Abstract: Short straw is a desired trait in oat germplasm (*Avena sativa* L.). Marker-assisted selection, a key tool for achieving this objective, is limited by the presence and number of available markers. Here, we have attempted to develop markers sufficiently linked to a gene specifying short straw so that marker-assisted selection could be applied. Bulked-segregant analysis was used to identify anonymous PCR-based markers associated with the dwarfing gene *Dw6* in an F2 population from the cross between *A. sativa* ‘Aslak’ and *A. sativa* ‘Kontant’. One random amplified polymorphic DNA (RAPD) and 1 retrotransposon-microsatellite amplified polymorphism (REMAP) marker were found to be associated with height. These were converted into codominant single-nucleotide polymorphism (SNP) markers. The SNP–REMAP and the SNP–RAPD markers were located 5.2 and 12.6 cM from *Dw6*, respectively. They can be used in future efforts both to enhance oat germplasm by application of molecular markers and to determine the nature of the gene through positional cloning.

Key words: *Avena sativa*, short straw, marker-assisted selection, RAPD, REMAP, SNP.

Résumé : Une paille courte est un caractère souhaitable chez l’avoine (*Avena sativa* L.). La sélection assistée de marqueurs, un outil important pour atteindre cet objectif, est toujours limitée par le nombre de marqueurs. Les auteurs ont tenté de développer des marqueurs suffisamment proches d’un gène contribuant à une courte paille pour que la sélection assistée puisse se pratiquer. Une approche BSA (analyse des ségrégants en mélange) a été employée pour identifier des marqueurs PCR anonymes liés au gène de nanisme *Dw6* au sein d’une population F2 issue d’un croisement entre *A. sativa* L. ‘Aslak’ et *A. sativa* ‘Kontant’. Un marqueur ADN polymorphe amplifié au hasard (RAPD) et un marqueur polymorphisme amplifié microsatellite-rétrotransposon (REMAP) étaient associés à la hauteur. Ces marqueurs ont été convertis en polymorphisme mononucléotidique (SNP) codominants. Les marqueurs SNP–REMAP et SNP–RAPD ont été situés à 5,2 et 12,6 cM de *Dw6*, respectivement. Ils pourront servir à la fois pour améliorer les ressources moléculaires et pour déterminer la nature du gène via un clonage positionnel.

Mots clés : *Avena sativa*, paille courte, sélection assistée de marqueurs, RAPD, REMAP, SNP.
nously applied gibberellic acid (GA) (Reid 1986). Eight dwarfing genes have been genetically classified in oat (Marshall and Murphy 1981; Milach et al. 1998), of which the first 5 are of no practical use because they confer either extreme dwarfing or meiotic irregularities (Marshall and Murphy 1981). The 3 mutants of practical value, Dw6, Dw7, and Dw8 (Marshall and Murphy 1981; Milach et al. 1998), are GA sensitive (Milach 1995). Generally, most GA-sensitive mutants are recessive (Herskowitz 1987), although Dw6 (Brown et al. 1980) and Dw8 (Milach et al. 1998) are dominant and Dw7 is partially dominant (Marshall and Murphy 1981). Only Dw6 and Dw7 have been used in cultivar development (Milach and Federizzi 2001).

Marker-assisted selection (MAS), the use of molecular markers linked to a desired gene, as an alternative to phenotyping could facilitate germplasm enhancement, particularly for the dominant dwarfing genes. A MAS approach with codominant markers obviates the need for progeny testing to distinguish homozygotes and heterozygotes for short straw, and opens the door to genetic characterization of the mutation. Markers that are closely linked to the gene of interest are prerequisites for MAS; such markers can be established with linkage mapping or bulked-segregant analysis (BSA; Michelmore et al. 1991).

The purpose of the present study was to identify anonymous PCR-based markers associated with the dwarfing gene Dw6 in oat (Avena sativa L.) using BSA. We aimed to convert the best markers into easily scorable and reproducible PCR markers with specific primers.

The nucleotide sequence data reported here have been deposited in GenBank under the accession Nos. AY737795 (REMAP sequence from Avena sativa ‘Kontant’) and AY737796 (REMAP sequence from ‘Kontant’).

**Material and methods**

**Plant material**

To obtain a population varying in height, the Finnish early, high-quality, spring oat Avena sativa ‘Aslak’ (Boreal Plant Breeding Ltd., Jokionen, Finland) was crossed with the short Dutch oat Avena sativa ‘Kontant’ (LWZ 95-208) (Zelder b.v., represented by Wiersum b.v., Dronten, Netherlands). The pedigree of ‘Kontant’ is ZE 90-214 × (CB 8853 × ‘Aintree’), the dominant dwarfing gene being inherited from the line CB 8853 (Cebeco Seeds Group b.v., Vlijmen, Netherlands). This dwarfing gene was produced by irradiating line OT 184 with fast neutrons, resulting in the mutant line OT 184D (McKenzie and Brown 1973), which was later renamed OT 207, and the mutant gene Dw6 (Brown et al. 1980). Dw6 is GA sensitive (Burrows 1978). The mutant line is 34%–37% shorter than the progenitor line (Brown et al. 1980; Milach et al. 2002). The number of internodes is the same in the mutant as in the progenitor line, but the 3 uppermost internodes are 30%–50% shorter (Milach et al. 2002). Panicles are also 16%–18% shorter than normal.

All plants in this study were grown under controlled conditions in a greenhouse. One individual from the cross between ‘Aslak’ and ‘Kontant’ was self pollinated to produce a mapping population of 111 F2 plants. DNA was extracted using the DNeasy® Plant Mini Kit (Qiagen, Valencia, Calif.). Fresh leaf material (approximately 100 mg) was crushed with a FastPrep™ FP120 (BIO 101, Thermo Savant, Waltham, Mass.). DNA concentrations were then measured using the GeneQuant II RNA/DNA Calculator (Pharmacia Biotech, Piscataway, N.J.).

**Height measurements**

The heights of 111 F2 individuals were measured from the soil surface to the top of the panicle on the main tiller of each plant at maturity. To discriminate between homozygotes and heterozygotes for short straw, 15–30 F3 seeds from each self-pollinated F2 individual were sown and the heights of the plants measured.

**Bulking of F2 individuals**

The BSA method (Michelmore et al. 1991) was used to examine molecular markers for association with plant height. For this purpose, DNA from the 9 shortest and 9 tallest F2 individuals were pooled. Marker screening of the bulks was initiated before the heights of the F3 progenies were measured. The results revealed that the short bulk contained 5 heterozygotes for Dw6 and, therefore, a new short bulk was created.

**Markers**

The random amplified polymorphic DNA (RAPD) primers consisted of random 10mers or 11mers from Operon Technologies (Alameda, Calif.). RAPD analysis was carried out as follows: the reaction volume of 25 µL contained 0.75 U Taq polymerase (MBI Fermentas, St. Leon-Rot, Germany), buffer containing (NH4)2SO4 as supplied by the enzyme manufacturer, 1.9 mmol/L MgCl₂, 100 µmol/L of each dNTP, 600–1200 nmol/L primer, and 20 ng DNA. The PCR program consisted of an initial denaturation step of 3 min at 94°C; 35 cycles of 30 s at 94°C, 30 s at 38°C, and 1 min at 72°C; and a final extension step of 5 min at 72°C. The amplifications were performed in an MJ Research PTC-100 thermal cycler (Waltham, Mass.).

Retrotransposon-based inter-retrotransposon amplified polymorphism (IRAP) markers detect polymorphisms in amplification products generated between retrotransposons; REMAPs (retrotransposon-microsatellite amplified polymorphism) detect polymorphisms between retrotransposons and microsatellites (Kalander et al. 1999; Schulman et al. 2004). Although retrotransposon markers have been applied to build a map of diploid Avena (Yu and Wise 2000), there has not yet been any attempt to apply them in MAS. Primers suitable for IRAP in oat were designed by cloning retrotransposon regions from the oat genome, identifying the long terminal repeats (LTRs), and choosing conserved motifs at or near their termini. The microsatellite-based primers contained sets of repeat units anchored at their 3′ ends by a nucleotide outside of the repeat. IRAP markers were amplified in a reaction volume of 20 µL, using 1 U Taq polymerase (MBI Fermentas), the enzyme manufacturer’s buffer containing (NH4)2SO4, 1.5 mmol/L MgCl₂, 200 µmol/L each dNTP, 200 nmol/L primer, and 20 ng DNA. The PCR program was run in an MJ Research PTC-100 thermal cycler and consisted of an initial denaturation step of 2 min at 94°C; 45 cycles of 20 s at 94°C, 20 s at 60°C, and 2 min at 72°C; a final extension step of 10 min at 72°C. The PCR program for REMAP was the same as for IRAP with the fol-
lowing exceptions: it was run for 30 cycles, the annealing temperature was 58 °C, and Failsafe™ 2× PCR PreMix D (Epitect, Madison, Wis.), containing MgCl₂ and dNTPs, was used.

The inter-simple sequence repeat (ISSR) technique involved amplification of DNA segments between 2 identical microsatellite repeat regions (Zietkiewicz et al. 1994). The PCR protocol and program were the same as for IRAP, except that 35 cycles and an annealing temperature of 56 °C were used. The primers contained a microsatellite repeat sequence anchored at the 3′ end.

The sequence-characterized amplified region (SCAR) and single-nucleotide polymorphism (SNP) PCRs contained 0.75 U Taq polymerase (MBI Fermentas) or Biotools DNA polymerase (Biotools B&H Labs, S.A., Reutlingen, Germany), buffer supplied by the enzyme manufacturer, 1.5–2 mmol/L MgCl₂, 100 µmol/L each dNTP, 200 nmol/L each primer, and 20 ng DNA in a total reaction volume of 25 µL. The amplification was carried out in a Masterscycler® gradient (Eppendorf, Hamburg, Germany) or a PTC-220 DNA Engine Dyad™ Peltier Thermal Cycler (MJ Research) using the following program: an initial denaturation of 2 min at 95°C; 35 cycles of 30 s at 95 °C, 30 s at 64 °C (SCAR and SNP–RAPD) or 60 °C (SCAR and SNP–REMAP) and 1 min at 72 °C; a final extension of 6 min at 72 °C. The SNP reaction was carried out using the MegaBACE S NuPe Genotyping Kit (Amersham Pharmacia Biotech).

Two markers showing linkage to the dwarfing gene were sequenced from the parent ‘Kontant’ and, after conversion into SCARs, also from the parent ‘Aslak’. The markers were excised from the gel or purified directly from the PCR with the GFX™ PCR DNA and Gel Band Purification Kit (Amersham Biosciences). The DNA fragments were ligated with the TOPO TA Cloning Kit for Sequencing (Invitrogen, Carlsbad, Calif.) into the pCR®4-TOPO® vector, and transformed into E. coli TOP10 cells. Plasmids were purified with the CONCERT Rapid Plasmid Miniprep System (Life Technologies, Carlsbad, Calif.), and the presence of the insert was ascertained by EcoRI restriction analysis. Four to 5 clones from each transformation were sequenced with the MegaBACETM 500 Sequencer (Amersham Biosciences).

Statistical analyses

The association of markers with height was tested using the non-parametric Kruskal–Wallis test (PROC NPAR1WAY, SAS Institute Inc., Cary, N.C.), based on a χ² test of rank sums (Wilcoxon scores). A significance threshold of P ≤ 0.05 was used. The goodness of fit between observed and expected F₂ segregation ratios was tested by a χ² analysis. MAPMAKER 3.0 software (Lander et al. 1987) was used for determining linkage between the SNP markers and the dwarfing gene. Genetic distances in centimorgans (cM) were calculated by Haldane’s mapping function (Haldane and Smith 1947). BLAST searches (Altschul et al. 1997) against GenBank and EMBL databases were performed using the National Center for Biotechnology Information (Bethesda, Md.) server (http://www.ncbi.nlm.nih.gov).

Results and discussion

The average height of ‘Aslak’ specimens was 86 cm and that of ‘Kontant’ was 54 cm. Plant heights in the F₂ population varied from 45 to 145 cm, with a mean (± standard deviation (SD)) of 71.6 ± 26.3 cm. The distribution of heights indicates a monogenic dominant inheritance of the dwarfing gene. The height of the 9 shortest individuals varied from 45 to 52 cm; the heights of the 9 tallest ranged from 127 to 145 cm. The height of the second short bulk varied from 45 to 62 cm. We were able to classify the F₂ individuals into homozygotes (non-segregating progenies) and heterozygotes (segregating progenies) by measuring the heights in F₂ progenies (Fig. 1). The segregation of the different genotype classes fits the 1:2:1 ratio expected for a monogenic trait.

A total of 500 RAPD primers were tested for polymorphisms in the parents of the cross, ‘Aslak’ × ‘Kontant’. Fifty-six primers (11.2%) produced 73 polymorphic markers, of which 31 were inherited from ‘Aslak’ and 42 from ‘Kontant’. These markers were analyzed in the individual plants of the bulks. Only 1 marker (OPI-111450) was associated with plant height (Fig. 2α), and the total F₂ population was analyzed with it. The 89 F₂ individuals carrying the OPI-111450 marker had a height of 92.5 ± 42.8 cm and the 22 F₂ individuals without it had a height of 108.1 ± 31.4 cm. The association of the marker with height was statistically significant (P < 0.0001, χ² = 24.4). The RAPD phenotype of 10 individuals tested (9%) did not correlate to height.

The RAPD marker OPI-111450 was sequenced from ‘Kontant’ (AY737795, 457 bp including primers) and converted into a SCAR marker. The SCAR primers contained the original 10 bases of the RAPD primer plus the next 10 internal bases (the forward primer: ACATGCGTTGCGGTCGGGATA and the reverse primer ACATGCGTGCCCTCAGTCTGTC). The SCAR marker was, however, monomorphic, probably because the longer SCAR primers enable amplification even if there are mismatches in the annealing site of the RAPD primer. To recover the polymorphism between ‘Aslak’ and ‘Kontant’, the SCAR fragment was sequenced from ‘Aslak’. It differed from that of ‘Kontant’ at 4 positions (results not shown), and an SNP primer (CCA- TGTACACATCTTTCCACT) was designed to reveal one of the differences (‘Kontant’, C; ‘Aslak’, G). The height in the different SNP classes was as follows: 59.1 ± 15.2 cm (23 homozygous for the allele derived from ‘Kontant’), 64.9 ± 16.6 cm (67 heterozygous), and 106.8 ± 31.6 cm (21 homozygous for the allele derived from ‘Aslak’). The differences are statistically significant at P < 0.0001 (χ² = 31.43). The number of individuals misclassified with SNP analysis when compared with Dw6 genotype was 18.9% (21/111). However, the effective error rate (Penner et al. 1995) for the SNP–RAPD is only 4.5% (5/111 individuals). This is a measure of misclassified individuals that are of concern to a plant breeder; in this case, plants erroneously classified as homozygous for short straw according to the SNP–RAPD result.

We searched for a marker closer to Dw6 than the SNP–RAPD marker using the reconstructed bulks (see Materials and methods). Forty-two IRAP primers, 325 REMAP primer combinations, and 24 ISSR primers were used. The parents displayed polymorphism for 14 IRAP primers (33.3%) producing 35 markers, 2 ISSR primers (8.3%, 3 markers), and 40 REMAP primer combinations (12.3%, 62 markers).
Three IRAPs and 1 REMAP amplified the same polymorphism in the bulks as in the parents. Analysis of the individual plants in the bulks revealed that only the REMAP marker, amplified with primers D004 (reverse primer, TCA-CCATGTTCGACGAAACG, corresponding to the antisense strand 137 nt from the 3' end of the LTR of retrotransposon OARE-1) and (CAC)7T, separated the short and tall individuals (Fig. 2b). Furthermore, the REMAP marker...
was shown to be significantly associated with height \( P < 0.0001, \chi^2 = 31.69 \); only 6 individuals (5.4\%) were wrongly classified in the whole population. Plant height in the 91 \( F_2 \) individuals carrying the REMAP band was 61.8 ± 13 cm and in the 20 \( F_2 \) individuals without it 116.2 ± 25.4 cm.

The REMAP marker was sequenced from ‘Kontant’ (AY737796, 697 bp including primers) and converted into a SCAR marker. The SCAR primers were designed forwards from the original primers (the forward primer: GGCTGT-GGGAAACGGATAATA and the reverse primer GCTGTG-CGAATCCTCAAGTG). These primers amplified the marker also from ‘Aslak’, thus enabling its sequencing. The sequence of the SCAR product of ‘Aslak’ differs from that of ‘Kontant’ at 3 nucleotide positions (results not shown). We designed a primer (GATGGAGGTGAATGCCACAA) to detect one of these differences (‘Kontant’, T; ‘Aslak’, C) as an SNP polymorphism. Plant height of the different SNP classes was as follows: 55.0 ± 4.1 cm (23 homozygous for the allele derived from ‘Kontant’), 64.1 ± 14.1 cm (68 heterozygous), and 116.2 ± 25.4 cm (20 homozygous for the allele derived from ‘Aslak’). The differences are statistically significant at \( P < 0.0001 (\chi^2 = 50.63) \). The amount of misclassification when compared with the \( Dw6 \) genotype was 8.1\% (9/111), but the effective error rate was only 0.9\%.

SNP markers together with \( Dw6 \) formed 1 linkage group (LOD score > 17) with the order \( Dw6 \) – SNP–REMAP – SNP–RAPD, and distances of 5.2 and 7.4 cm. Milach et al. (1997) had already mapped \( Dw6 \) to 3.3 ± 1.3 cm from the RFLP locus, Xumm145B, which thus lies closer to \( Dw6 \) than do our SNP markers. Nevertheless, because the SNP markers are PCR based and amenable to high-throughput analysis using the many new SNP-based methods, they are much easier to use than the RFLP marker.

The DNA sequences of the RAPD and REMAP markers from ‘Kontant’ were compared against the non-redundant EST DNA databases using BLASTn, and against the Swissprot and non-redundant databases using BLASTx and tBLASTx, respectively. The terminal segment of the REMAP sequence matched the minus strand of the \( A. sativa \) OARE-1 retrotransposon, from which it was derived, extending to the beginning of the LTR of this element. The RAPD sequence matches \( A. sativa \) genomic accessions (AY038004, AY038007, and AY038013), which were annotated as containing a transposable element (unclassified) and receptor kinase domains, as well as several unannotated \( A. sativa \) ESTs. In our analyses by BLASTn, BLASTx, and tBLASTx, although the RAPD sequence does not match the retrotransposon-containing region of these clones and does have similarity to \( A. sativa \) transcripts (e.g., CN814902), we fail to find similarity to a kinase or other enzyme within the sequence itself. Likewise, these EST accessions, when searched using BLASTx, form weak matches (E values of 2.8–8.1) to glucose-6-phosphatase, \( \beta \)-galactosidase, and a lipid-transfer protein. Hence, although the RAPD sequence appears to be transcribed and is not repetitive, its genic identity remains unclear. Physical association with a receptor kinase would be intriguing owing to the connection shown in other plants between dwarfing and receptor-kinase function (Li and Chory 1997; Choe et al. 2002; Chono et al. 2003).

The availability of SNP markers close to the dwarfing gene \( Dw6 \) provides an option for rapid identification of homozygotes and heterozygotes for short straw in marker-assisted programs of germplasm improvement. In the future, it will be possible to locate \( Dw6 \) to the oat linkage map because the SNP–RAPD marker is polymorphic in the parents of the doubled-haploid population that is used in an ongoing mapping project at MTT Agrifood Research Finland.

**Acknowledgements**

We wish to thank Boreal Plant Breeding Ltd., especially oat breeder Leena Pietilä, for co-operation and for \( F_2 \) seeds, Ms. Anna Virta for DNA sequencing, and Ms. Marja-Riitta Arajärvi, Ms. Sirpa Moisander, Ms. Liisa Syrjänen, Mr. Yrjö Karppinen, and Ms. Ritva Koskenoja for their excellent technical assistance. We also thank Leo van der Pligt from Wiersum b.v. and Hein de Jong from Cebeco Seeds b.v. for their kind help in determining the origin of the dwarfing gene of ‘Kontant’.

**References**


