specificity and allow reliable discrimination between alleles (Newton *et al.*, 1989; Cha *et al.*, 1992; Kwok *et al.*, 1994).

It has been shown that ARMS is a simple and rapid method to detect gene mutations in various research fields, e.g. in detection of drug resistance in *Mycobacterium tuberculosis* (Fan *et al.*, 2003) and in barley SNP genotyping (Chiapparino *et al.*, 2004). The technique is as follows: primers could be designed to amplify fragments of differing sizes for each of the alleles in order to easily resolve them as different bands using agarose gel electrophoresis. To increase the specificity of the reaction, a mismatch is introduced at the 3'-end of each of the two allele-specific primers. Ye *et al.* (2001) have used the primer design computer program (Primer1).

In this study, 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 program reduces artificial products. In this study, the gradient PCR was used in order to find the most suitable annealing temperature, while in other studies touchdown PCR was used (Ye *et al.*, 2001). Ye *et al.* (2001) showed that a 1:10 ratio between outer and inner primers increased the amplification of the allele-specific products and reduced artifacts. However, this was not observed in this study. In this study, the 1:10 ratio rather led to weak amplification of the outer fragment. This finding is in agreement with Chiapparino *et al.* (2004). They reported that 1:10 outer to inner primer ratio lead, in some cases, to failure in amplification of outer fragments. Therefore, a lower ratio (1:2) between the outer and inner primers was chosen in this study.

El-Mezawy *et al.* (2002) have presented 17 AFLP markers within a distance of 5 cM around the *B* gene. In this study, another eight BAC-derived markers in a region of 3.8 cM encompassing the gene *B* were mapped. These BAC-derived markers have the following advantages: they are easily and co-dominantly to score, more reliable than AFLP markers, and could be efficiently used in breeding program to identify haplotype structures around the *B* locus.

The SNAP can be used to select for either bolting or non-bolting alleles in coupling. Since any single SNAP marker produces a dominant signal with only a single band per lane, false negative reactions may occur (reactions that do not produce an amplification product because of technical error rather than template dependence). Alternatively, important breeding material may be screened with both markers to eliminate the error, and breeding material in early generations may be tested with both markers to identify lines that are still heterozygous for the trait (Eckstein *et al.*, 2002). With tetra-primer ARMS, only one PCR reaction is needed to amplify bolting and non-bolting alleles. Therefore, ARMS assays are reliable and suitable for marker assisted selection (MAS) and genotyping of wild and cultivated beets.

Although no sequence information is needed for developing AFLP markers, the dominant nature of AFLPs made them less desirable. With AFLP markers, the scoring of different alleles is not possible, thus indicating the presence of low genetic information (Abdel-Satar *et al.*, 2003). As a result, another types of markers, e.g. PASA (PCR amplification of specific alleles) with length polymorphism, CAPS (cleaved amplified polymorphic sequences), and SNPs (single nucleotide polymorphisms) were used for mapping and isolation of different genes (Perry and Bousquet, 1998; Wang *et al.*, 1998; Patocchi *et al.*, 1999; Deng *et al.*, 2001; Jamsari *et al.*, 2004).

#### 4.4. Copy number of BAC sequences

In this study, the determination of the copy number of BAC-derived sequences has been found to be an essential first step in marker development. In maize, it has been described that before the use of BAC-end sequences for identifying SNPs (between the mapping parents), BAC-end sequences have been screened for the absence of repetitive elements. It has been found that approximately 20% of BAC-end sequences were single or low copy numbers and could be used for SNP marker developments (Meyers *et al.*, 2001). In this study, the determination of BAC sequences copy number was done using two different methods, Southern hybridization and hybridization of the BAC-end sequence as a probe to BAC colonies spotted on high-density filters. From 104 BAC ends used, 19% were found to be repetitive as determined by the two methods. The repetitive BACs were discarded. Flavell *et al.* (1974) have estimated the percentage of repetitive DNA in sugar beet, using reassociation kinetics. They found 60% repetitive sequences. In other study, it has been reported that the majority of the sugar beet genome consists of repetitive DNA (Schmidt *et al.*, 1998). Significant correlation ( $r=0.75$ ) between the Southern bands (detected with Southern method) and signals found on high-density filters for each BAC end indicated that one method is sufficient to determine the copy number. PCR with BAC end-derived primers and DNA from super BAC pools was used as a fast method for estimating the copy number of BAC-end sequences from Asparagus (Jamsari, 2003). Although the PCR-based method is faster, the hybridization methods revealed more reliable results. This is because different fragments of identical sizes could be amplified from different loci in the genome. All the markers developed in this study (PASA, CAPS, or SNPs) were derived from single or low copy BAC sequences only.

#### 4.5. Fine scale mapping

For fine mapping the *B* gene, the recombinants in the vicinity of the *B* locus were highly informative. They are considered as essential resources for finding the *B* locus relative to the markers in the region. The required number of recombinants can be determined by the ratio of genetic to physical distance. Durrett *et al.* (2002) described a formula for estimation of the number of gametes required for a representative fine genetic map. However, gene density should be considered. From a population consisted of 1617 F<sub>2</sub> plants (i.e. 3234 gametes), 300 non-bolting plants were selected for this analysis. The 300 non-bolting plants should be high enough to be representative for performing a fine-scale genetic mapping. Tanksley *et al.* (1995) had proposed that a minimum of 1500 plants (i.e. 3000 gametes) is needed to have a 95% confidence for finding a single crossover at 0.1 cM in a species with a large genome. In this study, based on 300 non-bolting plants, 17 recombinants have been identified. In order to determine the position of recombination relative to the *B* locus and the two flanking markers, the selected recombinants were phenotypically analyzed. Depending on the phenotype of the individual and the marker constitution, the position of the recombination (either between the left marker and the *B* locus or between the right marker and the *B* locus) was defined. The markers (seven BAC- and one RFLP-derived markers) were ordered according to the number of recombination breakpoints positioned between each other and between them and the *B* locus. From this analysis, the *B* locus was found to be flanked by two markers, GJ01co36F4 and GJ131co15F6, with recombination values of 0.0033 and 0.0084, respectively. Based on linkage analysis with MAPMAKER, these two markers mapped at 0.2 and 0.6 cM to the *B* gene, respectively. Interestingly, the two linkage methods (Hühn's formula and MAPMAKER) were in accordance and it was clear that the gene *B* should be located in between marker GJ01co36F4 (Contig 3 (CK2), Figure 28) and GJ131co15F6 or GJ10SP6 (Contig 2 (CK3), Figure 28).

In this study, based on 300 non-bolting plants, there was an agreement between the number of recombinants detected by each marker and its recombination value as calculated based on Hühn's formula (see Figure 23). Although, recombinant mapping is more effective in fine mapping of genes than the linkage mapping programs, the MAPMAKER and recombinant mapping attained the same conclusion. However, four markers (GJ17SP6, GJ18T7, GJ29T7, and GJ44SP6) could not be resolved

as they showed the same number of recombinants (3 recombinants). Although the population used for mapping, including only the recessive non-bolting plants (300 F<sub>2</sub> plants), was a selection from a large population (1617 F<sub>2</sub> plants), these four markers were not resolved. The reason for this can likely be attributed to the position of the *B* gene. It has been found that the distribution of recombination in sugar beet genome is very uneven (Halldén *et al.*, 1996; Nilsson *et al.*, 1997; Nilsson *et al.*, 1999).

Alpert and Tanksley (1996) described a high-resolution map of the *fw2.2* region in tomato based on 51 recombinants. The distances were calculated manually by dividing the number of recombinants by total number of meioses (gametes) multiplied by 100. In a recent study, potato X virus resistance *Nb* gene was positioned between two AFLP markers based on seven recombinants high-resolution genetic map (Marano *et al.*, 2002). The marker distances were corresponding to the percentage of recombination between molecular markers or between molecular markers and *Nb* locus. The same approach was used by Cevik and King (2002) to precisely localize the aphid resistance *Sd-1* gene in *Malus* on two overlapping BAC clones using eleven recombinants.

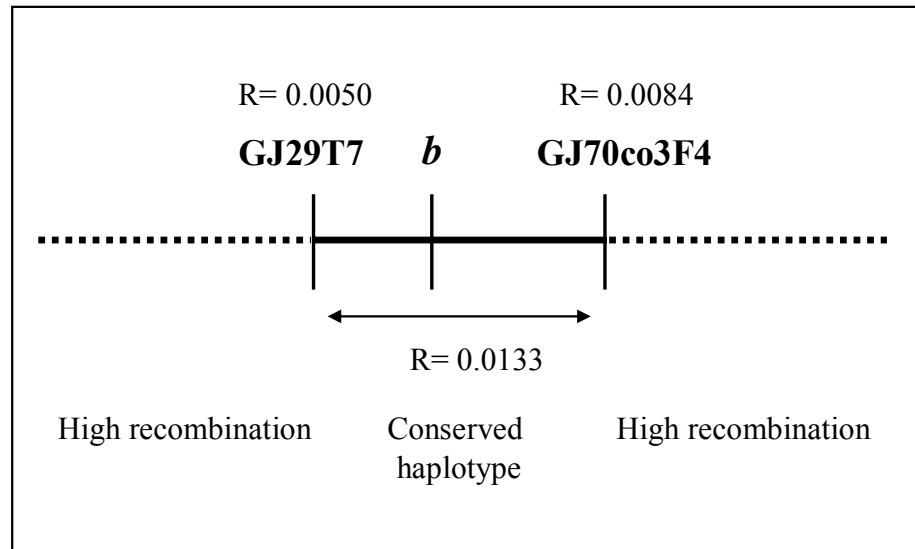
#### 4.6. Evaluation of *B* linked markers in wild *Beta* accessions

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 linkage disequilibrium (LD) between markers and *B* locus. Association studies have been done mainly in human disease studies (Risch and Merikangas, 1996). Recently, association mapping has provided an alternative method for mapping and fine mapping of different traits (simple traits and QTLs). LD was used to test qualitative traits (Pritchard *et al.*, 2000a). Later, other studies used this technique for quantitative traits (Thornberry *et al.*, 2001).

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 4 different subpopulations of wild beets. Although a large variation became obvious in this 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 that are always biennial. Since breeders have been strongly selected against the *B* allele, a marker tightly linked to this locus was expected to differentiate between the different groups of beet accessions. Interestingly, marker sequence GJ131co15F6 showed similarity to the *Myb* protein with unknown function in vernalization in *A. thaliana* and to a nuclear receptor corepressor in human which is involved in chromatin binding and is component of a *HDAC3* complex. The vernalization gene (*Vrn1*) of *Arabidopsis 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 are involved in chromatin remodeling and transition to flowering in *A. thaliana* (Bastow *et al.*, 2004; Sung and Amasino, 2004). One sugar beet haplotype was found reflecting strong selection pressure towards non-bolting 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 as a general rule for wild beets, mutations within the *B* gene resulting in biennial types still have the same haplotype. This could explain the presence of GJ29T7-GJ70co3F4 haplotypes in annual wild 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.*, 1994b; Abe *et al.*, 1997).

Taking together the data from recombination mapping (based on 300 non-bolting plants) and LD analysis, a map around the *B* gene can be depicted. The gene is flanked by markers GJ29T7 and GJ70co3F4 with distances of 0.0050 and 0.0084 recombination units, respectively (Figure 29). First studies on plants demonstrated that LD is depending on the outcrossing rate of a species and can range

between a few hundred base pairs in the case of maize and 250 kb in *Arabidopsis 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.



**Figure 29:** Graphical representation of marker order around the *B* gene on chromosome 2 of sugar beet. The flanking markers were found to be in complete LD with the *b* allele for all sugar beet varieties tested. Recombination values were calculated according to Hühn (1995).

Skinner *et al.* (2000) described a method for detecting marker-trait associations in alfalfa by comparing allele frequency shifts between populations selected for one trait and original unselected populations. It has been reported that association mapping approach is important for identifying regions which contain specific traits. Thornsberry *et al.* (2001) have successfully associated a deletion in the *Dwarf8* gene with flowering time variation in maize.

Kraft *et al.* (2000) reported that, the extent of LD was extended to less than 3 cM in sugar beet with AFLP markers. Studies on other plant species showed that LD decays in a range from a few 100 bp in maize (Remington *et al.*, 2001) and grape (Rafalski and Morgante, 2004) to 50 kb in soybean (Zhu *et al.*, 2003). Gebhardt *et al.* (2004) reported that the extent of LD in potato occurred across 0.3-1.0 cM.

In this study, comparing the results from linkage mapping (using MAPMAKER) and association mapping, some discrepancies were found. According to linkage mapping, markers GJ29T7 and GJ70co3F4 were mapped at 1.3 and 0.7 cM from the *B* locus (see Figure 22), respectively, while the distance between them was 0.6 cM. In contrast, association mapping revealed presence of LD between GJ29T7, GJ70co3F4, and the *b* allele. As *b* allele becomes fixed within the breeding lines due to the random action of genetic drift during selection. However, the recombination can start breaking up the linkage disequilibrium in following generations. This rapidly occurred for unlinked loci but much more slowly for loci that are tightly linked.

The disagreement between the linkage mapping using the MAPMAKER and association mapping was in accordance with the findings of Beer *et al.* (1997), who have analyzed 13 quantitative traits on 64 North American oat varieties and landraces grouped according to RFLP genotype at 48 loci. They

found associations, but there was not very good correspondence with results from conventional QTL mapping studies. This could be owing to the marker-trait disequilibrium may exist in the absence of linkage, and instead may have arisen simply as a consequence of population structure (Beer *et al.*, 1997; Jannink and Walsh, 2002).

The marker alleles associated with significantly different trait values observed by Beer *et al.* (1997) may have become associated with the phenotype through admixture of genetically divergent populations (for both markers and QTL), or through the effects of selection on both marker frequency and phenotype (Jannink and Walsh, 2002). One visible population structure in the Beer *et al.* (1997) data is the differentiation between spring and winter oat varieties, which vary in both phenotype and in marker frequencies (Souza and Sorrells, 1991). Beer *et al.* (1997) concluded that the effectiveness of LD mapping approach may be limited by genotype x environment interactions, multiple alleles of different value, and marker/QTL recombination.

Therefore, the structure of the experimental population has to be assessed, particularly the presence of subgroups with unequal allele distributions. In such structured populations, false-positive associations can be detected between a marker and a phenotype, even if the marker is not physically linked to the locus caused the phenotypic variation (Pritchard *et al.*, 2000b).

El-Mezawy (2001) has identified two minor QTLs for bolting beginning and vitality, which indicated the presence of two additional genetic factors affecting the bolting behavior.

Since the order of markers on the genetic map depends on their recombination frequencies which is consequently dependent on the phenotypic data, this disagreement between linkage mapping and LD mapping might be due to the environmental effects on the phenotype in case of linkage mapping.

#### **4.7. PCR mapping of sugar beet flowering time and vernalization orthologous genes**

Hohmann *et al.* (2003) used *FRI*, *FCA*, and *FLC* clones for hybridizing filters containing selected BAC clones from the *B* gene region. None of the candidate BACs gave a positive signal, which indicated that the gene *B* differs from the described Arabidopsis flowering regulators. Due to the fact that the bolting character is influenced by environmental factors like cold and day length, the attempt was to clone and map sugar beet orthologous genes related to the photoperiod and vernalization pathways.

Three sugar beet-derived amplicons were obtained using primer pairs derived from Arabidopsis *CO*, *VRN1* and *VRN2* gene sequences. These fragments have been genotyped in 299 F<sub>2</sub> plants with verified phenotype in F<sub>3</sub> families. Markers VRN1\_Bv and VRN2\_Bv were mapped on chromosome 2 of sugar beet with distances of 44.9 and 88.8 cM to the *B* locus, respectively. Marker CO\_Bv could not be assigned to the same linkage group. The results obtained with the candidate genes suggest that the early bolting gene (*B*) and these mapped sequences are different.

#### **4.8. Relationship between recombination rate and physical distance**

The genetic and physical maps provide a view of the linear arrangement of the genes and DNA markers on a chromosome. In this study, to orientate the physical map and estimate physical to genetic distance in the *B* gene region, eight BAC-derived markers (from four non-overlapping contigs, Contig 22, Contig 3, Contig 2, and Contig 1 and 4, see Figure 28) were mapped in 299 F<sub>2</sub> progeny (with verified phenotype in F<sub>3</sub>). Two of the markers (GJ01co36F4 and GJ10SP6) flanked the gene with a map distance of 0.2 cM. Unfortunately, these two markers (GJ01co36F4 and GJ10SP6) were derived from two non-overlapping contigs Contig 3 and Contig 2, respectively. The recombination frequencies of the BAC markers derived from the same contig provided a ratio of physical to genetic distance of 240 kb/cM for Contig 2 (CK3) and a ratio of 554 kb/cM for Contig 1 and 4 (CK4). The calculation was based on the physical and genetic distance between GJ10SP6 and GJ70co3F4 for Contig 2 (CK3)

and between GJ29T7 and GJ44SP6 for Contig 1 and 4 (CK4). The ratio between physical and genetic distance for Contig 3 (CK2) could not be calculated based on BACs GJ01 and GJ18. This is because no recombinants could be detected between markers GJ01co36F4 and GJ18T7. The physical distance between these two markers was 27 kb; therefore it is not likely for a crossover to occur. The ratio of Contig 2 (CK3) is comparable with the ratio of 200 kb/cM estimated by Hohmann *et al.* (2003).

An average ratio of 465 kb/cM could be calculated considering a whole physical map of 1,767 kb and a genetic interval of 3.8 cM. However, for the whole beet genome, 1 cM corresponds to 1.1 Mb of DNA (Hohmann *et al.*, 2003). The average ratios of physical to genetic distances varied between plant species, e.g. in *Arabidopsis*, 1 cM corresponds to 200 kb (Schmidt *et al.*, 1995), in rice 244 kb/cM (Chen *et al.*, 2002), in diploid *B. rapa* 700 kb/cM (Sadowski *et al.*, 1996), in maize 1460 kb/cM (Civardi *et al.*, 1994), and in barley and wheat ca. 4400 kb/cM (Faris *et al.*, 2000; Künzel *et al.*, 2000).

It is known that deviations between physical and genetic maps could exist as well as that the distances between maps are not directly proportional. In addition, they are not constant throughout a single genome, where recombination seems to be suppressed in gene-poor regions and occur more frequently in the gene-rich regions (Sandhu and Gill, 2002; Weng and Lazar, 2002). Segal *et al.* (1992) reported a ten-fold increase in recombination rate for the  $I_2$  resistance gene in tomato against *Fusarium oxysporum*, compared with the average of tomato genome, while Stirling *et al.* (2001) described more than 25-fold reduction in recombination rate in the *MX3* region in poplar.

Several factors have attributed to such disproportions. One factor referred to the centromere effect that limits crossing over in intervals close to the centromere. Suppressed recombination has been demonstrated for resistance loci *Tm-2a* and *Mi*, which are physically close to the centromere (Mahtani and Willard, 1998; Copenhaver *et al.*, 1999). Another factor causing disagreements between physical and genetic distances may be due to the size of the chromosome. Higher ratios between physical and genetic distances result from larger chromosomes (Bennetzen and Freeling, 1993; Gale and Devos, 1998b). Differences between chromosome length of a physical map and crossover values of a genetic map can be attributed to multiple crossing over and interference. As the distance between two loci on a chromosome segment increases, the possibility of a second crossover taking place is greater (Lichten and Goldman, 1995).

#### 4.9. Outlook

A comparison of the genetic and physical maps of the *B* locus demonstrated that *B* was flanked with two BAC-derived markers (originated from contigs Contig 2 and Contig 3) and an average of physical to genetic ratio of 465 kb/cM. Unfortunately, the two contigs are not overlapping so far. In this study, the gap between these two contigs could be estimated to 100-200 kb. However, using the contig distal BAC ends of BACs GJ403 (Contig 2) and DS 467 (Contig 3), two approaches could be done: first, chromosome walking experiments in both directions in order to close the gap, and second, conversion of these ends into genetic markers (CAPS or ARMS), which might be more tightly linked to the *B* locus. Chromosome landing approach with these two new BAC-derived markers may be successful to further resolve the location of the *B* gene. Nevertheless, due to the regular distribution of recombination events with respect to the *B* region, there might be no need for further resolution of the region, if the gap between these two contigs is closed.

FISH analysis could be used to determine directly the physical location of two flanking loci on chromosome 2. It may be particularly essential to perform this analysis, if the gap could not be closed as it might be only few kbs in length. This could not be cloned due to size selection and clustering of *HindIII* restriction sites.

Since the BAC-derived markers are genetically tightly linked to the *B* locus, it would be effective just to sequence 4-6 BAC clones completely (from the contigs representing the *B* gene region). These BAC clones should include the two markers (GJ29T7 and B70co3F4) that showed LD with the *b* allele as well as with the two flanking markers (GJ01co36F4 and GJ10SP6).

Availability of mutant plants which lost the function of the *B* gene, provide an important tool for comparing the candidate sequences from the sequenced BAC (s) as an essential step in cloning the *B* gene. As the bolting gene is cloned, its sequence can be used as a marker for identifying individuals contaminated with *B* allele. Genetic manipulation of early bolting could be obtained by transformation of beets with the *B* gene. Thus, the early bolting character could be employed to reduce the time required for breeding process.

Sequences of known genes can be used directly for identifying favorable alleles and follow their inheritance in segregating populations. In addition, the use of gene sequences as selectable markers has several advantages in comparison to other types of markers. The model plants *Arabidopsis* and rice have been completely sequenced and used to assist in positional cloning of genes from related species having much larger genomes and have no complete genome sequence.

In order to identify sugar beet vernalization and flowering orthologous genes, in this study a candidate genes approach has initiated. As known, most of the genes shared conserved functional domains among plant species. Therefore, *Arabidopsis*-derived primers were used to amplify conserved domains (vernalization and flowering time genes). Based on the sequences of cloned PCR fragments, three PCR-based markers were developed and mapped in the F<sub>2</sub> population segregating for bolting behavior. PCR amplicons (derived from sugar beet) of *VRN1* and *VRN2* were hybridized to the *b* and *B* BAC libraries. Positive BAC clones were identified.

For isolation of vernalization genes (*VRN1* and *VRN2*) from sugar beet, full length genomic sequences should be obtained by subcloning the positive BAC clones DS330, DS265, and DS476 for *VRN1*, and DS104, DS243, and DS426 for *VRN2*.

## 5. Summary

In sugar beet, bolting is considered to be an important trait. The presence of the dominant allele *B* of the bolting gene in annual beets causes bolting without vernalization. It is known that the bolting locus *B* is located on chromosome 2. In cultivated sugar beet, bolting is an undesirable phenomenon because it creates troubles during harvesting and sugar extraction as well as it reduces the yield. Owing to the incomplete penetrance of the *B* allele and its environmental dependence, closely linked molecular markers are needed to screen its presence in breeding lines.

The main goals of this study were to fine map the *B* gene using BAC-derived markers from the *B* gene region and to screen the breeding and wild accessions, in order to conduct a preliminary association mapping study of the *B* allele. From an uncomplete physical map (four different contigs), eighteen BACs were selected for marker development around the *B* gene. In the GABI BOLT project, three BACs were shotgun sequenced and 61 open reading frames (ORF) were identified. From 104 BAC ends, 37 BAC ends and 12 ORF single or low copy sequences were chosen for marker development. Thirty-one percent of the BAC-end sequences were found to be single copy and 24% were low. From all selected BAC sequences, 17 markers from 10 different BACs were developed. Ten polymorphisms were determined by simple agarose gel electrophoresis of either restricted or non-restricted PCR products. Another 7 markers were mapped by SNAP (single nucleotide amplified polymorphism) or tetra-primer ARMS (amplification refractory mutation system)-PCR.

A prerequisite for map-based cloning of the *B* gene is a high-resolution genetic map with two flanking markers tightly linked to the gene and a large insert genomic library. In order to select candidate BACs for cloning the *B* gene, genetic linkage was performed using two different populations:

- 1) A genetic linkage map (using MAPMAKER) surrounding the *B* locus was constructed with 8 BAC-derived, one RFLP-derived, and one YAC-derived marker locus. 4 previously mapped AFLP markers were also included. Based on 299 plants (with verified phenotype in F<sub>3</sub>) from a segregating population, the 14 marker loci mapped to a region spanning 28.3 cM around the *B* gene. Two markers (GJ01co36F4 and GJ10SP6) mapped at the same distance of 0.2 cM around the *B* gene.

- 2) Another genetic linkage analysis based only on the non-bolting plants from a F<sub>2</sub> population (1617 plants) segregating for early bolting was performed. 8 markers (7 BAC-derived and one RFLP-derived) were used and their distances to the bolting gene were calculated. The recombination values ranged between 0.0033 and 0.0201.

Both analyses revealed that the *B* locus is located between two markers (GJ01co36F4 and GJ10SP6). Unfortunately, these two markers originated from two non-overlapping BAC contigs (Contig 3 and Contig 2).

Comparison between the physical and genetic distances showed high recombination in the *B* locus region. Physical to genetic ratios of 240 and 554 kb/cM were estimated when comparing physical and genetic distances of markers in contigs Contig 2 and Contig 1 and 4, respectively.

In addition, a set of 41 wild and cultivated *Beta* accessions, differing in their character for early bolting, was genotyped with seven markers. A common haplotype encompassing two marker loci and the *b* allele found in all sugar beet varieties indicating linkage disequilibrium between these loci. This suggests that the bolting gene might be located in close vicinity to these markers and the corresponding BACs can be used for cloning the gene. However, there was a disagreement between results from association mapping and linkage mapping that might be due to population structure or environmental factors.

Using gene sequences from Arabidopsis, three PCR-based sugar beet markers for *VRN1*, *VRN2*, and *CO* were developed. After mapping of these markers in an F<sub>2</sub> population segregating for early bolting, *VRN1* and *VRN2* derived markers were found to be in a distance of 44.9 and 88.8 cM from the *B* locus,



---

respectively. These markers could be useful in the isolation of sugar beet vernalization and flowering time orthologous genes.

## 6. Zusammenfassung

Die Schossneigung ist ein wichtiges Zuchtmerkmal bei der Zuckerrübe. Bei den annualen Rüben wird das Schossen ohne vorherige Vernalisation durch die Anwesenheit des dominanten *B*-Allels des Schossgens bestimmt. Der *B*-Locus des Schossgens kartiert auf Chromosom 2. Unter den Kulturformen der Zuckerrübe stellt das frühzeitige Schossen eine unerwünschte Eigenschaft dar, die bei der Ernte und der Zuckergewinnung sowohl technische Probleme verursacht und den Zuckerertrag reduziert. Für den Nachweis des *B*-Allels in Zuchtlinien sind deshalb molekulare Marker, die eng mit dem Schoss-Phänotyp gekoppelt sind, von großer Bedeutung. Diese Marker sind umso wichtiger, da die Penetranz des *B*-Allels nicht immer vollständig ist und von Umweltfaktoren beeinflusst wird.

Die Hauptziele dieser Arbeit waren die Feinkartierung des Gen *B* unter Verwendung von Markern, die von BAC-Sequenzen aus der Gen *B*-Region abgeleitet wurden, und die Bestimmung der Allelkonstellationen von Zucht- und Wildformen der Zuckerrübe im Rahmen einer vorläufigen Assoziationskartierung. Insgesamt wurden 18 BACs aus vier verschiedenen *contigs* aus der noch unvollständigen physischen Karte der Gen *B*-Region für die Markerentwicklung herangezogen. Im GABI-BOLT-Projekt wurden drei BACs *shotgun* kloniert, je BAC ca. 50 kb nicht-redundante Sequenzinformation erzeugt sowie insgesamt 61 offene Leserahmen (ORFs) identifiziert. Für die Erzeugung von Markern wurden aus den *shotgun*-Sequenzen insgesamt 12 ORFs und aus den 104 BAC-Endsequenzen insgesamt 37 Sequenzen selektiert, die im Zuckerrübengenom entweder als einzeln (single copy) oder mit wenigen kopien (low copy) vorlagen. Es wurden 17 Marker aus 10 verschiedenen BACs entwickelt. Zehn Polymorphismen beruhten auf Größenunterschieden von PCR-Produkten (restringiert oder nicht restringiert) und sind nach einfacher Gelelektrophorese zu bestimmen. Die übrigen 7 Marker wurden mit Hilfe der SNAP- bzw. tetra-primer ARMS-PCR-Technik kartiert.

Eine Voraussetzung für die kartengestützte Klonierung des Schossgens *B* ist die hochdichte genetische Kartierung mit flankierenden, eng gekoppelten Markern und die Erstellung von genomischen *large insert*-Bibliotheken (BAC-Bibliotheken). Zur Selektion von Kandidaten-BACs und der Klonierung des Gens *B* wurde die genetische Kartierung mit zwei verschiedenen Verfahren durchgeführt:

1. Eine genetische Kopplungskarte (mit Hilfe von MAPMAKER) für die Gen *B*-Region wurde erstellt, basierend auf 8 BAC-abgeleiteten Markern sowie je einem YAC- und RFLP-abgeleiteten Marker. Zusätzlich wurden Markerdaten von vier bereits kartierten AFLPs einbezogen. Insgesamt 299 F<sub>2</sub>-Pflanzen, zu denen der Genotyp über eine phänotypische Bestimmung der F<sub>3</sub> überprüft wurde, wurden als Teilpopulation 7 zusammengefasst und mit 14 Markern genotypisiert. Die Markerloci kartierten in einer 28,3 cM-Region um das Gen *B* herum. Zwei Marker, GJ01co36F4 und GJ10SP6, kartierten in genetischen Abständen von jeweils 0,2 cM zum *B*-Locus.

2. Eine Kopplung wurde mit Hilfe einer Zweipunkt-Analyse an selektierten nicht-schossenden Pflanzen einer F<sub>2</sub>-Population (1617 Pflanzen), die für frühes Schossen spaltete, bestimmt. Es wurden 8 Marker verwendet (7 BAC-abgeleitete Marker sowie ein RFLP-abgeleiteter Marker) und deren genetische Abstände zum Schossgen errechnet. Die Rekombinationswerte lagen zwischen 0,0033 und 0,0201.

In beiden Analysen kartierte der Gen *B*-Locus zwischen den beiden Markern GJ01co36F4 und GJ10SP6. Leider wurden diese beiden Marker von BACs abgeleitet, die zu zwei verschiedenen, nicht-überlappenden *contigs* gehören (Contig 3 und Contig 2). Ein Vergleich der physischen und genetischen Abstände zeigte, dass in der Gen *B*-Region unterschiedlich hohe Rekombinationsraten auftraten. Für die *contigs* Contig 2 und Contig 1 & 4 wurden Verhältnisse von physischen zu genetischen Distanzen von jeweils 240 kb/cM bzw. 554 kb/cM ermittelt.

Zusätzlich wurde ein Satz von 41 Wild- und Kulturrüben, die sich hinsichtlich der Eigenschaft für frühes Schossen unterschieden, mit 7 Markern genotypisiert. In allen getesteten Zuckerrübensorten wurde ein gemeinsamer Haplotyp mit zwei Markerloci und dem *b*-Allel gefunden. Dies lässt auf ein Kopplungs-Ungleichgewicht (*linkage disequilibrium*) schliessen, und deutet auf eine Lokalisierung des Schossgens in der unmittelbaren Nähe der Marker hin. Entsprechende BACs aus dieser Region können für die Klonierung des Gens herangezogen werden. Die z. T. widersprüchlichen Ergebnisse aus der Assoziationskartierung und der Kopplungsanalyse können auf die Populationsstruktur oder Umwelteinflüssen beruhen.

Aus Sequenzen der Arabidopsis-Gene *VRN1*, *VRN2* und *CO*, die für die Umschaltung vom vegetativen zum generativen Wachstum von zentraler Bedeutung sind, konnten drei PCR-basierte Marker für die Zuckerrübe entwickelt werden. Nach Kartierung dieser drei Marker in einer F<sub>2</sub>-Population, die für frühes Schossen spaltete, zeigten diese Marker entweder keine Kopplung zum Gen oder lagen weiter vom *B*-Locus entfernt als die BAC-abgeleiteten Marker. Dennoch können diese Marker für die Isolierung von orthologen Zuckerrübenengen von Bedeutung sein.

## 7. References

- Abdel-Satar, M.A.; Khalil, M.S.; Mohamed, I.N.; Abd-El Salam, K.; Verreet, J.A. (2003). Molecular phylogeny of *Fusarium* species by AFLP fingerprint. *African Journal of Biotechnology* **2**: 51-55.
- Abe, J.; Yoshikawa, H.; Tsuda, C. (1987). Reproductive barriers in sugar beets and its relatives of the section *vulgare*, the genus *Beta*. *J. Fac. Agr. Hokkaido Univ.* **63**: 40-48.
- Abegg, F.A. (1936). A genetic factor for the annual habit in beets and linkage relationship. *J. Agric. Res.* **53**: 493-511.
- Adam, D. (2000). Now for the hard ones. *Nature* **408**: 792-793.
- Ahmad, M.; Cashmore, A.R. (1993). *HY4* gene of *Arabidopsis thaliana* encodes a protein with characteristics of a blue-light photoreceptor. *Nature* **366**: 162-166.
- Alpert, K.B.; Tanksley, S.D. (1996). High-resolution mapping and isolation of a yeast artificial chromosome contig containing *fw2.2*: A major fruit weight quantitative trait locus in tomato. *Proc. Natl. Acad. Sci.* **93**: 15503-15507.
- Alvarez, J.; Guli, C.L.; Yu, X-H.; Smyth, D.R. (1992). *Terminal flower*: a gene affecting inflorescence development in *Arabidopsis thaliana*. *Plant J* **2**: 103-116.
- Arumuganathan, K.; Earle, E.D. (1991). Nuclear DNA content of some important plant species. *Plant Mol. Biol. Rep.* **9**: 208-218.
- Aubert, D.; Chen, L.; Moon, Y.H.; Martin, D.; Castle, L.A.; Yang, C.H.; Sung, Z.R. (2001). *EMF1*, a novel protein involved in the control of shoot architecture and flowering in *Arabidopsis*. *Plant Cell* **13**: 1865-1875.
- Barzen, M.; Mechelke, W.; Ritter, E.; Kappert, E. S.; Salamini, F. (1995). An extended map of the sugar beet genome containing RFLP and RAPD loci. *Theor. Appl. Genet.* **90**: 189-193.
- Barzen, M.; Mechelke, W.; Ritter, E.; Seitzer, J. F.; Salamini, F. (1992). RFLP markers for sugar beet breeding: chromosomal linkage maps and location of major genes for rhizomania resistance, monogerm and hypocotyl colour. *Plant J.* **2**: 601-611.
- Bastow, R.; Mylne, J.S.; Lister, C.; Lippman, Z.; Martienssen, R.A.; Dean, C. (2004). Vernalization requires epigenetic silencing of *FLC* by histone methylation. *Nature* **427**: 164-167.
- Beer, S.C.; Siripoonwiwat, W.; O'Donoghue, L.S.; Sousza, E.; Matthews, D.; Sorrells, M.E. (1997). Associations between molecular markers and quantitative traits in an oat germplasm pool: can we infer linkages? *Journal of Agricultural Genomics* **3**. [online] URL: <http://www.ncgr.org/research/jag/papers97/paper197/indexp197.html>.
- Bennetzen, J.L.; Freeling, M. (1993). Grasses as a single genetic system: genome composition, colinearity and compatibility. *Trends in Genetics* **9**: 259-261.
- Bernier, G. (1988). The control of flower evocation and morphogenesis. *Annu. Rev. Plant Physiol.* **39**: 175-219.
- Bhatramakki, D.; Rafalski, A. (2001). The DNA fingerprinting of plants. In: R.J. Henry (ed.), *Plant genotyping*. CAB International Publishing, Wallingford, U.K. pp. 179-192.
- Bonierbale, M.W.; Plaisted, R.L.; Tanksley, S.D. (1988). RFLP maps based on a common set of clones reveal modes of chromosomal evolution of potato and tomato. *Genetics* **120**: 1095-1103.
- Borner, R.; Kampmann, G.; Chandler, J.; Gleissner, R.; Wisman, E.; Apel, K.; Melzer, S. (2000). A MADS domain gene involved in the transition to flowering in *Arabidopsis*. *Plant J.* **24**: 591-599.
- Botstein, D.; White, R.L.; Skolnick, M.; Davis, R.W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* **32**: 314-331.

- Boudry, P.; Broomberg, K.; Saumitou-Laprade, P.; Mörchen, M.; Cuguen, J.; Van Dijk, H. (1994a). Gene escape in transgenic sugar beet: what can be learned from molecular studies of weed beet populations. 3<sup>rd</sup> Int. Symp. on Biosafety Results of Fields Tests of Genetically Modified Plants and Microorganisms. Monterey, California, Published by the University of California, Oakland, California, U.S.A., pp. 45-88.
- Boudry, P.; McCombie, H.; Van Dijk, H. (2002). Vernalization requirement of wild beet *Beta vulgaris* ssp. *maritima*: among population variation and its adaptive significance. *J. Ecol.* **90**: 693-793.
- Boudry, P.; Mörchen, M.; Saumitou-Laprade, P.; Vernet, P.; Van Dijk, H. (1993). The origin and evolution of weed beets: consequences for the breeding and release of herbicide-resistant transgenic sugar beets. *Theor. Appl. Genet.* **87**: 471-478.
- Boudry, P.; Wieber, R.; Saumitou-Laprade, P.; Vernet, P.; Van Dijk, H. (1994b). Identification of RFLP markers closely linked to the bolting gene *B* and their significance for the study of the annual habit in beets (*Beta vulgaris* L.). *Theor. Appl. Genet.* **88**: 852-858.
- Bradeen, J.M.; Naess, S.K.; Song, J.; Haberlach, G.T.; Wielgus, S.M.; Buell, C.R.; Jiang, J.; Helgeson, J.P. (2003). Concomitant reiterative BAC walking and fine genetic mapping enable physical map development for the broad-spectrum late blight resistance region, *RB*. *Mol. Gen. Genomics* **269**: 603-611.
- Bradley, D.; Ratcliffe, O.; Vincent, C.; Carpenter, R.; Coen, E. (1997). Inflorescence commitment and architecture in *Arabidopsis*. *Science* **275**: 80-83.
- Bradshaw, H.D.; Stettler, R.F. (1994). Molecular genetics of growth and development in *Populus*. II. Segregation distortion due to genetic load. *Theor. Appl. Genet.* **89**: 551-558.
- Brumfield, R.T.; Beerli, P.; Nickerson, D.A.; Edwards, S.V. (2003). Single nucleotide polymorphisms (SNPs) as markers in phylogeography. *Trends Ecol. Evol.* **18**: 249-256.
- Büttner, G.; Märlander, B.; Manthey, R. (1995). Breeding for resistance to rhizomania in sugar beet (*Beta vulgaris* L.). *Plant Breed.* **14**: 160-164.
- Cai, D.; Kleine, M.; Kifle, S.; Harloff, H.J.; Sandal, N.N.; Marcker, K.A.; Kleinlanghorst, R.M.; Salentijn, E.M.J.; Lange, W.; Stiekema, W.J.; Wyss, U.; Grundler, F.M.W.; Jung, C. (1997). Positional cloning of a gene for nematode resistance in sugar beet. *Science* **275**: 832-834.
- Cardon, G.H.; Hohmann, S.; Nettessheim, K.; Saedler, H.; Huijser, P. (1997). Functional analysis of the *Arabidopsis thaliana* SBP-box gene *SPL3*: a novel gene involved in the floral transition. *Plant J.* **12**: 367-377.
- Cargill, M.; Altshuler, D.; Ireland, J.; Skالر, P.; Ardile, K.; Patil, N.; Lane, C.R.; Lim, E.P.; Kalyanaraman, N.; Nemesh, J.; Ziaugra, L.; Friedland, L.; Rolfe, A.; Warrington, J.; Lipshutz, R.; Daley, G.Q.; Lander, E.S. (1999) Characterization of single nucleotide polymorphism in coding regions of human genes. *Nat. Genet.* **22**: 231-238.
- Cevik, V.; King, G.J. (2002). Resolving the aphid resistance locus *Sd-1* on a BAC contig within a subtelomeric region of *Malus* linkage group 7. *Genome* **45**: 939-945.
- Cha, R.S.; Zarbl, H.; Keohavong, P.; Thilly, W.G. (1992). Mismatch amplification mutation assay (MAMA): application to the *c-H-ras* gene. *PCR Methods Appl.* **2**: 14-20.
- Chandler, J.; Dean, C. (1994). Factors influencing the vernalization response and flowering mutants of *Arabidopsis thaliana* (L.) Heynh. *J. Exp. Bot.* **45**: 1279-1288.
- Chandler, J.; Wilson, A.; Dean, C. (1996). *Arabidopsis* mutants showing an altered response to vernalization. *Plant J.* **10**: 637-644.
- Chao, S.; Baysdorfer, C.; Heredia-Diaz, O.; Musket, T.; Xu, G.; Coe, E.H. (1994). RFLP mapping of partially sequenced leaf cDNA clones in maize. *Theor. Appl. Genet.* **88**: 717-721.

- Chen, M.; Presting, G.; Barbazuk, W.B.; Goicoechea, J.L.; Blackmon, B.; Fang, G.; Kim, H.; Frisch, D.; Yu, Y.; Sun, S.; Higingbottom, S.; Phimphilai, J.; Phimphilai, D.; Thurmond, S.; Gaubdette, B.; Li, P.; Hatfield, J.; Main, D.; Farrar, K.; Henderson, C.; Barnett, L.; Costa, R.; Williams, B.; Walser, S.; Atkins, M.; Hall, C.; Budiman, M.A.; Tomkins, J.P.; Luo, M.; Bancroft, I.; Salse, J.; Regad, F.; Mohapatra, T.; Singh, N.K.; Tyagi, A.K.; Soderlund, C.; Dean, R.A.; Wing, R.A. (2002). An integrated physical and genetic map of the rice genome. *Plant cell* **14**: 537-545.
- 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, K.S.; Jung, M.; Dolan, M.; Smith, O.S.; Tingey, S.; Morgante, M.; Rafalski, A.J. (2002). SNP frequency, haplotype structure, and linkage disequilibrium in elite maize inbred lines. *BMC Genet.* **3**: 19.
- Civardi, L.; Xia, Y.; Edwards, K.J.; Schnable, P.S.; Nikolau, B.J. (1994). The relationship between genetic and physical distances in the cloned *al-sh2* interval of the *Zea mays* L. genome. *Proc. Natl. Acad. Sci. U.S.A.* **1994**: 8268-8272.
- Claverie, M.; Dirlewanger, E.; Cosson, P.; Bosselut, N.; Lecouls, A.C.; Voisin, R.; Kleinhentz, M.; Lafargue, B.; Caboche, M.; Chalhoub, B.; Esmenjaud, D. (2004). High-resolution mapping and chromosome landing at the root-knot nematode resistance locus *Ma* from Myrobalan plum using a large-insert BAC DNA library. *Theor. Appl. Genet.* **109**: 1318-1327.
- Collins, A. (2000). Mapping in the sequencing era. *Hum. Hred.* **50**: 76-84.
- Collins, F.S. (1995). Positional cloning moves from preditional to traditional. *Nat. Genet.* **9**: 347-350.
- Cook, R.J. (1998). Towards a successful multinational crop plant genome initiative. *Proc. Natl. Acad. Sci. U.S.A.* **95**: 1993-1995.
- Cooke, D.A.; Scott, R.K. (1993). *The Sugar Beet Crop*. Chapman and Hall, Publishers. 675 pp.
- Cooper, D.N.; Smith, B.A.; Cooke, H.J.; Niemann, S.; Schmidt, J. (1985). An estimate of unique DNA sequence heterozygosity in the human genome. *Hum. Genet.* **69**: 201-205.
- Copenhaver, G.P.; Nickel, K.; Kuromori, T.; Benito, M.-I.; Kaul, S.; Lin, S.; Bevan, M. *et al.*, (1999). Genetic definition and sequence analysis of Arabidopsis centromeres. *Science* **286**: 2468-2474.
- Covington, M.F.; Panda, S.; Liu, X.L.; Strayer, C.A.; Wagner, D.R.; Kay, S.A. (2001). *ELF3* modulates resetting of the circadian clock in Arabidopsis. *Plant Cell* **13**: 1305-1315.
- Curtis, G.J. (1964). Graft transmission of the flowering stimulus from a wild *Beta* species to a line of beet selected for resistance to bolting. *Nature* **202**: 1238-1964.
- Deng, Z.; Huang, S.; Ling, P.; Yu, C.; Tao, Q.; Chen, C.; Wendell, M.K.; Zhang, H.B.; Gmittle Jr, F.G. (2001). Fine genetic mapping and BAC contig development for the citrus tristeza virus resistance gene locus in *Poncirus trifoliata* (Raf.). *Mol. Genet. Genomics* **265**: 739-747.
- Desprez, M. (1980). Observasion et remarques sur la montée à graine chez la betterave sucriere. *CR Acad. Agric. Fr*: 44-53.
- Devlin, P.F.; Kay, S.A. (2000). Flower arranging in Arabidopsis. *Science* **288**: 1600-1602.
- Dong, S.L.; Wang, E.; Hsie, L.; Cao, Y.X.; Chen, X.G.; Gingeras, T.R. (2001). Flexible use of high density oligonucleotide arrays for single nucleotide polymorphism discovery and validation. *Genome Res.* **11**: 1418-1424.
- Drenkard, E.; Glazebrook, J.; Preuss, D.; Ausubel, F.M. (1997). Use of cleaved amplified polymorphic sequences (CAPS) for genetic mapping and typing. In: Caetano-Anolles G, Gresshoff P. M.

- (eds), DNA markers. Protocols, applications, and overviews. Wiley-Liss, New York, pp. 187-197.
- Drenkard, E.; Richter, B.G.; Rozen, S.; Stutius, L.M.; Angell, N.A.; Mindrinos, M.; Cho, R.J.; Oefner, P.J.; Davis, R.W.; Ausubel, F.M. (2000). A simple procedure for the analysis of single nucleotide polymorphisms facilitates map-based cloning in Arabidopsis. *Plant Physiol.* **124**: 1483-1492.
- Duke, J.A. (1983). Handbook of Energy Crops: *Beta vulgaris* L.  
[Http://www.hort.purdue.edu/newcrop/duke\\_energy/Beta\\_vulgaris.html](http://www.hort.purdue.edu/newcrop/duke_energy/Beta_vulgaris.html)
- Durrett, R.T.; Chen, K.-Y.; Tanksley, S.D. (2002). A simple formula useful for positional cloning. *Genetics* **160**: 353-355.
- Eckstein, P.E.; Krasichynska, N.; Voth, D.; Duncan, S.; Rossnagel, B.G.; Scoles, G.J. (2002). Development of PCR-based markers for a gene (*Un8*) conferring true loose smut resistance in barley. *Can. J. Plant Pathol.* **24**: 46-53.
- El-Mezawy, A. (2001). Fine mapping of the bolting gene from sugar beet (*Beta vulgaris* L.) with molecular markers. Christian Albrechts Universität.
- 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.
- Evans, A.; Weir, J. (1981). The evolution of weed beet in sugar beet crops. *Kulturpflanzen* **24**: 301-310.
- 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.
- Faris, J.D.; Haen, K.M.; Gill, B.S. (2000). Saturation mapping of a gene-rich recombination hot spot region in wheat. *Genetics* **154**: 823-835.
- Feinberg, A.P.; Vogelstein, B. (1983). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **132**: 6-13.
- Flavell, R.B.; Bennett, M.D.; Smith, J.B.; Smith, D.B. (1974). Genome size and proportion of repeated nucleotide sequence DNA in plants. *Biochem. Genet.* **12**: 257-268.
- Forster, V.; Marrese, D.; Staska, K.; Trinks, K.; Stander, J.R. (1997). Petition for determination of non regulated status: Glufosinate Tolerant Sugar Beet, Transformation Event T120-7. AgrEvo U.S.A. Company, Wilmington, Delaware. 67 pp.
- Fowler, S.; Lee, K.; Onouchi, H.; Samach, A.; Richardson, K.; Morris, B.; Coupland, G.; Putterill, J. (1999). *GIGANTEA*: a circadian clock-controlled gene that regulates photoperiodic flowering in Arabidopsis and encodes a protein with several possible membrane spanning domains. *EMBO J.* **18**: 4679-4688.
- Fridborg, I.; Kuusk, S.; Moritz, T.; Sundberg, E. (1999). The Arabidopsis dwarf mutant shi exhibits reduced gibberellin responses conferred by overexpression of a new putative zinc finger protein. *Plant Cell* **11**: 1019-1032.
- Frisch, M.; Quint, M.; Lübberstedt, T.; Melchinger A.E. (2004). Duplicate marker loci can result in incorrect locus orders on linkage maps. *Theor. Appl. Genet.* **109**: 305-316.
- Froelich, S.; Houlden, H.; Rizzu, P.; Chakraverty, S.; Baker, M.; Kwon, J.; Nowotny, P.; Isaacs, A.; Nowotny, V.; Wauters, E.; van Baren, M.J.; Oostra, B.A.; Hardy, J.; Lannfelt, L.; Goate, A.; Hutton, M.; Lendon, C.L.; Heutink, P. (1999). Construction of a detailed physical and transcript map of the FTDP-17 candidate region on chromosome 17q21. *Genomics* **60**: 129-136.

- Fullerton, S.M.; Clark, A.G.; Weiss, K.M.; Taylor, S.L.; Stengård, J.H.; Salomaa, V.; Boerwinkle, E.; Nickerson, D.A. (2000). Apolipoprotein E variation at the sequence haplotype level: implications for the origin and maintenance of a major human polymorphism. *Am. J. Hum. Genet.* **67**: 881-900.
- Gale, M.D.; Devos, K.M. (1998a). Comparative genetics in the grasses. *Proc. Natl. Acad. Sci. U.S.A.* **95**: 1971-1974.
- Gale, M.D.; Devos, K.M. (1998b). Plant comparative genetics after 10 years. *Science* **282**: 656-659.
- Gaudin, V.; Libault, M.; Pouteau, S.; Juul, T.; Zhao, G.; Lefebvre, D.; Grandjean, O. (2001). Mutations in *LIKE HETEROCHROMATIN PROTEIN 1* affect flowering time and plant architecture in *Arabidopsis*. *Development* **128**: 4847-4858.
- Gaut, B.S.; Long, A.D. (2003). The lowdown on linkage disequilibrium. *The Plant cell* **15**: 1502-1506.
- Gebhardt, C.; Ballvora, A.; Walkemeier, B.; Oberhagemann, P.; Schüler, K. (2004). Assessing genetic potential in germplasm collection of crop plants by marker-trait association: a case study for potatoes with quantitative variation of resistance to late blight and maturity type. *Mol. Breed.* **13**: 93-102.
- Gebhardt, C.; Ritter, E.; Salamini, F. (1994). RFLP map of potato. In: Phillips RL, Vasil IK (eds), DNA-based markers in plants. Kluwer Academic Publishers, Dordrecht Boston London, pp. 271-285.
- Gendall, A.R.; Levy, Y.Y.; Wilson, A.; Dean, C. (2001). The *VERNALIZATION 2* gene mediates the epigenetic regulation of vernalization in *Arabidopsis*. *Cell* **107**: 525-535.
- Germer, S.; Holland, M.J.; Higuchi, R. (2000). High-throughput SNP allele-frequency determination in pooled DNA samples by kinetic PCR. *Genome Res.* **10**: 258-266.
- Guo, H.; Yang, H.; Mockler, T.C.; Lin, C. (1998). Regulation of flowering time by *Arabidopsis* photoreceptors. *Science* **279**: 1360-1363.
- Gut, I.G. (2001). Automation in genotyping single nucleotide polymorphisms. *Hum. Mutat.* **17**: 475-492.
- Hacia, J.G.; Fan, J.B.; Ryder, O.; Jin, L.; Edgemon, K.; Ghandour, G.; Mayer, R.A.; Sun, B.; Hsie, L.; Robbins, C.M.; Brody, L.C.; Wang, D.; Lander, E.S.; Lipshutz, R.; Fodor, S.P.A.; Collins, F.S. (1999). Determination of ancestral alleles for human single nucleotide polymorphisms using high density oligonucleotide arrays. *Nat. Genet.* **22**: 164-167.
- Halldén, C.; Hjerdin, A.; Rading, I.M.; Säll, T.; Fridlundh, B.; Johannisdottir, G.; Tuveesson, S.; Åkesson, C.; Nilsson, N.-O. (1996). A high density RFLP linkage map of sugar beet. *Genome* **39**: 634-645.
- Hansen, M.; Kraft, T.; Christiansson, M.; Nilsson, N.-O. (1999). Evaluation of AFLP in *Beta*. *Theor. Appl. Genet.* **98**: 845-852.
- 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. *Genetical Research* **77**: 61-66.
- Hartmann, U.; Hohmann, S.; Nettesheim, K.; Wisman, E.; Saedler, H.; Huijser, P. (2000). Molecular cloning of *SVP*: a negative regulator of the floral transition in *Arabidopsis*. *Plant J.* **21**: 351-360.
- Heller, R.; Schondelmaier, J.; Steinrücken, G.; Jung, C. (1996). Genetic localization of four genes for nematode (*Heterodera schachtii* Schm.) resistance in sugar beet (*Beta vulagris* L.). *Theor. Appl. Genet.* **92**: 991-997.
- Hicks, K.A.; Albertson, T.M.; Wagner, D.R. (2001). *EARLY FLOWERING3* encodes a novel protein that regulates circadian clock function and flowering in *Arabidopsis*. *Plant Cell* **13**: 1281-1292.



- Hohmann, U.; Jacobs, G.; Telgmann, A.; Gaafar, R.M.; 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.
- Jack, T.; Brockman, L.L.; Meyerowitz, E. M. (1992). The homeotic gene *APETALA3* of *Arabidopsis thaliana* encodes a MADS box and is expressed in petals and stamens. *Cell* **68**: 683-697.
- Jamsari, A. (2003). Construction of high-density genetic and physical maps around the sex gene *M* of *Asparagus officinalis* L. Christian Albrechts Universität.
- Jamsari, A.; Nitz, I.; Reamon-Büttner S.M.; Jung, C. (2004). BAC-derived diagnostic markers for sex determination in asparagus. *Theor. Appl. Genet.* **108**: 1140-1146.
- Jannink, J.-L.; Walsh, J.B. (2002). Association mapping in plant populations. In: Kang, M.S. (ed.), *Quantitative Genetics, Genomics, and Plant Breeding*. CABI, Wallingford, UK, pp. 59-68.
- Jansen, R.C.; Nap, J.P. (2001). Genetical genomics: the added value from segregation. *Trends in Genetics* **17**: 388-391.
- Johanson, U.; West, J.; Lister, C.; Michaels, S.; Amasino, R.M.; Dean, C. (2000). Molecular analysis of *FRIGIDA*, a major determinant of natural variation in *Arabidopsis* flowering time. *Science*. **290**: 344-347.
- Jung, C.; Koch, R.; Fischer, F.; Brandes, A.; Wircke, G.; Herrmann, R.G. (1992). DNA markers closely linked to nematode resistance genes in sugar beet (*Beta vulgaris* L.) using chromosome additions and translocations originated from wild beets of the *Procumbents* species. *Mol. Gen. Genet.* **232**: 537-543.
- Kania, T.; Russenberger, D.; Peng, S.; Apel, K.; Melzer, S. (1997). FPF1 promotes flowering in *Arabidopsis*. *Plant Cell* **9**: 1327-1338.
- Kardailsky, I.; Shukla, V.K.; Ahn, J.H.; Dagenais, N.; Christensen, S.K.; Nguyen, J.T.; Chory, J.; Harrison, M.J.; Weigel, D. (1999). Activation tagging of the floral inducer *FT*. *Science*. **286**: 1962-1965.
- Kinoshita, T.; Harada, J.J.; Goldberg, R.B.; Fischer, R.L. (2001). Polycomb repression of flowering during early plant development. *Proc. Natl. Acad. Sci. U.S.A.* **98**: 14156-14161.
- Kobayashi, Y.; Kaya, H.; Goto, K.; Iwabuchi, M.; Araki, T. (1999). A pair of related genes with antagonistic roles in mediating flowering signals. *Science* **286**: 1960-1962.
- Konieczny, A.; Ausubel, F.M. (1993). A procedure for mapping of *Arabidopsis* mutations using co-dominant ecotype-specific markers. *Plant J.* **4**:403-410.
- Koornneef, M.; Alonso-Blanco, C.; Blankestijn-de Vries, H.; Hanhart, C.J.; Peeters, A. J. (1998). Genetic interactions among late-flowering mutants of *Arabidopsis*. *Genetics* **148**: 885-892.
- Koornneef, M.; Blankestijn-de Vries, H.; Hanhart, C.; Soppe, W.; Peeters, T. (1994). The phenotype of some late-flowering mutants is enhanced by a locus on chromosome 5 that is not effective in the Landsberg erecta wild-type. *Plant J.* **6**: 911-919.
- Koornneef, M.; Hanhart, C.J.; van der Veen, J.H. (1991). A genetic and physiological analysis of late flowering mutants in *Arabidopsis thaliana*. *Mol. Gen. Genet.* **229**: 57-66.
- Kosambi, D.D. (1944). The estimation of map distances from recombination values. *Ann. Eugenics* **12**: 172-175.

- Kraft, T.; Hansen, M.; Nilsson, N.-O. (2000). Linkage disequilibrium and fingerprinting in sugar beet. *Theor. Appl. Genet.* (2000) **101**: 323-326.
- Kruglyak, L. (1999). Prospects for whole-genome linkage disequilibrium mapping of common disease genes. *Nat. Genet.* **22**: 139-144.
- Kubo, H.; Peeters, A.J.; Aarts, M.G.; Pereira, A.; Koornneef, M. (1999). *ANTHOCYANINLESS2*, a homeobox gene affecting anthocyanin distribution and root development in *Arabidopsis*. *Plant Cell* **11**: 1217-1226.
- Künzel, G.; Korzun, L.; Meister, A. (2000). Cytologically integrated physical and restriction fragment length polymorphism maps for the barley genome based on translocation breakpoints. *Genetics* **154**: 397-412.
- Kurata, N.; Nagamura, Y.; Yamamoto, K.; Harushima, Y.; Sue, N.; Wu, J.; Antonio, B.A.; Shomura, A.; Shimizu, T.; Lin, S.-Y.; Inoue, T.; Fukuda, A.; Shimano, T.; Kuboki, Y.; Toyama, T.; Miyamoto, Y.; Kirihara, T.; Hayasaka, K.; Miyao, A.; Monna, L.; Zhong, H.S.; Tamura, Y.; Wang, Z.-X.; Momma, T.; Umehara, Y.; Yano, M.; Sasaki, T.; Minobe, Y. (1994). A 300 kilobase interval genetic map of rice including 883 expressed sequences. *Nat. Genet.* **8**: 365-372.
- Kwok, P.Y.; Deng, Q.; Zakeri, H.; Taylor, S.L.; Nickerson, D.A. (1996). *Genomics* **31**: 123-126.
- Kwok, S.; Chang, S.-Y.; Sninsky, J.J.; Wang, A. (1994). A guide to the design and use of mismatched and degenerate primers. *PCR Methods Appl.* **3**: 39-47.
- Ky, C.-L.; Barre, P.; Lorieux, M.; Trouslot, P.; Akkafou, S.; Louarn, J.; Charrier, A.; Hamon, S.; Noirot, M. (2000). Interspecific genetic linkage map, segregation distortion and genetic conversion in coffee (*Coffea* sp.). *Theor. Appl. Genet.* **101**: 669-676.
- Landegren, U.; Nilsson, M.; Kwok, P.Y. (1998). Reading bits of genetic information: methods for single-nucleotide polymorphism analysis. *Genome Res.* **8**: 769-776.
- Lander, E. S.; Schork, N.J. (1994). Genetic dissection of complex traits. *Science* **265**: 2037-2048.
- Lander, E.S.; Green, P.; Abrahamson, J.; Barlow, A.; Daly, M.J.; Lincoln, S.E.; Newburg, L. (1987). MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* **1**: 174-181.
- Lange, W.A.; Müller, J.; DeBock, T.S.M. (1993). Virulence in the beet cyst nematode (*Heterodera schachtii*) versus some alien genes for resistance in beet. *Fundam. Appl. Nematol.* **16**: 447-454.
- Lee, H.; Suh, S.S.; Park, E.; Cho, E.; Ahn, J.H.; Kim, S.G.; Lee, J.S.; Kwon, Y.M.; Lee, I. (2000). The *AGAMOUSLIKE 20* MADS domain protein integrates floral inductive pathways in *Arabidopsis*. *Genes Dev.* **14**: 2366-2376.
- Lee, I.; Aukerman, M.J.; Gore, S.L.; Lohman, K.N.; Michaels, S.D.; Weaver, L.M.; John, M.C.; Feldmann, K.A.; Amasino, R.M. (1994b). Isolation of *LUMINIDEPENDENS*: a gene involved in the control of flowering time in *Arabidopsis*. *Plant Cell* **6**: 75-83.
- Lee, I.; Michaels, S.D.; Masshardt, A.S.; Amasino, R.M. (1994a). The late-flowering phenotype of *FRIGIDA* and *LUMINIDEPENDENS* is suppressed in the Landsberg erecta strain of *Arabidopsis*. *Plant J.* **6**: 903-909.
- Letschert, J.P.W.; Lange, W.; Frese, L.; van den Berg, R.G. (1994). Taxonomy of *Beta* section *Beta*. *J. Sugar beet Res.* **31**: 69-85.
- Levy, Y.Y.; Mesnage, S.; Mylne, J.S.; Gendall, A.R.; Dean, C. (2002). Multiple roles of *Arabidopsis VRN1* in vernalization and flowering time control. *Science* **297**: 243-246.
- Lexander, K. (1980). Present knowledge of sugar beet bolting mechanism. 43<sup>rd</sup> Winter Congr Inst. Int. Rech. Betterav, Bruxelles, pp. 245-258.

- Lichten, M.; Goldman, A.S. (1995). Meiotic recombination hotspots. *Annu. Rev. Genet.* **29**: 423-444.
- Lichten, M.J.; Fox, M.S. (1983). Detection of non-homology containing heteroduplex molecules. *Nucl. Acids Res.* **11**: 3959-3971.
- Lin, X.; Kaul, S.; Rounsley, S.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; Mason, T.; Bowman, C.L.; Barnstead, M. et al. (1999). Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*. *Nature* **402**: 761-768.
- Lincoln, S.; Daly, M.; Lander, E. (1992). Constructing genetic maps with MAPMAKER/EXP 3.0. Whitehead Institute Technical Report, 3<sup>rd</sup> edn. Whitehead Institute, Cambridge, Massachusetts.
- Lincoln, S.E.; Lander, E.S. (1992). Systematic detection of errors in genetic linkage data. *Genomics* **14**: 604-610.
- Litt, M.; Luty, J.A. (1989). A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *Am. J. Hum. Genet.* **44**: 397-401.
- Little, S. (1997). ARMS analysis of point mutations. In: Taylor, G.R. (ed.) *Laboratory Methods for the Detection of Mutations and Polymorphisms in DNA*. CRC Press, Boca Raton, FL, pp. 45-51.
- Liu, B.-H. (1998). *Statistical genomics: linkage, mapping and QTL analysis*. CRC press New York.
- Liu, X.L.; Covington, M.F.; Fankhauser, C.; Chory, J.; Wagner, D.R. (2001). *ELF3* encodes a circadian clock regulated nuclear protein that functions in an Arabidopsis PHYB signal transduction pathway. *Plant Cell* **13**: 1293- 1304.
- Long, A.D.; Lyman, R.F.; Langley, C.H.; Mackay, T.F.C. (1998). Two sites in the *Delta* gene region contribute to naturally occurring variation in bristle number in *Drosophila melanogaster*. *Genetics* **149**: 999-1017.
- Longden, P.C.; Scott, R.K.; Tyldesley, J.B. (1975). Bolting of sugar beet grown in England. *Outlook on Agriculture* **8**: 188-193.
- Löptien, H. (1984). Breeding nematode-resistant beets. Development of resistant alien additions by crosses between *Beta vulgaris* L. and wild species of the section *Patellaris*. *Z. Pflanzenzücht* **92**: 208-220.
- Macknight, R.; Bancroft, I.; Page, T.; Lister, C.; Schmidt, R.; Love, K.; Westphal, L.; Murphy, G.; Sherson, S.; Cobbett, C.; Dean, C. (1997). *FCA*, a gene controlling flowering time in Arabidopsis, encodes a protein containing RNA-binding domains. *Cell*. **89**: 737-745.
- Mahtani, M.; Willard, H. (1998). Physical and genetic mapping of the human X-chromosome centromere: repression of recombination. *Genome Res.* **8**: 100-110.
- Mandel, M.A.; Gustafson-Brown, C.; Savidge, B.; Yanofsky, M.F. (1992). Molecular characterization of the Arabidopsis floral homeotic gene *APETALA1*. *Nature* **360**: 273-277.
- Marano, M.R.; Malcuit, I.; De Jong, W. (2002). High-resolution genetic map of *Nb*, a gene that confers hypersensitive resistance to potato virus X in *Solanum tuberosum*. *Theor. Appl. Genet.* **105**: 192-200.
- Margara, J. (1960). Recherches sur le déterminisme de l'élongation et de la floraison dans le genre *Beta*. *Ann. Amélior. Plant* **10**: 361-471.
- Martinez-Zapater, J.M.; Somerville, C.R. (1990). Effect of light quality and vernalization on late-flowering mutants of *Arabidopsis thaliana*. *Plant Physiol.* **92**: 770-776.
- Meksem, K.; Ruben, E.; Hyten, D.; Triwitayakorn, K.; Lightfoot, D.A. (2001). Conversion of AFLP bands into high-throughput DNA markers. *Mol. Genet. Genomics* **265**: 207-214.

- Meyers, B.C.; Tingey, S.V.; Morgante, M. (2001). Abundance distribution and transcriptional activity of repetitive elements in the maize genome. *Genome Res.* **11**: 1660-1676.
- Michaels, S.D.; Amasino, R.M. (1998). A robust method for detecting single-nucleotide changes as polymorphic markers by PCR. *Plant J.* **14**: 381-385.
- Michaels, S.D.; Amasino, R.M. (1999). *FLOWERING LOCUS C* encodes a novel MADS domain protein that acts as a repressor of flowering. *Plant Cell* **11**: 949-956.
- Morales, M.; Roig, E.; Monforte, A.J.; Arús, P.; Garcia-Mas, J. (2004). Single-nucleotide polymorphisms detected in expressed sequence tags of melon (*Cucumis melo* L.). *Genome* **47**: 352-360.
- Mullis, K.B.; Faloona, F.A. (1987). Specific synthesis of DNA in vitro via a polymerase catalysed chain reaction. *Meth. Enzymol.* **255**: 335-337.
- Munerati, O. (1931). L'eredità della tendenza alla annualità nella commune barbabietola coltivata. *Ztschr Züchtung, Reihe A. Pflanzenzüchtung* **17**: 84-89.
- Nadeau, J.H. (1989). Maps of linkage and syteny homologies between mouse and man. *Trends Genet.* **5**: 82-86.
- Nag, D.K.; White, M.A.; Petes, T.D. (1989). Palindromic sequences in heteroduplex DNA inhibit mismatch repair in yeast. *Nature* **340**: 318-320.
- Neale, D.B.; Savolainen, O. (2004). Association genetics of complex traits in conifers. *Trends Plant Sci.* **9**: 325-330.
- Neff, M.M.; Neff, J.D.; Chory, J.; Pepper, A.E. (1998). dCAPS, a simple technique for the genetic analysis of single nucleotide polymorphisms: experimental applications in *Arabidopsis thaliana* genetics. *The Plant J.* **14**: 387-92.
- Nelson, D.C.; Lasswell, J.; Rogg, L.E.; Cohen, M.A.; Bartel, B. (2000). FKF1, a clock-controlled gene that regulates the transition to flowering in *Arabidopsis*. *Cell* **101**: 331-340.
- Newton, C.R.; Graham, A.; Heptinstall, L.E.; Powell, S.J.; Summers, C.; Kalsheker, N.; Smith, J.C.; Markham, A.F. (1989). Analysis of any point mutation in DNA. The amplification refractory mutation system (ARMS). *Nucl. Acids Res.* **17**: 2503-2516.
- Ni, M.; Tepperman, J.M.; Quail, P.H. (1998). PIF3, a phytochrome-interacting factor necessary for normal photoinduced signal transduction, is a novel basic helixloop-helix protein. *Cell* **95**: 657-667.
- Nicolas, A.; Rossignol, J.L. (1983). Gene conversion: point-mutation heterozygosities lower heteroduplex formation. *EMBO J* **2**: 2265-2270.
- Nikaido, A.; Yoshimaru, H.; Tsumura, Y.; Suyama, Y.; Murai, M.; Nagasaka, K. (1999). Segregation distortion for AFLP markers in *Cryptomeria japonica*. *Genes Genet. Syst.* **74**: 55-59.
- Nikaido, A.M.; Ujino, T.; Iwata, H.; Yoshimura, K.; Yoshimura, H.; Suyama, Y.; Murai, M.; Nagasaka, K.; Tsumura, Y. (2000). AFLP and CAPS linkage maps of *Cryptomeria japonica*. *Theor. Appl. Genet.* **100**: 825-831.
- Nilsson, N.-O.; Halldén, C.; Hansen, M.; Hjerdin, A.; Säll, T. (1997) Comparing the distribution of RAPD and RFLP markers in a high density linkage map of sugar beet. *Genome* **40**: 644-651.
- Nilsson, N.-O.; Hansen, M.; Panagopoulos, A.H.; Tuvevsson, S.; Ehilde, M.; Christiansson, M.; Rading, I.M.; Rissler, M.; Kraft, T. (1999). QTL analysis of *Cercospora* leaf spot resistance in sugar beet. *Plant Breed.* **118**: 327-334.
- Oleo, M.; Lange, W.; D'Haeseleer, M.; De Bock, T.S.M.; Jacobs, M. (1993). Isozyme analysis of primary trisomics in beet (*Beta vulgaris* L.). Genetical characterization and techniques for

- chromosomal assignment of two enzyme-coding loci: leucine aminopeptidase and glutamate oxaloacetate transaminase. *Theor. Appl. Genet.* **86**: 761-768.
- Olsen, K.M.; Halldorsdottir, S.S.; Stinchcombe, J.R.; Weinig, C.; Schmitt, J.; Purugganan, M.D. (2004). Linkage disequilibrium mapping of *Arabidopsis* *CRY2* flowering time alleles. *Genetics* **167**: 1361-1369.
- Orita, M.; Iwahana, H.; Kanazana, H.; Hayashi, K.; Sekiya, T. (1989). Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. *Proc. Natl. Acad. Sci. U.S.A.* **86**: 2766-2770.
- Owen, F.V. (1954). The significance of single gene reactions in sugar beets. *Proc. Amer. Soc. Sugar Beet Technol.* **8**: 32-398.
- Owen, F.V.; McFarlane, J.S. (1958). Successive annual backcrosses to a nonbolting inbred line of sugar beets. *Proc. Amer. Soc. Sugar Beet Technol.* **10**: 124-132.
- Park, D.H.; Somers, D.E.; Kim, Y.S.; Choy, Y.H.; Lim, H.K.; Soh, M.S.; Kim, H.J.; Kay, S.A.; Nam, H.G. (1999). Control of circadian rhythms and photoperiodic flowering by the *Arabidopsis* *GIGANTEA* gene. *Science* **285**: 1579-1582.
- Patocchi, A.; Vinatzer, B.A.; Gianfranceschi, L.; Tartarini, S.; Zhang, H.B.; Sansavini, S.; Gessler, C. (1999). Construction of a 550 kb BAC-contig spanning the genomic region containing the apple scab resistance gene *Vf*. *Mol. Gen. Genet.* **262**: 884-891.
- Peng, J.; Carol, P.; Richards, D.E.; King, K.E.; Cowling, R.J.; Murphy, G.P.; Harberd, N.P. (1997). The *Arabidopsis* *GAI* gene defines a signaling pathway that negatively regulates gibberellin responses. *Genes Dev.* **11**: 3194-3205.
- Perry, D.J.; Bousquet, J. (1998). Sequence-Tagged-Site (STS) markers of arbitrary genes: development, characterization and analysis of linkage in black spruce. *Genetics* **149**: 1089-1098.
- Pharis, R.P.; King, R.W. (1985). Gibberellins and reproductive development in seed plants. *Annu. Rev. Plant Physiol.* **26**: 517-568.
- Pillen, K.; Steinrücken, G.; Herrmann, R.G.; Jung, C. (1993). An extended linkage map of sugar beet (*Beta vulgaris* L.) including nine putative lethal genes and their restorer gene X. *Plant Breed.* **111**: 265-272.
- Pillen, K.; Steinrücken, G.; Wricke, G.; Herrmann, R.G.; Jung, C. (1992). A linkage map of sugar beet (*Beta vulgaris* L.). *Theor. Appl. Genet.* **84**: 129-135.
- Plomion, C.; O'Malley, D.M.; Durel, C.E. (1995). Genomic analysis in maritime pine (*Pinus pinaster*). Comparison of two RAPD maps using selfed and open-pollinated seeds of the same individual. *Theor. Appl. Genet.* **90**: 1028-1034.
- Pocock, T.P.; Lenton, J.R. (1980). Potential use of retardants for chemical control of bolting in sugar beet. In: Clifford D.R., Lenton J.R. (eds), *Recent developments in the use of plant growth retardants*. *Br. Plant Growth Regul. Group Monogr.* **4**: 41-52.
- Pritchard, J.K. (2001). Deconstructing maize populations structure. *Nat. Genet.* **28**: 203-204.
- Pritchard, J.K.; Stephens, M.; Donnelly, P. (2000b) Inference of population structure using multilocus genotype data. *Genetics* **155**: 945-959.
- Pritchard, J.K.; Stevens, M.; Rosenberg, N.A.; Donnelly, P. (2000a). Association mapping in structured populations. *Am. J. Hum. Genet.* **67**: 170-181.
- Putterill, J.; Robson, F.; Lee, K.; Simon, R.; Coupland, G. (1995). The *CONSTANS* gene of *Arabidopsis* promotes flowering and encodes a protein showing similarities to zinc finger transcription factors. *Cell* **80**: 847-857.

- Qi, X.Q.; Bakht, S.; Devos, K.M.; Gale, M.D.; Osbourn, A. (2001). L-RCA (Ligation rolling circle amplification): a general method for genotyping of single nucleotide polymorphism (SNPs). *Nucl. Acids Res.* **29**: U68-U74.
- Quesada, V.; Macknight, R.; Dean, C.; Simpson, G.G. (2003). Autoregulation of *FCA* pre-mRNA processing controls Arabidopsis flowering time. *EMBO J.* **22**: 3142-3152.
- Quint, M.; Duñle, C.M.; Melchinger, A.E.; Lübberstedt, T. (2003). Identification of genetically linked RGAs by BAC screening in maize and implications for gene cloning, mapping, and MAS. *Theor. Appl. Genet.* **7**: 1171-1177.
- Rae, S.J.; Aldam, C.; Dominguez, I.; Hoebrechts, M.; Barnes, S.R.; Edwards, K.J. (2000). Development and incorporation of microsatellite markers into the linkage map of sugar beet (*Beta vulgaris* spp.). *Theor. Appl. Genet.* **100**: 1240-1248.
- Rafalski, A. (2002). Applications of single nucleotide polymorphisms in crop genetics. *Curr. Opin. Plant Biol.* **5**: 94-100.
- Rafalski, A.; Morgante, M. (2004) Corn and humans: recombination and linkage disequilibrium in two genomes of similar size. *Trends Genet.* **20**: 103-111.
- Ratcliffe, O.J.; Nadzan, G.C.; Reuber, T.L.; Riechmann, J.L. (2001). Regulation of flowering in Arabidopsis by an *FLC* homologue. *Plant Physiol.* **126**: 122-132.
- Ratcliffe, O.J.; Riechmann, J.S. (2002). Arabidopsis transcription factors and the regulation of flowering time: A genomic perspective. *Curr. Issues Mol. Biol.* **4**: 77-91.
- Reamon-Büttner, S.M.; Jung, C. (2000). AFLP-derived STS markers for the identification of sex in *Asparagus officinalis* L. *Theor. Appl. Genet.* **100**: 432-438.
- Reich, D.E.; Cargill, M.; Bolk, S.; Ireland, J.; Sabeti, P.C.; Richter, D.J.; Lavery, T.; Kouyoumjian, R.; Farhadian, S.F.; Ward, R.; Lander, E.S. (2001). Linkage disequilibrium in the human genome. *Nature* **411**: 199-204.
- Remington D.L.; Thornsberry, J.M.; Matsuoka, Y.; Wilson, L.M.; Whitt, S.R.; Doebley, J.; Kresovich, S.; Goodman, M.M.; Buckler IV, E.S. (2001). Structure of linkage disequilibrium and phenotypic associations in the maize genome. *Proc. Natl. Acad. Sci. U.S.A.* **98**: 11479-11484.
- Remington, D.L.; Whetten, R.W.; Liu, B.H.; O'Malley, D.M. (1999). Construction of an AFLP genetic map with nearly complete genome coverage in *Pinus taeda*. *Theor. Appl. Genet.* **98**: 1279-1292.
- Risch, N.; Merikangas, K. (1996). The future of genetic studies of complex human diseases. *Science* **273**: 1516-1517.
- Risch, N.J. (2000). Searching for genetic determinants for the new millennium. *Nature* **405**: 847-856.
- Ronaghi, M.; Uhlen, M.; Nyren, P. (1998). A sequencing method based on real-time pyrophosphate. *Science* **281**: 363-365.
- Sadowski, J.; Gaubier, P.; Delseny, M.; Quiros, C.F. (1996). Genetic and physical mapping in Brassica diploid species of a gene cluster defined in *Arabidopsis thaliana*. *Mol. Gen. Genet.* **251**: 298-306.
- Saghai-Marooif, M.A.; Soliman, K.M.; Jorgensen, R.A.; Allard, R.W. (1984). Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. *Proc. Natl. Acad. Sci. U.S.A.* **81**: 8014-8018.
- Samach, A.; Coupland, G. (2000). Time measurement and the control of flowering in plants. *Bioessays* **22**: 38-47.

- Samach, A.; Onouchi, H.; Gold, S.E.; Ditta, G.S.; Schwarz-Sommer, Z.; Yanofsky, M.F.; Coupland, G. (2000). Distinct roles of *CONSTANS* target genes in reproductive development of *Arabidopsis*. *Science* **288**: 1613-1616.
- Sambrook, J.; Fritsch, E.F.; Maniatis, T. (1989). *Molecular cloning: a laboratory manual* (2<sup>nd</sup> eds). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Sambrook, J.; Russell (2001). *Molecular Cloning: A Laboratory Manual*. Third Edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor NY, U.S.A..
- Sandhu, D.; Gill, K.S. (2002). Gene-containing regions of wheat and other grass genomes. *Plant Physiol.* **128**: 803-811.
- Sanger, F.; Nicken, S.; Coulson, A.R. (1977). DNA sequencing with chain-termination inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* **74**: 5463-5467.
- Schaffer, R.; Ramsay, N.; Samach, A.; Corden, S.; Putterill, J.; Carre, I.A.; Coupland, G. (1998). The late elongated hypocotyl mutation of *Arabidopsis* disrupts circadian rhythms and the photoperiodic control of flowering. *Cell* **93**: 1219-1229.
- Schmidt, R. (2002) Plant genome evolution: lessons from comparative genomics at the DNA Level. *Plant Mol. Biol.* **48**: 21-37.
- Schmidt, R.; West, J.; Love, K.; Lenehan, Z.; Lister, C.; Thompson, H.; Bouchez, D.; Dean, C. (1995). Physical map and organization of *Arabidopsis thaliana* chromosome 4. *Science* **270**: 480-483.
- Schmidt, T.; Kubis, S.; Katsiotis, A.; Jung, C.; Heslop-Harrison, J. S. (1998). Molecular and chromosomal organization of two repetitive DNA sequences with intercalary locations in sugar beet and other *Beta* species. *Theor. Appl. Genet.* **97**: 696-704.
- Schneider, K.; Borchardt, D.C.; Schafer-Pregl, R.; Nagl, N.; Glass, C.; Jeppsson, A.; Gebhardt, C.; Salamini, F. (1999). PCR-based cloning and segregation analysis of functional gene homologues in *Beta vulgaris*. *Mol. Gen. Genet.* **262**: 515-524.
- Schneider, K.; Weisshaar, B.; Borchardt, D.C.; Salamini, F. (2001). SNP frequency and allelic haplotype structure of *Beta vulgaris* expressed genes. *Mol. Breed.* **8**: 63-74.
- Schomburg, F.M.; Patton, D.A.; Meinke, D.W.; Amasino, R.M. (2001). *FPA*, a gene involved in floral induction in *Arabidopsis*, encodes a protein containing RNA-recognition motifs. *Plant Cell* **13**: 1427-1436.
- Schondelmaier, J.; Heller, R.; Pillen, K.; Steinrücken, G.; Jung, C. (1995). A linkage map of sugar beet. *Proc. 58<sup>th</sup> Congr. IIBR*, June 19-22, Beaune, France, 37-44.
- Schondelmaier, J.; Steinrücken, G.; Jung, C. (1996). Integration of AFLP markers into a linkage map of sugar beet (*Beta vulgaris* L.). *Plant Breeding* **115**: 231-237.
- Schuler, G.D. (1998). Electronic PCR: bridging the gap between genome mapping and genome sequencing. *Trends Biotechnol.* **16**:456-459.
- Schultz, T.F.; Kiyosue, T.; Yanovsky, M.J.; Wada, M.; Kay, S.A. (2001). A role for LKP2 in the circadian clock of *Arabidopsis*. *Plant Cell* **13**: 2659-2670.
- Schumacher, K.; Schondelmaier, J.; Barzen, E.; Steinrücken, G.; Bonchardt, D.; Weber, W.E.; Jung, C.; Salamini, F. (1997). Combining different linkage maps in sugar beet (*Beta vulgaris* L.) to make one map. *Plant Breed.* **116**: 23-38.
- Scortecci, K.C.; Michaels, S.D.; Amasino, R.M. (2001). Identification of a MADS-box gene, *FLOWERING LOCUS M*, that represses flowering. *Plant J.* **26**: 229-236.
- Scott, M. (1945). Translocation of the reproductive stimulus in sugar beets. *Bot. Gaz.* **107**: 86-95.

- Segal, G.; Sarfatti, M.; Schaffer, M.A.; Ori, N.; Zamir, D.; Fluhr, R. (1992). Correlation of genetic and physical structure in the region surrounding the *I<sub>2</sub>* *Fusarium oxysporum* resistance locus in tomato. *Mol. Gen. Genet.* **226**: 177-181.
- Setiawan, A.; Koch, G.; Barnes, S.R.; Jung, C. (2000). Mapping quantitative traits (QTLs) for resistance to *Cercospora* leaf spot disease (*Cercospora beticola* Sacc.) in sugar beet (*Beta vulgaris* L.). *Theor. Appl. Genet.* **100**: 1176-1182.
- Shannon, S.; Meeks-Wagner, D.R. (1991). A mutation in the Arabidopsis *TFL1* gene affects inflorescence meristem development. *Plant Cell* **3**: 877-892.
- Sharrock, R.A.; Quail, P.H. (1989). Novel phytochrome sequences in *Arabidopsis thaliana*: structure, evolution, and differential expression of a plant regulatory photoreceptor family. *Genes Dev.* **3**: 1745-1757.
- Sheldon, C.C.; Burn, J.E.; Perez, P.P.; Metzger, J.; Edwards, J.A.; Peacock, W.J.; Dennis, E.S. (1999). The *FLF* MADS box gene: a repressor of flowering in Arabidopsis regulated by vernalization and methylation. *Plant Cell* **11**: 445-458.
- Shi, M.M. (2001). Enabling large-scale pharmacogenetic studies by high-throughput mutation detection and genotyping technologies. *Clin. Chem.* **47**: 164-172.
- Shimamoto, Y.; Tanada, T.; Abe, J. (1990). Analysis for bolting of sugar beet by means of the test crosses of biennial lines with annual line. *Proc. Jpn Sugar Beet Technol.* **32**: 134-137.
- Silverstone, A.L.; Ciampaglio, C.N.; Sun, T. (1998). The Arabidopsis RGA gene encodes a transcriptional regulator repressing the gibberellin signal transduction pathway. *Plant Cell* **10**: 155-169.
- Simko, I. (2004). One potato, two potato: haplotype association mapping in autotetraploids. *Trends Plant Sci.* **9**: 441-448.
- Simpson, G.G. (2003). Evolution of flowering in response to day length: flipping the *CONSTANS* switch. *Bioessays.* **25**: 829-832.
- Simpson, G.G.; Dean, C. (2002). Arabidopsis, the rosetta stone of flowering time?. *Science* **296**: 285-289.
- Simpson, G.G.; Gendall, A.R.; Dean, C. (1999). When to switch to flowering. *Annu. Rev. Cell Dev. Biol.* **15**: 519-550.
- Skinner, D.Z.; Loughin, T.; Obert, D.E. (2000). Segregation and conditional probability of molecular markers with traits in autotetraploid alfalfa. *Mol. Breed.* **6**: 295-306.
- Smit, A.L. (1983). Influence of external factors on growth and development of sugar beet (*Beta vulgaris* L.). Centre for Agricultural Publishing and Documentation, Wageningen, The Netherlands.
- Smith, G.A. (1987). Sugar beet. In: Fehr, W.R. (ed.), Principles of Cultivar Development. MacMillan Publishing Company, pp.577-625.
- Somers, D.E.; Schultz, T.F.; Milnamow, M.; Kay, S.A. (2000). *ZEITLUPE* encodes a novel clock-associated PAS protein from Arabidopsis. *Cell* **101**: 319-329.
- Soppe, W.J.; Jacobsen, S.E.; Alonso-Blanco, C.; Jackson, J.P.; Kakutani, T.; Koornneef, M.; Peeters, A.J. (2000). The late flowering phenotype of *fwa* mutants is caused by gain-of-function epigenetic alleles of a homeodomain gene. *Mol. Cell* **6**: 791-802.
- Southern, E.M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**: 503-517.



- Souza, E.; Sorrells, M.E. (1991). Relationships among 70 North American oat germplasms: I. Cluster analysis using quantitative characters. *Crop Sci.* **31**: 599-605.
- Stam, P.; van Ooijen, J. W. (1995). Joinmap version 2.0: software for the calculation of genetic linkage maps. CPRO-DLO, Wageningen, The Netherlands.
- Stander, J.R. (1995). Letter to USDA/APHIS providing rationale for sugar beet to be a crop considered under notification, pp.6.
- Stirling, B.; Newcombe, G.; Vrebalov, J.; Bosdet, H.D.; Bradshaw, J. (2001). Suppressed recombination around the *MXC3* locus, a major gene for resistance to poplar leaf rust. *Theor. Appl. Genet.* **103**: 1129-1137.
- Stout, M. (1946). Relation of temperature to reproduction in sugar beets. *J. Agri. Res.* **72**: 49-68.
- Strayer, C.; Oyama, T.; Schultz, T.F.; Raman, R.; Somers, D.E.; Mas, P.; Panda, S.; Kreps, J.A.; Kay, S.A. (2000). Cloning of the Arabidopsis clock gene *TOC1*, an autoregulatory response regulator homolog. *Science* **289**: 768-771.
- Sugano, S.; Andronis, C.; Green, R.M.; Wang, Z.Y.; Tobin, E.M. (1998). Protein kinase CK2 interacts with and phosphorylates the Arabidopsis circadian clock associated 1 protein. *Proc. Natl. Acad. Sci. U.S.A.* **95**: 11020-11025.
- Sugano, S.; Andronis, C.; Ong, M.S.; Green, R.M.; Tobin, E.M. (1999). The protein kinase CK2 is involved in regulation of circadian rhythms in Arabidopsis. *Proc. Natl. Acad. Sci. U.S.A.* **96**: 12362-12366.
- Sun, T.-P.; Kamiya, Y. (1994). The Arabidopsis *GAI* locus encodes the cyclase ent-kaurene synthetase A of gibberellin biosynthesis. *Plant Cell* **6**: 1509-1518.
- Sung, S.; Amasino, R.M. (2004). Vernalization and epigenetics: how plants remember winter. *Curr. Opin. Plant Biol.* **7**: 4-10.
- Tanksley, S.D.; Ganai, M.W.; Martin, G.B. (1995). Chromosome landing: a paradigm for map-based gene cloning in plants with large genomes. *Trends Genet.* **11**: 63-68.
- Tartarini, S.; Gianfranceschi, L.; Sansavini, S.; Gessler, C. (1999). Development of reliable PCR markers for selection of the *Vf* gene conferring scab resistance in apple. *Plant breed.* **118**: 183-186.
- Telgmann, A. (2002). Analyse von BAC-Klonen zur Erstellung einer physikalischen Karte der Schossgenregion der Zuckerrübe (*Beta vulgaris* L.). Christian Albrechts Universität.
- Teutonico, R.A.; Osborn, T.C. (1994). Mapping of RFLP and qualitative trait loci in *Brassica rapa* and comparison to the linkage maps of *B. napus*, *B. oleracea*, and *Arabidopsis thaliana*. *Theor. Appl. Genet.* **89**: 885-894.
- Thiel, T.; Kota, R.; Grosse, I.; Stein, N.; Graner, A. (2004). SNP2CAPS: a SNP and INDEL analysis tool for CAPS marker development. *Nucl. Acids Res.* **32**: e5.
- Thomas, S.; Phillips, A.; Van Ragen, P.; Debenham, B.; Scott, R.; Matthews, A.; Hidden, P. (1998). The control of bolting in sugar beet, through the manipulation of gibberellin biosynthesis. I ACR-Long Ashton, L. A. Res. Station, Bristol, BS189AF.
- Thornsberry, J.M.; Goodman, M.M.; Doebley, J.; Kresovich, S.; Nielsen, D.; Buckler, E.S. (2001). *Dwarf8* polymorphisms associate with variation in flowering time. *Nat. Genet.* **28**: 286-289.
- Tian, Y.; Fan, L.; Thureau, T.; Jung, C.; Cai, D. (2004). The absence of TIR-type resistance gene analogues in the sugar beet (*Beta vulgaris* L.) genome. *J. Mol. Evol.* **58**: 40-53.

- Törjék, O.; Berger, D.; Meyer, R.C.; Müssig, C.; Schmid, K.J.; Rosleff Sörensen, T.; Weisshaar, B.; Mitchell-Olds, T.; Altmann, T. (2003). Establishment of a high-efficiency SNP-based framework marker set for Arabidopsis. *The Plant J.* **36**: 122-140.
- Ugozzoli, L.; Wallace, R.B. (1991). Allele-specific polymerase chain reaction. *Methods Enzymol.* **2**: 42-48.
- Valverde, F.; Mouradov, A.; Soppe, W.; Ravenscroft, D.; Samach, A.; Coupland G. (2004). Photoreceptor regulation of *CONSTANS* protein in photoperiodic flowering. *Science.* **303**: 1003-1006.
- Van Dijk, H.; Boudry, P.; McCombie, H.; Vernet, P. (1997). Flowering time in wild beet (*Beta vulgaris* ssp. *maritima*) along a latitudinal cline. *Acta. Oecol.* **18**: 47-60.
- Van Dijk, K.; Boudry, P. (1992). Genetic variability for life-histories in *Beta maritima*. In: Frese L (ed.) International *Beta* Genetic Resources Network. A report on 2<sup>nd</sup> Int. *Beta* Genetic Resources Workshop held at the Institute for Crop Science and Plant Breeding, Braunschweig, Germany, 24-28 June 1991. Int. Crop Network Series No 7 Int. Board for Plant Genetic Resources, Rome: 4-16.
- Vignal, A.; Milan, D.; Sancristobal, M.; Eggen, A. (2002). A review on SNP and other types of molecular markers and their use in animal genetics. *Genet. Sel. Evol.* **34**: 275-305.
- Voorrips, R.E. (2002). MapChart: Software for the graphical presentation of linkage maps and QTLs. *J. of Heredity* **93**: 77-78.
- Vos, P.; Hogers, R.; Bleeker, M.; Reijans, M.; Lee, T.; Hornes, M.; Frijters, A.; Pot, J.; Peleman, J.; Kuiper, M.; Zabeau, M. (1995). AFLP: a new technique for DNA fingerprinting. *Nucl. Acids Res.* **23**: 4407-4414.
- Wagner, H.; Wricke, G. (1991). Genetic controls of five isozyme systems in sugar beet. *Plant Breed.* **107**: 124-130.
- Wang, D.G.; Fan, J.-B.; Siao, C.-J.; Berno, A.; Young, P.; Sapolsky, R.; Ghandour, G.; Perkins, N.; Winchester, E.; Spencer, J.; Kruglyak, L.; Stein, L.; Hsie, L.; Topaloglou, T.; Hubbell, E.; Robinson, E.; Mittmann, M.; Morris, M.S.; Shen, N.P.; Kilburn, D.; Rioux, J., Nusbaum, C.; Rozen, S.; Hudson, T.J.; Lipshutz, R.; Chee, M.; Lander, E.S. (1998). Large-scale identification, mapping, and genotyping of single nucleotide polymorphisms in the human genome. *Science* **280**: 1077-1082.
- Wang, Z.Y.; Kenigsbuch, D.; Sun, L.; Harel, E.; Ong, M.S.; Tobin, E.M. (1997). A Myb-related transcription factor is involved in the phytochrome regulation of an Arabidopsis Lhcb gene. *Plant Cell* **9**: 491-507.
- Wang, Z.Y.; Tobin, E.M. (1998). Constitutive expression of the *CIRCADIAN CLOCK ASSOCIATED 1 (CCA1)* gene disrupts circadian rhythms and suppresses its own expression. *Cell* **93**: 1207-1217.
- Watson, L.; Dallwitz, M.J.(1992). "The Families of Flowering Plants: Descriptions, Illustrations, Identification, and Information Retrieval." Version: 15<sup>th</sup> October 1998.  
<http://biodiversity.uno.edu/delta/>.
- Weber, J.L.; May, P.E. (1989). Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am. J. Hum. Genet.* **44**: 388-396.
- Weigel, D.; Ahn, J.H.; Blázquez, M.A.; Borevitz, J.O.; Christensen, S.K.; Fankhauser, C.; Ferrándiz, C.; Kardailsky, I.; Malancharuvil, E.J.; Neff, M.M. *et al.* (2000). Activation tagging in Arabidopsis. *Plant Physiol.* **122**: 1003-1013.

- Weissenbach, J. (1998). The Human Genome Project: from mapping to sequencing. *Clin. Chem. Lab Med.* **36**: 511-514.
- Welsh, J.; McClelland, M. (1990). Fingerprinting genomes using PCR with arbitrary primers. *Nucl. Acids Res.* **18**: 7213-7218.
- Weng, Y.; Lazar, M.D. (2002). Comparison of homoeologous group-6 short arm physical maps of wheat and barley reveals a similar distribution of recombinogenic and gene-rich regions. *Theor. Appl. Genet.* **104**: 1078-1085.
- Wilson, R.N.; Heckman, J.W.; Somerville, C.R. (1992). Gibberellin is required for flowering but not for senescence in *Arabidopsis thaliana* under short days. *Plant Physiol.* **100**: 403-408.
- Ye, S.; Dhillon, S.; Ke, X.; Collins, A.R.; Day, I.N.M. (2001). An efficient procedure for genotyping single nucleotide polymorphisms. *Nucl. Acids Res.* **29**: 17e88.
- Yoshida, N.; Yanai, Y.; Chen, L.; Kato, Y.; Hiratsuka, J.; Miwa, T.; Sung, Z.R.; Takahashi, S. (2001). *EMBRYONIC FLOWER2*, a novel polycomb group protein homolog, mediates shoot development and flowering in *Arabidopsis*. *Plant Cell.* **13**: 2471-2481.
- Youil, R.; Kemper, B.W.; Cotton, R.G.H. (1995). Screening for mutation by enzyme mismatch cleavage with T4 endonuclease VII. *Proc. Natl. Acad. Sci. U.S.A.* **92**: 87-91.
- Zhao, Z.M.; Fu, Y.X.; 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, X.Y.; Wolff, D.W.; 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.
- Zhu, Y.L.; Song, Q.J.; Hyten, D.L.; Van Tassell, C.P.; Matukumalli, L.K.; Grimm, D.R.; Hyatt, S.M.; Fickus, E.W.; Young, N.D.; Cregan, P.B. (2003). Single nucleotide polymorphisms (SNPs) in soybean. *Genetics* **163**: 1123-1134.

## 8. Appendix

### 8.1. Copy number determination of BAC-derived sequences

**Table 21:** Results of copy number determination of the selected BACs. (n. d.: not determined).

BAC number	Copy number (SP6-ends)	Copy number (T7-ends)
GJ01	Low copy	Single copy
GJ03	Multi-copy	Repetitive
GJ06	Low copy	Low copy
GJ08	Low copy	Repetitive
GJ10	Low copy	Single copy
GJ11	Repetitive	Repetitive
GJ12	Repetitive	Single copy
GJ13	Repetitive	Multi-copy
GJ15	Single copy	n. d.
GJ17	Multi-copy	Single copy
GJ18	Multi-copy	Single copy
GJ21	Single copy	Multi-copy
GJ22	Low copy	Low copy
GJ23	Repetitive	Repetitive
GJ26	Single copy	Low copy
GJ28	Multi-copy	Repetitive
GJ29	Single copy	Single copy
GJ33	Single copy	Single copy
GJ37	Repetitive	Low copy
GJ38	Multi-copy	Repetitive
GJ42	Single copy	Multi-copy
GJ43	n. d.	Low copy
GJ44	Single copy	Single copy
GJ45	Low copy	Low copy
GJ47	Low copy	Repetitive
GJ49	Low copy	Single copy
GJ55	Single copy	Single copy
GJ63	Low copy	Multi-copy
GJ65	n. d.	Multi-copy
GJ67	Single copy	n. d.
GJ68	n. d.	Repetitive
GJ70	Single copy	Single copy
GJ73	Multi-copy	Multi-copy
GJ76	Single copy	Low copy
GJ79	Low copy	Low copy
GJ80	Multi-copy	Multi-copy
GJ83	Single copy	Low copy
GJ90	Repetitive	n. d.
GJ131	Low copy	Multi-copy
GJ132	Multi-copy	Multi-copy

**Table 22:** Copy number of different contig sequences from shotgun libraries analyzed by Southern hybridization.

<b>BAC1</b>		<b>BAC70</b>		<b>BAC131</b>	
<b>Sequence Contig</b>	<b>Copy number</b>	<b>Sequence Contig</b>	<b>Copy number</b>	<b>Sequence Contig</b>	<b>Copy number</b>
1	Multi-copy	3	Single copy	7	Single copy
3	Low copy	7	Repetitive	7	Repetitive
15	Multi-copy	9	Repetitive	8	Single copy
33	Low copy	11	Single copy	9	Repetitive
36	Single copy	17	Repetitive	15	Single copy
38	Multi-copy	46	Low copy	28	Single copy
43	Multi-copy	53	Repetitive	-	-

## 8.2. Markers and populations that have been mapped in this study

**Table 23:** List of markers developed from BAC-Ends, ORF-, RFLP-, and cloned-PCR sequences (n. d.: not determined).

Marker name	Clone bank No.	Source	Mapping population (number of plants)	Annealing temperature (°C)	Genotyping method	Distance to B (cM)
GJ01T7	3104	BAC-end	960701, subpop4 (1359)	60	PASA	n. d.
GJ10SP6	3116	BAC-end	960701, subpop7 (299)	61.1	SNAP	0.3
GJ10T7	3109	BAC-end	n. d.	60	CAPS/ <i>Hinf</i> I	n. d.
GJ15SP6	3117	BAC-end	n. d.	n. d.	Sequencing	n. d.
GJ17SP6	3110	BAC-end	960701, subpop4 (1359), subpop5, 6 & pop2 (300 bb)	60	CAPS/ <i>Hinf</i> I	5.8
GJ17T7	3105	BAC-end	n. d.	60	PASA	n. d.
GJ18SP6	3118	BAC-end	n. d.	n. d.	Sequencing	n. d.
GJ18T7	3119	BAC-end	960701, subpop7 (299), subpop5, 6 & pop2 (300 bb)	60	ARMS	1.1
GJ21SP6	3120	BAC-end	n. d.	n. d.	Sequencing	n. d.
GJ22T7	3106	BAC-end	n. d.	60	PASA	n. d.
GJ29T7	3111	BAC-end	960701 subpop4 (1359), subpop5, 6 & pop2 (300 bb)	60	CAPS/ <i>Alu</i> I	4
GJ37T7	3107	BAC-end	n. d.	60	PASA	n. d.
GJ44SP6	3121	BAC-end	960701, subpop7 (299), subpop5, 6 & pop2 (300 bb)	62	ARMS	1.4
GJ49SP6	3112	BAC-end	960701, subpop4 (46)	60	CAPS/ <i>Sau</i> 3AI	n. d.
GJ70SP6	3122	BAC-end	n. d.	59.5	SNAP	n. d.
GJ73T7	3123	BAC-end	n. d.	n. d.	Sequencing	n. d.
GJ131SP6	3113	BAC-end	n. d.	60	CAPS/ <i>Alu</i> I	n. d.
GJ131T7	3108	BAC-end	n. d.	60	PASA	n. d.
GJ01co15F3		ORF-	n. d.	n. d.	Sequencing	n. d.
	3124	sequence				
GJ01co36F4		ORF-	960701, subpop7(299), subpop5, 6 & pop2 (300 bb)	62	ARMS	0.8
	3125	sequence				
GJ70co3F4		ORF-	960701, subpop7(299), subpop5, 6 & pop2 (300 bb)	57	ARMS	2.1
	3126	sequence				
GJ131co15F2		ORF-		n. d.	Sequencing	n. d.
	3127	sequence				
GJ131co15F6		ORF-	960701, subpop7(299), subpop5, 6 & pop2 (300 bb)	62	ARMS	0.1
	3128	sequence				
Köln60	3129	RFLP-clone	960701, subpop4 (46)	n. d.	Sequencing	n. d.
pKP374		RFLP-clone	960701, subpop4 (1359), subpop5, 6 & pop2 (300 bb)	58	CAPS/ <i>Sau</i> 3AI	10.4
	3114					
pKP591	3130	RFLP-clone	960701, subpop4 (46)	n. d.	Sequencing	n. d.
pKP730	3115	RFLP-clone	n. d.	58	CAPS/ <i>Sau</i> 3AI	n. d.
pKP814	3131	RFLP-clone	960701, subpop4 (46)	n. d.	Sequencing	n. d.
Y102H12R	2387	YAC-end	960701, subpop4 (1359)	58	PASA	4.8
CO_Bv		Seq. cloned	960701, subpop7(299)	62	PASA	n. d.
	3080	PCR				
VRN1_Bv		Seq. cloned	960701, subpop7(299)	61	PASA	44.9
	3083	PCR				
VRN2_Bv		Seq. cloned	960701, subpop7(299)	61	PASA	88.8
	3094	PCR				

---

**8.3. Sequences of BAC ends, PCR products (non-bolting, A906001 and bolting, 93167P) that have been used in the development of markers.**

All sequences are available from Prof. C. Jung, Plant Breeding Institute, Christian-Albrechts University of Kiel, Olshausenstrasse 40, 24098 Kiel ([c.jung@plantbreeding.uni-kiel.de](mailto:c.jung@plantbreeding.uni-kiel.de)).

## 9. Acknowledgments

I would like to express my deep gratitude to my supervisor Prof. Dr. C. Jung for giving me a precious chance to conduct my doctoral research in his laboratory, as well as his excellent support, kind encouragement and valuable advices during my PhD. I would like to give my gratitude to Prof. Dr. F. Kempken for correcting my thesis as a second referee.

I would like also to express my sincere thanks to Dr. U. Hohmann for his kind guidance in lab work, as well as in writing up the dissertation. I am also very grateful to Prof. M. Hühn for his kind help in discussing linkage analysis. It was a great pleasure to discuss with him, and I appreciate his generosity in every concern.

My special thank goes to Dr. D. Cai for his help, advises and discussions. A special thank is also addressed to all colleagues and technicians in the Plant Breeding Institute, Kiel, Germany. Special mention deserves to be made of my appreciation to Jessica Lied and Susane Werner for their reading and comments on the thesis as well as for those I did not mention by name and assisted me.

The financial support from Egyptian government is gratefully acknowledged.

Here, I need to express my deepest gratitude and respect to my parents and my brothers and sister for their constant prayers, moral support and encouragement.

Finally, my gratefulness is forwarded to my wife and my daughter whom their love and continuous support tremendously helped me to finish my PhD work.



## 10. Curriculum vitae

Name: Reda Mohamed Gaafar  
Date of birth: 15/10/1970  
Place of birth: El-Gharbia, Egypt  
Nationality: Egyptian  
Marital status: Married

### Education:

1977-1983: Elementary school  
1983-1986: Prep school  
1986-1989: Senior-High school  
1989-1993: Undergraduate student at Botany Department, Fac. of Science, University of Tanta, Tanta, Egypt (B.Sc. degree in Botany, May 1993, Very Good with Honor)  
1994-1997: M.Sc. (Dec. 1997), "Evaluation of gene ecological variations among some *Cyperus* species in Egypt" Botany Dept., Fac. of Sci., Univ. of Tanta, Tanta, Egypt  
Aug. 2001-Feb. 2005: PhD student at Institute of Plant Breeding, University of Kiel, Kiel, Germany

### Professional background:

1993-1997: Demonstrator and Research student at Botany Dept., Fac. of Sci., Univ. of Tanta, Tanta, Egypt  
Since 1998: Assistant Lecturer at Botany Dept., Fac. of Sci., Univ. of Tanta, Tanta, Egypt

## **11. Erklärung**

Hiermit erkläre ich, dass die von mir eingereichte Dissertation abgesehen von der Beratung durch meinen akademischen Lehrer Prof. Dr. Christian Jung nach Inhalt und Form meine eigene Arbeit ist, dass sie nicht ganz oder zum Teil an einer anderen Stelle im Rahmen eines Prüfungsverfahrens vorgelegen hat, und dass sie weder veröffentlicht worden ist noch zur Veröffentlichung eingereicht wurde.

Kiel, den

Reda M. Gaafar