Theor Appl Genet (2000) 100:1240–1248

S.J. Rae · C. Aldam · I. Dominguez · M. Hoebrechts
S.R. Barnes · K.J. Edwards

Development and incorporation of microsatellite markers into the linkage map of sugar beet (*Beta vulgaris* spp.)

Received: 14 July 1999 / Accepted: 27 October 1999

Abstract A set of informative simple sequence repeat markers has been identified for use in the marker-assisted breeding of *Beta vulgaris*. Highly enriched small insert genomic libraries were constructed, consisting of 1536 clones (with inserts of between 250–900 bp). Screening the clones with CA, CT, CAA, CATA and GATA nucleotide-repeat probes revealed positive hybridisation to over 50% of the clones. Of these 340 were sequenced. Primer pairs were designed for sequences flanking the repeats and, of these, 57 pairs revealed length polymorphism with 12 *Beta* accessions. Heterozygosity levels of the SSR loci ranged from 0.069 to 0.809. Heterozygosity levels were found to be similar to those detected employing RFLP probes with the same accessions. Phenetic analysis using the markers, indicated relationships in accordance with known pedigrees. Twenty three of the SSR markers were polymorphic in one or both of two F_2 mapping populations, and were placed relative to a framework of RFLP probes. The markers are distributed over all nine linkage groups of sugar beet.

Key words Simple sequence repeats · Linkage maps · Sugar beet · Marker-assisted selection

Introduction

The transfer of desirable alleles to create or improve elite cultivars can be aided by the use of molecular markers (Sorrells and Wilson 1997), which allow individuals to be selected on the basis of their genotype, both for specific regions associated with qualitative or quantitative traits and for genetic background. This marker-assisted selection (MAS) can increase selection efficiency by allowing the identification of optimal genotypes within populations of manageable size (Staub et al. 1996).

The marker systems most commonly used in plants are restriction fragment length polymorphisms (RFLPs), randomly amplified polymorphic DNAs (RAPDs), amplified fragment length polymorphism (AFLP) and simple sequence repeats (SSRs or microsatellites) (Powell et al. 1996). SSRs are tandemly repeated units of between 1–6 nucleotides and are found dispersed throughout eukaryotic genomes (Tautz 1989). The frequencies of different SSR classes vary between plant species, with $(AT)_n$ and $(AA)_n$ being the most prevalent (Akkay et al. 1992; Lagercrantz et al. 1993; Powell et al. 1996). SSRs are frequently highly polymorphic, largely due to variation in the number of repeat units, in turn thought to be caused by errors in replication (Tautz 1989), DNA repair (Sai et al. 1997), or recombination processes (Innasn et al. 1997).

Dense linkage maps for sugar beet (*Beta vulgaris*) have been published, with lengths ranging between 540 cM (Barzen et al. 1992) and 1057 cM (Pillen et al. 1993). Some subsequently developed markers have been integrated into these maps, e.g. AFLPs into RAPD or RFLP maps (Pillen et al. 1992, 1993; Barnes et al. 1996). SSRs have previously been found in sugar beet, with Mörchen et al. (1996) developing four sequence-tagged SSRs. The integration of SSR markers into linkage maps has been reported for a number of crops, including barley (Lui et al. 1996), *Sorghum* (Taramino et al. 1997), rice (Cho et al. 1998), and wheat (Stephenson et al. 1998).

Sugar beet is an outbreeding commercial crop, and the co-dominant nature of SSR markers makes them particularly well-suited for application in such species. This paper describes the development of SSRs as sequence tagged sites for *Beta* species. Twelve *Beta* accessions, representative of material used in breeding programmes, have been employed to detect useful SSR markers and

Communicated by J.W. Snape

S.J. Roe · C. Aldam · K.J. Edwards (✉)
IACR-Long Ashton Research Station,
Institute of Arable Crop Research,
Department of Agricultural Sciences, University of Bristol,
Long Ashton, Bristol BS41 9AF, UK
e-mail: Keith-Edwards@bbsrc.ac.uk

I. Dominguez · M. Hoebrechts · S.R. Barnes
ADVANTA Biotechnology Department, SES Europe N.V.,
Industriepark, Soldatenplein Z2 nr. 15, B-3300 Tienen, Belgium

measure the levels of polymorphism and heterozygosity. The markers selected have been shown to distinguish the accessions according to known pedigrees; and heterozygosity levels have been compared with those of selected RFLP probes. The informative SSRs have been incorporated into the genetic map of sugar beet, with markers identified on each of the nine chromosome pairs (Schmidt et al. 1993).

Materials and methods

Sugar beet germplasm

Accessions used for polymorphism analysis were provided from the Advanta germplasm collection (shown in Table 1) and includes commercial pollinators, maintainers (O-type), fodder beet accessions as well as *Beta maritima* and *Beta cicla* (Swiss Chard). The Lars K1 sugar beet genotype (*B. vulgaris*) was used for the construction of the highly enriched small insert genomic library.

Table 1 Germplasm used for the study of marker polymorphism

Accession	Type	Abbreviated name
BOA893	Tetraploid pollinator	T1
BOA113	Tetraploid pollinator	T2
S146	Diploid pollinator	D1
S111	Diploid pollinator	D2
D164	O-type (diploid)	S1
D214	O-type (diploid)	S2
R119–24	*Beta maritima* (diploid)	Bm
RS93615F9/2	Doubled-haploid	DH
DBB5	Fodder beet (diploid)	F1
DBC3	Fodder beet (diploid)	F2
DBD3	Fodder beet (diploid)	F3
BC3–3	*Beta cicla* (diploid)	Bc

DNA extraction

DNA was extracted from plant leaf material using either the Nucleon Phytopure kit (Amersham) or the CTAB technique (Dellaporta et al. 1983).

Genomic libraries highly enriched for SSRs

Enriched genomic DNA libraries were prepared according to Edwards et al. (1996). Briefly, 1 µg of genomic DNA was digested with *Ssp*I or *Rsa*I. Filter-immobilised oligonucleotides representing the CA, CT, CAA, CATA and GATA SSR marker classes were used to select for genomic fragments containing SSRs, and enriched fragments were amplified by PCR and ligated into a modified pUC19 vector, pJV1. Plasmids were transformed into DH10B (Life Technologies) and plated on to L-agar plates containing 100 µg ml^{-1} of ampicillin. After overnight incubation at 37°C, single colonies were transferred to microtitre plates. Following an overnight incubation at 37°C, glycerol was added to a final concentration of 25% for long-term storage at –70°C.

Hybridisation screening of enriched libraries

The genomic libraries were screened by transferring clones onto nylon membranes and probing with radiolabelled oligonucleotides (Sambrook et al. 1989). A Beckmann Biomek 2000 was used to replicate 1536 colonies onto an 8×12-cm membrane [Hybond -N (Amersham)], which was placed on L-agar containing 100 µg ml^{-1} of ampicillin, and incubated for 16 h at 37°C. The membranes were transferred onto Whatman 3mm paper, soaked in denaturation buffer (1.5 M NaCl, 0.5 M NaOH) for 5 min, and then twice onto paper soaked in neutralisation buffer (1.5 M NaCl, 0.5 M Tris-HCl, pH 7.5) for 3 min. The membranes were then washed vigorously in 2×SSC (0.3 M NaCl, 0.03 M trisodium citrate) for 5 min. After air-drying the filters, DNA was fixed by crosslinking to the membrane using the optimal crosslink function of a Spectrolinker XL-1000 UV Crosslinker (Spectronics Corporation).

Filters were screened using each of the following oligonucleotides probes: $(CA)_{15}$, $(CT)_{15}$, $(CATA)_{10}$, $(CAA)_{10}$ and $(GATA)_{10}$, each end-labelled using γ-^{32}P-ATP (DuPont-NEN) and T4 polynucleotide kinase (Pharmacia Biotech). Membranes were pre-washed with hybridisation buffer (6×SSC, 0.25% dried milk powder and 0.01% SDS) and incubated in a OV5 (Biometra) rotary oven at 50°C for 1 h. The wash buffer was replaced by 25 ml of hybridisation buffer containing 100 ng of the radiolabelled oligonucleotide and incubated at 50°C for 16 h. Membranes were washed at 50°C, four times for 5 min, with 200 ml of 2×SSC containing 0.1%SDS, then air dried and exposed to X-ray film overnight.

Plasmid purification

Plasmid DNA was extracted using the Wizard Minipreps D Purification System (Promega) from 5-ml overnight cultures in L-broth (1% w/v NaCl, 5% w/v Yeast Extract, 1% w/v Bacteriological Peptone) containing 50 µg ml^{-1} of ampicillin.

Sequencing

Plasmid DNA was added to the Dye Terminator Cycle Sequencing Ready Reaction mix (Perkin-Elmer) and cycle sequenced using the following conditions: 25 cycles of 96°C for 10 s, 50°C for 5 s, 60°C for 2 min, and then held at 25°C. The amplified product was precipitated with ethanol and sequenced using an ABI PRISM 377 DNA Sequencer (Perkin Elmer) according to the manufacturer's instructions.

Primer design

Before primers were designed, each sequence was examined, using the program Autoassemble version 1.4 (Applied Biosystems, Perkin Elmer Corporation), for sequence similarity to other clones. For those clones containing three or more repeats of a SSR element, pairs of primers of 21 nucleotides were designed from the regions flanking the repeat, following the guidelines described by Bej et al. (1991). Oligonucleotide primers were synthesised by Genosys.

Microsatellite primer characterisation

The forward primer of each pair was end-labelled with 20 nCi [γ-^{33}P-ATP (DuPont-NEN] per 30 pmol of primer. The SSR loci were then amplified in 25-µl reactions containing: 50 ng of DNA template, 1×PCR buffer (GibcoBRL), 1 U of *Taq* DNA polymerase (Perkin Elmer), 200 mM of each dNTP, and 45 pmol of both primers. Thermocycling was performed using a Perkin Elmer 9600, using 35 cycles of 94°C for 30 s, 56°C for 1 min, 72°C for 1 min, and a final extension at 72°C for 10 min prior to holding at 25°C.

An equal volume of formamide containing 0.4% w/v bromophenol blue and 0.25% w/v xylene cyanol FF, was added to each reaction, prior to electrophoresis on 6% denaturing polyacrylamide gels [19:1 acrylamide:bis-acrylamide – EASIgel (Scotlab)]. The gels contained 7 M urea and 1×TBE (89 mM Tris, 89 mM boric acid, 2 mM EDTA, pH 8.4). Electrophoresis was at 55-W constant power for between 3–4 h (Sambrook et al. 1989). The gels were dried onto Whatmann 3MM paper and exposed to X-ray films.

Phenetic analysis

Autoradiograms were scored for the presence or absence of "alleles" (defined operationally as bands of the same size) generated by each primer pair. The information content of each SSR marker was calculated using its expected heterozygosity (H_e), the probability that two individuals taken at random from the population considered will have different alleles at a locus (Nei 1987), and defined as:

$H_e=1-\sum(p_i)^2$,

where (p_i) is the frequency of the *i*th allele in the population studied for a particular SSR locus.

Cluster analyses were based on similarity matrices using the unweighted pair group method arithmetic average (UPGMA) (Nei 1987) with the Jaccard coefficient (presence of shared bands). Analysis was performed using Genstat V3.2 (Grant 1997) and relationships between accessions were visualised using a dendrogram.

Linkage analysis and map construction

Two separate F_2 mapping populations were used. Population B was produced by selfing a single plant which was the F_1 between cloned plants BUF1 (from the genotype S111.217, a diploid sugar beet breeding line) and DBC5 (from the genotype K9S30857, a fodder beet line). Population C also resulted from the selfing of a single F_1 plant made between male-sterile (genotype SBMS501) and male-fertile (SBDP6353) sugar beet plants. The SSR markers were pre-screened for polymorphism in each mapping population, and polymorphic loci applied to 96 and/or 178 plants for populations B and C, respectively. Amplification was performed as detailed in the characterisation section, except that unlabelled primers were used, and the polymorphisms were detected by electrophoresis in 3% MetaPhor (FMC) agarose gels.

A framework of RFLP probes, chosen to give approximately even coverage of the nine sugar beet linkage groups, was applied to the mapping populations, as described previously (Barnes et al. 1996); 32 RFLP probes were used in population 'B' and 48 in population 'C'. A third reference population "A" (described in Barnes et al. 1996), which was polymorphic for common RFLP markers, was also used to help merge the information obtained from populations B and C.

The segregation of all markers was checked by a χ^2 test. Markers showing distorted segregation at $P<0.05$ (i.e. not 3:1 or 1:2:1 for dominantly or co-dominantly scored loci, respectively: see Table 3) were excluded from the initial framework map construction. Linkage analysis for all the markers was performed using the program JoinMap (Stam and Van Ooijen 1995), as suggested in the manual, using a minimum LOD value of 3.0 and the Kosambi mapping function. These parameters were maintained with the exception of the 'Jump' value that needed to be increased to 5.0 for linkage group 5 to be fully combined. Fixed orders were not applied when combining linkage groups unless discrepancies with the known RFLP probe orders had already been encountered.

Results

Abundance of SSRs within the enriched sugar beet genomic libraries

A total of 1536 clones were screened for the presence of CA, CT, CAA, CATA and GATA repeats. The *Rsa*I library revealed 763 (49.7%) positive clones, and, the *Ssp*I library 918 (59.7%).

The frequencies of the dinucleotide classes found here (Table 2) are similar to those published Mörchen et al. (1996); who reported that CA repeats are most common in sugar beet genomic libraries. A number of clones (19.8%) contained two different repeat motifs and 0.6% contained three different repeat classes. The results also confirm the observation that SSRs with a large number of repeat units (>20) are less frequent than SSRs with a small number of repeat units (Lagercrantz et al. 1993; Morgante and Olivieri 1993; Wang et al. 1994). Thus we can assume that at a first approximation the loci analysed here are roughly representative of the SSRs found within the sugar beet genome.

Table 2 Frequency of SSR motifs in two sugar beet genomic libraries revealed by colony hybridisation

SSR Classes	*Rsa*I library	*Ssp*I library
CA	617 (40.2%)	884 (57.6%)
CT	53 (3.5%)	34 (2.2%)
CAA	49 (3.2%)c	N/A
CATA	39 (2.5%)	N/A
GATA	5 (0.3%)	N/A
Total Positive Clones	763 (49.7%)	918 (59.7%)

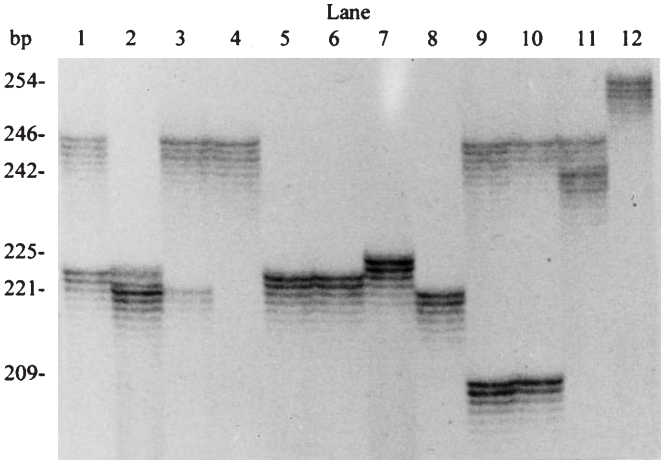

Fig. 1 Analysis of *Beta* species accessions using the microsatellite 6D4. *Lane numbers 1–12* refer to the amplification products from the following accessions; T1, T2, S1, S2, D1, D2, *B. maritima*, DH, F1, F2, F3, *B. cicla*. The band sizes (bp) of the observed alleles are indicated on the left and were extrapolated by electrophoresing the sequencing products of the pBSMB sequencing control on the same gel

Sequencing of SSR-containing clones

From the library screening, 340 clones were sequenced. Of these, 66.4% were unsuitable for designing primers due to redundancy, no clear SSR being found, or else the SSR was too close to the vector linker site. A total of 114 clones (33.6%) were used to design primers.

Screening of the primers to reveal SSRs

Primer pairs were first tested by amplification of the cloned fragment from which their sequence was derived:

Table 3 Characterised and mapped SSRs; including heterozygosity level and Chi-square value from the segregation of their alleles in two mapping populations

Clone accession	Repeat	Major Repeat	Observed size range (bp)	Number of alleles	He	Expected segregation ratio	Probability value Chitest
Ms3A4	$(TT)_3$–$(AC)_{14}$	CA	154–174	3	0.472	1:2:1	0.145[a]
Ms3D2	$(CAA)_3$–$(CAA)_3$	CAA	308–349	5	0.704	1:3	0.790[b]
Ms3G8	$(CCG)_4$–$(GA)_4$	CCG	75–97	3	0.165	1:2:1	0.153[a]
Ms4D3	$(AC)_{17}$–$(CA)_4$	CA	120–146	4[c]	0.694	1:2:1	0.845[b]
Ms5A1a	$(TA)_3$–$(AC)_{18}$	CA	152–180	6[c]	0.790	1:3	0.450[b]
Ms5A1b	(secondary loci)		173–232	8[d]	0.515	1:3	0.307[b]
Ms6B7	$(AC)_{23}$ (with mapping population C)	CA	146–152	4	0.564	1:3	0.481[b]
Ms6B7	(with mapping population B)					1:2:1	0.055[a]
Ms6D4	$(TC)_8$–$(CT)_3$–$(TT)_3$	CT	209–246	7	0.763	1:2:1	0.074[a]
Ms9B11	$(AC)_{16}$–$(AA)_8$–$(AA)_4$	CA	238–275	6	0.779	1:2:1	0.791[b]
Ms10B9	$(AT)_4$–$(CA)_5(CACAA)_3(CATA)_{22}$	CATA	301–320	7[c]	0.796	1:2:1	0.789[a]
Ms11A5	$(CA)_{18}$	CA	273–308	5	0.737	1:2:1	0.933[b]
Ms12G10	$(TC)_6$–$(CT)_{17}$	CT	110–159	4[c]	0.615	1:2:1	0.767[a]
Ms13D2	$(TC)_{13}(CA)_8$	CT	208–231	4	0.651	1:2:1	0.129[a]
Ms13E10	$(CAA)_{14}$–$(CT)_3$–$(AA)_3$	CAA	363–395	5	0.718	1:2:1	0.661[a]
Ms14A2	$(AC)_{15}$-$(TA)_5$	CA	171–226	6	0.742	1:3	0.003[b]
Ms14H6	$(TT)_6$–$(GT)_3$–$(CT)_{26}$ $(GT)_{13}$–$(GT)_5$	CT	239–283	4[d]	0.698	1:2:1	0.207[b]
Ms15A11	$(CA)_{10}$–$(CA)_6$–$(CA)_3$	CA	117–170	4[d]	0.658	1:3	0.2×10^{-3}[b]
Ms15G8	$(TA)_3(AC)_3$–$(AC)_{16}$–$(AC)_{11}$	CA	157–217	6[c]	0.450	1:2:1	0.096[a]
Ms16C1	$(AA)_6$–$(CA)_{16}$	CA	205–240	9[c]	0.809	1:2:1	0.175[a]
Ms17B7B	$(CT)_{17}(CA)_{20}$–$(CT)_3$–$(CT)_8$–$(TC)_3$ –$(CCT)_3$–$(CT)_4$–$(CT)_3$	CA	242–341	5[c]	0.584	1:2:1	0.482[a]
Ms17C7	$(AT)_3$–$(TT)_3$–$(AC)_6$–$(AC)_{14}$–$(AC)_{27}$	CA	213–279	5	0.692	1:2:1	0.992[a]
Ms17F2	$(CAA)_3$-$(CA)_{28}$	CA	120–221	7	0.773	1:2:1	0.673[b]
Ms17F11	$(AT)_3(AC)_{23}$–$(AG)_3$	CA	199–207	3	0.357	1:3	0.431[b]
Ms18A7	$(CA)_3$–$(AC)_9$–$(CA)_6$–$(CA)_4(CA)_{12}$ –$(GA)_4$	CA	241–317	4[c]	0.556	1:3	1.000[b]

[a] Segregation pattern observed with mapping population b
[b] Segregation pattern observed with mapping population c
[c] In at least one of the accession lines used in the polymorphism screen no amplification was observed. When an accession has such a result it was excluded from the calculation of the SSR's heterozygosity
[d] On amplification the number of bands observed from one or more lines was greater than that expected given the accessions known ploidy levels. Lines giving such results were excluded from the calculation of heterozygosity levels

91 of the 114 pairs produced a product of the base-pair size calculated from their nucleotide sequencing. These 91 primer pairs were then tested for their ability to reveal polymorphism across the set of lines (see Fig. 1 for an example), following electrophoresis on sequencing gels. Twenty four of the primer pairs failed to give simple scorable patterns and a further ten SSRs were monomorphic, reducing the number of the primer pairs investigated to 57; a second distinct locus was observed in five cases, giving a total of 62 SSR loci.

Table 3 summarises the polymorphisms detected by the mapped loci, along with their SSR motifs. It is noteworthy that the allele sizes do not always differ by simple multiples of the repeat length (Fig. 1). As expected, loci that have the highest number of alleles tend to give higher levels of heterozygosity; with this limited data set, however, there is no relationship between repeat number and heterozygosity. The accessions of different ploidy are capable of revealing different numbers of alleles – only one for doubled-haploids and up to four for tetraploids (though note that from the 62 loci, none revealed more than three alleles in the tetraploid accessions). In six cases, amplification resulted in more bands than expected from the known ploidy levels. For 19 of the SSR loci, no amplification occurred with one or more of the accessions, indicating the presence of null alleles. Since some null alleles may have gone undetected in plants where another allele has amplified this estimated null frequency must be considered a minimum.

Of the characterised SSR markers, ten could be described as perfect (using the nomenclature of Weber and

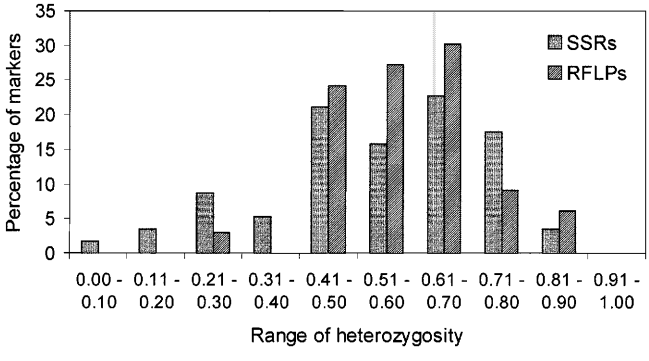

Fig. 2 Comparison of the heterozygosity levels between SSR and RFLP markers. The histogram displays SSR markers and RFLP probes sorted by the heterozygosity levels they reveal

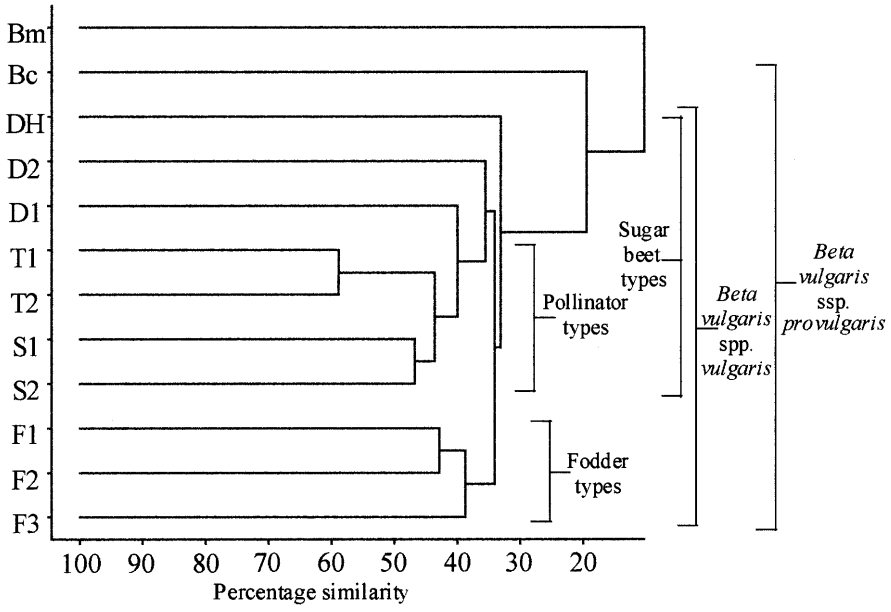

Fig. 3 The phenetic relationship of *Beta* germplasms. The genetic distance tree (UPGMA) of ten sugar beet breeding lines and *B. cicla* and *B. maritima*. The indicated groupings show agreement with the evolutionary affinities found by Ford-Lloyd and Williams (1975)

May 1989), of which five also contained a run of five or more single nucleotides (SNR), e.g. A_6. Fourteen of the SSRs were interrupted, and 32 were classified as compound (Table 3).

Comparison of SSR marker polymorphism to that produced by RFLP probes

The levels of polymorphism detected by the SSR markers developed here were compared with that produced by a set of 33 RFLP markers from which data was for the same 12 accessions (Barnes et al. 1996). Figure 2 compares the heterozygosity revealed by the two marker systems, and shows that SSR markers produce a similar expected heterozygosity (0.545) to that produced by the RFLP probes (0.583).

Phenetic analyses based on the SSR polymorphism

Genotypic data from the 62 SSR loci were used to analyse the phenetic relationship of the 12 *Beta* spp. accessions, which is shown as a dendrogram in Fig. 3. The relationships conform to expectations for the sections of *Beta* previously described (Ford-Lloyd and Williams 1975; Hansen et al. 1999; McGrath et al. 1999). *B. maritima* and *Beta provulgaris* have evolutionary affinities to the ancestral maritime form, section *B. provulgaris* having diverged into the *B. cicla* (Swiss chard and spinach beet) and *B. vulgaris* (sugar beet, fodder beet and red beet) sections. From Fig. 3, the fodder beet accessions form a group distinct from the sugar beet lines, and within the sugar beet group the pollinators form a cluster in which the tetraploids are more similar to one another. The diploid O-type accessions are less similar than the pollinators within the sugar beet grouping, whereas the doubled-haploid accessions appears to be the most distinct of the *B. vulgaris* section (see Discussion).

Linkage mapping

SSR polymorphism screening revealed 12 and 14 markers that were polymorphic in populations ‘B’ and ‘C’, respectively. Chi-sqaure-values for SSR marker segregation are shown in Table 3. SSRs which had a segregation ratio of approximately 1:2:1 were scored as co-dominant markers, whereas SSRs which had a segregation ratio of approximately 3:1 were scored as dominant markers (Table 3).

Linkage mapping was first performed with each population individually. The 32 RFLP marker framework (covering 351 cM) of population B covered all nine linkage groups of markers, and 12 SSR loci could be placed at a LOD-threshold of 3.0 (Stam and Van Ooijen 1995).

Similarly, 14 SSRs were mapped in population C, from which one linkage group was excluded because of severe segregation distortion. The total size of the framework for population C, comprising 48 RFLP probes was 401 cM. One SSR marker (Ms5A1) revealed two loci in this population.

Data from the two populations were merged with data from a third population (A: Barnes et al. 1996) to create the combined map, with nine linkage groups (note that the data for linkage group 8 are derived solely from population B), without any statistical problems, using JoinMap (Fig. 4). The total size for the joined framework is 458 cM, with the largest gap between two markers being 23 cM. Twenty three SSR markers have been positioned on the joined map, although this includes one SSR marker (Ms14A2) which has a distorted segregation. There are clusters of SSR markers, within 1 cM of each other, on

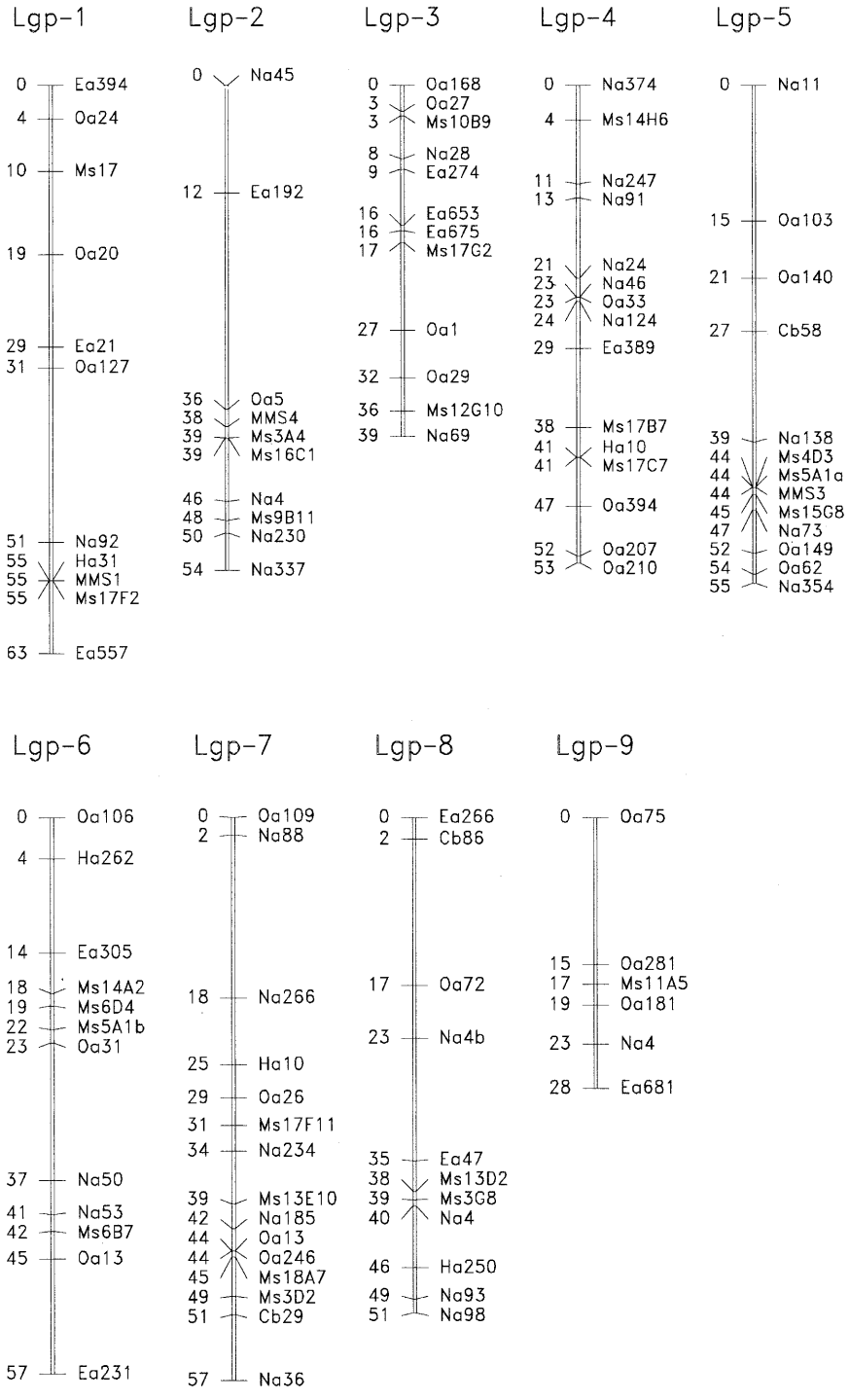

Fig. 4 Sugar beet genetic linkage map. Markers prefixed with *Ms* are the SSR markers developed in this study. The three markers prefixed with *MMs* were developed by Morchen et al. (1996). Markers not prefixed with either Ms or MMs are RFLP framework markers chosen from the Advanta linkage map (Barnes et al. 1996)

linkage groups 2, 5 and 8. Sequence alignment conducted on the clustered SSR markers found no homology.

Discussion

The use of enriched libraries for efficient SSR isolation has been demonstrated for a number of species, and the proportion of SSR-containing clones reported here is typical of that reported elsewhere (Edwards et al. 1996). Although such enrichment may introduce bias between microsatellite classes and select for longer than average SSR loci, the frequency of polymorphic loci of different sequence classes is in broad agreement with that reported in other plant species (Kresovich et al. 1995; Weising et al. 1996; Jarret et al. 1997), and within *Beta* itself (Mörchen et al. 1996).

Sequence analysis of 340 clones has resulted in 62 polymorphic loci, of which 23 have been mapped in the present study. Only a small number of clones (24) lacked

any SSR motif, showing that the hybridisation selection is an efficient procedure. Two major factors affected our ability to design primers for the remaining clones. Firstly, a number were rejected due to apparent duplication within the library, presumably caused by selective amplification of certain sequences during the enrichment, or possibly by the presence of SSRs within families of repeated sequences (Schmidt et al. 1991). Secondly, a number of clones contained insufficient flanking sequences for the design of suitable primers. Thus from the original 340 clones sequenced only 114 primer pairs could be designed. Of these only 57 pairs proved useful in the accessions tested. A similar success rate has been recently reported by Rossetto et al. (1999) for Tea Tree.

A striking feature of the sequence motifs detected in these enriched libraries is the proportion of clones that contain compound SSRs – several different runs of the same repeat, separated by other sequences, or contiguous runs of different repeats. This explains the presence of certain repeats (e.g. CGG) in the SSRs, even though these were not used to select the sequences during enrichment. The fact that the great majority of SSRs are compound appears to have little influence on the levels of polymorphism detected in the limited germplasm screen that we have carried out. Levels of heterozygosity are comparable to those detected by RFLPs (which had been pre-selected for their polymorphism during the construction of the RFLP map). The complex structure of the loci makes it difficult to explore the relationship between repeat number and levels of polymorphism, although no obvious trend (comparing either the sum of the repeat units or only the number of repeats of the largest repeat) can be seen with the loci described here. Our work has produced further evidence that even short repeats can be highly polymorphic (Mörchen et al. 1996), e.g. 15B2 with its repeat $(CGG)_3$ and possessing a H_e of 0.426.

The majority of primer pairs appear to amplify a single locus, although five gave a clear second locus. Six other markers, although apparently single loci, gave more bands than theoretically possible for at least one plant, having regard to its ploidy (e.g. two bands in a doubled haploid, or three in a diploid), indicating the possible presence of secondary loci. We presume that this is due to the occurrence of duplicated chromosome regions (Barnes et al. 1996; Pillen et al. 1996). There are other reasons to suggest that the estimates of allele number and heterozygosity are low. In the screen of primer pairs on the 12 *Beta* accessions used here, a number of pairs gave no amplification in some of the accessions, indicating that "null" alleles can exist. Any plant heterozygous for such alleles would appear to be homozygous for the other allele. As expected, the nulls were seen most frequently in the doubled-haploid lines (because their presence cannot be masked by heterozygosity), and in the "wild" *B. maritima* and *B. cicla* accessions. The conservation of primer regions will be highest in accessions most similar to that used to construct the genomic libraries (Westman and Kresovich 1998). It has been reported in other species that sequence conservation is often low in regions immediately flanking the SSRs (Edwards 1999), which could lead to the absence of primer sites and to non-amplification, and hence to null alleles.

The observations that sequence variation occurs in a locus, without necessarily being reflected in a length polymorphism, and that null alleles may be relatively frequent, could lead us to question the utility of microsatellite markers as a phenetic tool. It could be argued that simple identity in amplicon length is not a guarantee of identity by descent. The results presented here, however, show that, at least within the accessions used (cultivated beets and some near relatives), the variation revealed correlates well with the expectations from the known relationships between the accessions (breeder information and see Ford-Lloyd and Williams 1975) and predictions from RFLP polymorphism. The observation that the tetraploid lines clustered the closest among the sugar beet germplasm, and that the doubled-haploid line was relatively distant, is related in part to their biology, and to the statistical method used to calculate their similarity. The use of the Jaccard coefficient, where only the presence of shared bands are accounted for in the simple matrix-cluster analysis, is advocated because of the varying ploidy levels of the accessions and because null alleles and/or extra bands were encountered. However, because a tetraploid has more possible alleles at a given locus, and a doubled-haploid less, this gives a double-haploid fewer chances than a tetraploid of finding an allele in common with another line. Thus the dendrogram reflects in part the known genetic relationships (wild accessions being placed most distantly) and in part the levels of heterozygosity (with most heterozygous lines tending to cluster together).

For molecular markers to be useful not only for phenetic but also for breeding applications, it is important to establish their coverage of the genome, and to anchor their positions on genetic maps. The latter are essential to permit comparability between programmes, and to enable the selection of a set of markers for genome scanning and background genome selection. In the present study we have mapped 23 of the SSRs in two mapping populations, and found markers on each of the nine beet linkage groups. The markers that have been mapped in the F_2 populations (F_2s were chosen because of their high information content; Staub et al. 1996) were those showing the clearest polymorphism in the two populations, which could be mapped easily using agarose-gel electrophoresis. There is, however, no reason to suppose that these represent a biased subset of beet SSRs.

The uneven distribution of some types of molecular marker in the genetic linkage map of sugar beet has already been described (Barnes et al. 1996), and is likely to be due to a combination of effects. One major influence is the possibility of localisation of recombination, so that equally spaced markers will appear to cluster in regions of low recombination (Barzen et al. 1992; Schondelmaier et al. 1996; Schumacher et al 1997). A second effect could be due to the type of sequence that

generates the marker; where these can be derived from repeated sequences (for example, some types of AFLP marker) polymorphisms may be concentrated in heterochromatic or other repeat-rich regions of the genome. The distribution of the SSR markers over the linkage groups in this study, however, appears to be fairly even, although major conclusions cannot be drawn with this sample size. The numbers of markers in each sequence class are even smaller, although so far there is no evidence to support the suggestion that CA repeats, being located mainly around centromeric regions by in situ hybridisation, are likely to be unsuited for use (Schmidt and Heslop-Harrison 1996). Whilst markers of this major repeat class are responsible for two of the three SSR marker clusters, they are the most common (17 out of 23) and are distributed throughout this map. DNA sequences have been checked to ensure that these clustered markers are not from the same physical locus.

Acknowledgements IACR – Long Ashton is funded by the Biotechnology and Biological Sciences Research Council of the United Kingdom. S.J.R. was supported by an Industrial CASE award from the BBSRC to K.J.E, and by Advanta. The microsatellite markers are available for academic research applications. Applications should be addressed to Dr. S. R. Barnes. The experiments carried out in this study comply with current UK and EU laws.

References

Akkay MS, Bhagwat AA, Cregan PB (1992) Length polymorphism of simple sequence repeat DNA in soybean. Genetics 132:1131–1139

Barnes S, Massaro G, LefPbvre M. Kuiper M, Verstege E (1996) A combined RFLP and AFLP genetic map for sugar beet. Proc 59th IIRB Congress:555–560

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

Bej AK, Mahbubani MH, Atlas RM (1991) Amplification of nucleic acids by polymerase chain reaction (PCR) and other methods and their application. Crit Rev Biochem Mol Biol 26: 301–334

Cho YS, Hong SK, Song MT, Moon HP, Lee JH, Kim NS (1998) Comparison of genetic variation among rice varieties detected by RAPD, AFLP and SSRP. Korean Journal of Genetics 20:117–128

Dellaporta SL, Wood J, Hicks JB (1983) A plant DNA minipreparation: version II. Plant Mol Biol 4:19–21

Edwards KJ (1999) Method for determining the genotype of an organism using allele-specific oligonucleotide probes, which hybridise to microsatellite flanking sequences. Patent No. WO99/01576

Edwards KJ, Barker JHA, Daly A, Jones C, Karp A (1996) Microsatellite libraries enriched for several microsatellite sequences in plants. BioTechniques 20:758–760

Ford-Lloyd BV, Williams JT (1975) A revision of *Beta* section *Vulgares* (Chenopodaiceae), with new light on the origin of cultivated beets. Bot J Linn Soc 71:89–102

Grant F (1997) Genstat and the quantitative revolution. Scientific Computing World, October 46–47

Hansen M, Kraft T, Christiansson M, Nilsson N-O (1999) Evaluation of AFLP in *Beta*. Theor Appl Genet 98:845–852

Innasn H, Terauchi R, Miyashita T (1997) Microsatellite polymorphism in natural populations of the wild plant *Arabidopsis thaliana*. Genetics 146:1441–1452

Jarret RL, Merrick LC, Holms T, Evans J, Aradhya MK (1997) Simple sequence repeats in watermelon [*Citrullus lanatus* (Thunb.) Matsum. and Nakai]. Genome 40:433–441

Kresovich S, Szewc-McFadden AK, Bliek SM, McFerson JR (1995) Abundance and characterization of simple sequence repeats (SSRs) isolated from a size-fractionated genomic library of *Brassica napus* L. (Rapeseed). Theor Appl Genet 91: 206–211

Lagercrantz U, Ellegren H, Andersson L (1993) The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. Nucleic Acids Res 21:111–1115

Liu Z, Biyashev R, Saghai Maroof M (1996) Development of simple sequence repeat DNA markers and their integration into a barley linkage map. Theor Appl Genet 93: 869–876

McGrath JM, Derrico CA, Yu Y (1999) Genetic diversity in selected, historical US sugarbeet germplasm and *Beta vulgaris* ssp. *maritima*. Theor Appl Genet 98:968–976

Mörchen M, Cuguen J, Michaelis G, Hanni C, Laprade PS (1996) Abundance and length polymorphism of microsatellite repeats in *Beta vulgaris* L. Theor Appl Genet 92:326–333

Morgante M, Olivieri AM (1993) PCR – amplified SSRs as markers in plant genetics. Plant J 3:175–182

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York

Pillen K, Steinrucken 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, Steinrucken G, Wricke 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

Pillen K, Schondelmair J, Jung C, Herrmann R (1996) Genetic mapping of genes for 12 nuclear – encoded polypeptides associated with the thylakoid membranes in *Beta vulgaris* L. FEBS Lett 395:58–62

Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S, Rafalski A (1996) The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. Mol Breed 2:225–238

Rossetto M, McLauchlan A., Harriss FCL, Henry RJ, Baverstock PR, Lee LS, Maguire TL, Edwards KJ (1999) Abundance and polymorphism of microsatellite markers in the tea tree (*Melaleuca alternifolia, Myrtaceae*). Theor Appl Genet 98: 1091–1098

Sai EA, Kokoskaa RJ, Dominska M, Greenwell P, Petes TD (1997) Microsatellite instability in yeast: dependence on repeat unit size and DNA mismatch repair genes. Mol Cell Biol 17:2851–2858

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd edn. cold Spring Harbor Laboratory, Cold Spring Harbor, New York

Schmidt T, Heslop-Harrison JS (1996) The physical and genomic organisation of microsatellites in sugar beet. Proc Natl Acad Sci USA 93:8761–8765

Schmidt T, Jung C, Metzlaff M (1991) Distribution and evolution of two satellite DNAs in the genus *Beta*. Theor Appl Genet 82:793–797

Schmidt T, Bllobez K, Metzlaff, Kaemmer D, Weising K, Kahl G (1993) DNA fingerprinting in sugar beet (*Beta vulgaris*) – identification of doubled – haploid breeding lines. Theor Apply Genet 85:653–657

Schondelmaier J, Steinrucken G, Jung C (1996) Integration of AFLP markers into a linkage map of sugar beet (*Beta vulgaris* L.). Plant Breed 115:231–237

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

Sorrells ME, Wilson WA (1997) Direct classification and selection of superior alleles for crop improvement. Crop Sci 37: 691–697

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

Staub JE, Serquen FC, Gupta M (1996) Genetic markers, map construction, and their application in plant breeding. HortScience 31:729–740

Stephenson P, Bryan G, Kirby J, Collins A, Devos K, Busso C, Gale M (1998) Fifty new microsatellite loci for the wheat genetic map. Theor Appl Genet 97:946–949

Taramino G, Tarchini R, Ferrario S, Lee M, Pe ME (1997) Characterisation and mapping of simple sequence repeats (SSRs) in *Sorghum bicolor*. Theor Appl Genet 95:66–72

Tautz D (1989) Hypervariability of simple sequence repeats as a general source for polymorphic DNA markers. Nucleic Acids Res 17:6463–6471

Wang Z, Weber JL, Zhong G, Tanksley SD (1994) Survey of plant short tandem DNA repeats. Theor Appl Genet 88:1–6

Weber JL, May PE (1989) Abundant classes of human DNA polymorphism which can be typed using the polymerase chain reaction. Am J Hum Genet 44:388–396

Weising K, Fung RWM, Keeling DJ, Atkinson RG, Gardner RC (1996) Characterisation of microsatellites from *Actinidia chinensis*. Mol Breed 2:117–131

Westman AL, Kresovich S (1998) The potential for cross-taxa simple sequence repeat (SSR) amplification between *Aribidopsis thaliana* L. and crop brassicas. Theor Appl Genet 96:272–281