A major gene for grain cadmium accumulation in oat (Avena sativa L.)

Pirjo Tanhuanpää, Ruslan Kalendar, Alan H. Schulman, and Elina Kiviharju

Abstract: Cadmium (Cd) is a nonessential heavy metal that is highly toxic to living cells at very low concentrations. Most of the Cd in plants derives from soils. Owing to the large amounts consumed, cereals are the major source of dietary Cd, and Cd content in oat can exceed accepted limits. Plants have a set of mechanisms that control the uptake, accumulation, trafficking, and detoxification of Cd and other metals. Genetic factors affect the variation in Cd level between plant species and cultivars, and the development of cultivars that poorly accumulate Cd is a worthwhile goal. Because of the expense of Cd screening, the use of molecular markers linked to low Cd accumulation could be an alternative to phenotyping for selection. In this study, such markers were sought using bulked-segregant analysis in an F2 population from the cross between oat cultivars ‘Aslak’ and ‘Salo’, the second of which is known to be a high Cd accumulator. Four markers associated with grain Cd concentration were found: 2 RAPDs (random amplified polymorphic DNAs), 1 REMAP (retrotransposon-microsatellite amplified polymorphism), and 1 SRAP (sequence-related amplified polymorphism). The first 3 were converted into more reproducible SCAR (sequence-characterized amplified region) markers. The 4 markers were assigned to 1 linkage group that exhibited a QTL (quantitative trait locus) representing a major gene for grain Cd concentration.

Key words: Avena sativa, cadmium, marker-assisted selection, RAPD, REMAP, SCAR.

Introduction

Cadmium (Cd) is a nonessential heavy metal that is highly toxic to living cells at very low concentrations. In humans it can damage kidneys, causing loss of calcium and thereby osteoporosis (Kazantzis 2004), and it is also a suspected carcinogen (Lemen et al. 1976). Cereals are consumed in large amounts and are therefore the major source of dietary Cd. Most of the Cd in plants derives from soils. The concentration of Cd in soil may vary (Sippola and Mäkelä-Kurto 1986) and depends on both atmospheric deposition and the use of animal manures, phosphate fertilizers,
and sewage sludge (Adriano 2001). In addition, Cd concentrations are affected by climatic conditions, with warm, dry growing seasons tending to increase it (Eurola et al. 2003). In the European Union, the maximum permitted level of Cd for cereals is 0.1 mg/kg fresh mass (European Commission 2001). Oat, a crop often marketed for the healthful effects of its grain beta-glucans (Food and Drug Administration 2003; Katz et al. 2001), may exceed this limit in some circumstances (Eurola et al. 2003; Karlsson 1994).

Plants, like other organisms, have evolved mechanisms that control the uptake, accumulation, trafficking, and detoxification of metals. Various cation transporters, which also transport Cd, have been identified, including ZIP and Nramp proteins (reviewed in Clemens 2001). Following transport, metal ions are bound by chelators and chaperones. Chelators such as phytochelatins (Cobbett 2000) contribute to metal detoxification, whereas chaperones deliver metal ions to organelles. Excess metal ions have to be removed from the cytosol, and the main storage compartment in plant cells for toxic compounds including metals is the vacuole. Although metal translocation from roots to shoots has been studied, its control is poorly understood, it appears to vary between monocotyledonous and dicotyledonous plants, and it is influenced by the many factors mentioned above (Page et al. 2006; Clemens 2001).

Just as Cd uptake, translocation, and sequestration are mediated by many proteins, the ultimate differences in Cd accumulation between species or even cultivars are unlikely to be monogenically determined. Nevertheless, genetic tools for the development of low-accumulating cultivars are needed, because an important source of Cd contamination is phosphate fertilizer (He et al. 2005), particularly when applied to acid soils (Tiessen 1995). Furthermore, supplies of high-quality, low-Cd phosphate rock are limited and have economic trade-offs (Cupit et al. 2002). Earlier studies have pointed to genetic factors that determine differences in Cd accumulation (Eurola et al. 2003 for oat; Clarke et al. 2002 for durum wheat).

Because determination of Cd concentration is laborious, time-consuming, and expensive, marker-assisted selection (MAS), the use of molecular markers linked to a desired gene, could be an alternative to phenotyping. Such markers can be established with linkage mapping or bulked-segregant analysis (BSA; Michelmore et al. 1991). In durum wheat, a single gene controlling grain Cd concentration, with low Cd being dominant, has been found (Clarke et al. 1997), and BSA was successfully used to find 2 RAPD (random amplified polymorphic DNA) markers linked to this gene (Penner et al. 1995). The purpose of the present study was to identify anonymous PCR-based markers associated with a gene affecting grain Cd accumulation in oat (Avena sativa L.) using BSA. We also aimed to convert the best markers into easily scorable and reproducible PCR markers using specific primers.

Materials and methods

Plant material

All plants used in the study were raised in a greenhouse. A population of 150 F₂ plants was derived from a cross between 2 spring oat individuals, 1 from cultivar ‘Askal’ (Boreal Plant Breeding Ltd., Finland) and the other from cultivar ‘Salo’ (Svalöf-Weibull AB, Sweden). The parent ‘Salo’ was known to have the tendency to accumulate Cd in the field (Eurola et al. 2003). The parents were also tested for Cd accumulation in greenhouse pots by a method described below to confirm the difference in this trait before the cross was made.

DNA was extracted using the DNeasy® Plant Mini Kit (QIAGEN GmbH, Hilden, Germany). Fresh leaf material (approx. 100 mg) was crushed with a FastPrep™ FP120 Cell Disrupter (BIO 101, Thermo Savant, Waltham, Massachusetts). DNA concentrations were measured using the GeneQuant II RNA/DNA Calculator (Pharmacia Biotech Ltd., Cambridge, England).

Screening test for Cd accumulation

A greenhouse pot test was developed for the screening of Cd accumulation in oat. Seeds of the 150 F₂ individuals were sown in 12 cm diameter pots containing a peat-soil mix (1 individual per pot). The pots were put in boxes and placed randomly on a table in a greenhouse. On the basis of the preliminary experiment with the parental genotypes (data not shown), we decided to apply 0.5 mg Cd per kg soil as an aqueous solution of Cd(NO₃)₂·4H₂O (Fluka, Buchs, Switzerland). Four replicates of the parents were used and the parents were also tested without the addition of Cd. The parents were placed randomly among the F₂ plants. No replicates were available for the F₂ plants. The plants were grown under controlled conditions, under a 16 h photoperiod provided by fluorescent lamps, and watered manually. At the 3-leaf stage, a solution containing Superrex Nursery Stock fertilizer (N 19%, P 4%, K 20%, Kekkilä, Finland) was used for watering.

Cd analysis

The Cd content was determined in the fully mature seeds of the main panicle of each individual. In addition, straw samples and the roots of the parents were analysed. The roots (washed with deionized water) of the 4 replicates were combined before analysis. The Cd was measured by the accredited technique using inductively coupled plasma mass spectrometry (ICP-MS) (PerkinElmer ELAN 6000) as described by Kumpulainen and Paakkki (1987).

Bulking of F₂ individuals

Bulked-segregant analysis (Michelmore et al. 1991) was used to search for molecular markers associated with Cd accumulation. For this purpose, DNAs from the 9 F₂ individuals that accumulated the least Cd (from 650 to 820 µg/kg seed) and the 9 that accumulated the most Cd (from 2540 to 3490 µg/kg seed) were pooled.

Markers

All PCR amplifications were performed in a PTC-220 DNA Engine Dyad™ Peltier Thermal Cycler (MJ Research, Waltham, Massachusetts) using Biotools DNA polymerase (Biotools B&M Labs, S.A., Madrid, Spain) and the buffer supplied by the enzyme manufacturer, containing 2 mmol/L MgCl₂. RAPD analysis was carried out basically as described...
in Tanhuanpää et al. (2006) with the above-mentioned exceptions in the PCR protocol. RAPD primers consisted of random 10- or 11-mers from Operon Technologies (Alameda, California).

In the REMAP (retrotransposon-microsatellite amplified polymorphism) method, amplification of the region between the long terminal repeats (LTRs) of retrotransposons and nearby simple sequence repeats generates products that serve as markers (Kalander et al. 1999, Kalander and Schuman 2006). Primers were designed by cloning retrotransposon regions from the oat genome, identifying the long terminal repeats, and choosing conserved motifs at or near their termini. The microsatellite-based primers contained repeat units (composed of 2 or 3 bases) anchored at their 3’ ends by a nucleotide. REMAP markers were amplified as described by Tanhuanpää et al. (2006) using Biotools polymerase.

Primers for amplifying SRAP (sequence-related amplified polymorphism) markers target intron–exon boundaries (Li and Quiros 2001). The PCR programme for SRAPs was as follows: the first 5 cycles were run at 94 °C for 1 min, 35 °C for 1 min, and 72 °C for 1 min. The subsequent 40 cycles were carried out with an annealing temperature of 47 °C. The PCR reaction (25 μL) consisted of 0.5 U Biotools DNA polymerase, 100 μmol/L each dNTP, 400 nmol/L each primer, and 20 ng DNA. The forward primer was labelled with a fluorescent dye (FAM [5-carboxyfluorescein] or TET [6-carboxytetrachlorofluorescein]), thus enabling the scoring of markers with a MegaBACE™ 500 Sequencer (GE Healthcare, Buckinghamshire, UK).

The PCR reactions for the SCAR (sequence-characterized amplified region) markers contained, in 25 μL, 0.5 U Biotools DNA polymerase, 50 μmol/L each dNTP, 200 nmol/L each primer, and 20 ng DNA. The PCR protocol consisted of an initial denaturation of 2 min at 95 °C, 35 (SCAR AF20 and SCAR A002+(AC)G) or 38 cycles (SCAR AF15) of 30 s at 95 °C, 30 s at the annealing temperature (SCAR AF20: 58 °C; SCAR A002+(AC)G: 62 °C; and SCAR AF15: 66 °C), and 1 min at 72 °C, and a final extension of 5 min at 72 °C.

The 2 RAPD markers and 1 REMAP marker that showed linkage to Cd accumulation were sequenced. The PCR products were excised from the gel and purified with the GFX™ PCR DNA and Gel Band Purification Kit (GE Healthcare, Buckinghamshire, UK). The DNA fragments were ligated with the TOPO TA Cloning® Kit for Sequencing (GE Healthcare, Buckinghamshire, UK) into the pCR®4-TOPO® vector and chemically transformed into TOP10 E. coli cells. A FastPlasmid Mini kit (Eppendorf, Hamburg, Germany) was used for plasmid purification, and the presence of the insert was ascertained by EcoRI restriction analysis. One (for RAPD marker OPAF15190) to 5 clones from each transformation were sequenced from both directions with the MegaBACE™ 500 Sequencer. The sequence of OPAF15190 was later verified by sequencing the corresponding SCAR marker directly from 3 different PCR reactions.

Statistical analyses

The Cd accumulation in the plants from the greenhouse pot test and the association of markers with Cd accumulation were analyzed using the general linear model (GLM

![Fig. 1. Cd concentration in roots, straw, and grains of oat cultivars ‘Aslak’ (□) and ‘Salo’ (■) after Cd-feeding pot tests in the greenhouse. Means are shown with standard error bars, except in roots, where no replicates existed. The values from control samples without Cd added were as follows: for ‘Aslak’, 120 (roots), 148 ± 81 (straw), and 59 ± 24 μg/kg (grains), and for ‘Salo’, 220 (roots), 109 ± 24 (straw), and 75 ± 15 μg/kg (grains).](image-url)
roots (83% of the Cd of ‘Salo’), the discrimination against Cd accumulation was greater in the straw (63%) and still greater in the grains (43%).

The concentration of Cd in the grains of the ‘Aslak’ × ‘Salo’ F₂ progeny varied from 650 to 3490 μg/kg, the mean being 1575 μg/kg with a standard deviation of 550 and standard error of 45 (Fig. 2). The distribution fits the hypothesis of single-gene inheritance with the allele for low accumulation being dominant. However, transgressive segregation (especially phenotypes more extreme than those of the high-accumulating parent) suggests that some minor genes also influence Cd accumulation in the ‘Aslak’ × ‘Salo’ cross, either directly or indirectly.

### Table 1. DNA markers associated with Cd accumulation (general linear model analysis) in the F₂ population of the A. sativa cross ‘Aslak’ × ‘Salo’.

<table>
<thead>
<tr>
<th>Marker</th>
<th>Type</th>
<th>Size (bp)</th>
<th>Amplified from</th>
<th>Accession No.</th>
<th>Original primer(s)</th>
<th>SCAR primers</th>
</tr>
</thead>
<tbody>
<tr>
<td>OPAF20:1350</td>
<td>RAPD</td>
<td>1354</td>
<td>Aslak</td>
<td>DQ397884</td>
<td>CTCCGGACACAG</td>
<td>AGGAAGACCTGACCAGACA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CGAAGCGAAGCAGAAGCA</td>
</tr>
<tr>
<td>OPAF15:960</td>
<td>RAPD</td>
<td>957</td>
<td>Salo</td>
<td>DQ397885</td>
<td>CACGAACCTC</td>
<td>CACGAACCTCCTGGAAAAAG</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CACGAACCTCGCCAAGAGTAAG</td>
</tr>
<tr>
<td>A002+(AC)9:1380</td>
<td>REMAP</td>
<td>1378</td>
<td>Aslak</td>
<td>DQ397886</td>
<td>GTGTACGACTACATCATCCACGAG</td>
<td>CCGCTGTCGTTAAACACTT</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A002+ACACACACACACACACACG</td>
</tr>
<tr>
<td>me1em6490:486</td>
<td>SRAP</td>
<td>486</td>
<td>Aslak</td>
<td>—</td>
<td>TGAGTCCAAACCGGATA</td>
<td>GACCTGACTAACGAAATTGCA</td>
</tr>
</tbody>
</table>

*Fig. 2. Frequency distribution for grain Cd concentration in the F₂ population of the A. sativa cross ‘Aslak’ × ‘Salo’ based on markers SCAR AF20 (a) and SCAR AF15 (b). Marker genotypes are presented as fills in the columns: open, homozygous for ‘Aslak’ allele; hatched, homozygous for ‘Salo’ allele; solid, heterozygous (SCAR AF20) or homo- and heterozygous (SCAR AF15) for ‘Salo’ allele.*

© 2007 NRC Canada
Table 2. Associations between DNA markers and grain Cd accumulation (mean ± SD, with number of individuals in parentheses) by variance analysis in the F2 population of the A. sativa cross ‘Aslak’ × ‘Salo’.

<table>
<thead>
<tr>
<th>Marker</th>
<th>Cd accumulation (µg/kg)</th>
<th>P value</th>
<th>F</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>SCAR AF20</td>
<td>1151±332 (37)</td>
<td>1463±384 (69)</td>
<td>2107±508 (44)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>SCAR AF15</td>
<td>1702±539 (118)</td>
<td>—</td>
<td>1103±257 (32)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>SCAR A002+(AC)₉G</td>
<td>1371±411 (109)</td>
<td>—</td>
<td>2117±506 (41)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>SRAP me1em6930</td>
<td>1385±421 (104)</td>
<td>—</td>
<td>2051±549 (44)</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

Note: R², the proportion of total phenotypic variance explained by the markers, A, individuals with the visible marker allele; C, individuals without the marker. For SCAR AF20, which is codominant, A and C represent individuals with different-sized alleles and B represents heterozygotes.

‘Aslak’ and ‘Salo’ were analyzed with 218 RAPD, 31 SRAP, and 337 REMAP primers and primer combinations, of which 78 (36%), 26 (84%), and 73 (22%) produced polymorphisms, respectively. These were further tested in the bulks. Two RAPDs, 5 SRAPs, and 4 REMAPs showed the same polymorphisms in the bulks and were screened against the individuals in the bulks. Four markers (Table 1) seemed to be associated with Cd accumulation and were further analyzed in the total F2 population. The markers segregated according to Mendelian ratios.

The RAPD markers OPbAF201350 and OPbAF15980, and the REMAP marker A002+(AC)₉G1380 were sequenced and converted into SCAR markers. SCAR A002+(AC)₉G and SCAR AF15 amplified the same dominant polymorphisms as the original REMAP and RAPD markers but SCAR AF20 was codominant, i.e., it exhibited length polymorphism. All 4 markers discovered were highly significantly associated with grain Cd concentration (Table 2). SCAR AF20 seemed to best explain the phenotypic variance in this trait. This is partly due to its codominance even though the R² value for the corresponding RAPD marker was also the highest among the markers (0.39). The most suitable markers for selection were SCAR AF20 and SCAR AF15, owing to their ability to identify homozygotes for the desired low-accumulation allele. The effective error rate (Penner et al. 1995) for these markers is very low: 0.7% (1/150) for SCAR AF20 and 0% for SCAR AF15. The effective error rate measures detrimental (from the breeder’s point of view) misclassifications, i.e., individuals erroneously classified as being homozygous for the low-accumulation allele (Fig. 2).

All 4 markers were assigned to 1 linkage group (LOD (logarithmic odds) score > 17). The correct order of the loci could not be resolved with MAPMAKER because the probabilities of the 2 best orders (SCAR AF15, SCAR AF20, me1em6, A002+(AC)₉G and SCAR AF20, me1em6, A002+(AC)₉G, SCAR AF15) were almost the same. Only the position of SCAR AF15 varies between these 2 orders, and the difficulty in localizing it is due to the marker being in repulsion with the 2 other dominant loci. In an F2 population, linkage statistics and recombination frequencies are poorly estimated between closely linked dominant markers in repulsion (Mather 1936). Therefore, we chose to use the order suggested by JoinMap for the QTL analysis. The length of the group was 10.2 cM (determined with MAPMAKER) and the order of loci was SCAR AF15, A002+(AC)₉G, me1em6, SCAR AF20. The QTL affecting Cd accumulation was situated in this group but its precise location remained unresolved because the maximum LOD score (19.9) was situated at the end of the group and, in addition, the LOD score curve was quite flat. However, the QTL was situated nearest to SCAR AF15, and the variance explained at this point was 54.5%.

The 3 sequenced markers associated with Cd accumulation were compared against the non-redundant DNA databases using BLASTn. There were no meaningful matches for SCAR AF20. The strongest match for SCAR A002+(AC)₉G was to the Hordeum vulgare MLA1-2 gene (E value of 5e⁻⁶⁶, identity 116/143 = 81%). However, the match was not to the MLA1-2 gene itself, and BLASTx and tBLASTx both at the NCBI site and on the TREP server confirmed that the SCAR corresponds to a LINE retrotransposon most similar to the LINEs Karin (GenBank accession No. AF325197) of wheat, E = 5e⁻⁵⁷, and Persephone of barley (GenBank acc. No. AY643843), E = 2e⁻⁵⁵, in the reverse transcriptase domain. Hence, this polymorphic SCAR is due to a retrotransposon insertion.

SCAR AF15 matched the Triticum aestivum WAK3 gene, a wall-associated kinase, with an E value of 1e⁻⁴⁸ and an identity of 225/267 = 84% at the DNA level. The BLASTx tool confirmed the match to WAK3, showing 83% identity and 73% similarity over 86 residues, as well as high similarity to protein-kinase domains in rice and in Brachypodium sylvaticum. The 225 bp sequence of SCAR AF15 that is homologous to WAK3 corresponds to the exon sequence of the gene. The WAKs are a subset of a large class belonging to the receptor-like kinase (RLK) superfamily of genes. They are associated with the cell wall and are putative signaling molecules between the cell wall and the cytoplasm (Zhang et al. 2005; He et al. 1996). As such, they are involved in a variety of functions in plants, including biotic and abiotic stress response as well as cell elongation and development (Hou et al. 2005). In Arabidopsis, overexpression of a WAK confers aluminum tolerance (Sivaguru et al. 2003), whereas a WAK-like protein (WAKL4) is induced by copper, nickel, and zinc (Hou et al. 2005).

In conclusion, 4 DNA markers associated with a major gene affecting Cd accumulation in oat were found in this study. Even though the exact location of the gene could not be resolved, all the markers are located sufficiently near to it because in breeding, even 15–20 cM may be an acceptable distance (Lee 1995). However, the most suitable marker in breeding will probably be SCAR AF15. First, the visible allele of SCAR AF15 is derived from ‘Salo’ and thus it is possible to identify homozygotes for the low-accumulation...
allele. Although the $R^2$ value for this marker is quite low (0.23), this is most probably due to its dominance. Second, the QTL for Cd accumulation is located nearest to this marker. Finally, the sequence of the marker matches a wall-associated kinase (WAK3), which hints that WAK3 is a candidate gene for Cd accumulation. Whether or not we have identified the correct candidate gene, the DNA markers found can be used in future breeding programs to select for low Cd accumulators in oat. This will become increasingly important as high-quality phosphate rock is consumed. The chance that even one of them will be polymorphic in the de-\nessed population that is being used to construct an oat linkage map at MTT Agrifood Research Finland. As a consequence, the gene for grain Cd accumulation will be localized on this map. Ultimate verification of the candidate gene sequence, the gene for grain Cd accumulation will be localized in the haploid population that is being used to construct an oat linkage map at MTT Agrifood Research Finland. As a consequence, the gene for grain Cd accumulation will be localized on this map. Ultimate verification of the candidate gene will help in transfer of the low Cd accumulation trait to other cereals and crops.

Acknowledgements

The authors wish to thank Boreal Plant Breeding Ltd. for the plant material, Ms. Anneli Virta for DNA sequencing, Ms. Merja Eurola for Cd analyses, biometrician Lauri Jauhiainen for help in experimental design, and Ms. Tarja Hovivuori, Ms. Leena Ramstedt, Mr. Yrjö Karpipinen, and Ms. Ritva Koskenoja for their excellent technical assistance. Raisio Research Foundation and Ministry of Agriculture and Forestry in Finland are acknowledged for their financial support.

References

Michel, R.W., Paran, I., and Kesseli, R.V. 1991. Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific geno-


