Theor Appl Genet (2001) 102:801–809

M. Macaulay · L. Ramsay · W. Powell · R. Waugh

A representative, highly informative 'genotyping set' of barley SSRs

Received: 17 May 2000 / Accepted: 5 September 2000

Abstract We have developed a 'genotyping set' of 48 SSR-based genetic markers for application in genetical studies of barley. The SSRs are a subset of a collection of approximately 600 SSRs available to the barley research community. They have been specifically chosen according to the following criteria: (1) they are single locus; (2) their product quality is good under standard assay conditions; (3) they are distributed across the barley genome; and (4) they exhibit reasonably high polymorphic information content (PIC) values in the cultivated barley gene-pool. To maximise genotyping throughput, one of each SSR primer pair was 5´ end-labelled with either fam, hex or tet fluorochromes to allow automated data capture after running the samples on a DNA sequencer. SSR product sizes were assembled from a reference set of 24 barley genotypes which allowed the construction of 'graphical genotypes' of each of the individual lines. The graphical genotypes provide a convenient tool for interrogating genetic similarity in the individuals surveyed. The product sizes were compared to those obtained from end-labelling one of the primers with ^{33}P and separating the products by denaturing PAGE followed by autoradiography. Although inconsistencies in size were common, they could generally be easily resolved. A reference manual for use of the 'genotyping set' has been produced and is available as a PDF download file at http://www.scri.sari.ac.uk/ssr/pdf. These well-characterised barley SSRs, for the first time, provide a common set of robust PCR-based tools which can be used to integrate and compare information collected from fundamental and/or applied genetic studies on barley in different laboratories across the world.

Communicated by P. Langridge

M. Macaulay · L. Ramsay · W. Powell · R. Waugh (✉)
Genomics Unit, Division of Genetics,
Scottish Crop Research Institute, Invergowrie, Dundee DD2 5DA, Scotland, UK
e-mail: rwaugh@scri.sari.ac.uk
Tel.: +44-01382-568584, Fax: +44-01382-568587

Keywords Barley · SSRs · Genotyping · Mapping · Automation

Introduction

The development and widespread adoption of molecular markers for genetical studies has heralded a fundamental shift from the analysis of phenotype to genotype. However, the expanding number of molecular assays capable of determining genotype has resulted in a situation where data is seldom directly comparable between different studies. This is particularly true of popular generic assays such as RAPD (Williams et al. 1991) and AFLP (Vos et al. 1995) but is also a feature of species-specific assays such as RFLP (Botstein et al. 1980) and STS-based approaches (e.g. CAPS, SCARS, SSCP etc.). While recognising the importance of individual studies which exploit a single marker system, for wider and longer-term value it would be prudent if genetic data were generated using a common approach and recorded in a way that would allow the integration and amalgamation of data sets gathered at different times and in different laboratories around the world. Arguably, only two types of assay presently have the potential to generate such a common data format. These are the analysis of: (1) single nucleotide polymorphisms (SNPs) (Coryell et al. 1999); and (2) length variation in Simple Sequence Repeats (SSRs) (Powell et al. 1996).

SNPs are bi-allelic markers which are routinely discovered by direct comparison of homologous DNA sequences. The allele pool at a given locus is determined empirically but may need to be updated by additional sequence information from cultivated, wild or exotic germplasm. Deployment on a practical scale requires a routine, robust and accessible detection system. A number of approaches for SNP allele detection, based on either allele-specific hybridisation, primer extension or ligation, have been demonstrated or are currently under development (Kuklin et al. 1997; Chen et al. 1998; Cho et al. 1999; Erdos et al. 1999). In anticipation of a suit-

able method being developed, SNP discovery programs are being actively pursued by a number of research groups. However, the likely time-frame before SNP technology is routinely deployed in the community remains unclear. In short, a gap exists between SNP discovery and reduction to practice.

In contrast, SSR-based markers have been around for a considerable time, particularly in human genetics where they have proven to be an exceptionally valuable, well-utilised technology. The main reasons for this are their reproducibility and ease of use, their locus specificity and high information content. Large numbers of SSRs are now becoming available for a selection of crop plants (Roeder et al. 1995; Bryan et al. 1997; Milbourne et al. 1998; Ramsay et al. 2000) and these provide valuable sets of tools for crop-specific research programs. Unlike SNPs, SSR allele characterisation is straightforward, with alleles revealed as a change in the length of the products generated by PCR-amplification across an SSR-containing region using specific primers designed to the unique SSR-flanking sequences. Separation on high-resolution acrylamide gels allows product sizing with base-pair accuracy. A PCR product of a defined molecular size, generated by sequence-defined primer pairs, is diagnostic for a given allele. Given the simplicity and specificity of the assay, and the ease with which a widely reproducible definition of an allele at a locus can be described, a collection of SSR-based markers provides an opportunity to develop a set of coherent and integrated descriptors of a genetic locus, which should have long-term relevance and, therefore, impact upon both data comparison and integration.

However, not all SSRs are equal! They can vary in robustness, quality of the amplification products, amplification of single or multiple loci; they may be clustered on the genome; they may have a low information content in the gene-pool of interest. Choosing the most informative and robust set from those available would be difficult for those not intimately associated with the generation and evaluation of the markers, particularly as much additional information is almost certainly known about the SSRs but is not detailed in scientific reports. We have, therefore, developed a set of 48 SSRs which are of high quality, robust, highly informative in the cultivated barley gene-pool (i.e. have high PIC values), and are distributed across the barley genome. To aid throughput, we have fluorescently labelled one of each primer pair with a different fluorochrome, which allows simultaneous examination and detection of multiple SSR loci on a single run with an automated DNA sequencer.

Materials and methods

The accessions used in the current study, along with their immediate pedigrees, are given in Table 1 (the 'reference genotypes'). DNA was extracted from leaf material from glasshouse-grown plants. The isolation of the microsatellite markers and their location on the barley genetic map has been described previously in detail (Saghai Maroof et al.1994; Becker and Heun 1995; Liu et al. 1996; Struss and Pleiske 1998; Ramsay et al. 2000; Russell et al. 1997a,b; 2000). For detection on an ABI377 DNA sequencer running Genescan and Genotyper software, one of the primers from each pair was synthesised with the 5´ end-nucleotide labelled with fam, hex or tet. The labelled primers are indicated in Table 2. For comparison, the same primer was end-labelled with ^{33}P, as described previously (Russell et al. 1997a,b), and the products separated on PAGE gels and detected by autoradiography.

The alleles detected by each microsatellite were recorded and converted into a numeric scale ranging from one to the number of different microsatellite alleles detected over all 24 reference geno-

Table 1 Genotypes used in this study

Cultivar	Pedigree	Comments
Alexis	Breun 16220×Triumph	Spring, 2-row, 1986, Germany
Aramir	Volla×Emir	Spring, 2-row, 1972, Netherlands
Atem	(L92×Minerva)×Emir)×Zepher	Spring, 2-row, 1979, Netherlands
Banba	Spratt Archer×Kenia	Spring, 2-row, 1960, Ireland
Banteng	Platen 2349×(Vinesco)×Dea)×(Jumbo)	Winter, 6 row, 1976, Germany/Netherlands
Baronesse	(343/6×34/6)×J-427×(Oriol×6153 P40)	Spring, 2-row, 1989, Germany
Beartrice	(Centra×Iris)×(Piroline×Jet)	Spring, 2-row, ?, France
Betina	Mutant of Vada (*Hordeum laevigatum*×Gull)	Spring, 2-row, 1969, France
Blenheim	Triumph×Egmont	Spring, 2-row, 1985, UK
Claret	[(Proctor×HP5466)×Armelle]×Abacus	Spring, 2-row, 1980, UK
Corniche	Beindorf 99991/70×640	Spring, 2-row, 1983, Germany
Criewener 403	Selection from Proskowetz Hanna	Spring, 2-row, 1910, Germany
Golf	(Armelle×Lud)×Luke	Spring, 2-row, 1982, UK
Grammos	Mutant of cv Rivale	Winter, 2-row, 1971, Greece
Grete	Stamm 732219×(Pella×Dura)	Winter, 6 row, 1989, Germany
Grit	(S5474/67×46459/68)×480/68	Spring, 2-row, 1977, Germany
Hana	Selection from Landrace Hanna	Spring, 2-row, 1945, Austria
HS92	*H. spontaneum*	*H. spontaneum*
Igri	Malta×(Carlsberg 1427×Ingrid)	Winter, 2-row, 1976, Germany
Ingrid	Balder×(Binder×Opal)	Spring, 2-row, 1958, Sweden
Lina	(Mari *5×Multan)×A6564/3×Lofa	Spring, 2-row, 1982, Sweden
Morex	Cree×Bonanza	Spring, 6 row, 1979, USA
Plaisant	Ager×Nymphe	Winter, 6 row, 1979, France
Steptoe	WA3564×Unitan×(Muir×Nilan)	Spring, 6 row, 1973, USA

types. From the frequencies of the 1 to *n* alleles detected by each microsatellite, polymorphic information content (*PIC*) values were calculated as:

$$PIC = 1 - \left(\sum_{i=1}^{n} p_i^2\right) - \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} 2 p_i^2 p_j^2$$

(Weber 1990); where p_i and p_j are the frequencies of the i^{th} and j^{th} alleles in a given population.

Results and discussion

Choice of SSRs for the genotyping set

Liu et al. (1996), Struss and Pleiske (1998) and Ramsay et al. (2000) have isolated a total of >600 SSRs which can be deployed in genetical studies in barley. Of these, well over half have been intra-chromosomally mapped in segregating doubled-haploid experimental populations. For many studies – particularly those involving germplasm evaluation for academic or statutory purposes – large numbers of markers are not required. Rather, a well-distributed, highly informative and robust set of markers would be of particular value. This requirement has already been recognised for several animal species, and internationally approved sets of highly informative SSRs have been developed and are widely and reliably used, e.g. for parentage testing and other applications. There are obvious advantages to having an agreed set of standard genotyping tools, the greatest being the direct comparability it affords between different laboratories and different studies. As such, a common set of tools facilitates the assembly of an SSR allele database, which can be used as a reliable reference for future applications such as the protection of plant variety rights. Recognising the potential benefits, we have identified a set of 48 SSRs, which should satisfy many of the potential applications of molecular markers in barley. We would like to encourage the use of this preliminary 'genotyping set' in order that the barley research community establish a platform from which it can reap the benefits of comparability and integration.

The SSRs were chosen on the basis of a number of criteria. First, markers were identified which gave representative genome coverage. Second, autorads of the original mapping population were checked to confirm that the chosen SSRs were single-copy. Third, data from germplasm screens of the 24 reference genotypes (Table 1) were analysed to derive PIC values and provide an estimate of their level of informativeness. PIC values ranged from 0.08 (HvLOX) to 0.94 (Bmac0040) in the 24 genotypes surveyed, with an average PIC value of 0.64. SSRs mapped on the Lina×*Hordeum spontaneum* genetic map (an interspecific cross, chosen originally to map SSRs because it was very diverse) exhibit pronounced clustering around the centromeric regions of each linkage group (Ramsay et al. 2000). While some expansion of the map distances are observed in intraspecific populations (Graner et al. 1991; Kleinhofs et al. 1993; our unpublished results), the biased genetic distribution of the SSRs imposed some limits on the development of a set with a uniform genome distribution. For example, on chromosome 6H, a marker interval of approximately 70 cM exists between BMac0316 and BMag0173. While other markers are mapped within this region, none exhibited the dual demands of single-copy and high quality required for the development of a robust genotyping marker. Additional markers (in subsequent genotyping sets) will be required to fill such gaps. In contrast, the centromeric regions of all linkage groups contain several very good quality single-locus markers. Largely by default, we have included at least a pair of markers which are genetically tightly linked in the centromeric portions of each linkage group.

Development of the genotyping set

Having identified a suitable set of SSRs, one primer from each pair was re-synthesised with a fam, hex or tet 5′ end-labelled nucleotide to allow data capture on an automated DNA sequencer. The labels were chosen for each of the primers to allow multiple SSRs to be separated on a single gel in order to increase throughput. Each primer pair was used to screen the reference set for allelic diversity. With this approach, separation of the alleles was easily and reliably achieved. As an example, Fig. 1 shows an electropherogram of the BMac0032 locus on six of the reference genotypes. The characteristic stuttering of dinucleotide SSR-based markers is clearly evident. All 48 primer pairs were used to screen the reference DNAs. Table 2 presents the primer information for the 48 SSRs in the genotyping set, the SSR repeat motif and fluorescent label. Their map position (on the Lina×*Hordeum spontaneum* (HS92) map), allele sizes on the reference genotypes, and PIC values are given in Fig. 2.

Running fluorescently labelled SSRs on an automated DNA sequencer has many advantages over conventional analysis on denaturing PAGE with radiolabelled primers. On a 96-lane ABI377 we are currently able to multiplex three samples in single run but can also re-load the same gel up to two additional times, i.e. we are able to obtain up to 900 genotypic datapoints per run with automatic data capture and recording (although 300–600 is more routine).

Comparison between isotopic- and fluorescence-based allele detection and sizing

Allele sizing on the ABI using 'Genescan' and 'Genotyper' software is to two decimal places. One consequence of this is that alleles are best placed into 'bins', the resolution of which may be arbitrarily assigned, but is usually 1 bp or 2 bp depending on the SSR. Visual inspection of the tracks is also necessary to determine whether two alleles are different or the same, especially when stuttering is pronounced. In contrast, SSRs on autorads are by default given an exact bp call. Perhaps unsurprisingly, changing the format from isotopic to fluorescence detec-

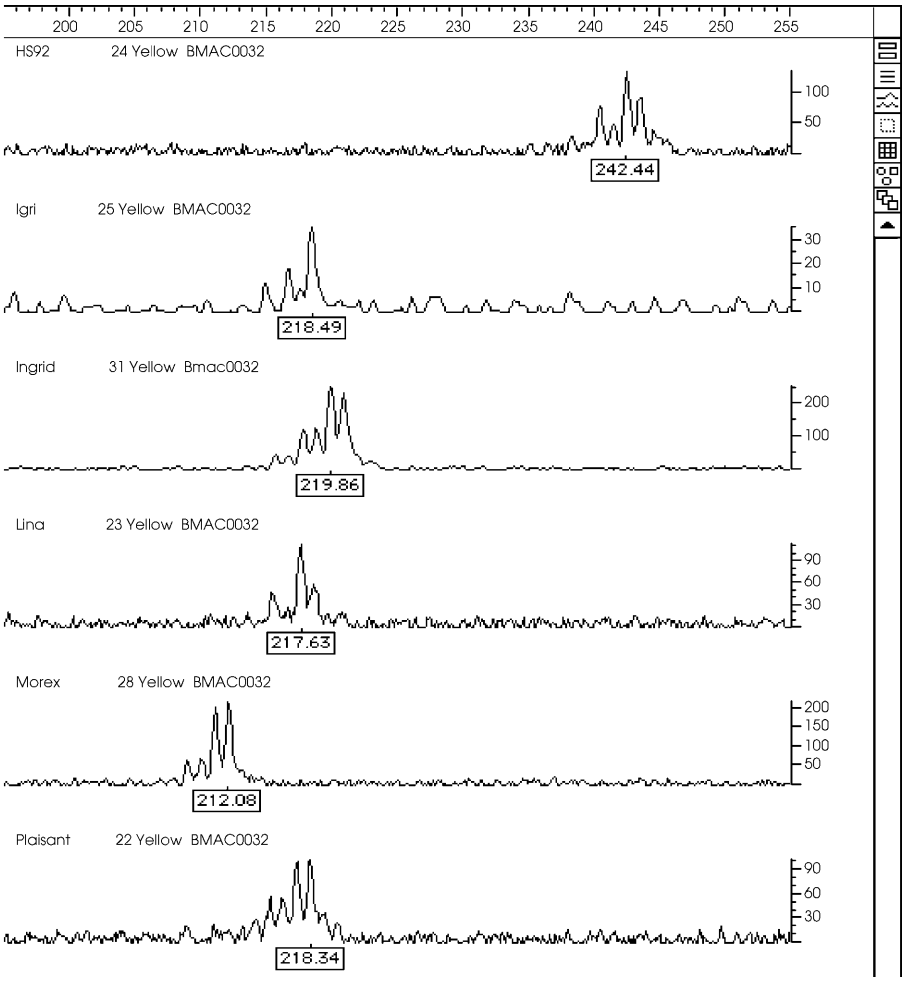


Fig. 1 Electropherogram of alleles scored at the Bmac0032 locus in six barley cultivars

tion revealed inconsistencies in estimated allele sizes. Thus, while 43 of the SSRs revealed a pattern of alleles that was the same on both platforms, the estimated sizes were out by as much as 4 bp up or down (consistently for all genotypes). As a consequence, the inclusion of 'reference genotypes' is essential when setting up the parameters for any new SSR genotyping assay or any new platform. Along with the use of allele 'bins', we have found that 'reference genotypes' eliminate discrepancies between gels and between laboratories and support the smooth integration of allelic data (Williams et al. 1999). On a cautionary note, within this set of 43 SSRs, on 11 occasions, certain genotypes had clearly a 'wrong' allele-size called (Morex for Bmac0209, Bmac0136, HVM67 and HvLOX; Lina for Bmag0007, Bmag0211, HvHVA1; Alexis and Plaisant for Bmag0211; Igri for Bmag0225 and HS92 for Bmac0273). In the majority of cases, after close examination, the allele scored on the ABI at each of these loci was the same as that in an adjacent track. Apparently, PCR products from the adjacent tracks had leaked into the track containing what must have been a failed PCR reaction for the above genotypes. While this was verified by repeating the relevant assays, it highlights a potential problem of automated data capture which we would not have noticed had we not been comparing the allele-calls across platforms. Sensitivity of detection and 'tracking' problems contributed to the 'false calls'. However, a combination of two observations, equal allele sizes in adjacent lanes and low signal strength in one of them, should flag potential problems of this kind.

Five of the 48 SSRs were more difficult to compare. For these (Bmac0040, Bmag0135, Bmag0222, Bmag0382 and EBmac0415) the pattern of allele sizes across the genotypes appeared superficially similar but the estimated allele sizes, and thus the total number of alleles detected, had several inconsistencies. Close examination of the autorads (and ABI trace files) suggested that the difficulties in estimating sizes came from a combination of: (1) a large PCR product size causing inaccurate size estimates; (2) stuttering; and (3) a large allele size range (which meant that scoring alleles on autorads was particularly difficult). As Genescan and Genotyper software is considered to provide a greater level of accuracy in estimating allele size (Kimpton et al. 1994), in these cases we took the ABI score as being correct.

One of the original criteria we applied when developing this set of markers was to use those which revealed a maximum number of alleles in the germplasm we were primarily interested in studying (e.g. Bmac0399, Bmac0040, and Bmag0225). Retrospectively, these markers required most

Table 2 SSRs in the genotyping set: primer sequences and repeat motif

SSR	Forward primer	Reverse primer (labelled)	Repeat
Bmac0399	CGATGCTTTACTATGAGAGGT	GGGTCTGAAGCCTGAAC (5′ fam)	(AC)21
Bmac0032	CCATCAAAGTCCGGCTAG	GTCGGGCCTCATACTGAC (5′ hex)	(AC)7 T(CA)15(AT)9
Bmag0211	ATTCATCGATCTTGTATTAGTCC	ACATCATGTCGATCAAAGC (5′ hex)	(CT)16
Bmag0382	TGAAACCCATAGAGAGTGAGA	TCAAAAGTTTCGTTCCAAATA (5′ fam)	(AG)7AA(AG)7
HvHVA1	CATGGGAGGGGACAACAC	CGACCAAACACGACTAAAGGA (5′ tet)	(ACC)5
WMC1E8	TCATTCGTTGCAGATACACCAC	TCAATGCCCTTGTTTCTGACCT (5′ hex)	(AC)24
Bmac0134	CCAACTGAGTCGATCTCG	CTTCGTTGCTTCTCTACCTT (5′ fam)	(AC)28
HVM36	TCCAGCCGACAATTTCTTG	AGTACTCCGACACCACGTCC (5′ fam)	(GA)13
Bmag0378	CTTTTGTTTCCGTAGCATCTA	ATCCAACTATAGTAGCAAAGCC (5′ tet)	(AG)14
Bmac0093	CGTTTGGGACGTATCAAT	GGGAGTCTTGAGCCTACTG (5′ hex)	(AC)24
Bmag0125	AATTAGCGAGAACAAAATCAC	AGATAACGATGCACCACC (5′ tet)	(AG)19
HVM54	AACCCAGTAACACCTGTCCTG	AGTTCCCTGACCCGATGTC (5′ hex)	(GA)14
EBmac0415	GAAACCCATCATAGCAGC	AAACAGCAGCAAGAGGAG (5′ fam)	(AC)17
HvLTPPB	AGACGCTGAGTACGTTGAG	CAAAGTACAACAAACTCACGA (5′ tet)	(AC)10(AT)5
Bmac0067	AACGTACGAGCTCTTTTTCTA	ATGCCAACTGCTTGTTTAG (5′ fam)	(AC)18
Bmac0209	CTAGCAACTTCCCAACCGAC	ATGCCTGTGTGTGGACCAT (5′ tet)	(AC)13
Bmag0136	GTACGCTTTCAAACCTGG	GTAGGAGGAAGAATAAGGAGG (5′ tet)	(AG)6-(AG)10-(AG)6
Bmag0225	AACACACCAAAAATATTACATCA	CGAGTAGTTCCCATGTGAC (5′ hex)	(AG)26
Bmag0013	AAGGGGAATCAAAATGGGAG	TCGAATAGGTCTCCGAAGAAA (5′ tet)	(CT)21
HVM62	TCGCGACCAGACGAGAAG	AGCTAGCCGACGACGCAC (5′ fam)	(GA)11
HVM40	CGATTCCCCTTTTCCCAC	ATTCTCCGCCGTCCACTC (5′ hex)	(GA)6(GT)4(GA)7
Bmag0384	TGTGAGTAGTTCACCATAGACC	TGCCATTATCATTGTATTGAA (5′ fam)	(AG)18
HVM03	ACACCTTCCCAGGACAATCCATTG	AGCACGCAGAGCACCGAAAAAGTC (5′ tet)	(AT)29
Bmag0353	ACTAGTACCCACTATGCACGA	ACGTTCATTAAAATCACAACTG (5′ fam)	(AG)21
EBmac0701	ATGATGAGAACTCTTCACCC	TGGCACTAAAGCAAAAGAC (5′ fam)	(AC)23
HvMLO3	CTTCCATGTCACCTACAG	CGAACTGGTATTCCAAGG (5′ hex)	(CTT)6
HVM67	GTCGGGCTCCATTGCTCT	CCGGTACCCAGTGACGAC (5′ tet)	(GA)11
EBmac0970	ACATGTGATACCAAGGCAC	TGCATAGATGATGTGCTTG (5′ hex)	(AC)8
Bmac0113	TCAAAAGCCGGTCTAATGCT	GTGCAAAGAAAATGCACAGATAG (5′ tet)	(AT)7(AC)18
EBmac0684	TTCCGTTGAGCTTTCATACAC	ATTGAATCCCAACAGACACAA (5′ hex)	(TA)7(TG)11-(TG)11(TTTG)5
Bmag0223	TTAGTCACCCTCAACGGT	CCCCTAACTGCTGTGATG (5′ tet)	(AG)16
HVLEU	TTGGAAGTGTACAGCAATGGAG	TGAAAGGCCCCACAAGATAG (5′ fam)	(ATTT)4
Bmag0222	ATGCTACTCTGGAGTGGAGTA	GACCTTCAACTTTGCCTTATA (5′ fam)	(AC)9(AG)17
HvLOX	CAGCATATCCATCTGATCTG	CACCCTTATTTATTGCCTTAA (5′ fam)	(AG)9
Bmac0316	ATGGTAGAGGTCCCAACTG	ATCACTGCTGTGCCTAGC (5′ fam)	(AC)19
Bmag0173	CATTTTTGTTGGTGACGG	ATAATGGCGGGAGAGACA (5′ hex)	(CT)29
Bmag0218	CATAGAGAGGGAGGGAGAG	TCAACCTTACTGCATCTTTG (5′ tet)	(AG)6(AG)6
Bmac0018	GTCCTTTACGCATGAACCGT	ACATACGCCAGACTCGTGTG (5′ hex)	(AC)11
Bmag0009	AAGTGAAGCAAGCAAACAAACA	ATCCTTCCATATTTTGATTAGGCA (5′ fam)	(AG)13
EBmac0806	ACTAAGTCCTTTCACGAGGA	GTGTGTAGTAGGTGGGTACTTG (5′ tet)	(CA)4(GA)(CA)8-(CA)5
Bmac0040	AGCCCGATCAGATTTACG	TTCTCCCTTTGGTCCTTG (5′ fam)	(AC)20
Bmag0021	ATTTTTATCAGAACGTCTCTCTC	CTAACTTCTCTCTCCCTCTCC (5′ tet)	(CA)10AA(GA)28
Bmag0206	TTTTCCCCTATTATAGTGACG	TAGAACTGGGTATTTCCTTGA (5′ hex)	(GT)5(AG)14
HVCMA	GCCTCGGTTTGGACATATAAAG	GTAAAGCAAATGTTGAGCAACG (5′ fam)	(AT)9
Bmac0273	ACAAAGCTCGTGGTACGT	AGGGAGTATTTCACCCTTG (5′ tet)	(AC)20(AG)20
Bmag0120	ATTTCATCCCAAAGGAGAC	GTCACATAGACAGTTGTCTTCC (5′ hex)	(AG)15
Bmac0156	AACCGAATGTATTCCTCTGTA	GCCAAACAACTATCGTGTAC (5′ tet)	(AC)22(AT)5
Bmag0135	ACGAAAGAGTTACAACGGATA	GTTTACCACAGATCTACAGGTG (5′ tet)	(AG)10GG(AG)12

time for interpretation when comparing data between runs or from a very large sample set. It is likely they will also be troublesome when comparing data between laboratories. As a result, we are currently choosing SSRs which reveal five or six alleles, preferably exhibiting step-wise mutations, to develop a second genotyping set of 48 SSRs to complement that described here. This should be available in the near future.

Graphical genotyping

Mapping major genes or QTLs onto genetic linkage maps provides primary information on a gene or locus which may be involved in a given trait. The application of this type of information is, however, frequently restricted to crosses or populations involving the actual parent which was used to map the favourable alleles in the first place. In many respects, determining the map location of a gene or QTL can be considered a two-dimensional exercise as it fails to evaluate the potential allelic variation for that trait within the gene-pool.

Genotypic evaluation, using mapped-SSRs, provides a convenient way of extending the results of linkage studies to a much broader germplasm base (i.e. it adds a third dimension). If we consider that portions of the genome between SSRs represent ‘linkage blocks’ that may or may not have been disturbed during the development

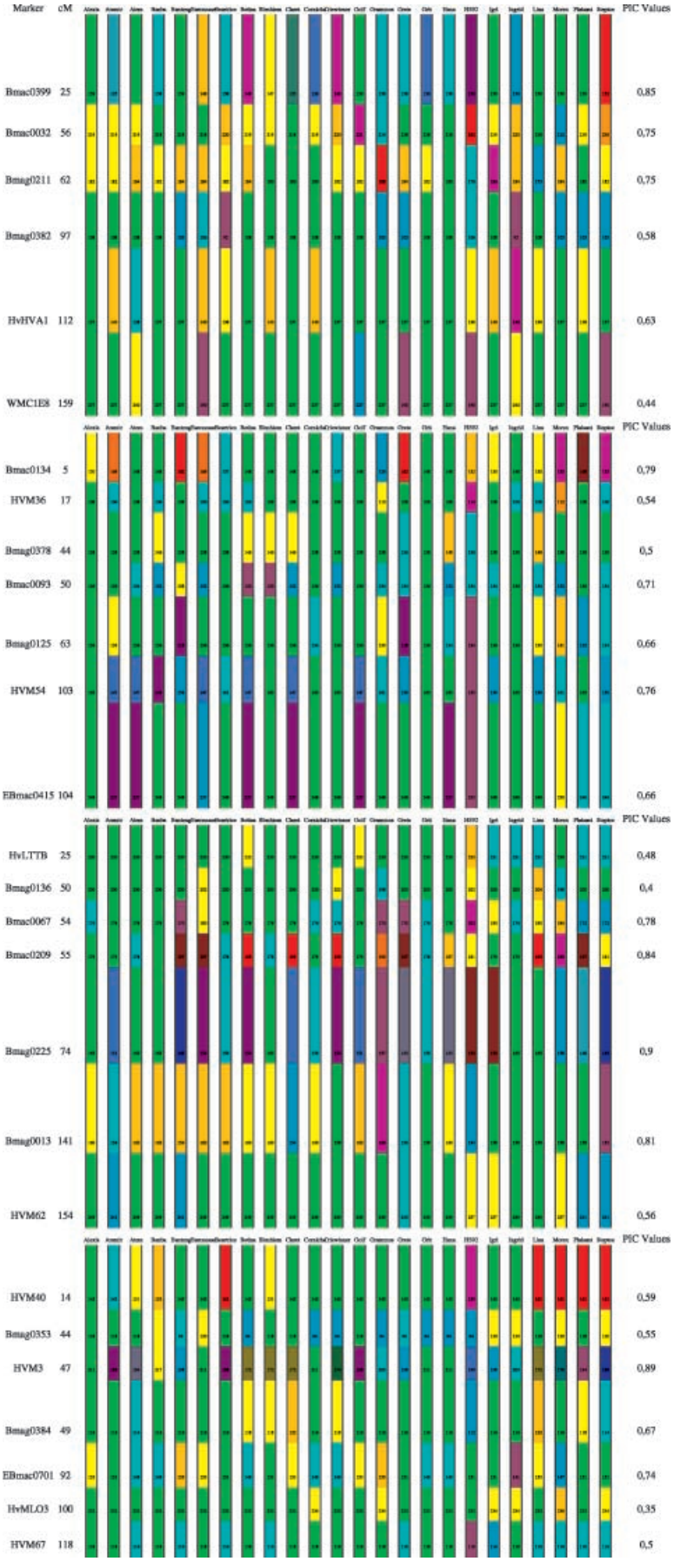
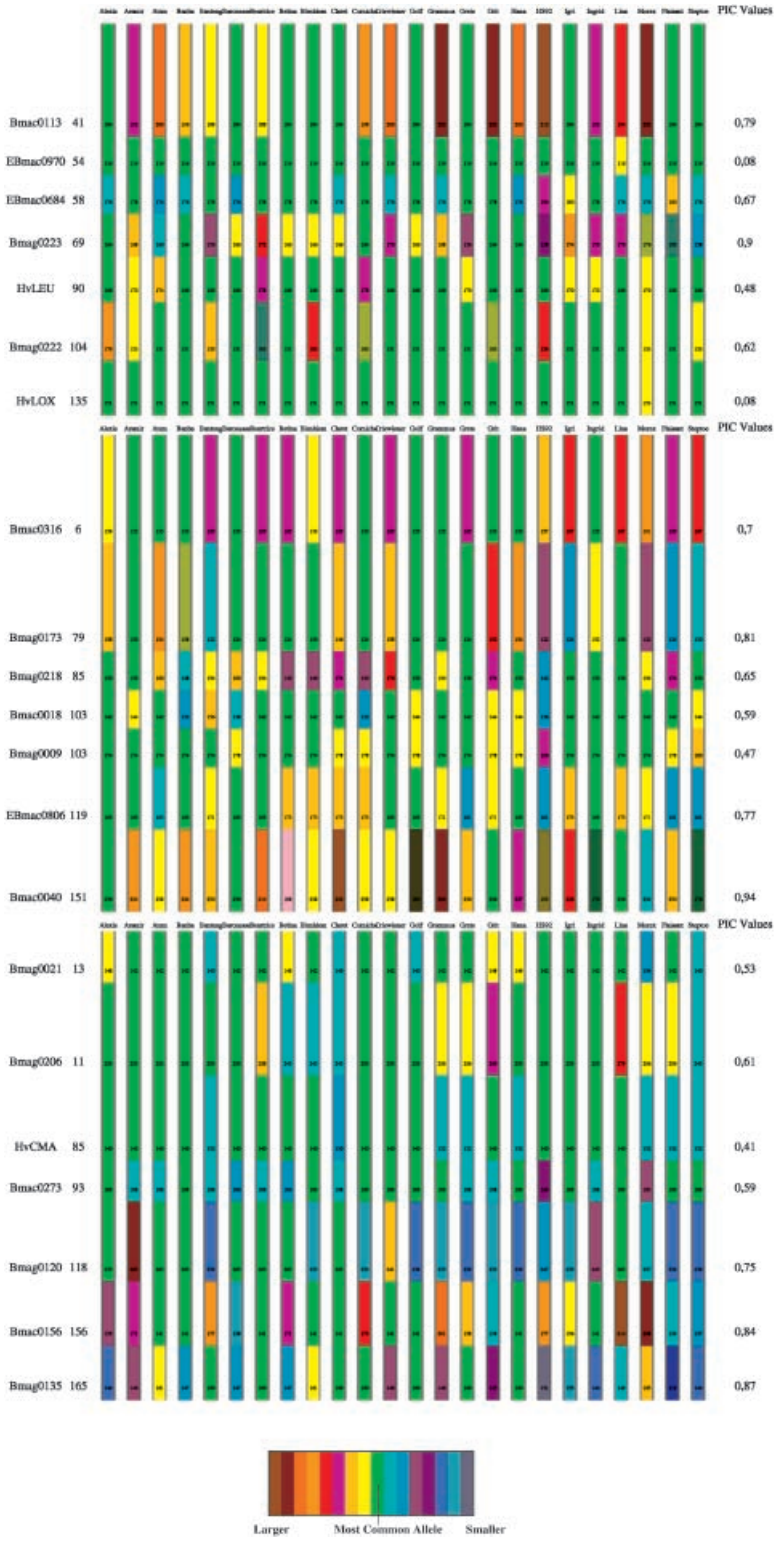


Fig. 2 Graphical genotypes of the 24 accessions surveyed in the development of the genotyping set of SSRs. Chromosomes 1H–7H are arranged in columns with markers in rows. Marker boxes are scaled according to the genetic distance between the SSRs in the Lina×*H. spontaneum* population (Ramsay et al. 2000). The position of each marker (in cM) from the top of the short arm of each barley linkage group is given next to the marker name (from the Lina×*H. spontaneum* population). The allele size scored for each SSR locus is given in each *'box'* and PIC values for each of the loci are given at the *right hand side* of each chromosome. The most common allele at each locus is *coloured green* with *yellow* and *light blue* representing the most-closely sized larger or smaller alleles, respectively. Other alleles were assigned different colours, according to the *'scale'* at the bottom

Fig. 2 (continued)



of new varieties, then the same SSR alleles flanking a given 'linkage block' in different accessions may be good evidence for considering the region to be identical by descent. If this assumption is correct, then similarly (SSR) defined linkage blocks will contain the same genetic (allele) information. Linkage blocks defined by different SSR alleles will, therefore, by definition, comprise a different set of gene alleles. In this case, a higher resolution of markers will be required to define smaller linkage blocks. The second set of genotyping SSRs mentioned above will help further-define the marker resolution required to maximise our ability to interpret genetic

information scrambled by plant breeding. Figure 2 shows graphical genotypes, constructed manually in MS Excel, of the 24 accessions used in this study. Without further interpretation, this information could be immediately exploited in a number of ways: e.g. it could form the basis of parental choice in crossing schedules; it could be used to infer the position of favourable alleles affecting major genes or quantitative traits (through association genetics); or to mine the gene-pool for novel alleles. To exploit mapped SSRs for these types of application, we are currently constructing a 'genotypic database' of a large set of barley cultivars and accessions (>500) with the markers described in this report. The database will be mounted on the World Wide Web (as part of the UK CropNet – http://synteny.nott.ac.uk/barley.html) and be accessible by all interested parties who will be able to interactively compare SSR-derived genotypes of barley from their laboratory with the existing catalogue.

Conclusions

The SSRs described here represent a highly informative set of molecular markers which are robust, easy to use, easy to interpret and record. By adopting the 'allele binning' concept and routine comparison to reference genotypes, they will facilitate the generation of information which will allow the integration of comparable data from different sources and different studies. This represents a significant improvement over current genotyping methods, which rarely allow remote comparison or integration. While we have used fluorescently labelled primers for automation detection in our laboratory, for low-tech deployment isotopic labelling or silver staining will also suffice, providing that comparison with the reference genotypes is routinely performed.

In the longer term, it is likely that SNP data will ultimately provide higher resolution and afford higher mutational stability than SSRs, along with the potential for automated, high-throughput genotyping systems. However, the development of a representative SNP collection is resource-demanding and, as yet, no SNP detection platform has been generally agreed by the community. Thus, it is likely that it will be some time before SNP-based genotyping is widely deployed. In contrast, the SSR genotyping set described here, and a subsequent set under development, provide a resource which can be deployed immediately in laboratories with widely differing technical capabilities, to generate data which can be integrated with, and interpreted in, a global context.

Acknowledgements The authors acknowledge grants received from the EUCT95–0003 (L. Ramsay) and BBSRC (M. Macaulay), which enabled this research. R. Waugh and W. Powell are funded by SERAD. W. Thomas and J. Russell are thanked for critically reading the manuscript.

References

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

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in using restriction fragment length polymorphisms. Am J Hum Genet 32:314–331

Bryan GJ, Collins AJ, Stephenson P, Smith JB, Orry A, Gale MD (1997) Isolation and characterisation of microsatellites from hexaploid bread wheat. Theor Appl Genet 94:557–563

Cho RJ, Mindrinos M, Richards DR, Sapolsky RJ, Anderson M, Drenkard E, Dewdney L, Reuber TL, Stammers M, Federspiel N, Theologis A, Yang WH, Hubbell E, Au M, Chung EY, Lashkari D, Lemieux B, Dean C, Lipshutz RJ, Ausubel FM, Davis RW, Oefner PJ (1999) Genome-wide mapping with biallelic markers in *Arabidopsis thaliana.* Nature Genet 23:203–207

Chen XN, Livak KJ, Kwok PY (1998) A homogeneous, ligase-mediated DNA diagnostic test. Genet Res 8:549–556

Coryell VH, Jessen H, Schupp JM, Webb D, Keim P (1999) Allele-specific hybridization markers for soybean. Theor Appl Genet 98:690–696

Erdos MR, Braun A, Mohike K, Silander K, Ghosh S, Collins FS (1999) MALDI-TOF mass spectrometry for high-throughput SNP detection: application to type-2 diabetes. Am J Hum Genet 65:SS A248

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

Kimpton C, Fisher D, Watson S, Adams M, Urquhart A, Lygo J, Gill P (1994) Evaluation of an automated DNA profiling system employing multiplex amplification of 4 tetrameric STR loci. Int J Legal Med 106:302–311

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

Kuklin A, Munson K, Gjerde D, Haefele R, Taylor P (1997) Detection of single-nucleotide polymorphisms with the WAVE(TM) DNA fragment analysis system. Genet Testing 1:201–206

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

Milbourne D, Meyer RC, Collins AJ, Ramsay LD, Gebhardt C, Waugh R (1998) Isolation, characterisation and mapping of simple sequence repeat loci in potato. Mol Gen Genet 259: 233–245

Powell W, Machray GC, Provan J (1996) Polymorphism revealed by simple sequence repeats. Trends Plant Sci 1:215–222

Ramsay L, Macaulay M, McLean K, Fuller J, Edwards K, Turesson S, Morgante M, degli-Ivanissivich S, Marmiroli N, Maestri E, Massari A, Powell W, Waugh R (2000) A simple sequence repeat-based linkage map of barley. Genetics 156: 1997–2005

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

Russell JR, Fuller JD, Macaulay M, Hatz BG, Jahoor J, Powell W, Waugh R (1997a) Direct comparison of levels of genetic variation among barley accessions detected by RFLPs, AFLPs, SSRs and RAPDs. Theor Appl Genet 95:714–722

Russell JR, Fuller JD, Young G, Thomas B, Taramino G, Macaulay M, Waugh R, Powell W (1997b) Discriminating between barley genotypes using microsatellite markers. Genome 40:442–450

Russell JR, Ellis RP, Thomas WTB, Waugh R, Provan J, Booth A, Fuller JD, Lawrence P, Young G, Powell W (2000) A retro-

spective analysis of spring barley germplasm development from ‘foundation genotypes’ to currently successful cultivars. Mol Breed (in press)

Struss L, Plieske P (1998) The use of microsatellite markers for detection of genetic diversity in barley populations. Theor Appl Genet 97:308–315

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

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414

Weber JL (1990) Informativeness of human $(dC–dA)_n \cdot (dG–dT)_n$ polymorphisms. Genomics 7:524–530

Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1991) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res 18:6531–6541

Williams LL, Hedge MR, Herrera G, Stapleton PM, Love DR (1999) Comparative semi-automated analysis of (CAG) repeats in the Huntington disease gene: use of internal standards. Mol Cell Probes 13:283–289