Theor Appl Genet (1996) 93: 869–876

Z.-W. Liu · R. M. Biyashev · M. A. Saghai Maroof

Development of simple sequence repeat DNA markers and their integration into a barley linkage map

Received: 2 October 1996 / Accepted: 1 March 1996

Abstract Simple sequence repeats (SSRs), or microsatellites, are a new class of PCR-based DNA markers for genetic mapping. The objectives of the present study were to develop SSR markers for barley and to integrate them into an existing barley linkage map. DNA sequences containing SSRs were isolated from a barley genomic library and from public databases. It is estimated that the barley genome contains one $(GA)_n$ repeat every 330 kb and one $(CA)_n$ repeat every 620 kb. A total of 45 SSRs were identified and mapped to seven barley chromosomes using doubled-haploid lines and/or wheat-barley addition-line assays. Segregation analysis for 39 of these SSRs identified 40 loci. These 40 markers were placed on a barley linkage map with respect to 160 restriction fragment length polymorphism (RFLP) and other markers. The results of this study demonstrate the value of SSRs as markers in genetic studies and breeding research in barley.

Key words Microsatellites · *Hordeum vulgare* · Molecular marker · Linkage map

Introduction

Linkage mapping is a prerequisite for many genetic analyses. Linkage maps of barley (*Hordeum vulgare*) have been constructed in several populations using restriction fragment length polymorphisms (RFLPs) as genetic markers (Graner et al. 1991; Heun et al. 1991). Recently, Kleinhofs et al. (1993) constructed a barley linkage map containing 295 RFLP, RAPD, isozyme and morphological marker loci. Despite these extensive mapping efforts, large gaps are still present in the barley linkage maps. Moreover, a lack of common loci among the different barley linkage maps has hindered the development of a composite barley map (Kleinhofs and Kilian 1994). Given the low level of RFLPs in barley (Graner et al. 1990), construction of a highly saturated linkage map would benefit from the development of new genetic markers.

Simple sequence repeats (SSRs), also known as microsatellites, are ideal DNA markers for genetic mapping and population studies because of their abundance (Weber 1990), high level of polymorphism (Cregan et al. 1994; Saghai Maroof et al. 1994), wide dispersion in diverse genomes (Wang et al. 1994), ease of assay by the polymerase chain reaction (PCR), and ease of dissemination among laboratories. Identification and characterization of SSRs, based on screening DNA libraries and/or searching public databases, have been reported for a number of plant species (Akkaya et al. 1992; Morgante and Olivieri 1993; Wu and Tanksley 1993; Zhao and Kochert 1993; Morgante et al. 1994; Saghai Maroof et al. 1994; Becker and Heun 1995; Liu et al. 1995). The utility of SSR markers has already been demonstrated in several plant genetic studies. These include the linkage with a virus resistance gene in soybean (Yu et al. 1994), the identification of chromosomal regions with significant effects on yield in rice (Zhang et al. 1994), and in the germ plasm assessment of rice and soybean (Yang et al. 1994; Maughan et al. 1995; Rongwen et al. 1995). Several of the aforementioned studies have reported on initiating the construction of SSR-based linkage maps. The most extensive SSR mapping studies in plant genomes are those on *Arabidopsis thaliana* (Bell and Ecker 1994) and soybean (Akkaya et al. 1995), where 30 and 34 SSR loci were placed in the linkage maps of the two species, respectively. In a survey of 207 accessions of wild and cultivated barley (Saghai Maroof et al. 1994) as many as 37 alleles were observed at a single SSR locus. The high level of allelic diversity displays the great potential of SSR markers for the genetic mapping of barley.

The objectives of the present study were (1) to develop SSR markers for barley from a genomic DNA library and from the GenBank and EMBL database sequences, (2) to integrate these markers into an existing barley linkage map.

Communicated by A. L. Kahler

Z.-W. Liu · R. M. Biyashev · M. A. Saghai Maroof (✉)
Department of Crop and Soil Environmental Sciences, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA

and (3) to determine the frequencies of $(GA)_n$ and $(CA)_n$ repeats in the barley genome.

Materials and methods

Isolation of SSRs from a barley genomic library, DNA sequencing and primer design

The procedures for constructing a small-insert plasmid library and for screening of the SSR-containing clones were as previously described (Liu et al. 1995). Briefly, genomic DNA from the barley cultivar Old ZZ was extracted (Saghai Maroof et al. 1984) and digested with the restriction enzyme *Taq*I. DNA fragments of 200–500 bp in length were size-fractionated and cloned into the plasmid vector pGEM-7zf(+) (Promega, Madison, Wis.) which had been digested with the restriction enzyme *Cla*I. Ligation, transformation, and colony hybridization were as described by Sambrook et al. (1989). Two synthetic dinucleotide repeats, $(GA)_{10}$ and $(CA)_{10}$, were used to screen the library for clones containing these two types of SSRs. To reduce the number of false positives, clones that hybridized with the two synthetic probes were picked onto new filters for a second round of hybridization with the same two probes.

Plasmid DNA from the (GA)- and (CA)-positive clones was sequenced either by an automated DNA sequencer using the *Taq* Dye-Deoxy Terminator kit at the Molecular Genetics Facility of the University of Georgia, or manually using the dideoxynucleotide chain-termination method with Sequenase (USB, Cleveland, Ohio). Flanking sequences of the repeats were used to design primers using the computer program PRIMER 0.5 (M.J. Daly, S.E. Lincoln and E.S. Lander, unpublished). Primers were selected to have a melting temperature of 60°C, to be 18–22 bases in length, and to amplify DNA fragments between 100 and 250 bp in size. Primers were synthesized by Research Genetics, Inc (Huntsville, Alabama).

SSRs isolated from the genomic library were designated as HVM (for *Hordeum vulgare* microsatellite) and numbered consecutively. A letter was added after the number if an SSR corresponded to more than one locus.

SSRs obtained from the public databases

The GenBank and EMBL databases were searched and a total of 12 barley SSR-containing sequences was selected. Four of the twelve have previously been studied by Saghai Maroof et al. (1994), including the ribulose-1,5-bisphosphate carboxylase activase gene (referred to as HVM3), the starch synthase gene (HVM4), an acyl carrier protein I gene (HVM7) and a seed imbibition protein gene (HVM9). Two other SSRs included in this study were from a gene for a pathogenesis-related protein (M. Hahn, H. Lehnackers and W. Knogge, unpublished) and a gene for a glucose-regulated protein (grp94 homologue, heat-shock protein 90, Walther-Larsen et al. 1993). We designate these two primer pairs as HVM5 and HVM11, respectively. The remaining six SSRs were selected from those described by Becker and Heun (1995). These six were derived from a β-ketacyl-acyl carrier protein synthase I isoenzyme (referred to as HVBKASI), an α-amylase inhibitor (HVCMA), a chalcone synthase gene (HVCSG), a gene for dehydrin-7 (HVDHN7), an mRNA for dehydrin-9 (HVDHN9) and a thiol protease aleurin (HVLEU). Two of the six, HVCMA and HVLEU, have been mapped by Becker and Heun (1995) to chromosomes 1 and 7, respectively. Finally, a wheat primer pair WMS6 (Röder et al. 1995) was also used in this study, bringing the total number of primers to 13.

PCR amplification of SSRs

PCR amplification with each primer pair was first tested using DNA from six barley cultivars: Steptoe, Morex, Harrington, TR306, Igri and Franka. These cultivars were the parental lines of the four doubled-haploid (DH) populations used in this mapping study. DNA from the cultivar Old ZZ (the source of the genomic library) and from the recombinant plasmids containing the inserts from which the primers were designed, were also used as two positive controls in the evaluation of the primers. Inclusion of these two controls gave us more confidence that the observed PCR products were the expected fragments (i.e., the PCR amplification products from Old ZZ and the corresponding plasmid were the same in length).

Depending on the primer pairs used, amplification of the SSRs was performed using one of the following five PCR conditions: (1) A "touchdown" PCR consisting of 18 cycles of 94°C for 1 min denaturing and 72°C for 1 min extension. Annealing (30 s) temperatures were progressively decreased by 1°C every second cycle from 64°C to 55°C. The PCR reaction continued for 30 additional cycles at 94°C for 1 min, 55°C for 1 min and 72°C for 1 min. The reaction ended with a 5-min extension at 72°C. (2) A similar "touchdown" PCR condition except that the annealing temperatures were decreased from 69°C to 60°C over 18 cycles, at which temperature the reaction continued for 20 additional cycles. (3) A regular PCR profile consisting of one cycle of 94°C for 3 min, 55°C for 2 min and 72°C for 1.5 min, followed by 30 cycles at 94°C for 1 min, 55°C for 2 min and 72°C for 1.5 min. The PCR conditions 4 and 5 were the same as those reported by Becker and Heun (1995) and Röder et al. (1995), respectively.

PCR was performed in a Perkin Elmer Cetus 480 thermal cycler. The reaction mixture and the polyacrylamide gel electrophoresis were as described by Saghai Maroof et al. (1994) except that the reaction mixture contained 2 mM $MgCl_2$.

Chromosomal assignment, linkage analysis and map construction

SSRs were initially mapped to barley chromosomes using a set of wheat-barley chromosome addition lines (Islam et al. 1981). These materials included a barley cultivar Betzes, a wheat cultivar Chinese Spring and six addition lines representing chromosomes 1, 2, 3, 4, 6, and 7 of barley. Each addition line contains the 21 pairs of wheat chromosomes from Chinese Spring plus one pair of barley chromosomes from Betzes.

Four mapping populations were used in this study. The first population was constructed from the cross of SteptoexMorex (S/M), and consisted of 96 doubled haploids (Kleinhofs et al. 1993). The second population contained 94 DHs derived from the cross of HarringtonxTR306 (H/T) (K. Kasha, personal communication). The third population included 119 DHs from the cross of HarringtonxMorex (H/M) (P. Hayes, personal communication). These are the three principal mapping populations of the North American Barley Genome Mapping Project. The 71 DHs of the fourth mapping population used in this study were derived from the cross of IgrixFranka (I/F) (Graner et al. 1991). The number of markers in the four maps of S/M, I/F, H/T and H/M are 434, 370, 212 and 53, respectively.

Segregation of each SSR locus was tested for goodness-of-fit to a 1:1 ratio using the computer program LINKAGE-1 (Suiter et al. 1983). The SSR data were then entered into the existing linkage data for each of the respective mapping populations. Linkage analysis was performed using the MAPMAKER computer program 3.0 b (Lander et al. 1987) at LOD=3.0, with a maximum Haldane distance of 50 cM.

Results

Identification of SSRs from a barley genomic library

Approximately 36 000 recombinant clones of a barley genomic library were screened with $(GA)_{10}$ and $(CA)_{10}$ probes. After colony hybridization and DNA sequencing of the plasmids isolated from the positive clones, a total of 55 microsatellites including 36 $(GA)_n$ and 19 $(CA)_n$ repeats

was identified. Flanking DNA sequences of 41 of the 55 SSRs were then used to design primers for PCR amplification. The remaining 14 plasmids, including 12 $(GA)_n$- and two $(CA)_n$-containing clones, were discarded as they contained only one flanking region that was long enough for primer design.

All the primers were initially evaluated using the "touchdown" PCR condition in which the final annealing temperature was 55°C. Some primer pairs generated a smear of PCR products, produced unexpected fragments with multiple lengths, or amplified no product. Such poor amplifications could be greatly improved by re-designing the primers from the original sequences and/or by varying PCR profiles and the magnesium concentration. Table 1 lists 45 primer sequences that successfully amplified PCR products of the expected size. These 45 included 32 derived from the barley genomic library, 12 from the public databases and one from a wheat genomic library. PCR conditions, a survey of polymorphism and the chromosomal locations of the SSRs are also presented in Table 1.

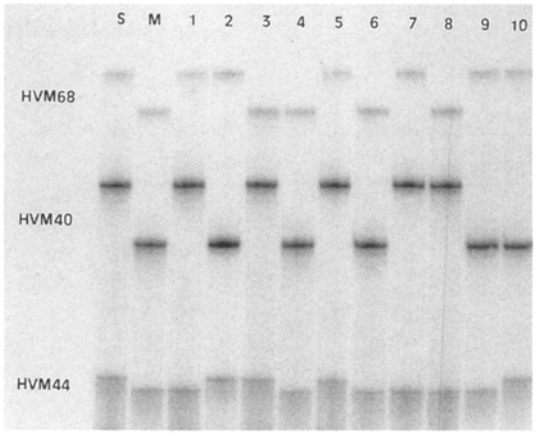

Fig. 1 Segregation patterns of three SSRs (HVM40, HVM44 and HVM68) in Steptoe (*S*)/Morex (*M*) doubled-haploid lines (DHs) assayed by multiplex PCR. The three primer pairs were used simultaneously in the same PCR reaction. *Numbers 1–10* represent a set of DH lines from the S/M cross

Chromosomal assignment and polymorphism of SSRs

SSRs amplified with the 45 primer pairs were initially mapped by wheat-barley addition-line analysis. Thirty nine SSRs were unambiguously assigned to specific barley chromosomes. Because similar-size DNA fragments were amplified both from barley and wheat genomes, the chromosomal locations of the remaining six SSRs could not be determined using the addition line assay.

SSR polymorphism, presumably resulting from the variation in the number of the nucleotide repeats, was surveyed among the six parental lines of the four mapping populations. Out of 45 primer pairs, 39 detected variation between two parental lines in at least one cross (Table 1) and subsequently were placed in a barley RFLP map. The remaining six SSRs were monomorphic across all parental lines. Figure 1 shows SSR polymorphism between two parents and segregation patterns among ten DH lines resulting from a multiplex PCR reaction using three primer pairs (HVM40, HVM44 and HVM68).

Linkage analysis and map construction

The majority of the primer pairs amplified single PCR products within the expected size ranges, which were presumably alleles of single loci. PCR amplifications with several primer pairs, however, produced minor bands in addition to the expected fragments. They either varied in length following the major bands or were invariable across the DNA samples examined. These minor bands were considered as PCR artifacts, probably resulting from either slippage events or chance homologies with the primers at other sites (Tautz 1989; Senior and Heun 1993). The primer pair HVM2, however, amplified major products in two distinct size ranges. Inheritance analysis using DH lines from the Steptoe/Morex cross indicated that the variants of the two size ranges segregate independently. Thus, SSR HVM2 corresponds to two separate loci, designated as HVM2A and HVM2B, which were mapped to chromosomes 7 and 1, respectively. The Morex allele of the HVM2A locus had the same size as those amplified from Old ZZ and the corresponding plasmid. The Steptoe allele of this locus was approximately 4 bp shorter than that of Morex. We are confident that the two alleles at HVM2A result from variation in the number of repeated core sequences. However, it is uncertain whether the polymorphism at HVM2B is due to variation at an SSR locus or to an amplified sequence length polymorphism resulting from random insertion/deletion mutations (Maughan et al. 1995). Further investigation will be needed to determine the cause of the polymorphism observed at the HVM2B locus.

A given SSR was often polymorphic in more than one of the four mapping populations. In these cases, mapping was carried out in only one population with an order of priority of S/M, H/T, H/M, or, I/F for the following reasons: (1) The S/M linkage map (Kleinhofs et al. 1993) contains the largest number of markers among the existing barley maps, and (2) There are more common markers between the S/M and H/T maps than among other maps of barley. Consequently, 22 SSRs (23 loci) were integrated into the S/M linkage map (Kleinhofs et al. 1993) by segregation analysis with a set of 160 previously mapped markers of this population (Fig. 2). Similarly, ten SSRs were placed on the existing map of H/T, three on H/M, and four on I/F. These 17 SSRs were then integrated into the S/M map based on the marker loci held in common among these four DH populations. Thus, a total of 40 SSR loci were integrated into the existing S/M RFLP map of barley (Fig. 2).

Of the 40 loci mapped by linkage analysis in this study, three exhibited distorted segregation ratios. The *P* values of the three loci were 0.01, 0.02 and 0.028, respectively.

Table 1 Primer sequences, PCR conditions, polymorphism and chromosomal locations of barley SSRs

SSR	Primer sequence (5′ → 3′)	Repeat	Size (bp)	PCR[a]	Polymorphic in[b]				Chromosome[c]
					S/M	H/T	H/M	I/F	
HVM2	CAGGTGTCTAGTGGGTGCCT TTACATCCAAGGAGCAATCCC	$(GA)_{11}$	201	(1)	+	–	+	–	7[d]
HVM3	ACACCTTCCCAGGACAATCCATTG AGCACGCAGAGCACCGAAAAAGTC	$(AT)_{29}$	188	(3)	+	+	+	+	4
HVM4	AGAGCAACTACCAGTCCAATGGCA GTCGAAGGAGAAGCGGCCCTGGTA	$(AT)_9$	198	(3)	+	+	+	+	1
HVM5	AACGACGTCGCCACACAC AGGAACGAAGGGAGTATTAAGCAG	$(GT)_6$, $(AT)_{16}$	202	(3)	+	–	+	–	1
HVM6	CATGAATGAATGATTGGTTTTG CGCATCCGTATGTATGAGTAA	$(GA)_9$	175	(1)	+	+	–	–	7
HVM7	ATGTAGCGGAAAAAATACCATCAT CCTAGCTAGTTCGTGAGCTACCTG	$(AT)_7$	174	(3)	–	–	–	–	7
HVM9	CTTCGACACCATCACCCAG ACCAAAATCGCATCGAACAT	$(TCT)_5$	221	(3)	–	–	+	–	3
HVM11	CCGGTCGGTGCAGAAGAG AAATGAAAGCTAAATGGGCGATAT	$(GGA)_3$, (GGA), $(GAA)_2$	181	(3)	+	–	–	–	6
HVM13	AGTAGCTATGTGTTTGGATCGC CATCAAGGGCATCCTCATG	$(GA)_6$, $(GA)_6$, $(GA)_6$	249	(1)	–	–	–	+	4
HVM14	CGATCAAGGACATTTGGGTAAT AACTCTTCGGGTTCAACCAATA	$(CA)_{11}$	158	(1)	+	–	–	+	6
HVM15	TCATAACCACGGCGTCCT CGTGACTGGAAACCCTGC	$(GA)_8$	166	(1)	–	–	–	–	3
HVM20	CTCCACGAATCTCTGCACAA CACCGCCTCCTCTTTCAC	$(GA)_{19}$	151	(1)	–	–	+	–	5
HVM22	TTTTGGGGGATGCCTACATA TTTCAAATGGTTGGATTGGA	$(AC)_{13}$	167	(1)	+	–	–	+	6
HVM23	TCGGTGAAGAAATACGAGGC TCTTTGTGACCTACCGGTCC	$(GA)_9$	246	(2)	+	–	+	–	2
HVM26	GGCTATCACATTTGGTACCATC GCATGTGTAGGTGTTGGTGG	$(CA)_{11}$	206	(2)	–	+	+	–	2
HVM27	GGTCGGTTCCCGGTAGTG TCCTGATCCAGAGCCACC	$(GA)_{14}$	192	(1)	+	+	+	+	3
HVM30	AGTGGGGAATGAGAGAATGG TGCTTGTGGGTCATCACAC	$(CA)_8$	150	(2)	+	–	+	–	7
HVM31	CGGTTTCTGGTTGCTTGG CGAAGGTCTCAGGCTTCATG	$(AC)_9$	163	(1)	–	–	–	–	6
HVM33	ATATTAAAAAAGGTGGAAAGCC CACGCCCTCTCCCTAGAT	$(CA)_7$	157	(1)	+	–	+	+	3
HVM34	ACCATGTTGCGTGTTGCTT CGGTTCGAAATCGAGTGG	$(GA)_{10}$	222	(2)	+	–	–	+	6
HVM36	TCCAGCCGAACAATTTCTTG AGTACTCCGACACCACGTCC	$(GA)_{13}$	114	(1)	+	–	+	–	2
HVM40	CGATTCCCCTTTTCCCAC ATTCTCCGCCGTCCACTC	$(GA)_6(GT)_4$ $(GA)_7$	160	(1)	+	+	+	–	4
HVM43	GGATTTTCTCAAGAACACTT GCGTGAGTGCATAACATT	$(CA)_9$	239	(1)	+	+	+	–	5
HVM44	AAATCTCAGGTTCGTGGGCA CCACGGAGACCACCTCACTT	$(GA)_8$	114	(1)	+	–	–	–	3
HVM49	CTCTATAGGCACGAAAAATTCC TTGCACATATCTCTCTGTCACA	$(CA)_{12}$	105	(1)	–	+	+	–	1
HVM50	TGAAGAAGCCCTCCGTATTC TGCTCAAGGCTCTAGGCTGA	$(GA)_9$	213	(3)[30]	+	+	–	–	2
HVM51	TCTAAATTACCTTCCCAGCCA AAGCAGACATGTAGGAGGTCA	$(GA)_3$ $(GGGA)_3$, $(GA)_8$	151	(1)	–	–	–	–	1

Table 1 Continued

SSR	Primer sequence (5′ → 3′)	Repeat	Size (bp)	PCR[a]	Polymorphic in[b]				Chromosome[c]
					S/M	H/T	H/M	I/F	
HVM54	AACCCAGTAACACCGTCCTG AGTTCCCTGACCCGATGTC	$(GA)_{14}$	159	(1)	+	–	+	–	2
HVM60	CAATGATGCGGTGAACTTTG CCTCGGATCTATGGGTCCTT	$(AG)_{11}$, $(GA)_{14}$	115	(1)	+	–	+	+	3
HVM62	TCGCGACCAGACGAGAAG AGCTAGCCGACGACGCAC	$(GA)_{11}$	251	(1)	+	+	+	–	3
HVM63	CGCGCAAGCATGAATACTC ACTCACAAGTGGCGCGTAC	$(GA)_9$	124	(1)	–	–	–	–	5
HVM64	GATGTGAAGGCTGCCTG ACACGCCCTATTACCCAGTG	$(GA)_4(GT)_7$ $(CT)_2(GT)_4(GA)_8$	253	(1)	–	–	–	–	5
HVM65	AGACATCCAAAAAATGAACCA TGGTAACTTGTCCCCCAAAG	$(GA)_{10}$	129	(1)	–	+	–	+	6
HVM67	GTCGGGCTCCATTGCTCT CCGGTACCCAGTGACGAC	$(GA)_{11}$	116	(1)	–	–	+	+	4
HVM68	AGGACCGGATGTTCATAACG CAAATCTTCCAGCGAGGCT	$(GA)_{22}$	204	(1)	+	+	+	+	4
HVM70	CCGCCGATGACCTTCTC ACCCACGACCTATGGCAC	$(CA)_8$	154	(3)	+	+	+	–	5
HVM74	AGGAAGTCATTGCGTGAG TGATCAAGAATGATAACATGG	$(GA)_{13}$	162	(3)	+	–	–	+	6
HVM77	GAAATTTGGTGTATGATGGTT CAAATCTTAAATCTCTCTGTTT	$(CA)_7$	199	(1)	–	–	–	+	4
HVBKASI*	ATTGGCGTGACCGATATTTATGTTCA CAAAACTGCAGCTAAGCAGGGGAACA	$(C)_{10}$, $(A)_{11}$	197	(4)	+	+	+	–	2
HVCMA*	GCCTCGGTTTGGACATATAAAG GTAAAGCAAATGTTGAGCAACG	$(AT)_9$	141	(4)	–	+	+	+	1
HVCSG*	CACTTGCCTACCTCGATATAGTTTGC GTGGATTCCATGCATGCAATATGTGG	$(CA)_4$, $(C)_{17}$	196	(4)	+	+	+	–	2
HVDHN7*	TTAGGGCTACGGTTCAGATGTT ACGTTGTTCTTCGCTGCTG	$(AAC)_5$	177	(4)	–	+	+	–	7
HVDHN9*	CATGGACAAGATCAAGGAGAAG CCCATTATTTATCTGTAGGAACGC	$(AC)_6$	128	(4)	–	+	+	+	7
HVLEU*	TTGGAAGTGTACAGCAATGGAG TGAAAGGCCCCACAAGATAG	$(ATTT)_4$	166	(4)	+	–	+	+	7
WMS6**	CGTATCACCTCCTAGCTAAACTAG AGCCTTATCATGACCCTACCTT	$(GA)_{40}$	205	(5)	–	+	+	–	4

[a] The numbers in the parentheses represent one of the five PCR conditions used in this study. See Materials and methods for details of the PCR conditions. The superscripts refer to the concentration of $MgCl_2$ in mM when it was not 2.0 mM
[b] S=Steptoe, M=Morex, H=Harrington, T=TR306, I=Igri, F=Franka, +=polymorphic, –=monomorphic
[c] Chromosomal locations of SSRs determined by wheat-barley chromosome addition line assay and/or linkage analysis
[d] This is the chromosomal location of locus HVM2A only. Locus HVM2B is on chromosome 1
* Primer sequences and PCR condition according to Becker and Heun (1995)
** Primer sequences and PCR condition according to Röder et al. (1995)

All of the three loci were mapped using the S/M DH lines and were skewed towards the Steptoe alleles. Figure 2 shows the composite linkage map of barley constructed from this study. This figure includes six monomorphic SSRs (shown at the bottom of the corresponding chromosomes) whose chromosomal assignments were determined only by the wheat-barley addition-line assay. Altogether, of the 45 SSRs reported in this study, five were located on chromosome 1, seven on chromosome 2, eight on chromosome 3, seven on chromosome 4, four on chromosome 5, and seven on each of chromosomes 6 and 7.

Discussion

The barley genomic library of this study consisted of approximately 36000 recombinant clones. The average in-

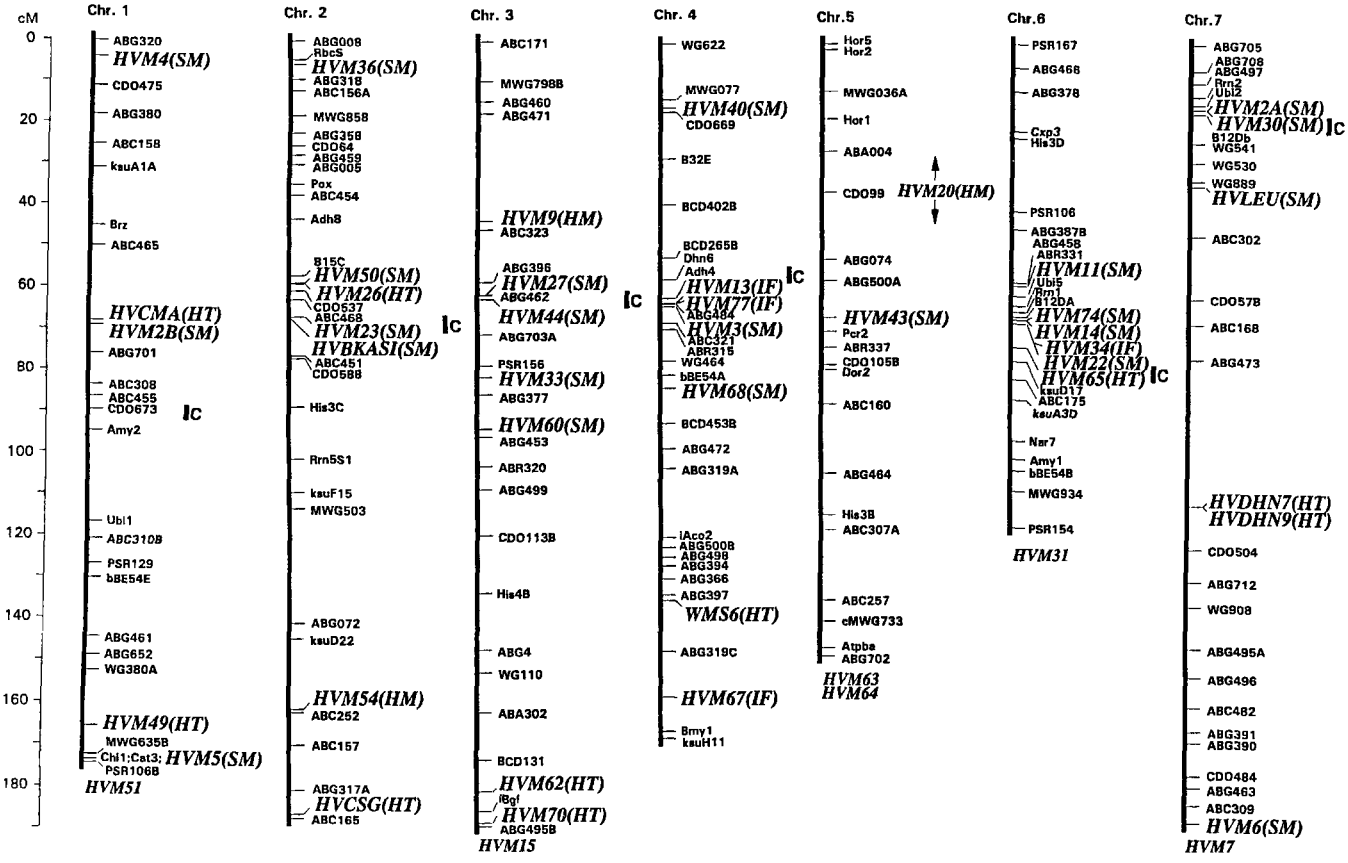

Fig. 2. Linkage map of barley based on the linkage data from the S/M cross (Kleinhofs et al. 1993). The SSR loci mapped in this study are shown in *bold and italics*. The scale to the left of the chromosomes shows map distances in centiMorgans (cM). The approximate centromere positions are shown by a *vertical bar with a C*. The abbreviations in parentheses following the SSR loci designate the DH populations in which the SSRs were mapped: *SM* – Steptoe/Morex, *HT* – Harrington/TR306, *HM* – Harrington/Morex and *IF* – Igri/Franka. SSRs displayed at the bottom of the chromosomes were assigned to chromosomes based only on the wheat-barley addition-line assays (they were monomorphic among the parental lines of the four populations). HVM20 mapped 36 cM away from CDO99 on chromosome 5; its orientation with respect to CDO99 could not be determined due to the absence of mapped markers in this chromosomal region of the H/M map

sert size of the clones was 335 bp. Therefore, the library contained 12 000 kb of barley DNA. Because the genome size of barley is 4.9×10^6 kb (Arumuganathan and Earle 1991), this library represented approximately 0.2% of the barley genome. After screening the library with $(GA)_{10}$ and $(CA)_{10}$ probes, 36 sequences containing $(GA)_n$ repeats and 19 sequences containing $(CA)_n$ repeats were identified. Thus, we estimate that, on average, the barley genome contains one $(GA)_n$ repeat every 330 kb and one $(CA)_n$ repeat every 620 kb, or a total of 1.5×10^4 $(GA)_n$ and 7.9×10^3 $(CA)_n$ repeats, respectively.

The frequencies of the $(GA)_n$ and $(CA)_n$ repeats based on DNA library screening have been reported for several plant genomes. There is one $(GA)_n$ repeat every 125–250 kb and one $(CA)_n$ repeat every 250–480 kb in *Arabidopsis thaliana* (Bell and Ecker 1994), *Brassica napus* (Lagercrantz et al. 1993), rice (Wu and Tanksley 1993) and in a turfgrass species (Liu et al. 1995). The occurrences of the two types of repeats in barley seem less frequent than those in the above species, but are similar to those reported in wheat where one $(GA)_n$ and one $(CA)_n$ repeat was observed every 440 kb and 704 kb, respectively (Röder et al. 1995). Despite the variation in these estimates, our results are consistent with the conclusion that $(GA)_n$ repeats are more abundant than $(CA)_n$ repeats in plant genomes and that both types of repeats are less frequent in plants than in mammals (Lagercrantz et al. 1993).

In *Arabidopsis* only one in 18 $(CA)_n$ repeats was found to be polymorphic (Bell and Ecker 1994). This did not appear to be the case in barley. Seventeen primer pairs were designed from a total of 19 $(CA)_n$-containing sequences isolated from our DNA library. Ten of the seventeen successfully amplified $(CA)_n$ repeats. Four of these $(CA)_n$ repeats were polymorphic in the H/T and I/F, and six were polymorphic in the S/M and H/M, populations of this study. Interestingly, most of the $(CA)_n$ sequences isolated from our DNA library were simple repeats. In contrast, most of the $(CA)_n$ repeat sequences identified in *Arabidopsis* (Bell and Ecker 1994) were complex, with adjacent repeats of other types. Whether there is a correlation between the complexity of the sequence structure and the level of polymorphism will be subject to further studies. Our results were similar to those of Bell and Ecker (1994) in that we encountered difficulty in amplifying $(CA)_n$ repeats. Out of the 17 $(CA)_n$ primer pairs, 7 failed to amplify any prod-

Table 2 Comparisons of RFLP and SSR polymorphisms as detected in four barley crosses

Cross	RFLPS[a]			SSRs	
	No. of probes	% Polymorphic, average[b]	% Polymorphic, combinded[c]	No. of SSRs	% Polymorphic
Steptoe × Morex	85	31.8	54.1[d]	45	60.0
Harrington × TR306	77	22.1	33.8[e]	45	42.2
Harrington × Morex	77	29.5	54.5[e]	45	62.2
Igri × Franka	163	NA	28.0[f]	45	37.8

[a] RFLP probes were from random genomic libraries except for Igri×Franka where the probes were from a genomic and a cDNA library. NA=not available
[b] Average percentage of polymorphic probes over the enzymes used
[c] RFLP detected with at least one of the enzymes used
[d] Heun et al. (1991). Five enzymes were used
[e] Unpublished data. Six enzymes were used
[f] Graner et al. (1991). Six enzymes were used

ucts, whereas only 2 out of 24 $(GA)_n$ primer pairs failed to amplify any relevant PCR products.

The level of polymorphism revealed by the 45 SSRs of this study was compared with that detected by RFLP markers in the same four barley crosses. Levels of polymorphism with SSRs were higher than those with RFLPs in all four crosses (Table 2). Overall 87% (39/45) of the SSRs detected polymorphism in at least one of the four crosses (Table 1). The informativeness of a marker system is one important parameter in genetic mapping. Our results show the unique value of SSRs as molecular markers for linkage mapping and other types of genetic analyses, particularly in species with low levels of polymorphism such as barley.

Distorted segregation ($P<0.05$) was observed at 3 (HVM2A, HVM30 and HVLEU) of the 40 loci mapped. The three distorted loci were located around the centromeric region of chromosome 7 (Fig. 2). Interestingly, in the published S/M map (Kleinhofs et al. 1993) nine of the ten loci on chromosome 7 that exhibited distorted segregation ratios were also located around the centromeric region. One possible explanation for this observation is that these markers might be linked to some loci around the centromere that cause segregation distortion, as reported by other researchers (Bronson et al. 1990; Brummer et al. 1993).

The 40 SSR loci mapped by linkage analysis were dispersed on all seven barley chromosomes, with a range of two to seven loci per chromosome. Five pairs of loci that were analyzed in the same mapping population either cosegregated or were closely linked. They were HVM23/HVBKASI on chromosome 2, HVM13/HVM77 on chromosome 4, HVM14/HVM74 on chromosome 6, HVM2A/HVM30 and HVDHN7/HVDHN9 on chromosome 7 (Fig. 2). The loci in each pair were different from each other, as determined by comparing their repeat sequences and the sequences of the flanking regions. Therefore, it is unlikely that the cosegregation observed at these loci resulted from the same SSR being sampled twice. All pairs, except HVDHN7/HVDHN9, resided near the centromeres of their respective chromosomes implying the clustering of some SSR markers. HVDHN7 and HVDHN9 represent repeat sequences in or near the genes for dehydrin-7 and dehydrin-9 (Becker and Heun 1995). The close linkage between HVDHN7 and HVDHN9 may suggest that they are members of a cluster of dehydrin genes.

Several SSRs mapped to large gaps in the S/M linkage map. HVM2B and HVCMA mapped between RFLP markers ABC465 and ABG701, while HVM49 mapped between markers WG380A and MWG635B on chromosome 1. HVM67 mapped between ABG319C and Bmy1 on chromosome 4. HVDHN7 and HVDHN9 mapped between ABG473 and CDO504 on chromosome 7. Placement of these SSR markers on the barley map greatly improves the coverage of these chromosomal regions.

The abundance, high level of polymorphism, and ease of genotyping make SSRs an excellent molecular marker system for many types of genetic analyses, including linkage mapping, germ plasm surveys and phylogenetic studies. Wider use of this class of molecular markers, however, will depend on the speed and cost-effectiveness of developing SSR markers and on the ability to genotype large numbers of individuals. To speed up data collection we routinely conduct multiplex PCR, in which two or more primer pairs are used simultaneously in one reaction (Fig. 1). Double or multiple loadings of a single gel can also accelerate the process of genotyping individuals (Saghai Maroof et al. 1994). Recent advances in automated fluorescence-based DNA fragment detection should improve the efficiency and accuracy of genotype analysis. These techniques may be particularly useful in marker-assisted plant breeding assays where multiple loci need to be screened in a short period of time. Searching published sequence databases is a quick and easy way to obtain SSRs. Once this source has been exploited, however, the only alternative is to screen DNA libraries, which is a time-consuming process. The construction of libraries enriched for simple sequence repeats (e.g., Ostrander et al. 1992; Kandpal et al. 1994) and the use of methods based on a combination of microsatellites with AFLP (amplified fragment length polymorphism) technology, hold great promise to accelerate the development of SSR-based genetic markers for genome analysis.

Acknowledgments This research was supported in part by grants from the North American Barley Genome Mapping Project and the United States Department of Agriculture/NRI No. 93-37300-8840. We thank A. Graner and D. Hoffman for providing DNA samples of I/F and H/M doubled-haploid lines, respectively, and P. Hayes for supplying the RFLP data for the H/M population. We also thank P. J. Maughan for his critical review of an earlier version of this manuscript

References

Akkaya MS, Bhagwat AA, Cregan PB (1992) Length polymorphisms of simple sequence repeat DNA in soybean. Genetics 132:1132–1139

Akkaya MS, Shoemaker RC, Specht JE, Bhagwat AA, Cregan PB (1995) Integration of simple sequence repeat DNA markers into a soybean linkage map. Crop Sci 35:1439–1445

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

Becker J, Heun M (1995) Barley microsatellites: allele variation and mapping. Plant Mol Biol 27:835–845

Bell CJ, Ecker JR (1994) Assignment of 30 microsatellite loci to the linkage map of *Arabidopsis*. Genomics 19:137–144

Bronson CR, Taga M, Yoder OC (1990) Genetic control and distorted segregation of T-toxin production in field isolates of *Cochliobolus heterotrophus*. Phytopathology 80:819–823

Brummer EC, Bouton JH, Kochert G (1993) Development of an RFLP map in diploid alfalfa. Theor Appl Genet 86:329–332

Cregan PB, Akkaya MS, Bhagwat AA, Lavi U, Rongwen J (1994) Length polymorphisms of simply sequence repeat (SSR) DNA as molecular markers in plants. In: Gresshoff PM (ed) Plant Genome Analysis. CRC Press, Boca Raton, Florida, pp 47–56

Graner A, Siedler H, Jahoor A, Herrmann RG, Wenzel G (1990) Assessment of the degree and type of restriction fragment length polymorphism in barley (*Hordeum vulgare*). Theor Appl Genet 80:826–832

Graner A, Jahoor A, Schondelmaier J, Siedler H, Pillen K, Fischbeck G, Wenzel G, Herrmann RG (1991) Construction of an RFLP map of barley. Theor Appl Genet 83:250–256

Heun M, Kennedy AE, Anderson JA, Lapitan NLV, Sorrels ME, Tanksley SD (1991) Construction of a restriction fragment length polymorphism map for barley (*Hordeum vulgare*). Genome 34:437–447

Islam AKMR, Shepherd KW, Sparrow DHB (1981) Isolation and characterization of euplasmic wheat-barley chromosome addition lines. Heredity 46:161–174

Kandpal RP, Kandpal G, Weissman SM (1994) Construction of libraries enriched for sequence repeats and jumping clones, and hybridization selection for region-specific markers. Proc Natl Acad Sci USA 91:88–92

Kleinhofs A, Kilian A (1994) RFLP maps of barley. In: Phillips RL, Vasil IK (eds) Advances in cellular and molecular biology of plants. vol 1: DNA-based markers in plants. Klumer Academic Publishers, the Netherlands, pp 163–198

Kleinhofs A, Kilian A, Saghai Maroof MA, Biyashev RM, Hayes P, Chen FQ, Lapitan N, Fenwick A, Blake TK, Kanazin V, Ananiev E, Dahleen L, Kudrna D, Bollinger J, Knapp SJ, Liu B, Sorrells M, Heun M, Franckowiak JD, Hoffman D, Skadsen R, Steffenson BJ (1993) A molecular, isozyme and morphological map of the barley (*Hordeum vulgare*) genome. Theor Appl Genet 86:705–712

Lagercrantz U, Ellegren H, Andersson L (1993) The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. Nucleic Acids Res 21:1111–1115

Lander ES, Green P, Abrahamson J, Barlow A, Daly M, Lincoln S, Newburg L (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genetics 1:174–181

Liu, Z-W, Jarret RL, Kresovich S, Duncan RR (1995) Characterization and analysis of simple sequence repeat (SSR) loci in seashore paspalum (*Paspalum vaginatum* Swartz). Theor Appl Genet 91:47–52

Maughan PJ, Saghai Maroof MA, Buss GR (1995) Microsatellite and amplified sequence length polymorphisms in cultivated and wild soybean. Genome 38:715–723

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

Morgante M, Rafalski A, Bidle P, Tingey S, Oliveri AM (1994) Genetic mapping and variability of seven soybean simple sequence repeat loci. Genome 37:763–769

Ostrander EA, Jong PM, Rine J, Duyk G (1992) Construction of small-insert genomic DNA libraries highly enriched for microsatellite repeat sequences. Proc Natl Acad Sci USA 89:3419–3423

Röder MS, Plaschke J, Konig SU, Borner A, Sorrells ME, Tanksley SD, Ganal MW (1995) Abundance, variability and chromosomal location of microsatellites in wheat. Mol Gen Genet 246:327–333

Rongwen J, Akkaya MS, Bhagwat AA, Lavi U, Cregan PB (1995) The use of microsatellite DNA markers for soybean genotype identification. Theor Appl Genet 90:43–48

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

Saghai Maroof MA, Biyashev RB, Yang GP, Zhang Q, Allard RW (1994) Extraordinarily polymorphic microsatellite DNA in barley: species diversity, chromosomal locations, and population dynamics. Proc Natl Acad Sci USA 91:5466–5470

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

Senior ML, Heun M (1993) Mapping maize microsatellites and polymerase chain reaction confirmation of the targeted repeats using a CT primer. Genome 36:884–889

Suiter KA, Wendel JF, Case JS (1983) LINKAGE-1: a PASCAL computer program for the detection and analysis of genetic linkage. J Hered. 74:203–204

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

Walther-Larsen H, Brandt J, Collinge DB, Thordal-Christensen H (1993) A pathogen-induced gene of barley encodes a HSP90 homologue showing striking similarity to vertebrate forms resident in the endoplasmic reticulum. Plant Mol biol 21:1097–1108

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

Weber JL (1990) Human DNA polymorphisms based on length variations in simple-sequence tandem repeats. In: Davies KE, Tilghman SM (eds) Genome analysis, vol 1: genetic and physical mapping, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, pp 159–181

Wu KS, Tanksley SD (1993) Abundance, polymorphism and genetic mapping of microsatellites in rice. Mol Gen Genet 241:225–235

Yang GP, Saghai Maroof MA, Xu CG, Zhang Q, Biyashev RB (1994) Comparative analysis of microsatellite DNA polymorphism in landraces and cultivars of rice. Mol Gen Genet 245:187–194

Yu YG, Saghai Maroof MA, Buss GR, Maughan PJ, Tolin SA (1994) RFLP and microsatellite mapping of a gene for soybean mosaic virus resistance Phytopathology 84:60–64

Zhang Q, Gao YJ, Yang SH, Ragab R, Saghai Maroof MA, Li ZB (1994) A diallel analysis of heterosis in elite hybrid rice based on RFLPs and microsatellites. Theor Appl Genet 89:185–192

Zhao X, Kochert G (1993) Phylogenetic distribution and genetic mapping of a $(GGC)_n$ microsatellite from rice (*Oryza sativa* L.). Plant Mol Biol 21:607–614