Single-Base Substitution in P\textsubscript{1B}-ATPase Gene Is Associated with a Major QTL for Seed Cadmium Concentration in Soybean

EDUARDO R. BENITEZ, MAKITA HAJKA, AND RYOJI TAKAHASHI

Abstract

Cadmium (Cd) is a pollutant metal present in soils and toxic to biologic organisms. Previous studies using recombinant inbred lines derived from a cross between soybean (\textit{Glycine max} [L.] Merr.) cultivars Harosoy and Fukuyutaka revealed a major quantitative trait loci for seed Cd concentration (cd1) in chromosome 9. The genome sequence of Williams 82 suggested that a P\textsubscript{1B}-ATPase gene involved in the transport of metals was located in the vicinity of cd1. cDNA sequencing suggested existence of two types of transcripts: one (\textit{GmHMA1a}) consisting of 9 exons and 8 introns and the other (\textit{GmHMA1b}) consisting of 8 exons and 7 introns. The putative polypeptide, GmHMA1a, consisted of 885 amino acids, whereas premature termination of translation of \textit{GmHMA1b} generated a putative polypeptide with 559 amino acids. \textit{GmHMA1a} of Fukuyutaka differed from that of Harosoy by a single-base substitution that led to an amino acid substitution from E to G at amino acid position 608. A derived cleaved amplified polymorphic sequence (dCAPS) marker was developed to detect the base substitution, and this dCAPS marker was successfully associated with seed Cd concentration. Transgenic experiments may be necessary to verify that \textit{GmHMA1} actually corresponds to cd1.

Key words: \textit{Glycine max}, P-type ATPase, seed cadmium concentration, soybean

Cadmium (Cd) is a pollutant metal naturally present in soils and highly toxic to biologic organisms even at low concentrations. This trace metal can accumulate in the human body as a result of ingesting contaminated food over time, and excessive intake can lead to chronic toxicity (Jackson and Alloway 1992). The Codex Alimentarius Commission of the Food and Agriculture Organization (FAO) and the World Health Organization (WHO), which is responsible for the safety of food and human health, proposed an upper limit of 0.2 mg kg\textsuperscript{-1} for Cd concentration in soybean (\textit{Glycine max} [L.] Merr.) seeds (Codex Alimentarius Commission 2001). From 2008, the Commission of the European Communities also changed the maximum level for Cd to 0.2 mg kg\textsuperscript{-1} in soybean grain (Commission of the European Communities 2008).

There is genotypic variation in the accumulation of Cd in soybean seeds (Arao et al. 2003). The Cd concentration in the seed of certain genotypes was found to be consistently low under various field and soil conditions, suggesting a possibility to develop cultivars with genetically low seed Cd concentration. DNA marker–assisted selection may be an effective method to incorporate this trait into cultivars with otherwise excellent agronomic characters.

Benitez et al. (2010) performed a quantitative trait loci (QTL) analysis of seed Cd content using a total of 93 recombinant inbred lines (RILs) of \textit{F}\textsubscript{5}\textsubscript{6}, \textit{F}\textsubscript{6}\textsubscript{7} and \textit{F}\textsubscript{7}\textsubscript{8} generations derived from a cross between two cultivars: Harosoy, with high seed Cd content, and Fukuyutaka, with low Cd content. One major QTL, cd1, was identified in chromosome 9 (molecular linkage group K) across years and generations accounting for 82%, 57%, and 75% of the genetic variation. The effect of the QTL was ascertained using near-isogenic lines (NILs) developed from the cultivars. The peak of the QTL was located in the vicinity of two SSR markers, Gm09:4770663 and Gm09:4790483.
Further, separate studies revealed a major QTL for seed Cd content, *cd1* at a similar genomic location (Jegadesan et al. 2010), suggesting that *cd1* and *cdA1* may be identical. This study was conducted to clone a candidate gene for the *cd1* using information from the soybean genome.

**Materials and Methods**

**Plant Materials**

The Canadian cultivar Harosoy with high seed Cd content (maturity group II) and the leading cultivar of southern Japan, Fukuyutaka, with low Cd content (maturity group VI) were used for this research. A total of 93 RILs of *F*₅₀₆ and *F*₅₈ generations developed from a cross between these cultivars (Benitez et al. 2010) were used for linkage mapping and QTL analysis. Cd concentration in the soil of the experimental field plots in three successive years was reported in Benitez et al. (2010). In addition, a pair of NILs for *cd1*, #89-1 with low Cd concentration and #89-46 with high Cd concentration (Benitez et al. 2010) was used to confirm the association between the QTL and the genetic polymorphism. Seeds of Harosoy and Fukuyutaka were planted on 10 June 2010 in the field at the National Institute of Crop Science, Tsukuba, Japan (36°05′N, 140°05′E). N, P, and K were applied at 3.0, 4.4, and 8.3 g m⁻², respectively. Seeds of the NILs were planted in vermiculite on 31 July and transplanted to the same field on August 9.

**RNA Extraction and cDNA Cloning**

For cDNA cloning, roots of Harosoy and Fukuyutaka were harvested from the field on August 7. For gene expression assay, roots and leaves of the NILs at R4 stage (Fehr et al. 1971) were harvested on September 17. Total RNA was extracted from 200 mg of roots or leaves using the TRIZOL Reagent (Invitrogen) according to the manufacturer’s instructions. cDNA was synthesized by reverse transcription of 5 μg of total RNA using the Superscript III First-Strand Synthesis System (Invitrogen) and an oligo(dT) primer according to the manufacturer’s instructions. PCR primers for cloning of *P*₁₆-ATPase gene were designed based on the genome sequence of US cultivar Williams 82 deposited in the soybean genome database (Phytozome, http://www.phytozome.net/soybean.php). DNA sequence of the entire cDNA (about 4 kb) was determined by cloning two fragments overlapping each other (upstream fragment of about 1.2 kb and downstream fragment of about 2.8 kb) with nested PCR. The first PCR was performed using primers at the extreme 5'-end (5'-GGAAGTCTTTATTATGCTTTGT-3') and 3'-end (5'-ATTAGACCTTGTTTTAGAC-3'). The PCR mixture contained 0.5 μg of cDNA, 10 pmol of each primer, 5 pmol of dNTPs, and 1 unit of ExTaq in 1× ExTaq Buffer supplied by the manufacturer (Takara Bio) in a total volume of 25 μl. A 30-s denaturation at 94 °C was followed by 30 cycles of 30-s denaturation at 94 °C, 1-min annealing at 61 °C and 3-min extension at 72 °C. A final 7-min extension at 72 °C completed the program. The second PCR was conducted by using 1 μl of 100 times dilution of the first PCR product as a template. The thermal cycling profile was identical with the first PCR except that denaturation was performed for 5 min and annealing was performed at 55 °C. The PCR primers for the upper fragments were 5'-GGAAGTCTTTATTATGCTTTGT-3' and 5'-GCACTTGCTTTAATCCAAGA-3'. The primers for the lower fragment were 5'-GCAGAAATAAGTGGAATTACT-3' and 5'-AACCAAACACGAAATGGAAGA-3'. The above PCRs were performed in an Applied Biosystems 9700 thermal cycler. The PCR fragments with the expected sizes were cloned into pCR 2.1 vector (Invitrogen) and sequenced.

**Sequencing Analysis**

Nucleotide sequences of both strands were determined with the BigDye terminator cycle method using an ABI3100 Genetic Analyzer (Applied Biosystems). Primers used for sequencing of internal regions are shown in Table 1. Nucleotide sequences and the putative amino acid translations were analyzed with the BLAST program (Altschul et al. 1997). Gene structure was estimated based on the comparison between cDNA sequences and the corresponding genome sequences of Williams 82 deposited in the soybean genome database. Multiple sequence alignment was performed with ClustalW (http://clustalw.ddbj.nig.ac.jp/top-j.html) using default settings.

**Analysis of Alternative Splicing**

Genomic DNA of Harosoy and Fukuyutaka was isolated from trifoliolate leaves by CTAB (Murray and Thompson 1980). To investigate whether alternative splicing occurred in the 8th intron, the DNA fragment including the intron was amplified from genomic DNA (of Harosoy and Fukuyutaka) and cDNA (roots of Harosoy, Fukuyutaka, and the NILs and leaves of the NILs) using a pair of PCR primers (5'-CATGGACGTGAGCAATGACAGG-3' and 5'-ATGACATTCTCAATTAGCTTTC-3') flanking the intron. The PCR mixture contained 30 ng of genomic DNA, 25 pmol of each primer, 5 pmol of nucleotides, and 1 unit of ExTaq in 1× ExTaq Buffer in a total volume of 25 μl. The initial 30-s denaturation at 94 °C was followed by 30 cycles of 30-s denaturation at 94 °C, 1-min annealing at 59 °C, and 1-min extension at 72 °C. A final 7-min extension at 72 °C completed the program. The amplified products were separated on an 8% nondenaturing polyacrylamide gel in 1× Tris/Borate/EDTA buffer (90 mM Tris–borate, 2 mM EDTA, pH 8.0). After electrophoresis, the gel was stained with ethidium bromide, and the DNA fragments were visualized under ultraviolet (UV) light.

**Derived Cleaved Amplified Polymorphic Sequence Analysis**

Genomic DNA of the NILs and the RILs were isolated from trifoliolate leaves by CTAB. Derived cleaved amplified polymorphic sequence (dCAPS) PCR primers
presence/absence of the gle-base substitution. The base substitution would result in QTL analysis was performed by composite interval mapping (Zeng 1993) using the QTL Cartographer version 2.5 (Wang et al. 2010) according to Benitez et al. (2010). A likelihood mapping and 5' end of the primer was chosen not to exceed the mid-log phase of product yield (22 cycles for *actin* and 30 cycles for *GmHMA1*). The PCR mixture contained 125 ng of cDNA, 25 pmol of each primer, 5 pmol of nucleotides, and 1 unit of ExTaq in 1 × Ex Taq Buffer in a total volume of 25 µL. After an initial 30-s denaturation at 94 °C, there were 30 cycles of 30-s denaturation at 94 °C, 1-min annealing at 59 °C, and 1-min extension at 72 °C. A final 7-min extension at 72 °C completed the program. The primers for the *GmHMA1* were 5'-CAAGAAAGAGTGGGC-TAATAG-3' and 5'-GAAAGTCAGAAGATTTCAATTT-3'. The primers for the *actin* were 5'-GAGGCTCTGGAATATT-CAACC-3' and 5'-AGAAATCTGTGAGGTACAGA-3'. PCR products were loaded on a 2% agarose gel, stained by ethidium bromide and visualized under UV light.

**Accession numbers**

Sequence data from this article have been deposited with the DDBJ Data Libraries under accession numbers AB608748 (*GmHMA1*a of Harosoy) and AB608749 (*GmHMA1*b of Harosoy).

**Results**

**cDNA Cloning**

A survey of the compiled genome sequence of Williams 82 from Sat.119 (Gm09:3585480) to Satt178 (Gm09:5438776) revealed that P1W-ATPase gene (Glyma09g06170.1, Gm09:4918664 to Gm09:4926453), which had been implicated in the transport of a range of essential and also potentially toxic metals across cell membranes (e.g., Cu2+, Zn2+, Cd2+, Pb2+), was located in the vicinity of cd1 (Williams et al. 2000). Two overlapping fragments covering the entire cDNA were cloned by RT-PCR from Harosoy and Fukuyutaka. The gene was designated as GmHMA1 according to the designation of AtHMA8 genes AtHMA1 to AtHMA8. Two types of transcript candidates with different lengths (3719 and 3929 bp) were identified for this gene (Figure 2). The shorter and longer transcripts were designated as GmHMA1a and GmHMA1b, respectively. Comparison with the genome sequence of Williams 82 revealed that GmHMA1a consisted of 9 exons and 8 introns (Figure 2). In GmHMA1b, however, 210-bp nucleotides corresponding to the eighth intron were retained as part of an exon (i.e., were not spliced out), resulting in a structure with 8 exons and 7 introns. These results suggest the existence of alternative splicing.

**SSR Analysis, Linkage Mapping and QTL Analysis**

A novel SSR marker (Gm09:5495287) was constructed using the Simple Sequence Repeat Candidate Marker Search Tool (http://cpgr.plantbiology.msu.edu/cpgr_ssr_marker_pred.shtml). The primers were 5'-AGATGTTCAGGGCCAGTTAGT3' and 5'-TTCATCATATGAGGAATGGG-3'. SSR analysis was performed as described in Benitez et al. (2010). A linkage map was constructed by combining genotypes of Gm09:5495287, SSR markers previously examined (Benitez et al. 2010), and the dCAPS marker using the AntMap software (Iwata and Ninomiya 2006) according to Benitez et al. (2010). QTL analysis was performed by composite interval mapping (Zeng 1993) using the QTL Cartographer version 2.5 (Wang et al. 2007). The threshold LOD score was determined by a permutation test with 1000 replications corresponding to genome-wide 5% level of significance.

**Analysis of Association Between Genotype and Seed Cd Concentration**

To investigate the association between the genotype of the dCAPS marker and seed Cd concentration, Cd concentration of the RILs was subjected to analysis of variance using the Statistica software (StatSoft, Inc., Tulsa, OK). Analyses were conducted by considering RIL and generation as main effects. To evaluate the effect of genotype, the sum of squares (SS) of the RILs was separated into SS between genotype (Harosoy genotype and Fukuyutaka genotype) and SS within genotype and then compared by F-test.

### Table 1  PCR primers used for sequencing the upstream and downstream fragments of soybean gene *GmHMA1*

<table>
<thead>
<tr>
<th>Fragment</th>
<th>Sequence (5′ to 3′)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper fragment</td>
<td>ATTTCTGAGTCAAATAGCC GCGGAACTTATTAATTGGAA ATACACTGAACTCTAATCC ATACATGAGCTGCTTGGCA</td>
</tr>
<tr>
<td>Lower fragment</td>
<td>ATATGACAGTCTGAACTCTAC ATTTGTGTATTGGTTTACG GAAAAGTCAGAAGATTTCAAT TCATCTGACCATTCTCAAT GATCTGCAATGTACGCTAA ATATAGCCAATAGGAAATCG GACATACGGAAAATACCTGA ATTTCTGAGTCACAAATAGCC</td>
</tr>
</tbody>
</table>

(5′-TGACATCGGTATCTCAGTGG-3′, 5′-ATGACAT-TCTCAGTGGTCTTTC-3′) were designed to detect a single-base substitution. The base substitution would result in presence/absence of the *Bna1* restriction site in the amplified product to generate a polymorphism (Figure 1). The PCR mixture contained 30 ng of genomic DNA, 25 pmol of each primer, 5 pmol of nucleotides, and 1 unit of ExTaq in 1 × Ex Taq Buffer in a total volume of 25 µL. After an initial 30-s denaturation at 94 °C, there were 30 cycles of 30-s denaturation at 94 °C, 1-min annealing at 59 °C, and 1-min extension at 72 °C. A final 7-min extension at 72 °C completed the program. The primers for the *GmHMA1* were 5′-CAAGAAAGAGTGGGCTAATAG-3′ and 5′-GAAAGTCAGAAGATTTCAATTT-3′. The primers for the *actin* were 5′-GAGGCTCTGGAATATT-CAACC-3′ and 5′-AGAAATCTGTGAGGTACAGA-3′. PCR products were loaded on a 2% agarose gel, stained by ethidium bromide and visualized under UV light.

**Accession numbers**

Sequence data from this article have been deposited with the DDBJ Data Libraries under accession numbers AB608748 (*GmHMA1*a of Harosoy) and AB608749 (*GmHMA1*b of Harosoy).
The open reading frame of GmHMA1a contained 2655 nucleotides with 1064 bp of the 3′-untranslated region. The putative polypeptide of GmHMA1a consisted of 885 amino acids. The apparent molecular mass and isoelectric points were 95,135 and 5.86, respectively. For GmHMA1b, however, alternative splicing generated a stop codon around the middle of the region corresponding to the 8th intron. Translation was prematurely terminated, resulting in a polypeptide consisting of 559 amino acids (Figures 1 and 2).

GmHMA1a had 49.8% amino acid similarity with P1B-ATPase of Arabidopsis AtHMA3 (Figure 3). The two consensus motifs of P-type ATPase, DKTGTLT motif and GDGxNDx motif, are conserved in GmHMA1a. The GDGxNDx motif is involved in ATP binding (Williams et al. 2000), but it was absent in GmHMA1b (Figure 1).

The cDNA sequence of GmHMA1a in Fukuyutaka was identical to that in Harosoy except for one base substitution from G to A at nucleotide position 2095. The substitution changed an amino acid from glycine to glutamic acid at amino acid number 608 of GmHMA1a, but it did not affect the putative amino acid translation of GmHMA1b because alternative splicing generated a stop codon upstream of the base substitution.

**Figure 1.** Amino acid sequence of GmHMA1a and a schematic diagram of dCAPS analysis to detect the polymorphism. (a) Amino acid sequence of GmHMA1a in the soybean cultivar Harosoy. The amino acid position where a base substitution occurred in cultivar Fukuyutaka is exhibited with single black underline. Consensus sequences for P-type ATPases (DKTGTLT and GDGxNDx motifs) are shown with double black underlines. Partial amino acid sequence of GmHMA1b, the alternatively spliced transcript, is highlighted by black boxes. Asterisks represent stop codons. (b) Schematic diagram of dCAPS analysis to detect the base substitution. Annealing sites of PCR primers are underlined. The mismatched nucleotide for dCAPS analysis is highlighted by black boxes. The nucleotides different between the cultivars are shown in bold. **BmrI** site in the PCR product of Fukuyutaka is double underlined.
To investigate whether alternative splicing occurred in the 8th intron, a DNA fragment containing the intron was amplified from genomic DNA and cDNAs using PCR primers flanking the intron. In PCR products derived from genomic DNA, only the fragments of about 600 bp containing the intron were amplified in Harosoy and Fukuyutaka (Figure 4). In contrast, short fragments of about 400 bp were amplified in addition to faint 600 bp bands from cDNA obtained from roots and leaves of all genotypes. These results suggest that alternative splicing occurred in roots and leaves of all genotypes. Judging from the band intensity of RT-PCR products, the amount of GmHMA1b transcript may be much lower than the GmHMA1a transcript.

dCAPS Analysis, Linkage Mapping, and QTL Analysis

PCR using dCAPS primers produced a 95-bp fragment in Harosoy, Fukuyutaka, and the NILs (Figure 5). After Bmr digestion, the fragments of Fukuyutaka and the NIL of Fukuyutaka type were digested to produce a shorter band of 70 bp. The dCAPS marker was designated as Gm-dCAPS-HMA1. Linkage mapping revealed that the marker was assigned to a position identical to Gm09:4790483 (Figure 6). The QTL analysis, after the incorporation of Gm-dCAPS-HMA1, indicated that the position of the QTL peak was located around the three markers, Gm09:4770663, Gm09:4790483, and Gm-dCAPS-HMA1, spanning 0.6 cM (Table 2 and Figure 6).

Association Between Genotype of Gm-dCAPS-HMA1 and Seed Cd Concentration

Analysis of variance indicated that the effects of RIL and generation were significant at 1% level (Table 3). Further, the effect of genotype of Gm-dCAPS-HMA1 was significant at 1% level. Cd concentration of the RILs with Harosoy...
Transcript levels in roots and leaves were similar between control produced bands of 546 bp and 552 bp, respectively. Concentration cosegregated perfectly in the Fukuyutaka genotype. The genotype and seed Cd RILs with Harosoy genotype was generally higher than RILs with the Harosoy genotype (Figure 7). Cd concentration of the Fukuyutaka genotype was more variable compared with the RILs with other genotypes.

Results of dCAPS analysis of Figure 5.

Figure 4. Results of alternative splicing analysis of GmHMA1. DNA fragments flanking the eighth intron was amplified using following PCR templates: genomic DNA of Harosoy and Fukuyutaka, cDNA derived from roots of Harosoy, Fukuyutaka, and a pair of NILs for cd1, #89-1 having Fukuyutaka genotype and #89-46 having Harosoy genotype, and cDNA derived from leaves of #89-1 and #89-46. φ, molecular marker ϕ174/HaelIII; H, Harosoy; F, Fukuyutaka; Nh, NIL with Harosoy genotype; Nf, NIL with Fukuyutaka genotype. The migration of the size marker (bp) is shown to the left of the gel.

The genotype was more variable compared with the RILs with Fukuyutaka genotype (Figure 7). Cd concentration of the RILs with Harosoy genotype was generally higher than RILs with Fukuyutaka genotype. The genotype and seed Cd concentration cosegregated perfectly in the F6:7 generation. In contrast, Cd concentration of some of the RILs with different genotypes overlapped in the F7:8 and F8:9 RILs.

Semi-quantitative RT-PCR

Results of semiquantitative RT-PCR analysis are presented in Figure 8. Specific primers for GmHMA1 and the actin control produced bands of 546 bp and 552 bp, respectively. Transcript levels in roots and leaves were similar between the NILs, but it is clear that GmHMA1 is preferentially expressed in the roots.

Discussion

Metal ions such as Cu2+, Zn2+, Mn2+, Fe2+, Ni2+, and Co2+ are essential micronutrients for plant metabolism but these, along with nonessential metals such as Cd2+, Hg2+, Ag2+, and Pb2+, can become toxic to plants as well as humans when present in excess. Conversely, accumulation of toxic metals by plants is used for a soil cleanup technique known as phytoremediation. A comprehensive understanding of the metal transport mechanisms in plants may be essential for selective development of crop plants that do not accumulate specific metals for protection of human health or plants that efficiently accumulate undesirable metals for phytoremediation.

A survey of the soybean genome database suggested that a P1B-ATPase gene that we designated as GmHMA1 was located at a position similar to cd1. P1B-ATPases have been identified in a wide range of organisms and have been implicated in the transport of a range of essential and also potentially toxic metals across cell membranes. P1B-ATPase constitutes one of the five major subfamilies of P-type ATPases: P1B-ATPases, Ca2+-ATPases (P2A and P2B), H+/K+-ATPases (P3A), putative aminophospholipid ATPases (P4), and a branch with unidentified specificity (P5) (Baxter et al. 2003). Ion pumps belonging to the P-type ATPase superfamily share a common enzymatic mechanism in which ATP hydrolysis is used to transport ions across a membrane (Møller et al. 1996). The P1B-ATPases have considerable potential as metal transporters in plants. In Arabidopsis, eight P1B-ATPases, AtHMA1–AtHMA8, have been reported (Baxter et al. 2003). Among them, AtHMA1–AtHMA4 exhibit high similarity with Zn2+/Cd2+ ATPases previously characterized in prokaryotes (Axelsen and Palmgren 2001). Heterologous expression of AtHMA4 enhanced Cd tolerance in yeast and complemented a zinc-hypersensitive mutant strain of Escherichia coli (Mills et al. 2003). Furthermore, functional expression of AtHMA3 complemented the Cd/Pb-hypersensitive yeast strain (Gravot et al. 2004). GmHMA1a had an amino acid homology of 49.8% with AtHMA3, suggesting that GmHMA1a might have similar functions. The two consensus motifs of P-type ATPases, DKTGTLT motif and GDGxNDx motif, are conserved in GmHMA1a.

cDNA sequence analysis suggested that alternative splicing generated two transcripts and protein isoforms in GmHMA1 gene. Sequence analysis suggested that GmHMA1a may be the functional isoform of P1B-ATPase, whereas GmHMA1b is devoid of the GDGxNDx domain that is involved in ATP binding, suggesting a complete loss of function. One of the two protein isoforms, GmHMA1b, is much less abundant compared with the other, and the rare isoform may have no ATPase function. Hence, the alternatively spliced isoform, GmHMA1b, may have little effect on the function of the GmHMA1 gene.
A high incidence of alternative splicing events (40–60%) is present in the human genome, predominantly in the form of exon skip, whereas a minor form is intron retention (5–16%) (Kan et al. 2009). In contrast, plants exhibit less alternative splicing events, of which intron retention is the most common type. Wang and Brendel (2006) analyzed cDNA/EST of Arabidopsis and rice and found alternative splicing events of 21% for both species, of which 54–56% were intron retention. We found evidence of alternative splicing events of the intron retention type in GmHMA1 similar to the soybean blue light photoreceptors (cryptochrome multigene family genes) GmCRY1b, GmCRY1c, GmCRY1d, and GmCRY2a (Matsumura et al. 2009).

A single-base substitution between Harosoy and Fukuyutaka resulting in amino acid substitution (glycine in Fukuyutaka and glutamic acid in Harosoy) in GmHMA1a. No catalytic domains have been ascribed to the region of amino acid substitution, but it was located immediately

**Table 2** QTL for seed Cd concentration in the F$_{6:7}$, F$_{7:8}$, and F$_{8:9}$ RILs developed from a cross between the soybean cultivars Harosoy and Fukuyutaka at Yawara, Japan, in 2006, 2007, and 2008

<table>
<thead>
<tr>
<th>QTL name</th>
<th>Linkage group</th>
<th>Generation</th>
<th>Proximal markers</th>
<th>Position$^a$, cM</th>
<th>LOD score</th>
<th>Additive effect$^b$</th>
<th>Variance explained, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>cd1 K</td>
<td>F$_{6:7}$</td>
<td>Gm09:4770663</td>
<td>39.91</td>
<td>38.44</td>
<td>0.10</td>
<td>84.41</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Gm-dCAPS-HMA1/Gm09:4790483</td>
<td>40.46</td>
<td>38.27</td>
<td>0.10</td>
<td>84.41</td>
<td></td>
</tr>
<tr>
<td>F$_{7:8}$</td>
<td></td>
<td>Gm09:4770663</td>
<td>39.91</td>
<td>20.65</td>
<td>0.22</td>
<td>52.25</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Gm-dCAPS-HMA1/Gm09:4790483</td>
<td>40.46</td>
<td>20.48</td>
<td>0.22</td>
<td>52.25</td>
<td></td>
</tr>
<tr>
<td>F$_{8:9}$</td>
<td></td>
<td>Gm09:4770663</td>
<td>39.91</td>
<td>28.78</td>
<td>0.13</td>
<td>73.77</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Gm-dCAPS-HMA1/Gm09:4790483</td>
<td>40.46</td>
<td>28.61</td>
<td>0.13</td>
<td>73.83</td>
<td></td>
</tr>
</tbody>
</table>

$^a$ Distance from top of the linkage group.

$^b$ Additive effects of each QTL are those of Harosoy allele in contrast to Fukuyutaka allele.
downstream of the ATP-binding domain. The glycine residue at the site of amino acid substitution was fully conserved in AtHMA3, AtHMA4, AtHMA6, and AtHMA7 (Mills et al. 2003; Gravot et al. 2004), suggesting that GmHMA1a of Fukuyutaka is the wild type. Similar to AtHMA3 and AtHMA4, expression of GmHMA1 was substantially lower than actin and was predominant in roots compared with leaves (Mills et al. 2003; Gravot et al. 2004).

The previous study suggested that the QTL peak was located in the vicinity of two SSR markers located at Gm09:4770663 and Gm09:4790483 (Benitez et al. 2010). In this study, we developed a dCAPS marker Gm-dCAPS-HMA1 based on the base substitution in GmHMA1. The polymorphism was also observed among the NILs, suggesting that GmHMA1a of Fukuyutaka is the wild type. Similar to AtHMA3 and AtHMA4, expression of GmHMA1 was substantially lower than actin and was predominant in roots compared with leaves (Mills et al. 2003; Gravot et al. 2004). The genotype of Gm-dCAPS-HMA1 was significantly associated with seed Cd concentration. The genotype and Cd concentration completely cosegregated in the F6:7 RILs. However, the Cd concentration of some of the RILs with different genotypes overlapped in the F7:8 and F8:9 RILs. LOD score of cd1 (F6:7 = 38.44, F7:8 = 20.65, F8:9 = 28.78) was negatively correlated (r = −0.95) with the distribution of Cd concentration of the RILs (SD of F6:7 = 0.11, F7:8 = 0.26, F8:9 = 0.15) (Table 2 and Figure 7). Judging from the homogeneity of SSR marker genotypes, genetic variation might be almost similar among the F6:7, F7:8, and F8:9 generations. Instead, the interaction between environmental factors and genetic effects might be the cause of yearly fluctuation in cosegregation between the dCAPS marker genotype and seed Cd concentration. Environmental effects may be evident in some cases (years) in which effects of the QTL might be diminished, such as in 2007 and 2008, or environmental effects could be weak such that QTL effects are almost complete, such as in 2006.

The P1B-ATPases may be present at the plasma membrane and function as efflux pumps removing potentially toxic metals from the cytoplasm or they may also be present at various intracellular membranes and be responsible for compartmentation of metals, for example, sequestration in the vacuole, Golgi, or endoplasmic reticulum (Williams et al. 2000; Mills et al. 2003; Gravot et al. 2004). Understanding of the role of this pump will require various information including tissue and cellular localization and regulation by metals. We identified a candidate gene for seed Cd concentration in soybean using populations and

### Table 3

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>df</th>
<th>SS</th>
<th>Mean square</th>
<th>F value</th>
</tr>
</thead>
<tbody>
<tr>
<td>RIL</td>
<td>90</td>
<td>7.43</td>
<td>0.08</td>
<td>8.08*</td>
</tr>
<tr>
<td>Between genotype</td>
<td>1</td>
<td>6.06</td>
<td>6.06</td>
<td>404.00*</td>
</tr>
<tr>
<td>Within genotype</td>
<td>89</td>
<td>1.37</td>
<td>0.015</td>
<td></td>
</tr>
<tr>
<td>Generation</td>
<td>2</td>
<td>5.27</td>
<td>2.63</td>
<td>258.00*</td>
</tr>
<tr>
<td>Error</td>
<td>180</td>
<td>1.84</td>
<td>0.01</td>
<td></td>
</tr>
</tbody>
</table>

* Significant at the 1% level.

Figure 7. Association between genotype of Gm-dCAPS-HMA1 and seed Cd concentration in the F6:7, F7:8, and F8:9 recombinant inbred lines developed from a cross between the soybean cultivars Harosoy and Fukuyutaka at Yawara, Japan, in 2006, 2007, and 2008.
NILs derived from a single cross combination and developed a dCAPS marker based on the base substitution that existed in these populations. A survey of various genetic resources with different seed Cd levels may be necessary to ascertain the prevalence of the base substitution, existence of different genetic polymorphisms associated with seed Cd concentration, and usefulness of the dCAPS marker that we developed. Transgenic experiments may be necessary to determine the function of GmHMA1a and to verify whether the amino acid substitution affected transport and accumulation of Cd in seeds.

Acknowledgments

The authors are grateful to Dr J. Abe (Hokkaido University) for advice and to Dr Joseph G. Dubouzet (SCION Research, New Zealand) for critical reading of the manuscript.

References


Received May 9, 2011; Revised October 9, 2011; Accepted October 11, 2011

Figure 8. Results of semiquantitative RT-PCR analysis of GmHMA1. Total RNA from roots and leaves of a pair of NILs for cd1, #89-1 having Fukuyutaka genotype and #89-46 having Harosoy genotype was used for analysis. Soybean actin gene was used as a control. Nh, near-isogenic line with Harosoy genotype; Nf, near-isogenic line with Fukuyutaka genotype; λ, molecular marker λ/Hind III. The migration of the size marker (bp) is shown to the right of the gel.

[286]