Theor Appl Genet (2007) 115:601–615
DOI 10.1007/s00122-007-0591-4

ORIGINAL PAPER

Analysis of DNA polymorphisms in sugar beet (*Beta vulgaris* L.) and development of an SNP-based map of expressed genes

Katharina Schneider · Dagmar Kulosa · Thomas Rosleff Soerensen · Silke Möhring · Martin Heine · Gregor Durstewitz · Andreas Polley · Eberhard Weber · Jamsari · Jens Lein · Uwe Hohmann · Emma Tahiro · Bernd Weisshaar · Britta Schulz · Georg Koch · Christian Jung · Martin Ganal

Received: 25 January 2007 / Accepted: 7 June 2007 / Published online: 11 July 2007

Abstract A panel of 13 sugar beet lines and one genotype each of the *Beta vulgaris* cultivars red beet and Swiss chard, and *B. vulgaris* ssp. *maritima* were used to identify polymorphisms in alignments of genomic DNA sequences derived from 315 EST- and 43 non-coding RFLP-derived loci. In sugar beet lines, loci of expressed genes showed an average SNP frequency of 1/72 bp, 1 in 58 bp in non-coding sequences, increasing to 1/47 bp upon the addition of the remaining genotypes. Within analysed DNA fragments, alleles at different SNP positions displayed linkage disequilibrium indicative of haplotype structures. On average 2.7 haplotypes were found in sugar beet lines, and haplotype conservation in expressed genes appeared to exceed 500 bp in length. Seven different genotyping techniques including SNP detection by MALDI-TOF mass spectrometry, pyrosequencing and fluorescence scanning of labelled nucleotides were employed to perform 712 segregation analyses for 538 markers in three F_2 populations. Functions were predicted for 492 mapped sequences. Genetic maps comprised 305 loci covering 599.8 cM in population K1, 241 loci distributed over 636.6 cM in population D2, and 166 loci over 507.1 cM in population K2, respectively. Based on 156 markers common to more than one population an integrated map was constructed with 524 loci covering 664.3 cM. For 377 loci the genome positions of the most similar sequences from *A. thaliana*

Communicated by R. Waugh.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-007-0591-4) contains supplementary material, which is available to authorized users.

K. Schneider (✉) · T. R. Soerensen · S. Möhring · B. Weisshaar
Department of Plant Breeding and Yield Physiology, Max-Planck-Institute for Plant Breeding Research (MPIZ), Carl-von-Linné Weg 10, 50827 Cologne, Germany
e-mail: katharina.schneider@gsf.de

K. Schneider
GSF-National Research Center for Environment and Health, Institute for Radiation Protection, Ingolstädter Landstrasse 1, 85764 Neuherberg, Germany

D. Kulosa · M. Heine · G. Durstewitz · A. Polley · M. Ganal
TraitGenetics GmbH, Am Schwabeplan 1b, 06466 Gatersleben, Germany

T. R. Soerensen · B. Weisshaar
Institute of Genome Research, University of Bielefeld, 33594 Bielefeld, Germany

E. Weber · E. Tahiro
Institute for Plant Breeding and Plant Protection, Martin-Luther-University Halle-Wittenberg, 06099 Halle, Germany

Jamsari · J. Lein · U. Hohmann · C. Jung
Plant Breeding Institute, Christian-Albrechts-University of Kiel, Olshausenstr. 40, 24098 Kiel, Germany

B. Schulz
KWS SAAT AG, Grimsehlstrasse 31, 37555 Einbeck, Germany

G. Koch
Strube-Dieckmann, Saatzucht Dieckmann GmbH & Co. KG, Kirchhorster Str. 16, 31688 Nienstaedt, Germany

were identified, but little evidence for previously presented ancestral genome structures was found.

Introduction

The breeding of sugar beet, an important crop for sucrose production in the temperate climate zone, is increasingly supported by DNA marker analysis. Anonymous RFLPs (restriction fragment length polymorphisms) were the first molecular markers which were positioned on genetic maps and tested for linkage to phenotypic traits (Barzen et al. 1992; Pillen et al. 1992). With the advent of PCR technology, RAPD (randomly amplified polymorphic DNA), AFLP (amplified fragment length polymorphism) as well as a small number of SSR (simple sequence repeat) markers complemented by some conventional isoenzyme and morphological markers permitted establishment of a map of all nine linkage groups of sugar beet covering 688 cM (Schumacher et al. 1997).

More recently, SNPs (single nucleotide polymorphisms) were established as the most abundant codominant marker class and used as efficient and robust marker systems. A pilot study based on 37 gene fragments sequenced in two inbred sugar beet lines revealed 1 SNP every 130 bp with an average nucleotide diversity π of 7.6×10^{-3} (Schneider et al. 2001). The analysis revealed disequilibrium between the theoretically possible combinations of alleles at two or more SNP loci indicative of haplotype structures (Rafalski 2002).

SNP detection has a high potential for automatization and multiparallel analysis allowing a high throughput of markers and individuals (Syvanen 2002). According to a recent study, SNP detection by MALDI-TOF mass spectrometry appears to be most accurate, and adapted amplification protocols allow cost-effective genotyping of up to 30 loci in one assay (Ragoussis et al. 2006). The pyrosequencing technology implemented as a quantitative analysis platform reliably allows distinguishing allele doses in genetic mapping of tetraploid potato (Rickert et al. 2002). However, this technology is not amenable to multiplexing. For medium throughput the SNaPshot technology was previously applied to sugar beet to generate nine multiplexed linkage group-specific SNP marker sets (Möhring et al. 2004). Among the 52 SNP markers reported in that paper, 27 were developed from sequenced RFLP probes.

For the identification of loci influencing a trait the candidate gene approach is now a widely accepted strategy (Pflieger et al. 2001). Concerning sugar beet conserved genes in carbohydrate metabolism (Schneider et al. 1999) were selected as candidates and mapped in a QTL analysis to identify loci affecting sugar quality and yield related traits (Schneider et al. 2002). Similarly genetic markers were developed from resistance gene analogues as candidates for disease resistance traits (Hunger et al. 2003; Tian et al. 2004). Initially, DNA fragments of candidate genes were amplified using degenerate primers deduced from alignments of heterologous sequences, but now the availability of EST collections from sugar beet (Bellin et al. 2002; Herwig et al. 2002) largely facilitates template selection.

Analysis of complex crop genomes such as the one of sugar beet might benefit from the existence of an ancestral genome for plant species postulated by the syntenic concept (Delseny 2004). Colinearity of large genome regions is obvious among grasses (Ahn et al. 1993; Ahn and Tanksley 1993) although detailed sequence comparisons identified numerous rearrangements and breakages of synteny (Bennetzen and Ma 2003). So far, based on comparative RFLP mapping of 117 cDNAs 26 syntenic blocks were identified between sugar beet and *Arabidopsis thaliana* some of which were also shared by sunflower, *Prunus* ssp. and potato (Dominguez et al. 2003).

Here we (1) provide the analysis of a comprehensive data set on sugar beet SNPs derived from up to 16 *Beta vulgaris* lines, and (2) show the distribution of RFLP- and EST-derived SNP markers on the nine linkage groups in three F_2 populations of sugar beet. Based on the new EST-derived markers we additionally address the question whether and to what extent the orders of genes in sugar beet and of the most similar sequences in the *A. thaliana* genome are conserved at the macrosyntenic level. The work is intended to support marker assisted sugar beet breeding and to facilitate positional cloning of agronomically important genes.

Materials and methods

Plant material and DNA

For SNP discovery, genomic DNA from a panel of 13 partly selfed diploid breeding lines including the dihaploid line K1P1 and F_1 genotypes was used (for full details see Table S1). Among those were both parents of the K1 and K2 population (*K* = KWS SAAT AG, Einbeck, Germany), and one parent and the F_1 plant of the D2 population (*D* = Strube-Dieckmann, Sülbeck, Germany) used for segregation analyses later. The panel included five more unrelated breeding lines being part of segregating populations not analysed here, an F_1 line which was derived from a cross between one of the lines of the SNP panel and an unrelated line as well as a further F_1 line which was derived from a cross between two additional breeding lines. The latter F_1 lines were included to assess SNPs qualitatively in the two additional populations which were not analysed further in the presented work. The enlarged SNP screening panel also contained one genotype each of red beet, Swiss

chard, and *Beta vulgaris* ssp. *maritima*. Plant leaf material or DNA is available against material transfer agreement from either KWS SAAT AG or Strube Dieckmann.

Genomic DNA from three F_2 populations was used for segregation analyses. The K1 population had 97 F_2 individuals, the K2 population which was previously described as population 618 (Schneider et al. 1999, 2001, 2002; Hunger et al. 2003) had 108 lines, and 100 F_2 genotypes were available for the D2 population.

Selection of target sequences for SNP marker development and marker nomenclature

DNA sequences were determined from cloned genomic DNA fragments previously used in RFLP hybridization experiments (Barzen et al. 1992, 1995; Pillen et al. 1992, 1993; Schondelmaier et al. 1997; Schumacher et al. 1997). To simplify RFLP marker nomenclature pKP probes were renamed as KI_R, and markers developed at the MPIZ Cologne were labeled MP_R both followed by four digits which incorporate the previous marker number. Sequences of either full or partial RFLP probes used for SNP discovery and segregation analyses were deposited in the EMBL/GenBank database under the accession numbers ED018871–ED018916, ED982889–ED982906, ED828992–ED828995, ED829000–ED829001, ED829016, ED829013, ED829017, ED809022 and ED809032–ED809036 and are accessible via http://www.gabi.rzpd.de. RFLP sequences were compared to all non-redundant GenBank CDS translations, PDB, SwissProt, PIR and PRF entries excluding environmental samples using BLASTX 2.2.14 (7 May 2006), and those retrieving coding sequences with e-values lower than e^{-20} were considered as genuine genes. Accession numbers, fragment sizes and annotations are listed in Table S2.

Transcribed sugar beet genes were selected from the Sputnik database (Rudd 2005) comprising about 10,000 unique EST sequences (Herwig et al. 2002) at the time of the analysis. In line with the RFLP marker nomenclature, names of EST-derived markers were composed of a prefix indicating the laboratory of origin (TG for TraitGenetics, KI for Kiel, and MP for MPIZ Cologne) followed by "_E" and a four digit number.

DNA sequencing

Purification of DNA templates to be sequenced involved the removal of unincorporated dNTPs and primers from the PCR assays using the ExoSAP-IT kit (Exonuclease I/Shrimp Alkaline Phosphatase enzyme mixture, USB, Cleveland, Ohio, USA) or 96 well NUCLEOFAST filter plates (Macherey-Nagel, Düren, Germany). Comparative sequencing of PCR products was performed either on ABI Prism 377 and 3700 sequencers using Big Dye Terminator chemistry at the ADIS DNA core facility at MPIZ or on ABI3100 sequencers with 50 cm capillary arrays using the Big Dye Cycle sequencing Kit 2.0 at TraitGenetics (all from Applied Biosystems, Darmstadt, Germany). Sequencing was also used for genotyping of SNPs where indicated.

DNA sequence analysis for SNP identification

At TraitGenetics, batches of tracefiles representing distinct loci were evaluated by the PHRED/PHRAP/POLYPHRED/CONSED software suite (Ewing et al. 1998; Ewing and Green 1998; Gordon et al. 1998; Nickerson et al. 1997). The software performed base calling, assignment of quality values, sequence assembly, automatic SNP detection and sequence editing. Polymorphisms within sequence contigs were identified by coloured tags and verified through visual inspection. At MPIZ and the University of Kiel the GCG software package (Genetic Computer group, University of Wisconsin at Madison, version 10.2) and the software package of Seqman™ (DNASTAR, GATC Biotech, Konstanz, Germany) were used for SNP identification, respectively.

DNA sequence comparisons to the *A. thaliana* genome

Sugar beet EST sequences were compared to cDNA sequences from *A. thaliana* available from TIGR (ftp://ftp.tigr.org/pub/data/a_thaliana/ath1/SEQUENCES/ATH1.cdna, version 5.0). An expected value of e^{-20} or smaller was used as a threshold to select similar sequences.

Detection of DNA polymorphisms affecting multiple bases

Deletions as small as 2 bp were examined by fragment analysis. For target amplification oligonucleotide primer pairs were designed using the Primer 3.0 program (Steve Rozen and Helen J. Skaletsky, Whitehead Institute for Biomedical Research). Primers of approximately 20 bp length were selected for a GC-content between 20 and 80%, and a melting temperature between 57 and 63°C (optimum 60°C). Forward primers were labeled with the fluorescent dyes FAM, HEX or ROX. PCR amplification was performed with approximately 25 ng of genomic template DNA, 1 × PCR buffer (50 mM KCl, 10 mM Tris–HCl pH 8.3), 1.5 mM $MgCl_2$, 0.15 µM of each primer, 0.25 mM dNTPs and 0.3 U of *Taq* polymerase (Promega, Mannheim, Germany) in a total volume of 10 µl. Thermocycling started with a denaturation step for 3 min at 94°C followed by 45 cycles of 1 min at 94°C, 1 min at 55, 60 or 65°C depending on the marker, 2 min at 72°C.

Reactions were stopped after a final extension step of 72°C for 7 min. Samples containing 0.5–1 µl PCR product of up to three differently labelled markers, 1 µl internal size standard and 9 µl Hi-Di formamide were separated using 36 cm capillary arrays on an ABI3100 machine (Applied Biosystems, Darmstadt, Germany). Alleles were detected using the GeneScan/Genotyper® software package (Applied Biosystems, Darmstadt, Germany). In few assays more than one locus segregated; in these cases loci were distinguished by small letters following the marker name.

SSCP (single strand conformation polymorphism) was detected by non-denaturing polyacrylamide gel electrophoresis as previously described (Schneider et al. 1999).

SNP detection

Pyrosequencing assays were designed using PSQ Assay Design Software (Biotage AB, Uppsala, Sweden). PCR was performed in 25 µl volume with approximately 25 ng of genomic template DNA, 0.2 µM of each primer, 0.5 U of *Taq* polymerase (Promega GmbH, Mannheim, Germany) and the same concentration of other ingredients as in fragment analysis. Thermocycling started with a denaturation step for 3 min at 95°C followed by 50 cycles of 15 s at 95°C, 30 s annealing temperature (depending on the marker 50, 55 or 60°C), 15 s at 72°C and stopped after a final extension step of 72°C for 7 min. Samples were processed with the PSQ HS 96 SNP Reagent Kit (Biotage AB, Uppsala, Sweden) and analysed on a PSQ HS 96A pyrosequencer.

SNaPshot assays were performed essentially as described in Möhring et al. (2004). Primers extended for the SNP position by a fluorescently labelled nucleotide were analysed on an ABI Prism 310 sequencer (Applied Biosystems, Darmstadt, Germany).

MALDI TOF MS analysis was carried out by the external service provider GAG BioScience GmbH (Bremen, Germany) using an Autoflex mass spectrometer (Bruker Daltonik GmbH, Bremen, Germany) to analyse extended primers.

Statistical evaluation of SNPs

SNPs, raw data on frequencies, and haplotype structures were extracted with special scripts (details available upon request). Heterozygosity for a single SNP locus was calculated as $h_j = 1 - \sum_{i=1}^{m} p_{ij}^2$ in which m is the number of alleles, P_{ij} the frequency of the allele i at this locus in a randomly mating population in Hardy-Weinberg equilibrium (Nei 1987). The same algorithm describes the quantification of PIC, polymorphism information content (Anderson et al. 1993). The average heterozygosity over all loci, H, which provides an estimate of the heterozygosity expected in a random pair of lines was determined as $\hat{h}_j = n(1 - \sum p_{ij}^2)/(n-1)$ with n being the number of lines tested for a single locus, and $\hat{H} = \sum_{j=1}^{r} h_j/r$ with r being the number of loci in the analysis. Nucleotide diversity is a measure for DNA sequence variability at a locus, and it was calculated as $\pi = K/L$ with K being the number of differences per nucleotide site and L being the sequence length in bp (Nei and Li 1979). Concerning EST-derived loci the DNA sequence from the line K1P1, parent 1 of the mapping population K1 was taken as the reference in all comparisons. For RFLP loci DNA sequences were compared to the sequence of the original RFLP hybridization probe. The average nucleotide diversity across all DNA fragments corresponds to the total number of variant sites divided by the total sequence length analysed (Li and Sadler 1991).

Map calculation

Maps were calculated using JoinMapTM software, versions 2.0 and 3.0 (Stam 1993; Stam and van Ooijen 1995) using previously mapped markers for identification of the nine linkage groups of sugar beet according to Schondelmaier and Jung (1997). Markers were grouped at LOD = 4, and map positions were determined in each linkage group according to Kosambi (1944). All recombination rates up to $r = 0.4999$ within a linkage group were taken into account to allow the inclusion of more distant loci which are not closely linked, therefore also the LOD threshold for map calculation was set to the minimum of 0.0001. Further selected parameters include a RIPPLE value of 1, a JUMP THRESHOLD of 5, and a TRIPLE THRESHOLD of 5. Maps were drawn with the program MapChart (Voorrips 2002).

Availability of the sequence alignments and marker assay information

Requests for detailed information concerning the amplification primers for the investigated amplicons, the full sequence alignments and the available marker assay information should be directed to Christian Jung or Martin Ganal. Information will be released upon signature of a material transfer agreement.

Results

Detection of sugar beet SNPs in selected lines based on RFLP and EST sequences as targets

Target sequences for SNP marker development were selected from (1) around 200 previously mapped RFLP

markers, (2) previously mapped candidate genes, and (3) approximately 10,000 unique sugar beet EST sequences.

SNP statistics in sugar beet

SNP analysis was performed at the level of genomic DNA using 315 ESTs and 43 RFLP sequences as source sequences for which genomic alignments of at least 100 bp length were available including up to 16 diploid lines of the SNP panel (Table 1). Concerning individual partly selfed breeding lines that contain residual heterozygosity, heterozygous positions were identified in 16% of the sequences generated for the analysed EST and RFLP loci on average. Heterozygous loci from parental lines were excluded from SNP identification. The ambiguities generated by heterozygous SNPs could be resolved by amplicon cloning and sequencing. Also, heterogeneous patterns from direct sequencing could be attributable to paralogous sequence variants (PSVs) as well as heterozygous allelic SNPs. Among otherwise unrelated lines the panel included two F_1 genotypes of which one parental line may be present in the same sequence analyses. In those cases, the frequency of the allele present in the parental line and the F_1 was overestimated. However, the bias introduced was minor, and the results were not corrected further. As to the qualitative analysis of SNPs, the inclusion of the F_1 lines allowed a complete and correct evaluation in populations for which only one parental line was available. But only in case of the D2 population identified SNPs were used for marker development and mapping in the presented work. However, sequence analysis was not possible in the case of indels in the analysed fragments. Only those RFLP loci were selected which did not show similarity to any other sequence at an e-value of e^{-20}. Thus, our analysis comprised three target sequence classes, the RFLP sequences which were considered non-coding, and the genomic amplification products based on EST sequences which were separated in exons and introns. Separate analyses were performed either for sugar beet breeding lines only or for the full panel including the three additional *Beta* genotypes which might contain other alleles than the sugar beet breeding pool. As the non-sugar beet genotypes were

Table 1 Statistics on SNPs identified in 315 EST-derived and 43 RFLP-derived sugar beet loci

SNP analysis of	315 EST-derived loci	43 RFLP-derived loci	
Plant material tested	Sugar beet breeding lines	Sugar beet breeding lines	Sugar beet breeding lines + other *Beta* cultivars/*B. maritima*
Average number of genotypes in alignment	9.5	8.9	10.4
Average alignment length (bp)	286	397	378
Total alignment length (bp)	89,941	17,074	16,265
Total length of exon sequence (bp)	80,631	–	–
Polymorphic alignments	265	39	40
Number of single and multiple nucleotide polymorphisms	1,486	510	577
Frequency of single and multiple nucleotide polymorphisms (1/bp)	60.5	33.5	28.2
Number of SNPs	1,246	295	345
Frequency of SNP in total sequence (1/bp)	72.2	57.9	47.1
Frequency of SNP in exons only (1/bp)	74.7	–	–
Frequency of SNP in introns only (1/bp)	55.7	–	–
Ratio transition/transversion	1.64	1.27	1.37
Average frequency of minor allele	0.24	0.28	0.24
Fraction of biallelic SNPs with a minor allele frequency lower than 0.1	0.19	0.15	0.27
Percentage triallelic SNPs	0.80	1.02	1.17
PIC-value of SNPs	0.33	0.37	0.32
Average heterozygosity of SNPs	0.36	0.42	0.36
Nucleotide diversity π (only polymorphic alignments included/all alignments included	0.0055/0.0048	0.0074/0.0070	0.0071/0.0069
Frequency of single and multiple base indel (1/bp)	1,249.2	517.4	396.7
Average number of haplotypes (with/without loci represented by just one haplotype)	2.7/2.8	2.7/2.8	3.2/3.3
Average heterozygosity of haplotypes (with/without loci represented by just one haplotype)	0.52/0.56	0.56/0.59	0.58/0.60

present in only 120 of the 315 EST-based alignments we focus here on the results obtained for sugar beet lines only. For RFLP-derived loci sequences of the other genotypes were included in 36 of the 43 analysed alignments, and we present the results for sugar beet lines only and the enlarged panel (Table 1). As PCR amplification and DNA sequencing were less efficient in the three other *Beta* genotypes the complete set of sequences was only achieved in few cases.

Nearly 90 kb of EST-based alignments were sequenced in about 9.5 genotypes with an average alignment length of 286 bp amounting to a total sequence of 875 kb. The analysis of RFLP sequences was based on more than 16 kb compared in 10.4 genotypes of the complete panel with an average alignment length of 378 bp. Excluding the non-breeding lines, the average alignment length was nearly 400 bp, in total 17 kb were sequenced in on average 8.9 sugar beet lines. In the EST-based analysis the frequency of polymorphisms for single and multiple nucleotides as multibase substitutions or indels was 1/60.5 bp. Concerning only SNPs the frequency decreased to 1/72.2 bp with one SNP every 74.7 bp in exons and one SNP every 55.7 bp in introns. In RFLP-derived sequences this value was 1 SNP every 57.9 bp considering breeding lines only, and 1 SNP every 47.1 bp when other *Beta* genotypes were added. When polymorphisms affecting more than one nucleotide were included the frequency rose to 1/33.5 bp among breeding lines, and to 1/28.2 in the full panel. In both ESTs and RFLPs transitions (i.e. substitutions of a pyrimidine base by a pyrimidine base or a purine base by a purine base) outnumbered transversions (i.e. exchange between pyrimidine and purine bases), in RFLPs the transition:transversion ratio was 1.27 when breeding lines were compared, and in ESTs it was 1.64. The average frequency of the minor SNP allele ranged between 0.24 and 0.28. In less than 20% of all biallelic SNP markers, the frequency of the minor allele was below 0.1; the percentage of fragments having less common alleles increased to 27% when only RFLP markers were tested in the full genotype panel. The fraction of triallelic SNPs was around 1% for all markers. PIC-values of SNP markers, which measure the polymorphism information content (Anderson et al. 1993) indicating the value of the marker for segregation analyses, ranged between 0.32 and 0.37. The average heterozygosity of SNP markers was 0.36 except in RFLP-derived loci amplified from sugar beet lines in which it was 0.42. Nucleotide diversity π was 5.5×10^{-3} for EST-derived loci when only polymorphic alignments were analysed, and the value decreased to 4.8×10^{-3} when also monomorphic alignments were considered. For non-coding RFLP loci π rose to 7.4×10^{-3} in the sugar beet lines (7.0×10^{-3} including monomorphic alignments), both values changed little upon the inclusion of non-sugar beet lines. Single and multiple base indels occurred at a frequency of one every 1,249.2 bp in EST-derived alignments, and one every 517.4 bp in RFLP sequence comparisons of sugar beet lines. Upon addition of the other *Beta* genotypes the indel frequency rose to 1/396.7 bp.

Haplotypes define the linkage between the alleles at two or more SNP loci if the theoretically possible allele combinations are not in equilibrium, a situation already known from sugar beet (Schneider et al. 2001). Based on 260 alignments of EST-derived loci the average number of haplotypes was 2.7 with a minimum of one and a maximum of six haplotypes in a single case. When 19 monomorphic alignments were excluded the mean value of haplotype number rose to 2.8. In 37 polymorphic alignments of non-coding RFLP sequences also 2.7 haplotypes were identified on average. To establish a relation between the length of the sequenced alignments, the total number of polymorphisms and the number of haplotypes we selected 201 EST-derived alignments with at least two polymorphisms and two haplotypes and plotted the average number of polymorphisms and haplotypes against the alignment length (Fig. 1). As the sequences of the two heterozygous F_1 lines could not be used for the identification of haplotypes, but entered the calculation of polymorphisms the latter may be slightly overestimated. A continuous increase in the number of polymorphisms was observed with alignment lengths increasing from 100 bp to 763 bp. However, the haplotype number remained almost constant at nearly three.

For EST-derived markers, the average heterozygosity of haplotypes was 0.52 or 0.56 depending on whether loci represented by just one haplotype were included or not, for RFLP-derived markers the values were slightly higher (Table 1).

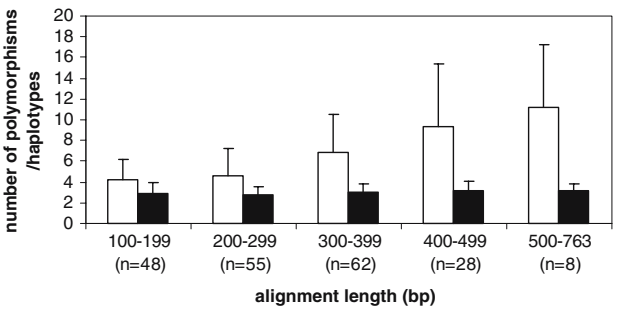


Fig. 1 Dependence of the number of polymorphisms and the number of haplotypes from the size of the amplified gene fragment. 201 EST-derived marker fragments were grouped into five classes according to their alignment length, and their average numbers of polymorphisms (*white columns*) and haplotypes (*black columns*) including standard deviations were indicated. *n* number of alignments per class

Establishment of marker assays by different methodologies using single nucleotide polymorphisms

For 538 sugar beet loci marker assays were developed, and 712 segregation analyses were performed. Information on which technique was applied for which marker in each of the three populations is available from Table S3. We developed both SNP marker systems for which the exact nature of the base substitution needs to be known, and markers for which the knowledge of the variant nucleotide(s) is not necessary.

The most common strategy used here involved initial amplification of the target sequence followed by annealing and extension of a sequence-specific primer just upstream from the SNP position. In 193 segregation analyses, differences in the extended primers were detected by MALDI-TOF mass spectrometry. Detection of different fluorescently labelled nucleotides at the SNP positions provided segregation data for 103 SNaPshot analyses. Additionally, in 121 segregation analyses, pyrosequencing was used which involves the direct determination of the nucleotides next to an SNP position. For 24 loci the allelic situation of the genotypes was determined by conventional sequencing.

Other techniques which exploit polymorphisms without knowledge of the variant nucleotide position(s) included CAPS (cleaved amplified polymorphic sequence, Konieczny and Ausubel 1993) with five segregation analyses, and single strand conformation polymorphism (SSCP) with further 140 segregation analyses. In 126 cases, length polymorphisms of amplified fragments were scored either conventionally in agarose gels or by capillary electrophoresis using fluorescently labelled primers.

All markers were scored co-dominantly except for TG_E0502, TG_E0502a, TG_E0502b, TG_E0518, TG_E0591, TG_E0578, TG_E0128, MP_E0050, MP_E0010, MP_E0139, MP_R0150, MP_E0036, MP_R0178, and MP_E0056.

Assembly of SNP-tagged loci in linkage maps of sugar beet

For map construction, 712 segregation analyses were considered of which 305 constitute the linkage map of the K1 population, 241 were used for the construction of the linkage map of the D2 population, and 166 for the K2 population. The three maps shared 138 markers which were polymorphic in two, and 18 markers which were polymorphic in all three populations (Fig. 2). For identification of linkage groups 80 previously mapped markers were integrated. Among these were 24 candidate genes related to carbohydrate metabolism and disease resistance (Schneider et al. 2002; Hunger et al. 2003), and 56 represented RFLP markers. Based on this marker framework

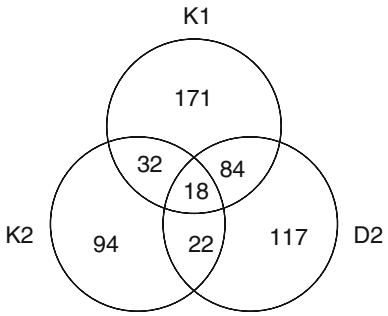


Fig. 2 Venn diagram showing the distribution of mapped markers in the three populations K1, K2 and D2

the map positions of 458 new EST-derived markers were determined.

The linkage map of the K1 population comprised 599.8 cM, the D2 map 636.6 cM, and the linkage map of the K2 population was 507.1 cM long (Fig. 3). This corresponded to theoretical marker densities of one marker per 2 cM in the K1 population, one marker per 2.6 cM in the D2 population, and one marker per 3 cM in the K2 population, respectively. However, the K1 map revealed 10 regions larger than 10 cM without marker, the K2 map had 7 such regions, and the D2 map 15 regions indicating an uneven marker distribution due to varying recombination frequencies along chromosomes. The three single maps were integrated to produce a map of 664.3 cM with 524 loci (Table S4).

Classification of expressed genes in the genetic map of sugar beet

Among 482 mapped EST sequences, 347 retrieved similar sequences from *A. thaliana* (see "Materials and methods"), which were classified manually with respect to their function (Table 2, Table S5). Most gene products were predicted to function as enzymes involved in primary or secondary metabolism, in regulation or signal transduction pathways, and in biotic or abiotic stress responses. Among mapped RFLP-derived markers ten more sequences with high similarities to coding genes were identified of which four were associated with particular functions (Table S2).

Comparison of similar loci in the sugar beet genetic map and the *A. thaliana* genome sequence

For 377 EST-derived loci present in the integrated sugar beet map (Table S4) the most similar *A. thaliana* sequence having an e-value of e^{-20} or smaller was determined. Chromosomal locations and nucleotide start positions of the respective genes were obtained from the AGI Map (http://www.arabidopsis.org/servlets/mapper).

According to Dominguez et al. (2003) three criteria had to be fulfilled as a threshold for non-random linkage

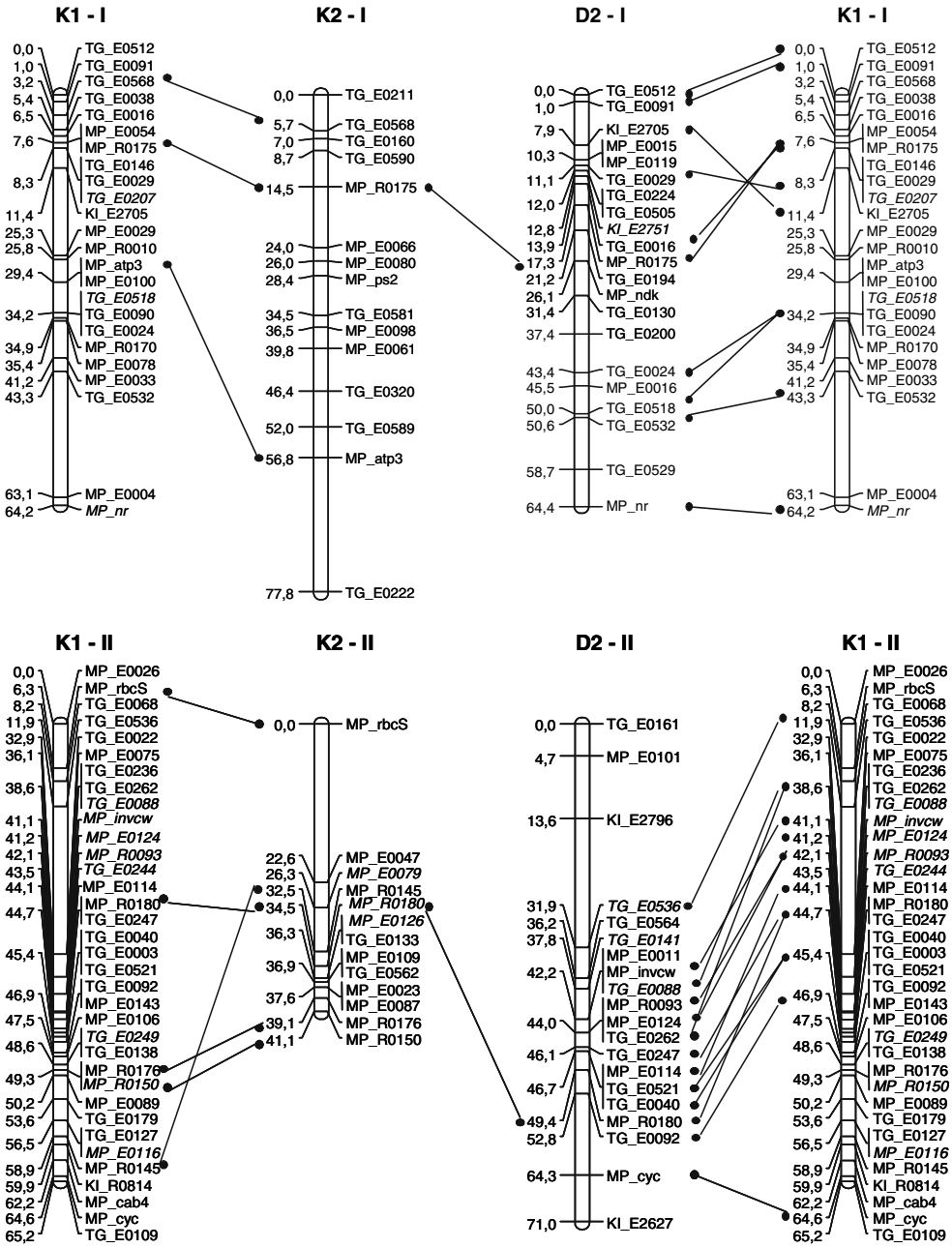
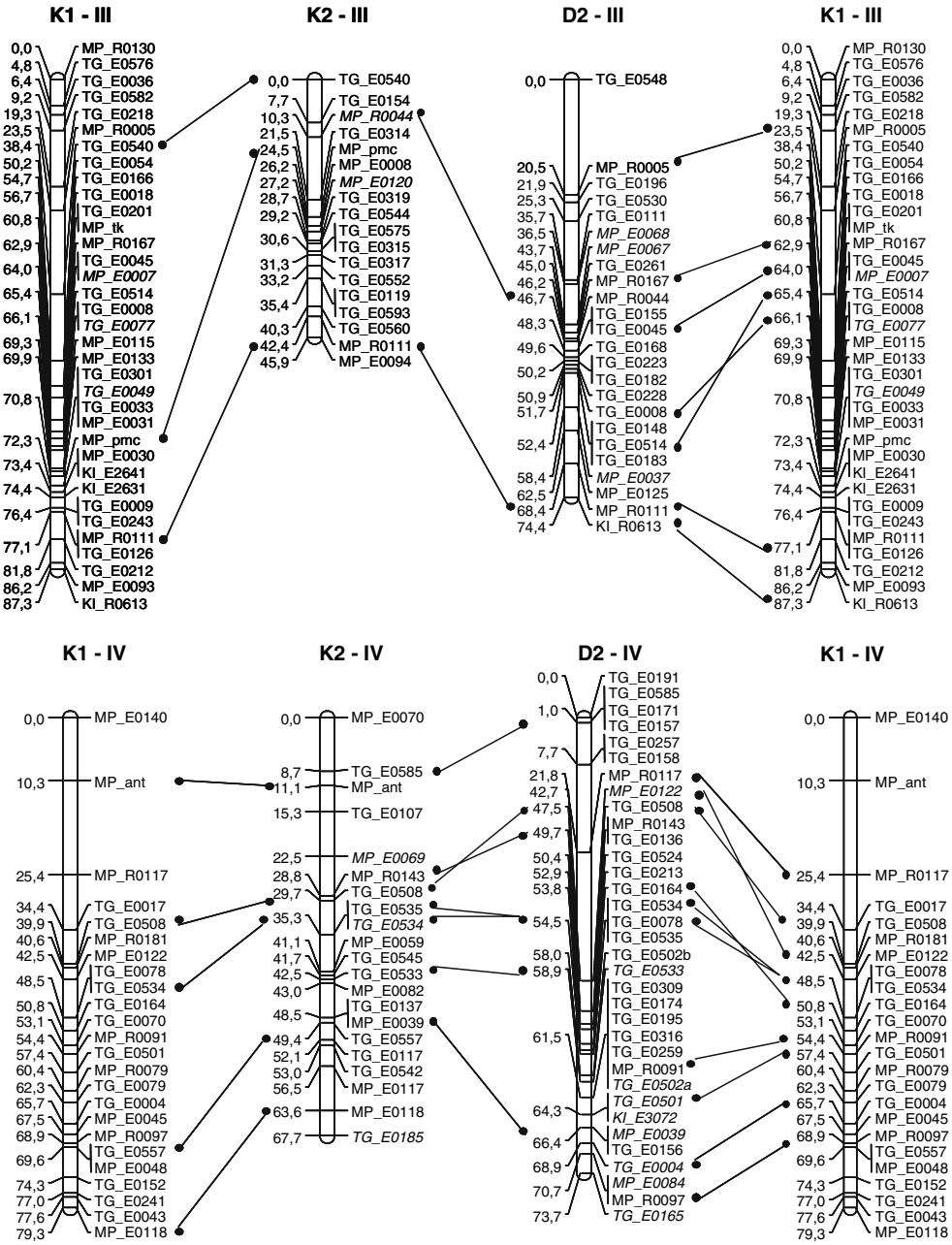


Fig. 3 Sugar beet genetic maps of the three populations *K1*, *K2* and *D2*. Marker names are given at the *right side* of each linkage group, at the *left side* cumulative genetic distances are indicated. Markers which could only be mapped in the third round of the Joinmap program are indicated in italics. Markers common between two populations are linked

between conserved sequences in two species. Loci forming a block in sugar beet had to cover no more than one percent of the genetic map corresponding to 6.6 cM. The related loci in *A. thaliana* were required to be within a distance of 1.2 Mbp which is one percent of the total genome sequence to our present knowledge. Thirdly, only blocks of at least three markers were considered.

Applying these criteria, twenty blocks comprising between three to six loci each in both sugar beet and *A. thaliana* were identified (Table 3; all markers in Table S6). Within the blocks the order of genes was co-linear with few exceptions. For each sugar beet linkage group at least one conserved block was found in one of the five chromosomes of *A. thaliana* except for chromosome 2 on which no similar sequences were found. *A. thaliana* chromosome 4 contained eight blocks deriving from six different sugar beet linkage groups with a total of 27 EST-derived markers. Some blocks in the sugar beet genome appeared to be in close proximity and/or were partly overlapping like blocks 1 and 2 of LGs V, VIII and IX. Whereas the two blocks of LG V were associated with two different *A. thaliana* chromosomes the two blocks of each LG VIII and IX were similar to different regions of *A. thaliana* chromosome IV. In total the conserved blocks identified here

Fig. 3 continued



cover 11.5% of the length of the integrated sugar beet map and 7.6% of the *A. thaliana* genome sequence.

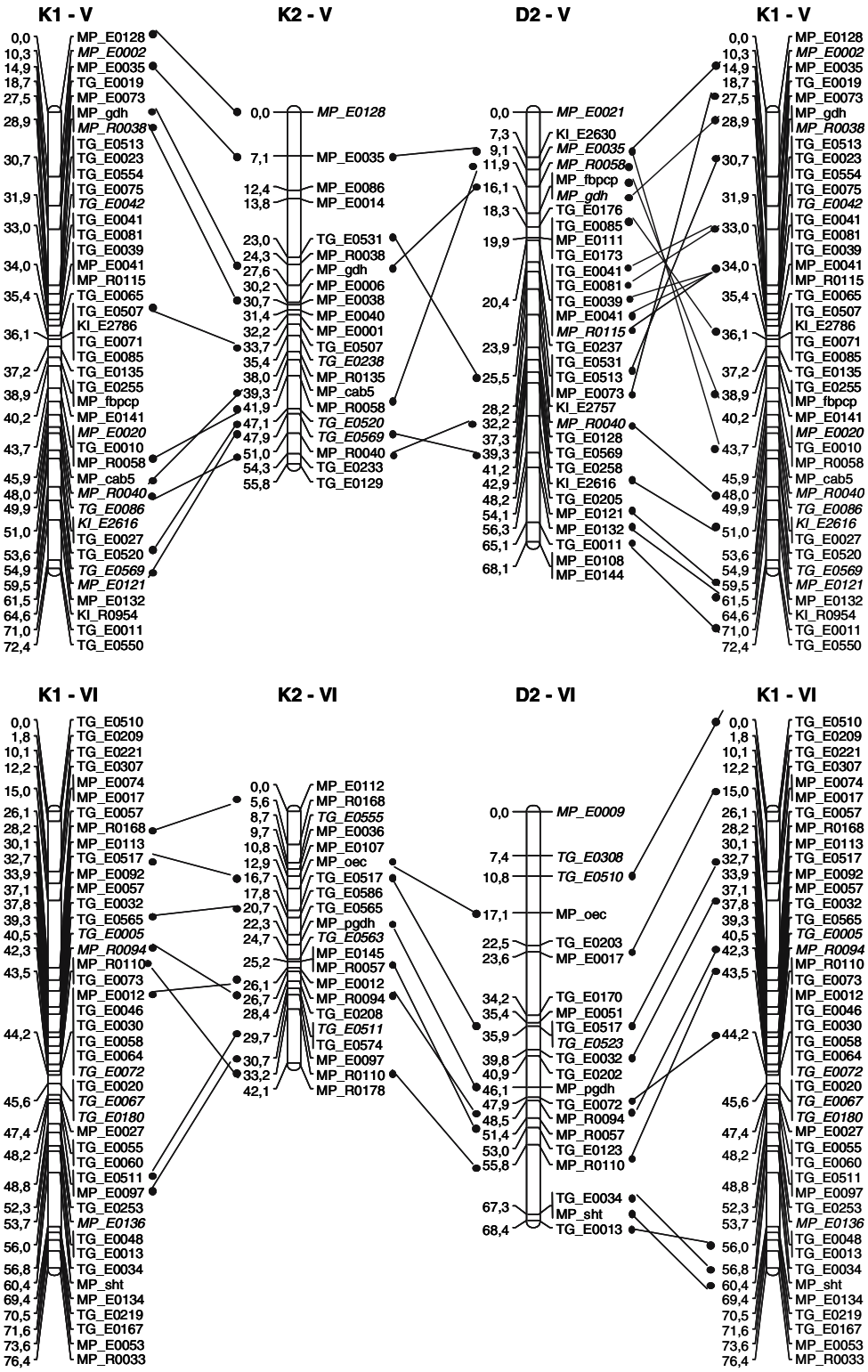
Discussion

Analysis of DNA sequence polymorphisms in sugar beet

The degree of DNA sequence polymorphism detected in a species depends largely on the type of germplasm used and the number of lines chosen for the analysis. For sugar beet a panel of 16 lines including 13 breeding lines and F_1 genotypes as well as two other *Beta vulgaris* cultivars, swiss chard and red beet, and one representative of *Beta vulgaris* ssp. *maritima* was selected. The three additional genotypes were intended to indicate the diversity outside the breeding pool.

A second factor influencing the degree of polymorphism detected is the type of loci selected for the analysis. Here, 315 EST-derived loci were taken as source to analyse fragments of expressed genes at the genomic level which are expected to be conserved at least in exons, and 43 RFLP sequences were chosen to represent non-coding

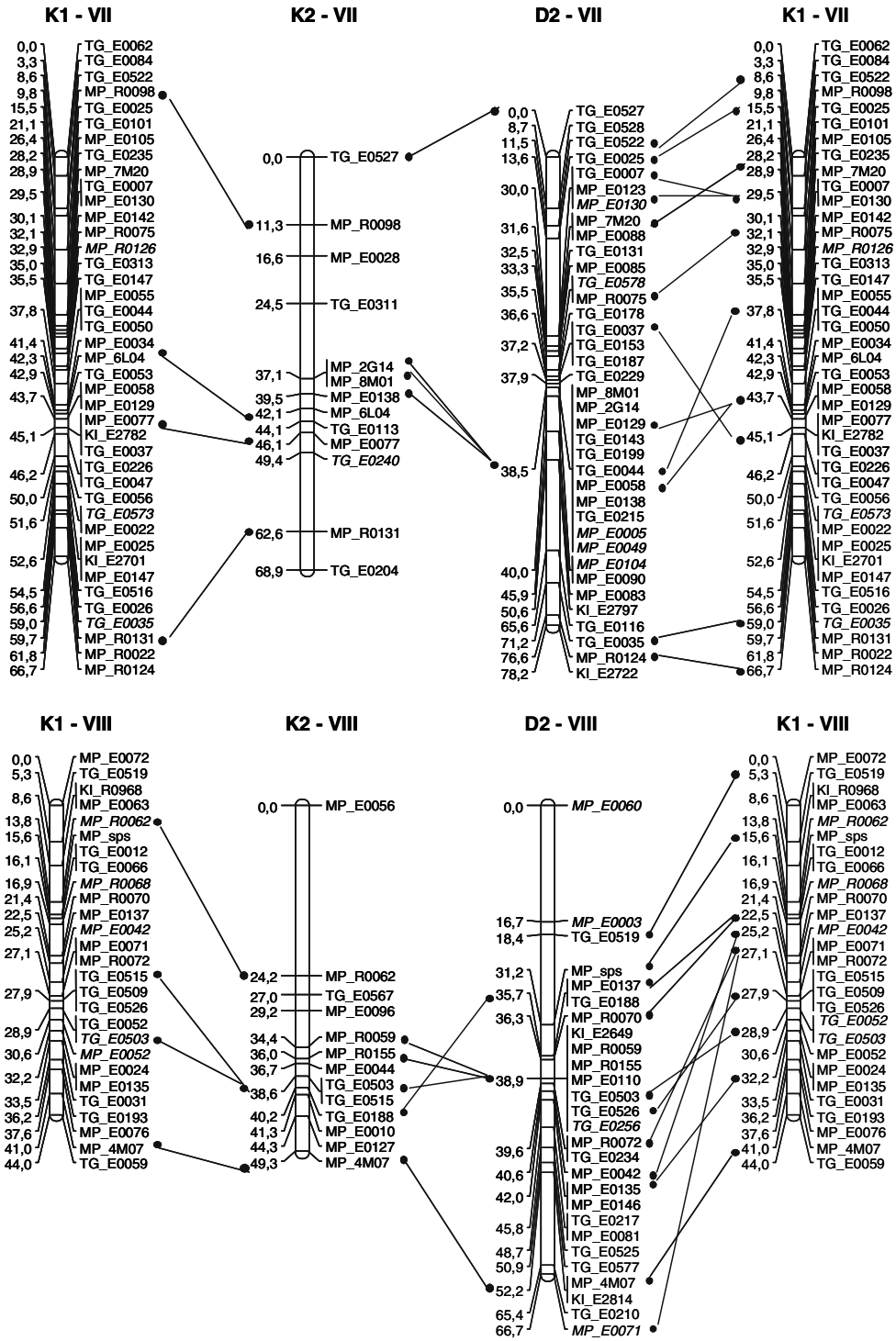
Fig. 3 continued



sequences which are expected to tolerate more mutational events leading to higher levels of polymorphism.

With regard to EST-derived loci for which we only considered sequences of sugar beet lines the frequency of polymorphisms affecting one or more nucleotides was determined to be 1/60.5 bp, for SNPs only it was 1/72.4 bp (Table 1). This frequency was higher than the previously published value of 1/130 bp (Schneider et al. 2001). However, in this pilot study, only two breeding lines and 37 gene fragments were investigated. In maize, an

Fig. 3 continued



outcrossing species like sugar beet, frequencies of one polymorphism (SNP or indel) per 41 bp and one SNP per 60.8 bp were observed when 18 gene fragments were analysed in 38 inbred lines (Ching et al. 2002). For the outbreeding forage grass species *Lolium perenne* a frequency of 1 SNP/54.1 bp was observed in the analysis of 100 candidate genes (Cogan et al. 2006). Thus, the polymorphism frequency determined here for sugar beet is in line with those of other outcrossing species. In inbreeding species polymorphisms were found at a much lower rate. SNP rates in coding regions of selfing species were 1/333 bp in rice (Yu et al. 2005), 1/200 bp in barley (Rostoks et al. 2005), 1/191 bp in soybean (Van et al. 2005) and 1/336 bp in *A. thaliana* (Schmid et al. 2003).

Fig. 3 continued

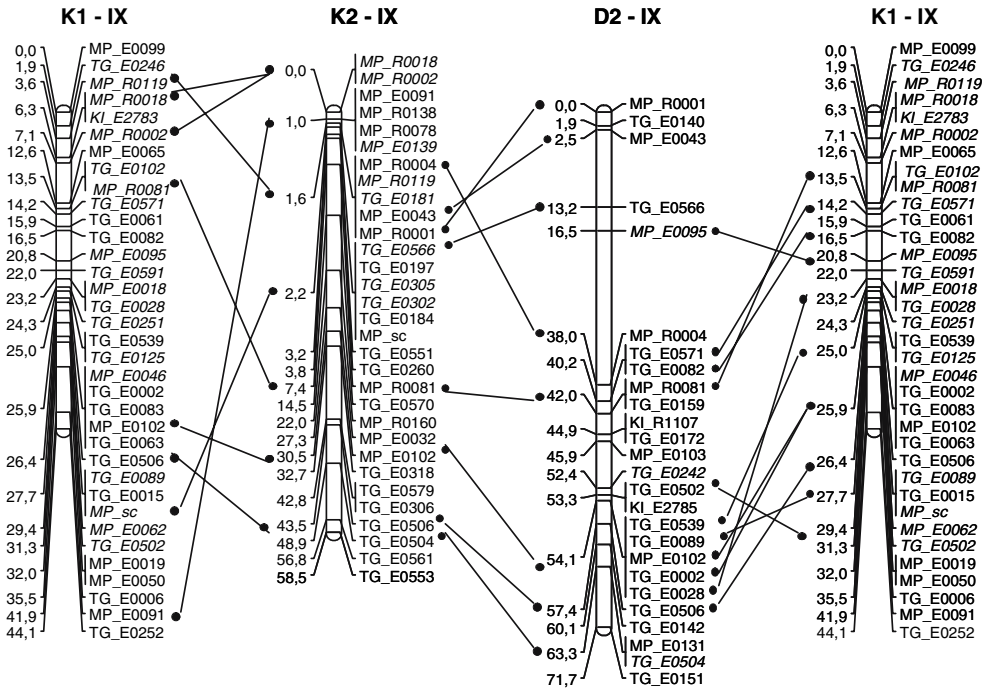


Table 2 Functional classification of 347 annotated loci mapped in the sugar beet genome

Functional classes of mapped sugar beet loci	Number of loci/class
Carbohydrate metabolism	54
Amino acid metabolism	18
Lipid metabolism	15
Nucleic acid metabolism	9
Cell cycle/cytoskeleton	10
Protein synthesis	15
Protein degradation	25
Cofactor biosynthesis	5
Secondary metabolism/cell wall biosynthesis	42
Energy/redox processes	36
Photosynthesis	9
Regulation/signal transduction	67
Disease resistance	11
Stress	14
Transport	17

The average frequency of the minor SNP allele of all biallelic loci was 0.24 except for the RFLP-derived loci analysed in sugar beet breeding lines in which it was 0.28. In case of biallelic SNPs with frequencies of 0.76 for the major and 0.24 for the minor allele the probability to find two different alleles in a random cross is $1 - (0.76^2 + 0.24^2) = 0.36$. This value expresses the average heterozygosity of SNPs (Table 1). Taking a minor allele frequency of at least 0.10 as a threshold for a common SNP, more than 80% of SNPs discovered here in the sugar beet breeding germplasm fall into this category.

The ratio of transitions to transversions was always in favour of transitions with values of 1.66 for EST-derived loci and 1.24–1.38 for RFLP-derived loci. This is comparable to the ratio of 1.53 in maize (Ching et al. 2002), but lower than the ratios of 2.12 in soybean (Van et al. 2005) and 2.39 in barley (Rostoks et al. 2005).

Concerning EST-derived sugar beet loci the average number of haplotypes detected was 2.7 or 2.8 depending on whether or not monomorphic loci represented by just one haplotype were included. In RFLP-derived non-coding loci a slightly higher value of 3.2 was only observed in the subset in which the non-sugar beet genotypes were included (Table 1) indicating the potential of germplasm from outside the sugar beet breeding pool. This result supported the previous sugar beet pilot study in which three haplotypes were identified for the majority of the ten analysed gene fragments (Schneider et al. 2001). The presented results also indicate that the linkage disequilibrium determining the haplotypes most likely exceeds the size of the tested fragments because the number of haplotypes remained stable in fragments of 100 up to 763 bp (Fig. 1). A similar situation was found in maize with 3.6 haplotypes on average per gene, and the size of haplotype structures was determined to be larger than 500 bp in elite breeding germplasm (Ching et al. 2002). In inbreeding barley there were 2.81 haplotypes when the monomorphic loci were excluded (Rostoks et al. 2005),

and also in soybean only between 1 and 5 different haplotypes were identified when 66 fragments were analysed in 25 genotypes (Zhu et al. 2003). These results indicate a deficiency in the haplotype variation irrespective of the fertilization system. For further breeding within the sugar beet germplasm displaying up to six haplotypes as revealed in our analysis no more than 3–4 SNP loci should be considered sufficient to unambiguously determine the haplotype in any given genotype. Three haplotypes reflecting the average number should be distinguishable by just two polymorphic sites. For haplotypes the average heterozygosity ranges between 0.52 and 0.6 depending on loci and germplasm used (Table 1). Thus, there is a probability of up to 60% that the haplotypes of any given locus differ in a random cross between lines of the respective *Beta* germplasm, a value similar to that found for alleles identified by SSCP and heteroduplex analysis (Schneider et al. 2001). The potential of haplotypes clearly exceeds that of single nucleotide polymorphisms to detect polymorphisms between lines. Apart from using techniques which detect multibase polymorphisms like SSCP analysis haplotype signatures are amenable to the high throughput level e.g. by multiplex SNP assays on mass spectrometry platforms (Ragoussis et al. 2006).

Construction of genetic maps based on SNPs

To our knowledge we present here the most comprehensive genetic maps of sugar beet based on co-dominantly scored markers mainly derived from sequences of expressed genes (Fig. 3). The order of anchor markers was largely conserved between the different maps. Except for a number of minor changes within small intervals of the linkage groups which may be due to the relatively small size of the populations and insufficient map resolution, obvious differences in map position assignments were only found for very few markers like MP_R0058 and MP_E0091 on linkage groups V and IX indicating potential genotyping errors. Whereas Schumacher et al. (1997) failed to integrate three maps due to different population structures, difficulties in the evaluation of common multilocus RFLP markers and incompatibilities in mapping programmes our strategy with three F_2 populations and only sequence-tagged robust marker systems allowed us to produce an integrated map based on 524 loci with a length of 664.3 cM (Table S4). The integrated map displayed only four regions larger than 10 cM without markers and has a theoretical average density of one marker per 1.4 Mbp of the sugar beet genome for which a physical size of 758 Mbp was estimated (Arumuganathan and Earle 1991).

Table 3 List of syntenic blocks identified in the genomes of sugar beet and *A. thaliana* based on 377 mapped EST-derived sugar beet markers for which similar sequences were identified in the *A. thaliana* genome with e-values smaller than e^{-20}

LG sugar beet	Block	Most proximal sugar beet locus	Number of loci/block	Genetic distance on integrated sugar beet map covered (cM)	LG Arabidopsis	Genomic sequence of A.th. covered (Mbp)
I	1	TG_E0194	3	24.1–29.9	III	21.32–22.46
	2	MP_E0078	3	37.1–42.7	III	22.75–22.81
II	1	TG_E0088	3	43.1–45.3	IV	15.86–15.93
III	1	TG_E0201	3	52.7–57.5	III	18.15–18.68
	2	MP_E0031	6	60.5–62.9	III	21.10–21.44
IV	1	TG_E0259	4	49.8–56.0	I	18.88–20.04
	2	TG_E0136	3	40.7–43.8	V	9.32–9.73
	3	TG_E0557	4	59.2–61.7	V	22.33–22.50
V	1	TG_E0513	4	26.9–32.0	I	11.46–11.87
	2	TG_E0042	3	29.8–31.0	IV	9.59–10.08
VI	1	TG_E0072	3	51.5–52.7	I	4.63–5.21
	2	MP_E0051	3	39.9–46.2	V	25.99–26.67
	3	TG_E0048	3	64.5–65.9	IV	16.96–17.83
VII	1	TG_E0056	4	47.3–49.8	III	19.64–19.70
	2	TG_E0187	5	40.0–46.1	IV	0.78–1.12
VIII	1	MP_E0042	3	42.9–48.4	IV	2.11–2.75
	2	MP_E0052	4	48.4–53.3	IV	8.84–9.65
IX	1	MP_E0139	3	31.2–33.7	IV	16.12–16.40
	2	TG_E0140	3	27.8–33.7	IV	17.92–18.22
	3	TG_E0318	3	53.3–54.4	V	24.37–24.50

Applications of SNP-based genetic maps using expressed genes as markers

Among the 538 loci presented in the genetic maps (Fig. 3) there were 347 EST-derived and 10 RFLP-derived sequences which were annotated with e-values smaller than e^{-20} indicating a potential function of the encoded gene product (Table 2). Plant function maps were generated for different traits in many species (reviewed in Gebhardt et al. 2005) following the assumption that genes associated with a particular function in a pathway relevant for the trait have an increased likelihood to control the trait. Phenotypic data on sugar yield and quality parameters were previously generated for the K2 population, and 21 QTLs were detected based on a genetic map of 192 loci of which 75 encode genes in related cellular metabolism (Schneider et al. 2002). A re-analysis of data incorporating the new functional markers would reveal whether there are candidate genes among them potentially affecting the investigated traits.

Due to DNA sequence conservation of expressed genes between species a comparison of their map positions reflect the genome organization of the respective species. Apart from syntenic regions between closely related species such as the grasses (Ahn et al. 1993) blocks of conserved gene order have also been reported for less related species such as sugar beet, sunflower, potato, *Prunus* spp. and *A. thaliana* (Dominguez et al. 2003). Whereas Dominguez et al. (2003) considered multiple redundant *A. thaliana* sequences per locus in the crop species we restricted ourselves to the most similar sequence which was required to have an e-value below e^{-20} and which was then considered an orthologue. For 377 sugar beet ESTs such *A. thaliana* sequences were identified. Under the same criteria for non-random linkage as used in Dominguez et al. (2003), we identified 20 blocks containing between three to six loci each in both sugar beet and *A. thaliana* (Table 3). Blocks from one sugar beet linkage group were not always found together on the same linkage group in *A. thaliana* indicating that the genomes of both species underwent rearrangements separately. There was no obvious co-localization of blocks with the regions described by Dominguez et al. (2003) which may be due to the different loci analysed, the exclusion of gene redundancy in our approach and the limited extension of the regions. The question whether or not the identified blocks really reflect synteny between sugar beet and *A. thaliana* can only be answered by investigating block structure in more detail in both species. However, our result suggests that co-linearity between the genomes of sugar beet and *A. thaliana* may be restricted to the microsyntenic level. Thus, in the absence of any other more closely related species from which molecular information could be transferred to sugar beet the development of comprehensive genetic and genomic tools like SNP markers, genetic and physical maps for this species are essential to understand the structure of the sugar beet genome, to identify the genes underlying agronomically important traits and to support breeding efforts.

Acknowledgment The project was supported by the GABI (Genomanalyse im biologischen System Pflanze) initiative funded by Bundesministerium für Bildung und Forschung (Germany). Funding by grants with the numbers 0312283A, 0312283B, 0312283E, 0312283H, and 0312283D/0 is gratefully acknowledged.

References

Ahn S, Tanksley SD (1993) Comparative linkage maps of the rice and maize genomes. Proc Natl Acad Sci 90:7980–7984

Ahn S, Anderson JA, Sorrells ME, Tanksley SD (1993) Homoeologous relationships of rice, wheat and maize chromosomes. Mol Gen Genet 241:483–490

Anderson JA, Churchill GA, Autrique JE, Tanksley SD, Sorrells ME (1993) Optimizing parental selection for genetic linkage maps. Genome 36:181–186

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

Barzen E, Mechelke W, Ritter E, Seitzer JF, Salamini F (1992) RFLP markers for sugar beet breeding: chromosomal linkage maps and location of major genes for rhizomania resistance, monogermy and hypocotyl colour. Plant J 2:601–611

Barzen E, Mechelke W, Ritter E, Schulte-Kappert E, Salamini F (1995) An extended map of the sugar beet genome containing RFLP and RAPD loci. Theor Appl Genet 90:189–193

Bellin D, Werber M, Theis T, Schulz B, Weisshaar B, Schneider K (2002) EST sequencing, annotation and macroarray transcriptome analysis identify preferentially root-expressed genes in sugar beet. Plant Biol 4:700–710

Bennetzen JL, Ma J (2003) The genetic colinearity of rice and other cereals on the basis of genomic sequence analysis. Curr Opin Plant Biol 6:128–133

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

Cogan NO, Ponting RC, Vecchies AC, Drayton MC, George J, Dracatos PM, Dobrowolski MP, Sawbridge TI, Smith KF, Spangenberg GC, Forster JW. (2006) Gene-associated single nucleotide polymorphism discovery in perennial ryegrass (*Lolium perenne* L.). Mol Genet Genomics 276:101–112

Delseny M (2004) Re-evaluating the relevance of ancestral shared synteny as a tool for crop improvement. Curr Opin Plant Biol 7:126–31

Dominguez I, Graziano E, Gebhardt C, Barakat A, Berry S, Arús P, Delseny M, Barnes S (2003) Plant genome archaeology: evidence for conserved ancestral chromosome segments in dicotyledonous plant species. Plant Biotechnol J 1:91–99

Ewing B, Green P (1998) Base-calling of automated sequencer traces using PHRED. 2. Error probabilities. Genome Res 8:186–194

Ewing B, Hillier L, Wendl MC, Green P (1998) Base-calling of automated sequencer traces using PHRED. 1. Accuracy assessment. Genome Res 8:175–185

Gebhardt C, Schmidt R, Schneider K (2005) Plant genome analysis: the state of the art. Int Rev Cytol 247:223–284

Gordon D, Abajian C, Green P (1998) CONSED: a graphical tool for sequence finishing. Genome Res 8:195–202

Herwig R, Schulz B, Weisshaar B, Hennig S, Steinfath M, Drungowski M, Stahl D, Wruck W, Menze A, O'Brien J, Lehrach H, Radelof U (2002) Construction of a "unigene" cDNA clone set by oligonucleotide fingerprinting allows access to 25000 potential sugar beet genes. Plant J 32:845–857

Hunger S, Di Gaspero G, Möhring S, Bellin D, Schäfer-Pregl R, Borchardt DC, Durel CE, Werber M, Weisshaar B, Salamini F, Schneider K (2003) Isolation and linkage analysis of expressed disease.resistance gene analogues of sugar beet (*Beta vulgaris* L.). Genome 46:70–82

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

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

Li WH, Sadler LA (1991) Low nucleotide diversity in man. Genetics 129:513–523

Möhring S, Salamini F, Schneider K (2004) Multiplexed, linkage group-specific SNP marker sets for rapid genetic mapping and fingerprinting of sugar beet (*Beta vulgaris* L.). Molecular Breed 14:475–488

Nei M (1987) Molecular Evolutionary Genetics. Columbia University Press, New York

Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci USA 76:5269–5273

Nickerson DA, Tobe VO, Taylor SL (1997) POLYPHRED: automating the detection and genotyping of single nucleotide substitutions using fluorescence-based resequencing. Nucleic Acids Res 25:2745–2751

Pflieger S, Lefebvre V, Causse M (2001) The candidate gene approach in plant genetics: a review. Mol Breed 7:275–291

Pillen K, Steinruecken G, Wricke G, Herrmann RG, Jung C (1992) A linkage map of sugar beet (*Beta vulgaris* L.). Theor Appl Genet 84:129–135

Pillen K, Steinrücken G, Herrmann RG, Jung C (1993) An extended linkage map of sugar beet (*Beta vulgaris* L.) including nine putative lethal genes and the restorer gene X. Plant Breed 111:265–272

Rafalski A (2002) Applications of single nucleotide polymorphisms in crop genetics. Curr Op Plant Biol 5:94–100

Ragoussis J, Elvidge GP, Kaur K, Colella S (2006) Matrix-assisted laser desorption/ionisation, time-of-flight mass spectrometry in genomics research. PLoS Genet 2:e100

Rickert AM, Premstaller A, Gebhardt C, Oefner PJ (2002) Genotyping of Snps in a polyploid genome by pyrosequencing. Biotechniques 32:592–3, 596–8, 600 passim

Rostoks N, Mudie S, Cardle L, Russell J, Ramsay L, Booth A, Svensson JT, Wanamaker SI, Walia H, Rodriguez EM, Hedley PE, Liu H, Morris J, Close TJ, Marshall DF, Waugh R (2005) Genome-wide SNP discovery and linkage analysis in barley based on genes responsive to abiotic stress. Mol Genet Genomics 274:515–527

Rudd S (2005) openSputnik—a database to ESTablish comparative plant genomics using unsaturated sequence collections. Nucleic Acids Res 33:D622–D627

Schmid KJ, Sorensen TR, Stracke R, Torjek O, Altmann T, Mitchell-Olds T, Weisshaar B (2003) Large-scale identification and analysis of genome-wide single-nucleotide polymorphisms for mapping in Arabidopsis thaliana. Genome Res 13:1250–1257

Schneider K, Borchardt DC, Schäfer-Pregl R, Nagl N, Glass C, Jeppson 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 DC, Salamini F (2001) SNP frequency and allelic haplotype structure of *Beta vulgaris* expressed genes. Mol Breed 8:63–74

Schneider K, Schäfer-Pregl R, Borchardt DC, Salamini F (2002) Mapping QTLs for sucrose content, yield and quality in a sugar beet population fingerprinted by EST-related markers. Theor Appl Genet 104:1107–1113

Schondelmaier J, Jung C (1997) Chromosomal assignment of the nine linkage groups of sugar beet (*Beta vulgaris* L.) using primary trisomics. Theor Appl Genet 95:590–596

Schumacher K, Schondelmaier J, Barzen E, Steinruecken G, Borchardt D, Weber WE, Jung C, Salamini F (1997) Combining different linkage maps in sugar beet (*Beta vulgaris* L.) to make one map. Plant Breed 116:23–38

Stam P (1993) Construction of integrated genetic linkage maps by means of a new computer package: JoinMap. Plant J 3:739–744

Stam P, Van Ooijen JW (1995) JoinMap (tm) version 2.0: Software for the calculation of genetic linkage maps. CPRO-DLO, Wageningen

Syvanen AC (2002) Accessing genetic variation: Genotyping single nucleotide polymorphisms. Nature Rev Genet 2:930–942

Tian Y, Fan L, Thurau 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

Van K, Hwang EY, Kim MY, Park HJ, Lee SH, Cregan PB (2005) Discovery of SNPs in soybean genotypes frequently used as the parents of mapping populations in the United States and Korea. J Hered 96:529–535

Voorrips RE (2002) MapChart: Software for the graphical presentation of linkage maps and QTLs. J Hered 93:77–78

Yu J, Wang J, Lin W, Li S, Li H, Zhou J, Ni P, Dong W, Hu S, Zeng C, Zhang J, Zhang Y, Li R, Xu Z, Li S, Li X, Zheng H, Cong L, Lin L, Yin J, Geng J, Li G, Shi J, Liu J, Lv H, Li J, Wang J, Deng Y, Ran L, Shi X, Wang X, Wu Q, Li C, Ren X, Wang J, Wang X, Li D, Liu D, Zhang X, Ji Z, Zhao W, Sun Y, Zhang Z, Bao J, Han Y, Dong L, Ji J, Chen P, Wu S, Liu J, Xiao Y, Bu D, Tan J, Yang L, Ye C, Zhang J, Xu J, Zhou Y, Yu Y, Zhang B, Zhuang S, Wei H, Liu B, Lei M, Yu H, Li Y, Xu H, Wei S, He X, Fang L, Zhang Z, Zhang Y, Huang X, Su Z, Tong W, Li J, Tong Z, Li S, Ye J, Wang L, Fang L, Lei T, Chen C, Chen H, Xu Z, Li H, Huang H, Zhang F, Xu H, Li N, Zhao C, Li S, Dong L, Huang Y, Li L, Xi Y, Qi Q, Li W, Zhang B, Hu W, Zhang Y, Tian X, Jiao Y, Liang X, Jin J, Gao L, Zheng W, Hao B, Liu S, Wang W, Yuan L, Cao M, McDermott J, Samudrala R, Wang J, Wong GK, Yang H (2005) The genomes of *Oryza sativa*: a history of duplications. PLoS Biol 3:e38

Zhu YL, Song QJ, Hyten DL, Van Tassell CP, Matukumalli LK, Grimm DR, Hyatt SM, Fickus EW, Young ND, Cregan PB (2003) Single-nucleotide polymorphisms in soybean. Genetics 163:1123–1134