(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
5 April 2012 (05.04.2012)

PCT

(10) International Publication Number
WO 2012/041913 A1

(51) International Patent Classification:
C12N 15/82 (2006.01) *C12N 9/14* (2006.01)

(21) International Application Number:
PCT/EP2011/066882

(22) International Filing Date:
28 September 2011 (28.09.2011)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
10290518.9 30 September 2010 (30.09.2010) EP

(71) Applicant *(for all designated States except US)*: **SOCIETÉ NATIONALE D'EXPLOITATION INDUSTRIELLE DES TABACS ET ALLUMETTES - S.E.I.T.A.** [FR/FR]; 143 Boulevard Romain Rolland, F-75014 Paris (FR).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **JULIO, Emilie** [FR/FR]; Le Bourg, F-24240 Cunèges (FR). **DORLHAC DE BORNE, Francois** [FR/FR]; 11 Rue St. Roch, F-24680 Lamonzie St. Martin (FR). **HERMAND, Victor** [FR/FR]; 12 Rue Rondelet, F-34000 Montpellier (FR).

(74) Agents: **UEXKÜLL & STOLBERG** et al.; Beselerstr. 4, 22607 Hamburg (DE).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT,

[Continued on next page]

(54) Title: TOBACCO WITH REDUCED CADMIUM CONTENT

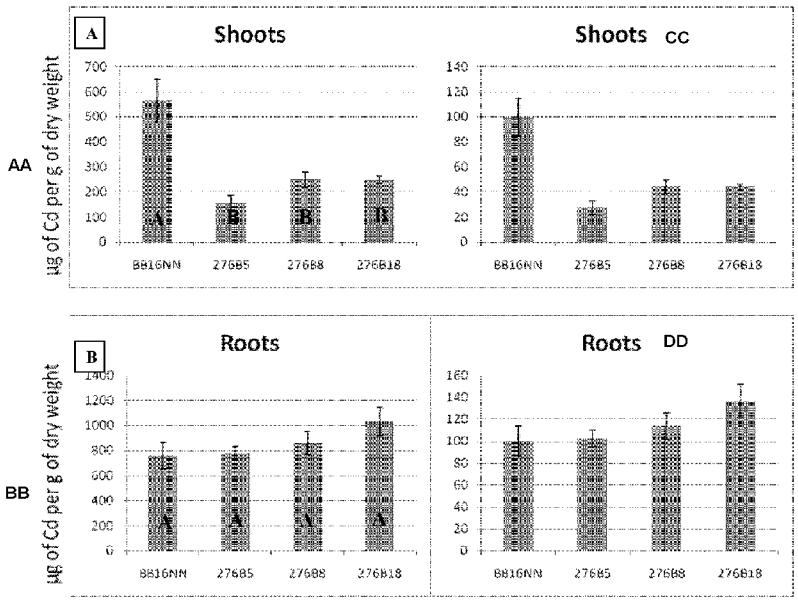


Figure 6

(57) Abstract: A tobacco plant comprising at least one mutation in a HMA gene, wherein the non-mutated HMA gene comprises the nucleotide sequence of SEQ ID NO:1 or a homolog sequence, wherein the mutation causes a substitution or a deletion or an insertion of at least one amino acid in the polypeptide encoded by the nucleotide sequence and wherein the mutation reduces the heavy metal uptake by the leaves of the plant by at least 30% in relation to the heavy metal uptake of plants comprising SEQ ID NO:1 or the homolog.

WO 2012/041913 A1

LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

— *with sequence listing part of description (Rule 5.2(a))*

Tobacco with Reduced Heavy Metal Content

Heavy metals are naturally present in soil and are taken up by plants to a different degree. Some heavy metals, such as manganese or zinc, are essential for plants, since they represent co-factors required for enzyme activity.

Other heavy metals are however not essential for plants and in some cases a reduction in the heavy metal concentration of plants or parts of plants would be advantageous. Cadmium (Cd) for example is a non-essential heavy metal present in the soil and naturally absorbed by plant roots. Cd is accumulated in leaves, for example in leaves of tobacco plants (Lugon-Moulin *et al.,* 2006). In some places Cd concentrations in soil are increased due to the use of Cd-rich phosphate fertilizers or Cd-contaminated sewage sludge (Karaivazoglou *et al.,* 2007). The mechanisms used by the plants for the uptake, partitioning and accumulation of heavy metals have been described (Verbruggen *et al.,* 2009).

It would be advantageous to reduce the heavy metal content of tobacco plant leaves and different strategies leading to this result have been developed.

For example, a mammalian metallothionein gene under the control of 35S promoter was expressed in tobacco, causing a decrease of 14% of cadmium in field grown plant (Yeargan *et al.,* 1992) or to altered cadmium tissue distribution with a 73 % cadmium reduction in lamina (Dorlhac *et al.,* 1998). More recently, high capacity divalent antiporters, AtCAX2 and AtCAX4, from *Arabidopsis thaliana* have been successfully used to enhance root vacuole Cd sequestration thus obtaining a 15-25% decrease of Cd in lamina of field grown tobacco (Korenkov *et al.,* 2009).

A family of Heavy Metal ATPases (HMA proteins) responsible for distributing heavy metals in the plant tissues after the metals have been taken up by the root have been identified and

genes encoding respective proteins have been cloned. Eight types of HMA genes involved in Cd transport to and Cd accumulation in leaves of *Arabidopsis thaliana* are known (Cobbet *et al.*, 2003). In *A. thaliana* HMAs from type 1, 5, 6, 7 and 8 are monovalent cation transporters (Cu+ and Ag+; Seigneurin-Berny *et al.*, 2006) and HMAs from type 2, 3 and 4 are bivalent cation transporters (Pb2+, Co2+, Zn2+, Cd2+; Hussain *et al.*, 2004). HMA3 is a vacuolar transporter involved in Cd, Pb and Zn tolerance (Morel et al. 2009).

HMA genes have also been identified in other plants and it has been suggested to generate genetically modified tobacco plants, wherein the expression of a HMA gene is inhibited by RNAi (WO2009/074325).

However, these approaches for reducing heavy metal in tobacco plants are based on the use of genetically modified organisms (GMOs). An approach that is not based on genetically modified plants would have significant advantages.

Mutation breeding has been used for a long time to modify existing traits or to create new valuable traits within a plant cultivar. Recently, this technique has been combined with a high-throughput mutation detection system and has proven to be efficient for the modification of properties in tobacco plants (Julio *et al.*, 2008).

Modern tobacco is an allotetraploid from *Nicotiana sylvestris* (maternal contributor) and *Nicotiana tomentosiformis* (paternal contributor; Yukawa *et al.*, 2006). The tetraploid nature of tobacco (2n=48) complicates the identification of traits in tobacco lines and the preparation of tobacco plants showing completely new traits.

There is thus still a need for tobacco plants with reduced heavy metal content in leaves.

In accordance with the present invention, this problem is now solved by tobacco plants comprising at least one mutation in a HMA gene, wherein the non-mutated HMA gene comprises the nucleotide sequence of SEQ ID NO:1 (nucleotide sequence of HMA 4 gene) or a homolog thereof, wherein the mutation causes a substitution or a deletion or an insertion of at least one amino acid in the polypeptide encoded by the nucleotide sequence and wherein the mutation reduces the heavy metal uptake by the leaves of the plant by at least 30% in relation to the heavy metal uptake of plants comprising SEQ ID NO:1 or the homolog.

The present inventors have surprisingly found that a tobacco HMA 4 gene can be modified such that the HMA 4 protein is inhibited without significant detrimental effects for the plants. In other words, the present invention provides normal tobacco plants that can be used for commercial purposes with significantly reduced heavy metal content in the leaves.

In accordance with the present invention, a homolog of the sequence of SEQ ID NO:1 refers to a sequence with at least 90%, preferably at least 95% sequence identity to the sequence of SEQ ID NO:1. Respective homolog or homologous sequences may represent differences between various tobacco plant lines or sequences derived from a common ancestor. *N. tabacum* is an allotetraploid plant and each plant comprises a genome from *Nicotiana sylvestris* and *Nicotiana tomentosiformis*. As it appears that SEQ ID NO:1 is derived from *N. sylvestris*, the HMA 4 sequence derived from *N. tomentosiformis* represents a homolog of the sequence of SEQ ID NO:1. Similar sequences derived present in different species are also identified in this application as ortholog sequences and represent a specific form of a homolog sequence. The % identity is preferably determined by the BLAST software for determining sequence identity.

In a preferred embodiment, the sequence of SEQ ID NO:1 or of the homolog only contains one or a small number of mutations

on the nucleic acid level, for example 1, 2, 3 or 4 nucleotide changes. The mutated sequence thus still has an identity of at least 90%, preferably at least 95% and most preferably at least 98% to the sequence of the HMA 4 gene. The mutation may for example be a miss-sense, a non-sense mutation or a splice mutation.

The tobacco plants of the present invention are preferably *Nicotiana tabacum* plants.

According to one aspect the present invention provides respective tobacco plants, wherein the reduction of heavy metal uptake is determined by growing tobacco plants with and without the mutation under identical conditions on a liquid medium containing the heavy metal and comparing the concentration of the heavy metal in the leaf, stem or shoot of the tobacco plant with the mutation to the concentration of the heavy metal in the leaf, stem or shoot of the tobacco plant that does not have the mutation. It is preferred that the plant and the reference plant are otherwise grown under the same conditions. The concentration of the heavy metal in the leaf, stem or shoot can be identified in amount of heavy metal in relation to amount of plant dry weight material.

The term "reduction in heavy metal uptake" is therefore used in the context of the present invention to describe a reduction in heavy metal concentration in a leaf, stem or shoot of a tobacco plant having a mutation in the HMA 4 gene in comparison to the corresponding tissue of a plant that does not have the mutation but was otherwise grown under the same conditions.

The invention encompasses plants which under these circumstances show a significant reduction in heavy metal concentration in tissues of a tobacco plant. It is particularly preferred that the mutation reduces the uptake of a heavy metal by at least 40%, at least 50%, at least 75% or at least 95%.

In accordance with the present invention, a reduction of uptake can be achieved for one or several but does not have to be achieved for all heavy metals. It is sufficient if a significant reduction is achieved for one heavy metal. It is preferred that the invention provides tobacco plants show a reduced uptake of cadmium, lead or arsenic. In its most preferred embodiment, the present invention provides tobacco plants which have a reduced uptake of cadmium.

Similarly, the reduction of the uptake need not completely eliminate the content of the heavy metal in the plant or plant tissue. In accordance with the present invention it is preferred that the reduction of heavy metal content is in the range of 40% to 70%, 50% to 80%, or 60% to 95%.

Several mutations have been identified in the HMA 4 gene that cause a significant reduction of the heavy metal uptake and are described below. Accordingly, the mutation may represent a miss-sense, a non-sense mutation or a splice mutation.

In one embodiment the tobacco plants of the present invention comprise a mutation selected from one of the mutations shown in Figure 5/1 and Figure 5/2. In a preferred embodiment, the tobacco plants comprise one of the mutations shown in Figure 5/1 or 5/2 with a SIFT score of less than 0,05. In the most preferred embodiment the tobacco plants of the present invention comprise a mutation selected from the group comprising the mutations: G294A, C576T, G406A, G347A, G363A, G553A, C374T, G290A, G964A, G1168A, G1211A, G1126A, C980T, G1195T, G1156A, G1070A, C2302T, G2208A, C2217T, G2190A, C2206T or C2277T in SEQ ID NO:1 or the homolog thereof.

According to one embodiment of the present invention, the tobacco plants may comprises a mutation in more than one HMA 4 gene. As indicated, tobacco plants contain HMA 4 genes from *N. Sylvestris* and from *N. tomentosiformis*. Plants that are

mutated in both HMA 4 genes are also identified as double mutants in the present invention.

The tobacco plants of the present invention can be homozygous or heterozygous for any one mutation in one of the HMA 4 genes. According to a particularly preferred embodiment, plants are provided that contain homozygous mutations in both HMA 4 genes.

In a further embodiment, a tobacco plant cell is provided, which may be derived from a tobacco plant as described above.

The present invention also relates to a part of a tobacco plant, wherein the part is a leaf, a lamina, a cut and/or a cured leaf, root, shoot, stem, flower or seed. Again, for all purposes of the present invention, the part of a *tobacco* plant is preferably a part of a *Nicotiana tabacum* plant. Seed of a tobacco plant as described above, represent a particularly preferred embodiment of the present invention.

The tobacco plants of the present invention or the parts thereof may be used to generate tobacco products well known in the art, including smokeless tobacco products like snus, snuff and cut tobacco, tobacco extract or reconstituted tobacco.

In an alternative embodiment, the tobacco plants of the present invention or the parts thereof are used to generate smoking articles which also represent an embodiment of the invention. Smoking articles are well known and include a cigarette, a small cigar, cigarillo or a cigar or simulated smoking articles containing tobacco.

In a further aspect, the present invention provides methods for generating a tobacco plant according to the present invention. Respective methods may comprise the following steps:

(a) screening a library of tobacco plants obtained by mutagenesis for at least one mutation in a HMA gene, wherein the non-mutated HMA gene comprises the nucleotide sequence of SEQ ID NO:1 (nucleotide sequence of HMA 4 gene) or a homolog thereof, wherein the mutation causes a substitution or a deletion or an insertion of at least one amino acid in the polypeptide encoded by the nucleotide sequence and wherein the mutation reduces the heavy metal uptake by the leaves of the tobacco plant by at least 30% in relation to the heavy metal uptake of plants comprising SEQ ID NO:1 or the homolog;

(b) crossing the tobacco plant having a mutation in the HMA gene with a commercial *Nicotiana tabacum* production plant; and

(c) identifying offspring with the mutation in the HMA gene;

(d) repeating steps (b) and (c).

As before, the reduction of heavy metal uptake can be determined by growing tobacco plants with and without the mutation under identical conditions on a liquid medium containing the heavy metal and comparing the concentration of the heavy metal in the leaf, stem or shoot of the tobacco plant with the mutation to the concentration of the heavy metal in the leaf, stem or shoot of the tobacco plant that does not have the mutation.

The methods of the present invention preferably reduce the uptake of the heavy metal by at least 30%, at least 50%, at least 75% or at least 95%.

The methods may be used to reduce the uptake of one or several heavy metals. It is preferred that the uptake of cadmium, lead or arsenic is reduced.

According to a preferred embodiment, steps (a) and/or (c) of the method use an assay analyzing the nucleotide sequence of

the tobacco (*Nicotiana tabacum)* plant. Respective assays for nucleotide analysis are well known in the art of molecular biology and include PCR-based techniques, DNA sequencing, hybridization and/or RFLP techniques.

The invention further provides methods for producing a tobacco product or a smoking article comprising the method of generating a tobacco plant as described above. A tobacco product or a smoking article can be obtained from the plants generated according to this method by further harvesting the leaves, stems and/or shoots and producing a tobacco product or smoking article from the same.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1: Partial sequence alignment of part A amplification in *N. tabacum*, *N. sylvestris* and *N. tomentosiformis*.
The first sequence is the reference contig. In this alignment three different sequences are present: sequences 1, 2, 3, 9, 10, 11 are shared between *N. tabacum* and *N. sylvestris*. Sequences 6, 7, 8, 12, 13, 14, 15, 16 are shared between *N. tabacum* and *N. tomentosiformis*. Sequences 4 and 5 are specific to *N. sylvestris*.

Figure 2: ds cDNA amplification of copies 1 and 2 of the HMA orthologs in *N. tabacum*.
Specific primers designed to anneal in exonic regions of part A were used to test the expression of the two copies in roots (R) and shoots (S). HmaAS-F/HmaA-RT-R primers pair amplify the *N. sylvestris* origin sequence (Copy S) and HmaAT-F/HmaA-RT-R primers pair amplify the *N. tomentosiformis* origin sequence (Copy T). Both copies are expressed in *N. tabacum*.

Figure 3: CODDLE analysis of the whole contig sequence (SEQ ID NO:1).
First line: amino-acid sequence of the HMA4 protein

Second line: DNA sequence of the HMA4 gene
Third line: below each nucleotide is an indication of the changes that can be detected with EMS mutagenesis.

Figure 4: CE-SSCP profile obtained by PCR amplification of the DNA mutant collection with primers Hma4-BF/Hma4-BR. The HmaA4-BF primer is labeled with FAM fluorophore, and the DNA strand appears in blue. The Hma-BR primer is labeled with VIC fluorophore, and the DNA strand appears in green.
1: DNA profile of a wild type DNA
2: DNA profile of a mutant DNA (family E1-276 of the collection), characterized by an additional peak corresponding to the mutated DNA strand (red arrow).

Figure 5: Table summarizing the mutations found in HMA targets after cloning and sequencing.
[a] = Identification number of the mutant family as described in the collection.
[b] = nucleotide or amino-acid change and position of the mutation as described with CODDLe analysis on contig 1 (original nucleotide/amino-acid+ position +new nucleotide/amino-acid).
"=" represents silent mutation, "*" represents stop codon.
[c] = SIFT score was obtained with the bioinformatics program SIFT (Sorting Intolerant from Tolerant) on contig 1. SIFT scores < 0.05 are predicted to be deleterious to the protein.
* non EMS mutation (G/C to A/T)

Figure 6: Cadmium content in three M3 E1-276 mutant lines (276B5; 276B8, 278B18) compared to the wild type (BB16NN).
Cadmium was measured in shoots (A) or roots (B). On the left, graphs represent cadmium accumulation in µg per g of dry weights. On the right, graphs represent

cadmium accumulation normalized in mutants with cadmium accumulation in wild type.
Cadmium content in shoots: the letters "A" in "B" in bars represent two distinct groups in a student "T" test performed on these samples.
Cadmium accumulation in roots: the letter "A" in bars indicates that all the means in this graph belong to the same statistical group.
Error bars represent standard deviation.

Figure 7: Picture of a F1 cross between the mutant E1-276-B18 and an industrial flue-cured tobacco plant showing the viability of the mutants.

Figure 8: Cadmium accumulation in shoots of EMS lines (M mutated plants; Htz heterozygous mutant plants; S wild-type plants; DW dry weight). The graphs on the left represent the amount of cadmium accumulated in the shoots. The error bars reflect the standard error. On the right side, the accumulation of cadmium in the mutated line is represented as a percentage with respect to its wild-type.

Figure 9: Statistal analysis of EMS lines. The p-values presented in this chart are derived from the student's t-test (http://www.graphpad.com). The differences observed between each mutated line and its wildtype is statistically relevant when the p-value is lower than 0,05 (95% confidence).

Figure 10: Accumulation of RNA transcripts in RNAi lines. The bars represents the mean of the accumulation of transcripts of either HMA alpha or HMA beta relative to the transcript accumulation of the reference gene cyclophilin and L2. Three plants were used per line. The error bars represent standard error. The letters A and B represent the two groups found by student's t-test realized with the open software "R".

Figure 11: Metal accumulation in shoots of RNAi lines. For each line, 4 plants were analyzed (only 3 for pGreen). The cadmium accumulation is represented at the top. On the left, the cadmium accumulation is expressed as a mean. On the top right, those means are expressed as percent of wild-type. At the bottom, the accumulation of iron and zinc in shoots is represented. The error bars represent standard error. The letters A and B represent statistical groups realized with the open software "R". The letters are independent in each graph.

Figure 12: Crosses performed between tobacco mutants to accumulate both mutated copies in one genome. Mutants were first backcrossed with wild-type and then crossed between them to obtain F1 generation. The F1 name is described as F1CD2, with the following numbers representing both mutant identifications.

Figure 13: EMS lines. The lines are presented with the collection from which they originate and the mutation that affects either one gene or the other.

Figure 14: amiRNA constructs. All of the targeted sequences and their corresponding primer sequences were obtained using online software (http://www.weigelworld.org/).

Figure 15: Targeted sequence of the hairpin construct. The primer pairs (A) were designed to amplify the same portion of *HMA* gene (B) but they carry different restriction sites.

Figure 16: pHannibal vector. This vector contains 2 sets of restriction sites separated by an intron.

Figure 17: Metal accumulation in shoots; graphical results obtained by Tukey multiple comparisons of means. The error bars represent standard error. The letters A, B, C and D represent statistical groups realized with the open software "R" for cadmium analyses.

Figure 18: Metal accumulation in shoots; graphical results are obtained by Tukey multiple comparisons of means.

The subsequent examples illustrate specific embodiments of the present invention:

EXAMPLE 1

Tobacco mutant collection

Plant material

Seeds of three tobacco lines BB16NN (Delon et al. 1999; Institut du Tabac de Bergerac, Accession N°1139), BY02 and V4K1, were used for creating *Nicotiana tabacum* mutant libraries. BB16NN and BY02 are burley type tobaccos, and V4K1 is a flue-cured type tobacco.

For copy number assessment, comparison of sequences with ancestors was done with tobacco lines ITB 645 (*Nicotiana tomentosiformis*) and ITB 626 (*Nicotiana sylvestris*).

EMS mutagenesis

The mutant collection has been described in Julio *et al.*, 2008. In short, two tobacco EMS (ethylmethane sulfonate) mutant libraries termed L1 and L2 were constructed by soaking tobacco BB16NN seeds (6000 seeds per library) overnight (16h) in 0.8% EMS (L1) or 0.6% EMS (L2) solutions, followed by 12 washings of 30 min in water under shaking. In addition, L1 library seeds were pre-germinated for 2 days before EMS treatment.

Two further collections were developed from BY02 (termed L3) and V4K1 (termed L4) seeds with a 0.7% EMS treatment and without pre-germination.

The mutagenized M1 seeds were grown to M1 plantlets in a greenhouse and transferred to the field to give M2 generation by self-pollination. M2 seeds were collected from each M1 plant and stored until use. Leaf material was collected from 8 M2 seeds sown in a single pot in greenhouse and pooled (two 8 mm diameter discs for each plant i.e. ~100 mg fresh weight per family) to constitute pooled M2 family. DNA was extracted from the leaf material using QIAGEN Dneasy 96 Plant Kit according to manufacturer's instructions.

EXAMPLE 2

Copy Number Assessment

PCR amplification

Part of the genomic sequence of an HMA 4 of *Nicotiana tabacum* gene is shown in SEQ ID NO:1.

N. tabacum is an allotetraploid plant comprising a genome from *Nicotiana sylvestris* and *Nicotiana tomentosiformis*. It appears that SEQ ID NO:1 is derived from *N. sylvestris*.

Four genomic regions of this HMA 4 gene, identified as part A, B, C and D, were amplified in *N. tabacum*, *N. sylvestris* and *N. tomentosiformis* by using the respective primer pairs:

PartA-F (CTACCGCTGCTATGTCATCAC) SEQ ID NO:2 and
PartA-R (TAGCACACTTGTCGATGTATC) SEQ ID NO:3;

PartB-F (GATACATCGACAAGTGTGCTA) SEQ ID NO:4 and
PartB-R (CTCCTTTAGTTATAGTCCCTG) SEQ ID NO:5;

PartC-F (AGTAAATACTGAATTGTCTAGTG) SEQ ID NO:6 and
PartC-R (GATGTTTTATCTCTACTATGAGC) SEQ ID NO:7;

PartD-F (GACCTGTTTAGCACTAATGCG) SEQ ID NO:8 and
PartD-R (TTATAATCATTTCAGCGTAATGCAG) SEQ ID NO:9.

PCR amplifications were carried out in a 20 µl volume containing 1 µl DNA, 10X AmpliTaq buffer (Applied Biosystems, Foster City, USA), 1 µl dNTPs (Applied Biosystems, 2.5 mM each), 50 ng of each primer and 0.05 U AmpliTaq Polymerase (Applied Biosystems). PCR was conducted using a thermal cycler (GeneAmp® PCR System 2700, Applied Biosystems) as follows: 35 cycles of 94°C for 30s, 62°C for 45s, 72°C for 1 min, followed by 7 min at 72°C for final extension.

Results

Four parts of the contig sequence shown in SEQ ID NO: 1 were amplified in *N. tabacum*, *N. tomentosiformis* and *N. sylvestris*. Part A corresponds approximately to exons 4 to 5, part B corresponds approximately to exon 6, part C corresponds to exon 7 and part D corresponds to exon 8. Six to ten clones were sequenced for each PCR product. An example of sequence alignment in part A is shown in Figure 1. Some sequences were common between *N. tabacum* and its ancestors (*N. sylvestris* and *N. tomentosiformis*); others are specific to one ancestor. Results are summarized in Table 1, below.

	Part A		Part B		Part C		Part D	
	Specific	Common with N.tabacum	Specific	Common with N.tabacum	Specific	Common with N.tabacum	Specific	Common with N.tabacum
N. sylvestris	2	1	n.f.	1	n.f.	1	n.f	1
N. tomentosiformis	n.f.	1	1	n.f.	1	n.f.	n.f.	1

Table 1: Results obtained after cloning and sequencing PCR products amplified with primers HmaA-F/R (partA), HmaB-F/R (partB), HmaC-F/R (partC) and HmaD-F/R (partD).

As 6 to 10 clones were sequenced per PCR product, results are expressed as a minimal number of existing copies.

n.f. = not found

These results show that *Nicotiana tabacum* contains at least two HMA4 loci, one locus from each of the two ancestors. As a consequence, any *Nicotiana tabacum* plant may contain at least four different HMA 4 alleles.

RNA extraction and ds cDNA testing

RNA was extracted from roots and shoot of two months old plants with RNeasy Plant Kit according to manufacturer's instructions (Qiagen). cDNA and ds cDNA was prepared according to manufacturer's instructions by using the MINT cDNA synthesis kit from Evrogen.

Expression of both loci was assayed by amplifying ds cDNA coming from *N. tabacum* roots and shoots with specific primers designed on the basis of part A exonic regions. HmaAS-F/HmaA-RT-R was used to amplify the *N. sylvestris* origin sequence and HmaAT-F/HMA-RT-R for the *N. tomentosiformis* origin sequence. Fragments were amplified for both copies (*N. sylvestris* and *N. tomentosiformis*) of the gene, as demonstrated in Figure 2, supporting the conclusion that the sequences present at both loci are expressed in *N. tabacum* roots and shoots.

Amplification of the two copies in expressed sequences was tested in region A, with specific non-labeled primers pairs:

- forward primers Hma-AS-F GGTGAATAGCATTCTTGCTGTG SEQ ID NO:19; and HmaAT-F, GTTGAATAGCATTCTTGCTGTT SEQ ID NO:21; and
- a new common reverse primer HmaA-RT-R, CTTGTTCTGAGCATCTTCGAC SEQ ID NO:20, designed to bind in the exonic region.

PCR was carried out in the same conditions than for copy number assessment. PCR products were visualized on a 1.5% agarose gel with EtBr (ethidium bromide) under UV.

EXAMPLE 3

Mutant Screening via PCR

New primers were designed to amplify specifically the regions to be used for mutant screening. Primers were selected according to three criteria:

- Positions of intron and exon: exonic regions are preferred, even if primers can be designed in intronic regions.
- High impact of EMS mutations on the protein sequence: the impact of the mutation on the protein function was assessed by the CODDLe software [Choosing codons to Optimize Discovery of Deleterious LEsions]. CODDLe identifies the region(s) of a user-selected gene and of its coding sequence [CDS] where the anticipated point mutations are most likely to result in deleterious effects on the gene's function (Till *et al.*, 2003). CODDLe results on the whole contig sequence is shown in Figure 3.
- Length of the sequence: the maximum length of the sequence for CE-SSCP screening is 500 bp. A large sequence is preferred to maximize the chance to discover mutants.

The following new primer pairs were designed to amplify:

Region A: HmaAS-F/HmaA-R primer pair amplifies the *N. sylvestris* origin sequence and HmaAT-F/HMA-R amplifies the *N. tomentosiformis* origin sequence.

Region B: HmaB-F/HmaB-R primer pair.

Region D: HmaDS-F/HmaDS-R primer pair amplifies the *N. sylvestris* origin sequence and HmaDT-F/HmaDT-R primer pair amplifies the *N. tomentosiformis* origin sequence.

The specificity of primer pairs was checked by cloning and sequencing ten PCR products.

Fluorescent labeled primers (6-FAM (blue), VIC (green), NED (yellow); all from Applied Biosystems) were used for Capillary Electrophoresis-Single Strand Conformation Polymorphism analysis (CE-SSCP). PCR reactions were carried out in the same conditions as used for copy number assessment.

Two different 410 bp copies were amplified in regions A with:

- forward primer HmaAS-F (VIC-GGTGAATAGCATTCTTGCTGTG) SEQ ID NO:10; or
- forward primer HmaAT-F (NED-GTTGAATAGCATTCTTGCTGTT) SEQ ID NO:12; and
- common reverse primer HmaA-R (6FAM-GCACAACATAAGATTCACTAAC) SEQ ID NO:11.

A unique 386 bp sequence was amplified in region B, with:

- forward primer HmaB-F 6FAM-GTCTGATTTCGACTGGTGATG SEQ ID NO:13; and
- reverse primer HmaB-R VIC-AAGAATATGTATGAGTGGTAACC SEQ ID NO:14.

Two 283 bp copies were amplified separately in region D, with:

- primer pair HmaDT-F, 6FAM-GAAATAGAGGGTGATAGTTTCC SEQ ID NO:15, and HmaDT-R, NED-CATTTCAGCGTAATGCAGAATTT SEQ ID NO:16, for the first copy; and
- HmaDS-F, VIC-GAAATAGAGGGTGATAGTTTCA SEQ ID NO:17, and HmaDS-R, 6FAM-CATTTCAGCGTAATGCAGAATTA SEQ ID NO:18, for the other copy.

Capillary electrophoresis

Fluorescent-labeled PCR products were diluted 1/20 in water before CE-SSCP analysis. Prior to loading on ABI Prism® 3130

(Applied Biosystems, Foster City, USA), 1µl of diluted sample was added to 10 µl formamide (Applied Biosystems) and 0.1 µl Genescan-500 LIZ Size Standard (Applied Biosystems). A denaturation step of 94°C for 3 min followed by cooling on ice was used for single strand conformation analysis.

Running conditions on ABI Prism® 3130 were as follows: 36 cm capillary array (16 capillaries), run temperature of 22°C, sample injection of 1 kV for 15 s and separation of 15 kV for 40 min. The non denaturing separation medium was POP Conformational Analysis Polymer (Applied Biosystems) 5%, glycerol (Sigma-Aldrich) 10% in 1X Buffer (10X) with EDTA (Applied Biosystems). The running buffer was glycerol 10% in 1X Buffer (10X) with EDTA. Results were analyzed with GeneMapper 4.0 (Applied Biosystems) software.

Cloning and sequencing

PCR products were cloned into pGEM-T vector Systems (Promega, Madison, USA) and transformed into *E. coli* according to manufacturer's instructions. Ten clones were sequenced for each family, using BigDye Terminator Sequencing Kit v3.1 (Applied Biosystems) and ABI Prism® 3130 (Applied Biosystems). Nucleotic sequences were aligned with Multalin (http://npsa-pbil.ibcp.fr/NPSA/npsa_multalinan.html).

Results

Mutant screening in the HMA4 orthologs in tobacco

Primers were fluorescently labeled and were used to amplify the DNA of the mutant collection. Fluorescent-labeled PCR products were analyzed by CE-SSCP as described above.

Mutants were detected by the presence of additional peaks on the analysis profile, compared to the control (DNA of a non mutated tobacco). An example with primers Hma4-BF/Hma4-BR is shown in Figure 4.

PCR products of mutants were cloned and sequenced to characterize the position of the mutation. Sequences obtained

were aligned and compared to the control. Impact of mutations on protein function was analyzed by the Sorting Intolerant From Tolerant (SIFT) program (Ng and Henikoff 2003).

Complete results of mutations obtained in Hma-AS, Hma-AT, Hma-B, Hma-DS and Hma-DT are summarized in the Table shown in Figure 5/1 and 5/2.

Nineteen mutations were obtained for Hma-AS target, 18 for Hma-AT, 20 for Hma-B, 11 for Hma-DT and 3 for Hma-DS.

Non-sense mutations could be obtained for Hma-AS and Hma-DT. Silent and miss-sense mutations represent respectively 9% to 33 % and 61% to 81% of the total number of mutations, as presented in table 2. One mutation in a splicing region could be found in Hma-B target. Mutations affecting the same amino-acid could be found. Two exactly redundant mutations were found in Hma-AT, one in Hma-AS, 3 in Hma-B and one in Hma-DT. Of the 71 mutations obtained, 2 transversions were observed (2.8%) (instead of G/C to T/A transitions expected with EMS treatment.)

	Total	Silent	Miss-sense	Non-sense	Intron	Splicing
Hma-AS	19	31,6	63,2	5,3	0,0	0,0
Hma-AT	21	22,2	57,1	0,0	9,5	0,0
Hma-B	20	20,0	70,0	0,0	5,0	5,0
Hma-DT	11	9,1	81,8	9,1	0,0	0,0
Hma-DS	3	33,3	66,7	0,0	0,0	0,0

Table 2: Percentage of mutation according to the target and the type of mutation.

EXAMPLE 4

Heavy Metals Translocation

Hydroponic culture

Selected M2 EMS lines were grown in soil in a greenhouse in order to obtain homozygous mutant seeds. DNA was extracted

from 2 month old plants for genotyping by CE-SSCP. Homozygous mutant lines were self-pollinated to obtain homozygous mutant seeds (M3). M3 seeds were germinated on Whatman paper soaked with a Hoagland-derived solution (KNO3 2,5mM ; NaH2PO 4 0,5 mM ; Ca(NO3)2 2,5mM ; MgSO4 0,5 mM ; FeNaEDTA 0,1 mM ; H3BO3 50 µM ; MnSO4 50 µM ; ZnSO4 15µM ; MoO4Na2 3µM ; KI 2,5µM ; CuSO4 50 µM ; CoCl2 44 µM). After 3 weeks, plants were transferred to the Hoagland-derived solution media were there growth is prolonged for 2 weeks. The solution media was then complemented with a final concentration of 10µM of CdCl2. After one week of treatments plants were cut in two parts: roots and shoots and were harvested independently.

Greenhouse testing

M2 lines were sown on Whatman paper until germination and 30 mutants plantlets were transferred in floating trays (20cm x 30cm), with five control plants randomly placed in the tray (non-mutagenized original lines). DNA was extracted from 1 month old plants for CE-SSCP genotyping to characterize homozygous, heterozygous and wild type plants for the HMA mutation. After one month and a half, the solution media was complemented with a final concentration of 1µM to 5µM of CdCl2. After one week of treatments, roots were rinsed with CaCl2 (1 mM) and plants were cut in three parts: roots, stem and shoots. Samples were pooled according to their genotype, with four plants per genotype.

Cadmium detection by ICP-MS

Samples were dried at 80°C for 72 hours and dry weight was measured. Metal contents were extracted in chloridric acid (1 mM) for 1 hour at 75°C.

Metals content in the chloridric acid were dosed by Inductively Coupled Plasma Mass Spectroscopy (ICP-MS) from Agilent (7500cx series). Ions are analyzed based on their mass to charge ratios (Newman, 1996). This technique has already been used for analysis of trace elements in many fields (e.g. Kelly *et al.*, 2002) and permit to detect cadmium traces as low as 10 ppt.

To obtain cadmium content the concentration of the dosed solution is multiplied by the dilution factor (this gives the amount of metal in the sample) and divided by the mass of the sample. To obtain translocation, the amount of metal in shoots is divided by the amount of metal in the whole plant (metal in roots plus metal in shoots).

Statistical analysis

For each line, mean value and standard deviation was calculated. The statistics software "R" was used to perform a student "T" test to compare those means (http://www.r-project.org/).

EXAMPLE 5

Backcrosses of mutation into elite lines

Mutants containing an interesting mutation were backcrossed with the original non-mutagenized plant in order to remove additional mutations present in the rest of the genome, not related with cadmium transfer. An example of a F1 cross of an industrial Flue-cured tobacco and E1-276-B18 mutant is shown in Figure 7. Mutants from the collection and elite lines were grown in greenhouse in floating beds. DNA of 30 plants per mutant family was extracted and analysed by CE-SSCP. Heterozygous plants were transferred in 5 litter's pots, along with elite lines. At flowering time, pollen of mutant was transferred on flower of the original line (BB16NN, BY02 or V4K1), cleared out of its stamen. Several cycles of backcrosses can be performed (BC1, BC2…) before fixing the mutation by two cycles of self-pollination of backcrossed mutant plants (BCxS1 and BCxS2).

EXAMPLE 6

Heavy Metals Translocation

Hydroponic culture

A. Mutation E200K

Experiments were conducted with the mutation E200K, in region B of the sequence, in the E1-276 line of BB16NN collection. Homozygous M2 mutant lines were self-pollinated to obtain homozygous mutant seeds (M3). Experiment was conducted on young M3 plantlets.

Results

Cadmium content in the homozygous M3 E1-276 mutant line is significantly reduced. Figure 6 describes the results obtained in terms of cadmium accumulation in roots and in shoots. In shoots cadmium content is reduced to different degrees and can be reduced by more than 50% in the lines of the present invention in comparison to wild type. As can be seen in Figure 6, cadmium content is reduced in shoots for example for more than 60% and in some plants even for more than 70%.

B. Mutation E3-277

Further experiments regarding the uptake of Cd and Zn were conducted with plants homozygous for the mutation E3-277.

Results

The results are shown in the Table 3, below and show that cadmium content in the homozygous E3-277 plants are significantly reduced, whereas the content of Zn is slightly reduced in comparison to control plants.

		Organ	Cd[µg/g d.w.b.]	Zn[µg/g d.w.b.]
E3-277-Wild type	WT-Plant1	Leaf	103	78,2
	WT-Plant1	Root	767	318
	WT-Plant2	Leaf	84,0	138
	WT-Plant2	Root	310	256
	WT-Plant3	Leaf	60,1	61,1
	WT-Plant3	Root	289	225
	WT-Plant4	Leaf	122	90,7
	WT-Plant4	Root	338,4	276
E3-277-Mutant	M-Plant1	Leaf	4,9	76,5
	M-Plant1	Root	33,3	53,7
	M-Plant2	Leaf	9,5	48,0
	M-Plant2	Root	66,7	625
BB16NN Control	Control-Plant1	Leaf	172,7	101
	Control-Plant1	Root	526	206
	Control-Plant2	Leaf	178	116
	Control-Plant2	Root	344	244
	Control-Plant3	Leaf	151	108
	Control-Plant3	Root	500	254
	Control-Plant4	Leaf	189	111
	Control-Plant4	Root	639	285

Table 3: Cadmium and Zinc content in homozygous E3-277 mutant line compared to wild-type line E3-277 and BB16NN Control line.

EXAMPLE 7

In this example cadmium uptake of shoots of *N. tabacum* plants was tested. These shoots are from EMS treated plant lines mutated either in the HMA 4 gene derived from *N. sylvestris* or in the HMA 4 gene derived from *N. tomentosiformis*.

Cadmium uptake of shoots of transgenic *N. tabacum* plants was also tested. These transgenic plant lines express an RNAi construct designed to silence either one or both of the two HMA 4 genes.

MATERIALS & METHODS

EMS lines

All EMS lines analyzed in this example are summarized in Figure 13. Among the tested lines, 3 are mutated in the HMA 4 gene derived from *N. sylvestris* (identified as alpha gene in Figure 13) and 2 carry mutations in the HMA 4 gene derived from *N. tomentosiformis* (identified as beta gene in Figure 13).

The lines 90 and 425 were backcrossed 2 times. Backcrossing reduced the amount of additional mutations by 75%.

The line 277 was backcrossed once but the mutation is at the heterozygous state for this line. All the other lines carry homozygous mutations. The other lines were not backcrossed.

All lines were grown in the presence of cadmium using the protocol described above, except that in this Example the cadmium concentration was 1µM instead of 10µM. Each line was cultivated in a bowl that also contain the corresponding wild type plant, *i.e.* a plant that carries the same set of additional mutations in its genome but that lacks the mutation in the corresponding *HMA* gene.

Transgenic lines

1) RNAi constructs

Five amiRNA constructs were made following the method described in Ossowski *et al*, 2008. An online resource (http://www.weigelworld.org/) was used to select the targeted sequences and design the corresponding primers. The result is presented in the Figure 14.

The amiRNAs were subcloned in a vector containing the "70S promoter" (35S promoter in which some enhancing regions are repeated) and a terminator (rbos). The promoter-amiRNA-terminator construct was inserted into a binary vector (pGreen) that permits expression *in planta*.

2) Hairpin constructs
The hairpin construct is designed to silence both of the *HMA* 4 genes. The targeted sequences as well as the primers used to amplify the same are described in Figure 15. The sequences were cloned into the pHannibal vector (Figure 16). pHannibal contains the appropriate restriction sites to clone the targeted sequences in both sense and antisense orientation.

The hairpin construct obtained was then subcloned twice and inserted into the pGreen vector with the same promoter and terminator used for the amiRNA constructs.

3) Plant transformation
Plants were transformed with the five amiRNA constructs, the hairpin construct and the empty pGreen vector as described in Horsch *et al*, 1984.

4) Growth conditions
The transgenic RNAi lines tested are the offspring of the plants regenerated after plant transformation.

RNAi lines were cultivated on the same Hoagland-derived media used for hydroponic culture but supplemented with 1% agar. The plants that lost their construct through segregation were eliminated by the addition of 200mg of hygromycin per liter of media. After two weeks of growth on selection media, plants were transferred to a media containing cadmium at a concentration of 1µM.

After two weeks of culture on the media containing cadmium, the plant material was collected. Approximately 1g of root was frozen in liquid nitrogen to analyze transcript accumulation by qPCR. The shoots were set aside to be analyzed for metal content by ICP-MS.

5) RNA extraction and cDNA synthesis

The roots were ground using a mortar and a pestle. RNA was obtained using the RNeasy kit (Qiagen) coupled with the DNase optional step. cDNA synthesis was performed with the M-MLV Reverse transcriptase (Promega) as described by the manufacturer.

6) Q-PCR analysis

The analyses were performed using a Roche 480 Lightcycler (Roche). The primer pairs used to amplify the alpha gene (Fw - ACAAAGTGCTCGGACACCAA ; Rev - CTTCTCGGTTGCAGAGTCCT) or the beta gene (Fw - ACAAAGTGCTCGGACACCAA ; Rev - CTTCTCGGTTGCAGAGTCTA) were designed to amplify specifically the targeted gene and not its homolog. This specificity and the efficiency of the primers were tested using a vector in which the region targeted by the primers was cloned.

The results were normalized using the cyclophilin gene (Fw - CTCTATGCCGACACCGTTCC ; Rev - TCACACGGTGGAAGGTTGAG) and the L2 gene (Fw - GGCGAAATGGGTCGTTTGATC ; Rev - CGTTCCGTTCGCCGAAGTCG). Those two reference genes were selected from the genes described in Nicot et al, 2005.

RESULTS

EMS lines

In 3 of the 5 tested lines, accumulation of cadmium in the shoots is reduced very significantly, namely by more than 50% (Figure 8 shows selected examples). These results were statistically validated for the lines 277 and 425 using the student t-test (Figure 9).

RNAi lines

1. Transcript accumulation in roots

The transcript accumulation measured by qPCR is presented in Figure 10. No diminution in the transcript level was observed for the hairpin line as compared to the line expressing the empty vector.

However, a diminution of accumulation of HMA 4 transcripts is observed for both genes in line 6a and for the HMA 4 gene derived from *N. tomentosiformis* in line 11b. The expression levels are in a different static group than the expression level of the wild type plants.

2. Metal accumulation in shoots
The accumulation of cadmium is reduced in all analyzed transgenic lines (Figure 11). It is reduced by more than 40% in lines 6b and hairpin. It is reduced by more than 60% in lines 6a and 11b. A student's t-test performed on those samples showed that the difference observed in lines 6a and 11b in comparison to the wild type plants is significant. No statistical differences between the different lines in terms of cadmium accumulation and zinc accumulation were determined.

CONCLUSION
This Example demonstrates that both *HMA* genes play a role in cadmium accumulation in shoots. Significant reduction of cadmium accumulation can already be achieved by silencing one of the *HMA* 4 genes. Further reduction of the cadmium accumulation can be obtained by silencing both genes.

EXAMPLE 8

In this Example the Cadmium accumulation in shoots of several plants was analyzed and compared. For this purpose, the lines 90, 416, 276 and 425 were backcrossed 2 times and fixed by two self-pollinations to obtain BC2S2 plants, selected for the mutation (M) or not (W). The plants identified in this Example as wild-type plants (or W) are thus plants that have also been subject to EMS treatment and carry the same set of additional mutations in their genome but lack the mutation in the corresponding HMA gene.

Seeds were first sterilized and sown directly on solid medium in vitro. Lines were grown on Hoagland-derived medium with

agar, containing 1µM of cadmium. The Hoagland-derived medium contains 0.1mM Fe and 15µM Zn.

Mutant lines (M), a heterozygous mutant line (Htz) and the corresponding wild-type (S) lines were analyzed. The experiments also included control plants BB16NN or BY02 (C) (wild-type, industrial seeds) and a tobacco RNAi line (obtained as described in Example 7).

Leaves were collected after one month and extracted for metals as previously described. Analyses were performed on 4 to 12 plants per genotype for Cd, Zn, Fe.

RESULTS

The results are shown in Figures 17 and 18 and confirm that these plants have a significantly reduced amount of cadmium in the shoots (Figure 17B). At the same time it could be shown that the plants are still able to take up metals such as Fe and Zn that are required for plant growth (Figure 18).

CONCLUSION

This Example demonstrates that the silencing of *HMA* genes in accordance with the present invention does not necessarily have a negative affect on the uptake of metals that are required for plant growth and development.

EXAMPLE 9

To develop commercial tobacco varieties with reduced cadmium content in leaf, crosses were performed between different tobacco mutants. In particular, plants with a mutation in the HMA 4 gene from *N. sylvestris* were crossed with plants with a mutation in the HMA 4 gene from *N. tomentosiformis*.

For this purpose both mutants were backcrossed with the elite lines to eliminate the mutation load in the rest of the genome, which may cause problems with fertility in the

progeny. The resulting BC1 plants were then crossed to get the F1 family, possessing two mutated copies at heterozygous state. All the F1 crosses are described in Figure 12.

The F1 plants resulting from the cross between mutants will be self-pollinated to obtain an F2 generation, in which homozygous plants for both mutated/or wild type copies will be present, including homozygous double mutant plants. These plants can be backcrossed into elite lines to obtain homozygous double mutant commercial plant lines.

REFERENCES

Lugon-Moulin N, Martin F, Marc R. Krauss. Ramey PB, Rossi L (2006) Cadmium concentration in tobacco (Nicotiana tabacum L.) from different countries and its relationship with other elements. Chemosphere Vol 63 1074-1086.

Karaivazoglou NA, Tsotsolis NC and Tsadilas CD (2007) Influence of liming and form of nitrogen fertilizer on nutrient uptake, growth, yield, and quality of Virginia (flue-cured) tobacco. Field Crops Research Vol 100, Issue 1:52-60.

Verbruggen N, Hermans C, and Schat H (2009) Mechanisms to cope with arsenic or cadmium excess in plants. Current Opinion in Plant Biology 12:364-372.

Yeargan R, Maiti IB, Nielsen MT, Hunt AG, Wagner GJ (1992) Tissue partitioning of cadmium in transgenic tobacco seedlings and field grown plants expressing the mouse metallothionein I gene. Transgenic Res. Nov 1 (6):261-7.

Dorlhac De Borne F, Elmayan T, De Roton C, De Hys L, Tepfer M (1998) Cadmium partitioning in transgenic tobacco plants expressing a mammalian metallothionein gene. Molecular Breeding 4 :83-90.

Korenkov V, King B, Hirschi K, Wagner GJ (2009) Root-selective expression of AtCAX4 and AtCAX2 results in reduced lamina cadmium in field-grown Nicotiana tabacum L. Plant Biotechnol J. 7 (3):219-26.

Cobbett CS, Hussain D, Haydon MJ (2003) Structural and functional relationships between type 1B heavy metal-transporting P-type ATPases in Arabidopsis. New Phytologist 159:315-321.

Seigneurin-Berny D, Gravot A, Auroy P, Mazard C, Kraut A, Finazzi G, Grunwald D, Rappaport F, Vavasseur A, Joyard J, Richaud P, Rolland N (2006) HMA1, a new Cu-ATPase of the chloroplast envelope, is essential for growth under adverse light conditions. The journal of biological chemistral 281:2882-2892.

Hussain D, Haydon M, Wang Y, Wong E, Sherson S, Young J, Camakaris J, Harper J, Cobbett C (2004) P-Type ATPase heavy metal transporters with roles in essential zinc homeostasis in Arabidopsis. The Plant Cell 16 :1327-1339.

Morel M, Crouzet J, Gravot A, Auroy P, Leonhardt N, Vavasseur A, Richaud P (2009) AtHMA3, a P1B-ATPase Allowing Cd/Zn/Co/Pb Vacuolar Storage in Arabidopsis. Plant physiology 149:894-904.

HAYES, Alec et al., TRANSGENIC PLANTS MODIFIED FOR REDUCED CADMIUM TRANSPORT, DERIVATIVE PRODUCTS, AND RELATED METHODS, WO2009/074325.

Julio E, Laporte F, Reis S, Rothan C and Dorlhac de Borne F (2008) Reducing the content of nornicotine in tobacco via targeted mutation breeding. Molecular Breeding, 21:1380-3743.

Yukawa M, Tsudzuki T, Sugiura M (2006) The chloroplast genome of Nicotiana sylvestris and Nicotiana tomentosiformis: complete sequencing confirms that the Nicotiana sylvestris progenitor is the maternal genome donor of Nicotiana tabacum. Molecular genetics and genomics 275:367-373.

Delon R, Poisson C, Bardon JC, Taillurat P (1999) Les nicotianées en collection à l'Institut du Tabac. SEITA, 86p.

Till BJ, Reynolds SH, Greene EA et al (2003) Large-scale discovery of induced point mutations with high-throughput TILLING. Genome Res. 13:524-530.

Ng PC and Henikoff S (2003) SIFT: predicting amino-acid changes that affect protein function. Nucleic Acid Res. 31:3812-3814.

Newman, Alan. 1996. "Elements of ICPMS." Analytical Chemistry. 68(Jan 1):46A-51A.

Kelly S, Baxter M, Chapman S, Rhodes C, Dennis J, and Brereton P (2002) The application of isotopic and elemental analysis to determine the geographical origin of premium long grain rice. European food research & technology, 214:72-82.

Ossowski et al., 2008, "Gene silencing in plants using artificial microRNAs and other small RNAs" The Plant Journal, February 2008, vol. 53(4): 674–690.

Horsch et al., 1984, "Inheritance of functional foreign genes in plants" Science 1984 Feb 3; 223(4635):496-8.

Nicot et al., "Housekeeping gene selection for real-time RT-PCR normalization in potato during biotic and abiotic stress", J Exp Bot. 2005 Nov; 56(421):2907-14.

Claims

1. A tobacco plant comprising at least one mutation in a HMA gene, wherein the non-mutated HMA gene comprises the nucleotide sequence of SEQ ID NO:1 or a homolog sequence with at least 90% identity to the sequence of SEQ ID NO:1, wherein the mutation causes a substitution or a deletion or an insertion of at least one amino acid in the polypeptide encoded by the nucleotide sequence and wherein the mutation reduces the heavy metal uptake by the leaves of the plant by at least 30% in relation to the heavy metal uptake of plants comprising SEQ ID NO:1 or the homolog.

2. The tobacco plant according to claim 1, wherein the tobacco plant is a *Nicotiana tabacum* plant.

3. The tobacco plant according to claim 1 or 2, wherein the reduction of heavy metal uptake is determined by growing tobacco plants with and without the mutation under identical conditions on a liquid medium containing the heavy metal and comparing the concentration of the heavy metal in the leaf, stem or shoot of the tobacco plant with the mutation to the concentration of the heavy metal in the leaf, stem or shoot of the tobacco plant that does not have the mutation.

4. The tobacco plant according to one of claims 1 to 3, wherein the mutation reduces the uptake of the heavy metal by at least 40%, at least 50%, at least 75% or at least 95%.

5. The tobacco plant according to one of claims 1 to 4, wherein the mutation reduces the uptake of cadmium, lead or arsenic.

6. The tobacco plant according to one of claims 1 to 5, wherein the mutation is a miss-sense, a non-sense mutation or a splice mutation.

7. The tobacco plant according to claim 6, wherein the mutation is selected from the group comprising the mutations: G294A, C576T, G406A, G347A, G363A, G553A, G290A, C374T, G964A, G1168A, G1211A, G1126A, C980T, G1195T, G1156A, G1070A, C2302T, G2208A, C2217T, G2190A, C2206T or C2277T in SEQ ID NO:1 or the homolog.

8. The tobacco plant according to one of claims 1 to 7, wherein the plant comprises a mutation in more than one HMA 4 gene.

9. The tobacco plant according to one of claims 1 to 8, wherein the plant is homozygous for the mutation in the HMA 4 gene.

10. The tobacco plant according to one of claims 1 to 9, wherein the plant is homozygous for mutations in both HMA 4 genes.

11. A tobacco plant cell derived from a tobacco plant according to one of claims 1 to 10.

12. Part of a *tobacco* plant according to one of claims 1 to 10, wherein the part is a leaf, a lamina, a cut and/or a cured leaf, root, shoot, stem, flower or seed.

13. The part of a *tobacco* plant according to claim 12, wherein the tobacco plant is a *Nicotiana tabacum* plant.

14. Tobacco product comprising a tobacco plant according to one of claims 1 to 10 or a part of a tobacco plant according to claim 12 or 13.

15. Tobacco product according to claim 14, wherein the tobacco product is a cut tobacco, tobacco extract or reconstituted tobacco or a smokeless tobacco product, like snus or snuff.

16. Smoking article comprising tobacco plants according to one of claims 1 to 10 or parts of a tobacco plant according to claims 12 to 13.

17. Smoking article according to claim 16, which is a cigarette, a small cigar, cigarillo, a cigar or a simulated smoking article containing tobacco.

18. Method for generating a tobacco plant comprising:

(a) screening a library of tobacco plants obtained by mutagenesis for at least one mutation in a HMA gene, wherein the non-mutated HMA gene comprises the nucleotide sequence of SEQ ID NO:1 or a homolog thereof, wherein the mutation causes a substitution or a deletion or an insertion of at least one amino acid in the polypeptide encoded by the nucleotide sequence and wherein the mutation reduces the heavy metal uptake by the leaves of the tobacco plant by at least 30% in relation to the heavy metal uptake of plants comprising SEQ ID NO:1 or the homolog;
(b) crossing the tobacco plant having a mutation in the HMA gene with a commercial *Nicotiana tabacum* production plant; and
(c) identifying offspring with the mutation in the HMA gene;
(d) repeating steps (b) and (c).

```
                    250       260       270       280       290       300
                    |         |         |         |         |         |
Contig              ACTACAAATCTAAATGGTAGTATAGTATTTCTTCATGCTTCATTTATTTAGTGCTGAAAC
N.sylvestris_seq1   ACTACAAATCTAAATGGTAGTATAGTATTTCTTCATGCTTCATTTATTTAGTGCTGAAAC
N.sylvestris_seq2   ACTACAAATCTAAATGGTAGTATAGTATTTCTTCATGCTTCATTTATTTAGTGCTGAAAC
N.sylvestris_seq3   ACTACAAATCTAAATGGTAGTATAGTATTTCTTCATGCTTCATTTATTTAGTGCTGAAAC
N.sylvestris_seq4   ACTACAAATCTAAATGGTAGTGTAGTATTCCTTCATGCTTCATTTATTTAGGGCCGAATC
N.sylvestris_seq5   ACTACAAATCTAAATGGTAGTGTAGTATTCCTTCATGCTTCATTTATTTAGGGCCGAATC
N.tabacum_seq6      ACTACAAATCTAAATGGTAGTGTAGTATTCCTTCATGCTTCATTTATTTAGTGCTGAAAC
N.tabacum_seq7      ACTACAAATCTAAATGGTAGTGTAGTATTCCTTCATGCTTCATTTATTTAGTGCTGAAAC
N.tabacum_seq8      ACTACAAATCTAAATGGTAGTGTAGTATTCCTTCATGCTTCATTTATTTAGTGCTGAAAC
N.tabacum_seq9      ACTACAAATCTAAATGGTAGTATAGTATTTCTTCATGCTTCATTTATTTAGTGCTGAAAC
N.tabacum_seq10     ACTACAAATCTAAATGGTAGTATAGTATTTCTTCATGCTTCATTTATTTAGTGCTGAAAC
N.tabacum_seq11     ACTACAAATCTAAATGGTAGTATAGTATTTCTTCATGCTTCATTTATTTAGTGCTGAAAC
N.tabacum_seq12     ACTACAAATCTAAATGGTAGTGTAGTATTCCTTCATGCTTCATTTATTTAGTGCTGAAAC
N.tabacum_seq13     ACTACAAATCTAAATGGTAGTGTAGTATTCCTTCATGCTTCATTTATTTAGTGCTGAAAC
N.tomentosiformis_seq14 ACTACAAATCTAAATGGTAGTGTAGTATTCCTTCATGCTTCATTTATTTAGTGCTGAAAC
N.tomentosiformis_seq15 ACTACAAATCTAAATGGTAGTGTAGTATTCCTTCATGCTTCATTTATTTAGTGCTGAAAC
N.tomentosiformis_seq16 ACTACAAATCTAAATGGTAGTGTAGTATTCCTTCATGCTTCATTTATTTAGTGCTGAAAC
```

Figure 1

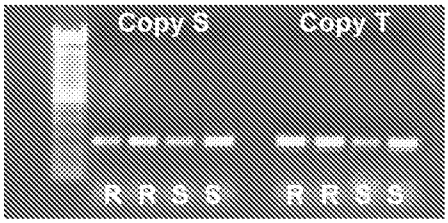
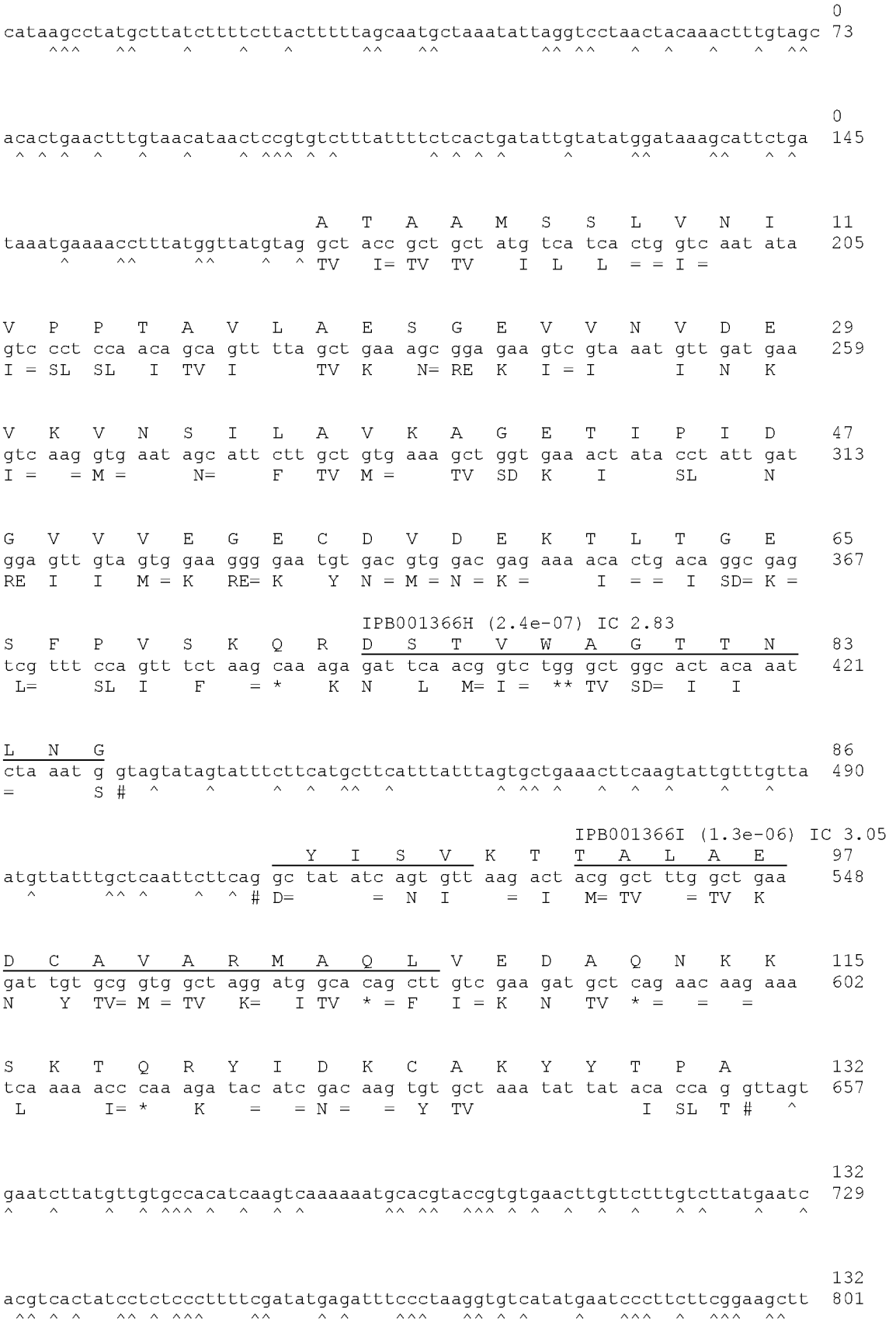


Figure 2


IPB001366H (2.4e-07) IC 2.83
IPB001366I (1.3e-06) IC 3.05

```
                                                                                                                                132
gccagcatagaagtctatcagtcctttcacttgacccgccctctcagcctgcctgcagtcatgggcgtcgca                                                        873
^^^ ^^   ^  ^ ^   ^ ^ ^^   ^ ^  ^ ^^^^^^^ ^ ^ ^^^ ^^^ ^^ ^ ^  ^^^^^ ^^^


                                                                                                                                132
ctactatattgctctttcgtttaaaactttttatttctaatactttcctgctctttgtgtatgtctgatttc                                                        945
^  ^      ^^ ^   ^^       ^         ^     ^   ^^ ^^ ^   ^ ^   ^ ^ ^    ^


                                    IPB001756J (2.2e-05) IC 2.17
                        I   V   A   I   S   A   S   L   A   I   V   P                                                         144
gactggtgatgttttgcag ca att gtg gct ata tca gct tct ttg gca att gtt cct                                                         1002
^ ^ ^^ ^  ^    ^^ # V      M = TV       L  TV   F    = TV      I   SL


                                IPB001756C (6.0e-05) IC 1.86
T   A   L   R   V   H   N   R   N   E   W   Y   R   L   A   L   V   T                                                         162
act gca tta aga gtt cac aat cga aat gaa tgg tat cgc ttg gct ttg gtc aca                                                         1056
 I  TV       K  I   Y =     *Q      K    **     CH=   = TV    = I =  I


L   V   S   A   C   P   C   A   L   V   L   S   T   P   V   A   M   C                                                         180
ttg gtg agt gca tgt ccg tgt gca ctt gtt cta tct aca cca gtt gcc atg tgt                                                         1110
  = M =  N  TV   Y  SL=  Y  TV  F   I   =    F   I  SL  I   TV=   I  Y


                                                IPB001756H (1.8e-11) IC 2.12
C   A   L   S   K   A   A   T   S   G   L   L   F   K   G   A   E   Y                                                         198
tgc gca ctt tca aaa gca gca acg tcc ggt ctt ctg ttt aaa gga gca gag tac                                                         1164
 Y= TV  F    L      TV  TV   M=  F= SD  F   = =         RE  TV  K =   =


                                        IPB006121B (5.0e-09) IC 0.70
L   E   T   L   A   K   I   K   I   M   A   F   D   K   T   G   T   I                                                         216
ctt gag act cta gct aaa atc aaa atc atg gct ttt gac aaa aca ggg act ata                                                         1218
F   K =  I  =   TV        =       =   I TV      N =      I  RE=  I


T   K   G   E   F   M   V   T   E   F   K   S   L   I   D   G   F   S                                                         234
act aaa gga gaa ttt atg gtg acc gag ttc aag tct ctg att gat ggt ttt agt                                                         1272
 I      RE  K         I M =  I= K =   =   =  F  = =     N   SD       N


L   N   T   L   L   Y   W                                                                                                      241
ctc aat aca ctg ctt tac tg gtaaaggttaccactcatacatattcttttatgttgccaaagaga                                                        1337
F =      I  = = F     =  * #    ^^   ^^ ^ ^   ^     ^      ^  ^^^   ^ ^


                                                                                                                                241
attcaaaatcttaactgattatctttcacggcacattgatagtgatataacatgattgatttatatcatata                                                        1409
   ^     ^    ^ ^     ^   ^ ^^^^ ^   ^   ^ ^      ^  ^   ^        ^


                                                                                                                                241
ttcataaaagatgaaatagggagtgccacattcacattctcatattgaagtttctgaaatggctctaatggt                                                        1481
  ^      ^  ^     ^^^ ^ ^^^ ^   ^ ^   ^ ^     ^  ^   ^ ^    ^^^ ^    ^^


                                                                                                                                241
tcaccatagagccaaaataacatatagacacaacgtcagccgtctgatattcagaacttagatggaatagtt                                                        1553
 ^ ^^   ^ ^^^       ^     ^ ^ ^  ^^ ^ ^^^^ ^ ^     ^ ^  ^   ^  ^^    ^


                                                                                                                                241
ggatcttatacattgaggacacataaaagtacttggtcatataaattttagaaacataatcaatgtattata                                                        1625
^^  ^     ^   ^ ^^ ^ ^      ^  ^  ^^ ^            ^   ^     ^   ^


                                                                                                                                241
atctaaaattcttcaaatattcttgatactgcaataacaaaagcacatgcacactgaatagaagccttgttc                                                        1697
  ^       ^  ^       ^  ^   ^ ^^     ^    ^^ ^  ^^ ^ ^ ^    ^  ^^^  ^  ^
```

```
                                                                           241
gtgtctaaaacatcgttgtggtctaaaacattcgtttagagtaaatactgaattgtctagtgaatattttca  1769
^ ^ ^     ^  ^^  ^ ^^ ^     ^   ^^    ^ ^      ^ ^    ^ ^  ^ ^        ^

    V   S   S   I   E   S   K   S   G   H   P   M   A   A   A   L   V      258
g g gtt tca agc att gag agc aag tca ggt cat ccg atg gca gcc gct ctg gtg   1822
# * I    L   N=      K =  N=    =  L  SD  Y   SL=   I TV  TV= TV   = = M =

D   Y   A   Q   S   N   S   V   E   P   K   P   D   R   V   E   Q   F      276
gac tat gca caa tca aat tcc gtt gag cca aag cct gat aga gtt gag cag ttt   1876
N =     TV  *    L       F= I   K = SL    = SL  N    K  I   K = * =

Q   N   F   P   G   E   G   I   F   G   R   I   D   G   M   E   I   Y      294
caa aat ttt cct ggt gaa ggg ata ttt gga aga att gat gga atg gaa atc tat   1930
*           SL  SD  K   RE=         RE   K      N   RE    I K     =

V   G   N   R   K   I   S   S   R   A   G   C   T   T   V                  309
gtc ggg aat agg aaa att tct tca aga gct gga tgt acc aca g gtaaatggttgaat  1987
I = RE=      K=          F   L   K  TV  RE   Y   I=  I  I #     ^^  ^

                                                                           309
catttcttatgctcatagtagagataaaacatcagagttataattataagtatatgatttctccagttaatt  2059
^    ^    ^^ ^   ^  ^ ^      ^  ^ ^ ^            ^     ^    ^ ^^ ^

                                                               P   E      311
ttgctgttagatttctttgacctgtttagcactaatgcggtggatgtttgaatttcag ta cca gaa     2126
  ^^ ^   ^      ^   ^ ^^ ^    ^^ ^    ^^^^ ^^  ^   ^      ^ #    SL  K

I   E   G   D   S   F   K   G   K   S   V   G   Y   I   F   L   G   S      329
ata gag ggt gat agt ttc aaa gga aag tct gtt gga tac ata ttt ttg gga tca   2180
    K = SD  N    N     =     RE    =  F  I   RE    =            = RE   L

                                                                   IPB001756I
(3.1e-05) IC 2.08
S   P   A   G   I   F   S   L   S   D   V   C   R   I   G   V   K   E      347
tct cca gct gga att ttc agt ctt tcc gat gtt tgt cga att ggt gta aaa gaa   2234
 F  SL  TV  RE        =  N  F    F= N   I    Y  *Q      SD  I       K

A   M   R   E   L   K   Q   M   G   I   K   T   A   M   L   T   G   D      365
gca atg aga gaa ctg aag cag atg ggt atc aaa acc gcg atg ctt act ggt gat   2288
TV    I  K  K   = =   = * =   I SD    =       I= TV=   I F    I  SD  N

C   Y   A   A   A   N   H   V   Q   D   Q                                  376
tgt tat gca gct gcc aac cat gtg cag gat cag gtatattaataattctgcattacgctga  2349
 Y      TV  TV  TV=   = Y   M = * = N    * = #              ^ ^^    ^^^ ^

                                                                           376
aatgattataaaaccctttggattattgtttagtcttaagaattttcactgaactcttattgtttccttctt  2421
   ^         ^^^   ^^      ^    ^ ^    ^      ^ ^ ^  ^ ^      ^   ^^  ^

                                                                           376
ctatcatcaacattggttaaacatttcatctaaatttagagaacgtattaccaagtaagtgctttatcttta  2493
^    ^  ^  ^   ^^      ^     ^  ^         ^ ^  ^^      ^^  ^    ^ ^^      ^
```

```
                                                                                                    376
cagggtcatataaaatacttaagacagtgtgatgtgaagatgaaggttaaatgttgatctggataaaccaag                           2565
^ ^^^ ^          ^    ^ ^ ^ ^ ^  ^ ^  ^  ^  ^^      ^  ^  ^ ^^     ^^  ^


                                                                                                    376
ttattatcacaactaatataagatatgctattgttctccaataattggacgattttcggacgtacgacgtac                           2637
       ^ ^  ^        ^    ^^    ^  ^ ^^       ^^ ^^     ^^^ ^^  ^^ ^^  ^


                                                                                                    376
aattcttcacatatgaaacctacatcagacgtacatgacacgctatgtttagcataaagagtcaagattagc                           2709
    ^  ^ ^    ^   ^^  ^  ^ ^ ^^  ^  ^ ^ ^^^   ^    ^^     ^ ^ ^  ^    ^^


                                                                                                    376
atgatgatttaagctgaatctgaatttcaagtatctattcttgtattgtacccaggggcggaactagtgttg                           2781
  ^  ^      ^^ ^   ^ ^     ^  ^   ^    ^  ^    ^  ^^^ ^^^^^^^  ^  ^ ^  ^


                                                                                                    376
tgcttagaggtctcaaacattgtatttgtgttaaaaaattcacttcatatgtatttaaataatttatccaga                           2853
 ^^   ^ ^^ ^ ^   ^   ^     ^ ^          ^ ^  ^    ^                ^^ ^


                                                                                                    376
gcagtgagccatatttttagaatccagaacccataaactcaaaatcatagatccacctctgattgtaagtc                           2925
^^ ^ ^ ^^^          ^   ^^ ^  ^^^     ^ ^     ^   ^  ^^ ^^ ^ ^   ^   ^ ^


                                                                    IPB001756I (3.9e-09) IC
2.08
                L   G   G   A   L   D   E   F   Q   A   E   L   L   P                           390
gaaacaattatgcag tta ggt gga gct ttg gat gaa ttt caa gca gaa ctc cta cca                           2982
^   ^      ^^ #     SD  RE  TV    = N   K       *   TV  K   F = =   SL


E   D   K   A   T   I   I   K   G   F   Q   K   E   A   P   T                                     406
gag gac aag gca aca atc atc aag ggt ttt cag aag gaa gct cca aca gc                                 3032
K = N =   = TV   I    =   =   = SD      * =   = K   TV  SL   I
```

Figure 3

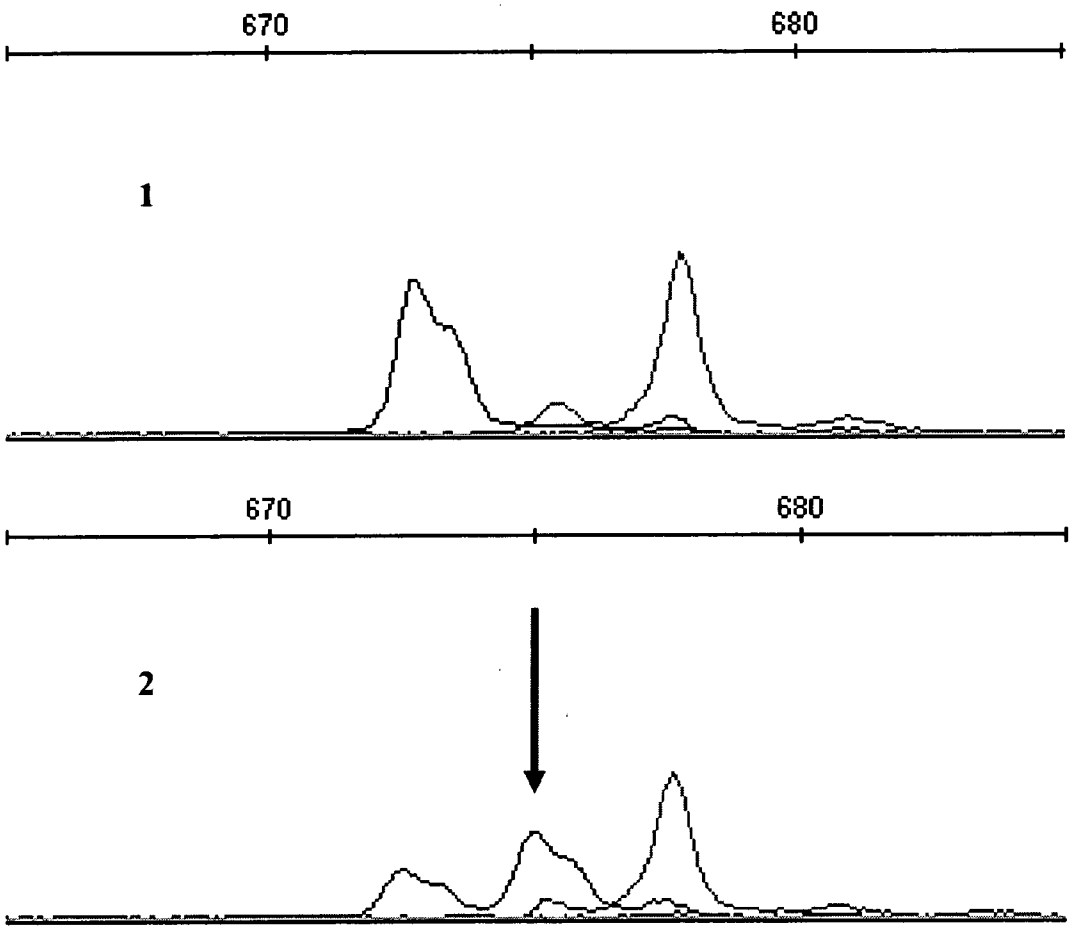


Figure 4

Figure 5/1

Target	Collection	Identification[a]	Nucleotide change[b]	Amino-acid change[b]	SIFT score[c]
Hma-AS	BB16NN	E1-110	C515T	G86=	1
Hma-AS	BB16NN	E1-397	G294A	G41D	0
Hma-AS	BB16NN	E2-90	C576T	L107F	0
Hma-AS	BB16NN	E2-1	C620T	Y121=	1
Hma-AS	BB16NN	E2-425	G406A	W78*	0
Hma-AS	BB16NN	E2-229	G290A	A40T	0,02
Hma-AS	BB16NN	E3-32	G579A	V108I	0,01
Hma-AS	BB16NN	E3-1071	C356T	L62=	1
Hma-AS	BB16NN	E3-176	G401A	V77I	0,01
Hma-AS	BB16NN	E3-217	C515T	G86=	1
Hma-AS	BB16NN	E3-271	G347A	E59K	0
Hma-AS	BB16NN	E3-422	G530A	K91=	1
Hma-AS	BB16NN	E3-309	G363A	G64D	0
Hma-AS	BB16NN	E3-372	G616A	R120K	0,09
Hma-AS	BB16NN	E3-1144	G579A	V108I	0,01
Hma-AS	BY02	450	G553A	C99Y	0
Hma-AS	BY02	695	G523A	S89N	0,03
Hma-AS	BY02	793	G582A	E109K	0,02
Hma-AS	BY02	802	C364T	G64=	1
Hma-AT	BB16NN	E1-9	G434A	intron	-
Hma-AT	BB16NN	E1-314	C556T	A100V	0,3
Hma-AT	BB16NN	E2-357	G290A	A40T	0
Hma-AT	BB16NN	E2-454	G570A	A105T	0,04
Hma-AT	BB16NN	E2-24	G370A	S66=	1
Hma-AT	BB16NN	E2-453	G542A	L95=	1
Hma-AT	BB16NN	E3-1211	G370A	S66=	1
Hma-AT	BB16NN	E3-1315	G555A	A100T	0,4
Hma-AT	BB16NN	E3-888	G537A	A94T	0,02
Hma-AT	BB16NN	E3-1038	C346T	D58=	1
Hma-AT	BY02	390	C374T	P68S	0
Hma-AT	BY02	493	C508T	intron	-
Hma-AT	BY02	769	G290A	A40T	0
Hma-AT	BY02	874	C571T	A105V	0,62
Hma-AT	BY02	1336	C576T	L107F	0
Hma-AT	V4K1	273	G367A	E65=	1
Hma-AT	V4K1	751	G555A	A100T	0,4
Hma-AT	V4K1	303	G546A	E97K	0,23
Hma-B	BB16NN	E1-84	C1001A*	P144H	0,06
Hma-B	BB16NN	E1-276	G1168A	E200K	0
Hma-B	BB16NN	E1-508	G1211A	G214E	0

Target	Collection	Identification[a]	Nucleotide change[b]	Amino-acid change[b]	SIFT score[c]
Hma-B	BB16NN	E1-521	G1168A	E200K	0
Hma-B	BB16NN	E2-204	C983T	A138V	0,01
Hma-B	BB16NN	E2-341	G1126A	A186T	0
Hma-B	BB16NN	E2-289	G1078A	A170T	0,03
Hma-B	BB16NN	E3-163	C980T	S137L	0
Hma-B	BY02	71	G1271A	S234N	0,08
Hma-B	BY02	320	G1170A	E200=	1
Hma-B	BY02	355	C983T	A138V	0,26
Hma-B	BY02	399	G1126A	A186T	0
Hma-B	BY02	505	C1203T	D211=	1
Hma-B	BY02	772	G1195T*	A210S	0
Hma-B	BY02	805	G1161A	E197=	1
Hma-B	BY02	936	G1156A	A196T	0
Hma-B	BY02	1024	G1062A	V164=	1
Hma-B	BY02	1445	G1070A	C167Y	0
Hma-B	BY02	1468	G964A	splicing	-
Hma-B	V4K1	767	C962T	intron	-
Hma-DT	BB16NN	E1-7	C2272T	A360V	0,4
Hma-DT	BB16NN	E3-640	C2302T	A370V	0
Hma-DT	BB16NN	E3-1370	G2190A	G333R	0
Hma-DT	BB16NN	E3-259	G2208A	D339N	0
Hma-DT	BB16NN	E3-277	C2217T	R342*	0
Hma-DT	BB16NN	E3-319	G2190A	G333R	0
Hma-DT	BB16NN	E3-1155	C2207T	S338=	1
Hma-DT	BB16NN	E3-1274	C2184T	P331S	0,39
Hma-DT	BB16NN	E3-1275	C2206T	S338F	0
Hma-DT	BY02	506	G2271A	A360T	0,03
Hma-DT	BY02	835	G2276A	G328E	0,14
Hma-DS	BB16NN	E3-906	G2242A	R350K	0,21
Hma-DS	BY02	416	C2277T	L362F	0
Hma-DS	BY02	560	G2252A	K353=	1

Figure 5/2

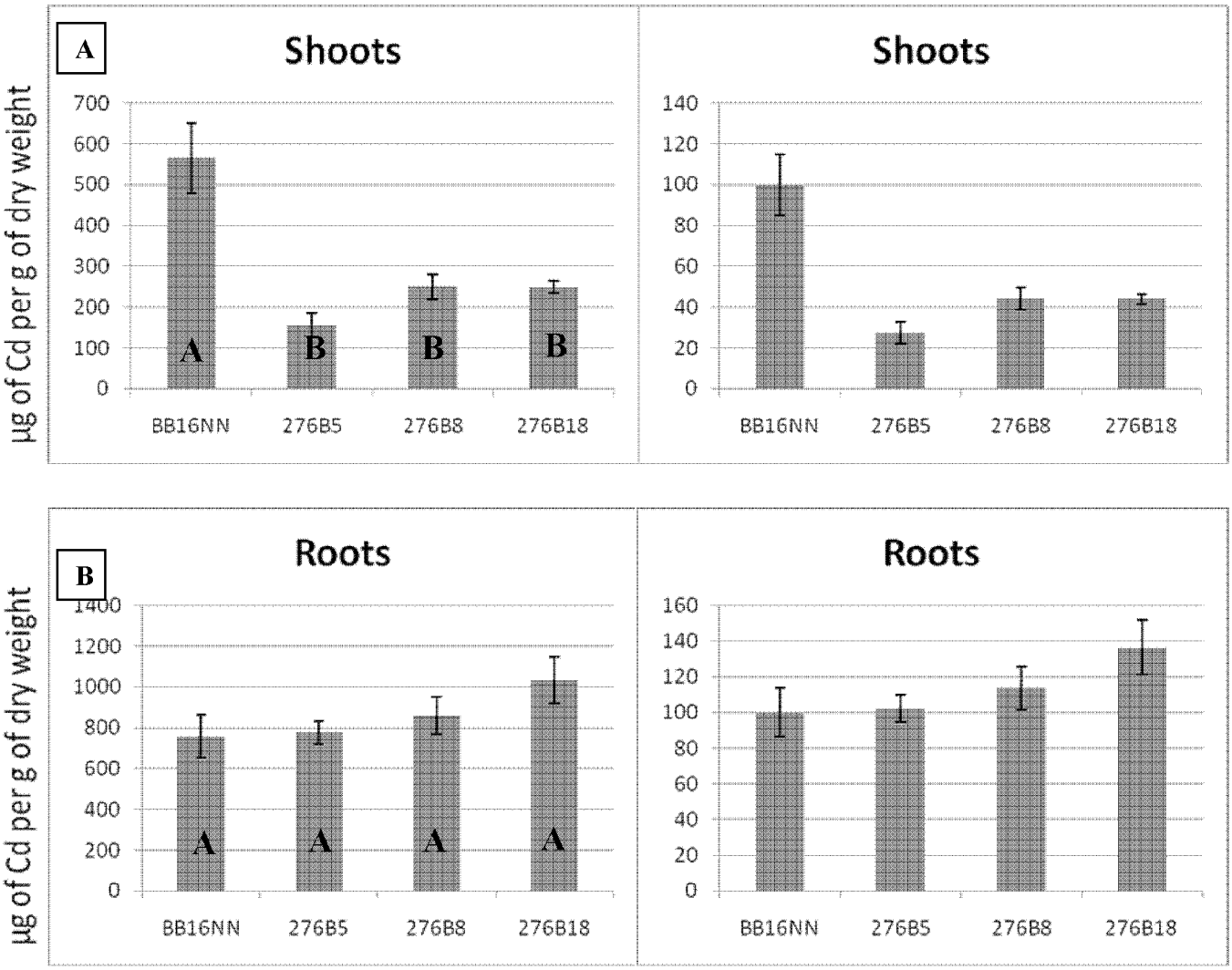

A
Shoots
µg of Cd per g of dry weight
700
600
500
400
300
200
100
0
A
B
B
B
BB16NN
276B5
276B8
276B18
Shoots
140
120
100
80
60
40
20
0
BB16NN
276B5
276B8
276B18
B
Roots
µg of Cd per g of dry weight
1400
1200
1000
800
600
400
200
0
A
A
A
A
BB16NN
276B5
276B8
276B18
Roots
160
140
120
100
80
60
40
20
0
BB16NN
276B5
276B8
276B18

Figure 6

Figure 7

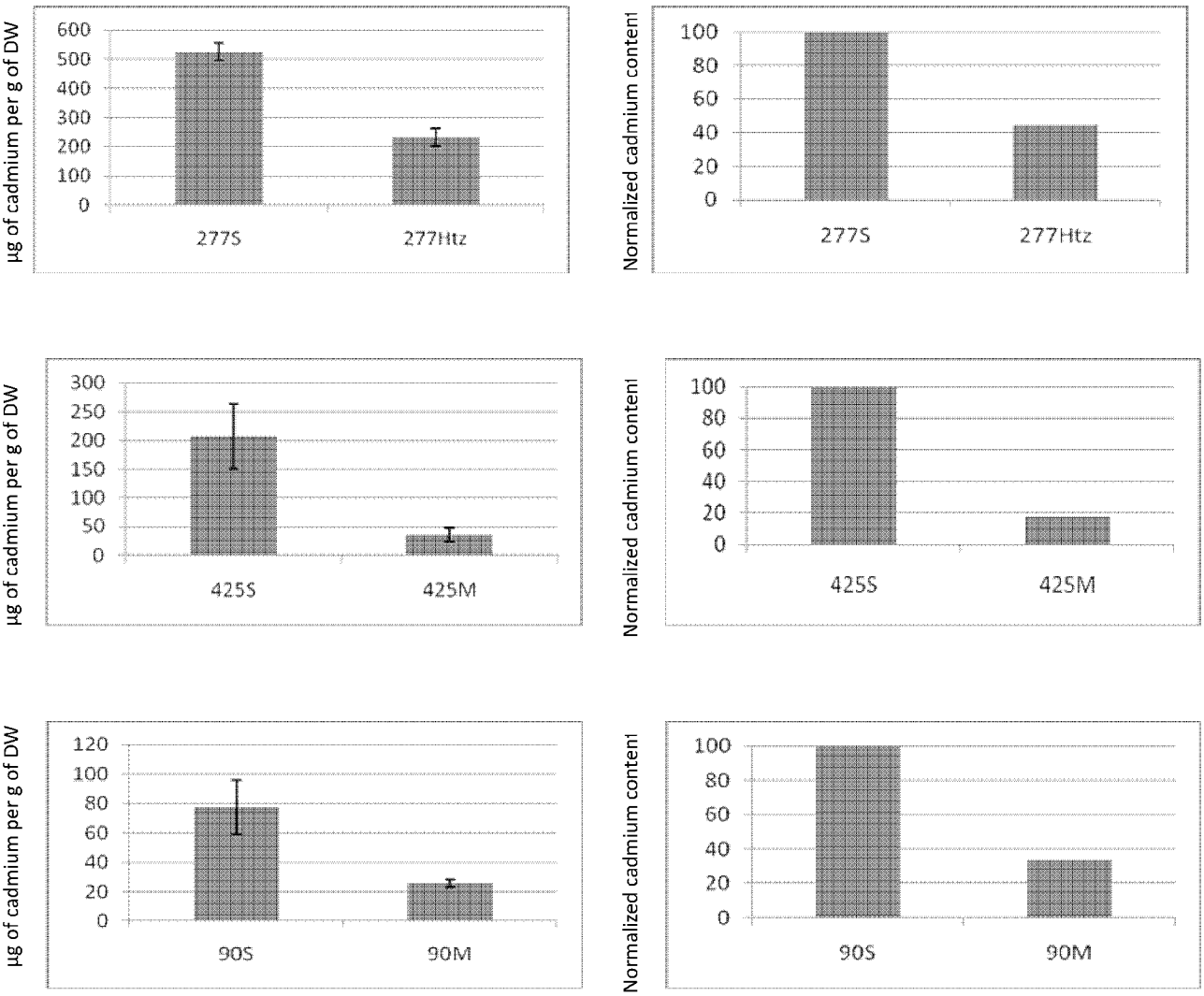

µg of cadmium per g of DW
600
500
400
300
200
100
0
277S
277Htz
Normalized cadmium content
100
80
60
40
20
0
277S
277Htz
µg of cadmium per g of DW
300
250
200
150
100
50
0
425S
425M
Normalized cadmium content
100
80
60
40
20
0
425S
425M
µg of cadmium per g of DW
120
100
80
60
40
20
0
90S
90M
Normalized cadmium content
100
80
60
40
20
0
90S
90M

Figure 8

	Mean cadmium accumulation (µg/g)	% of wild type	Effective	p-value	Significative difference
277S	525,0	***	2	0.0042	Yes
277Htz	232,6	44,3	4		
425S	207,0	***	4	0.0247	Yes
425M	36,1	17,4	4		
90S	77,2	***	3	0.0513	No
90M	25,7	33,3	3		

Figure 9

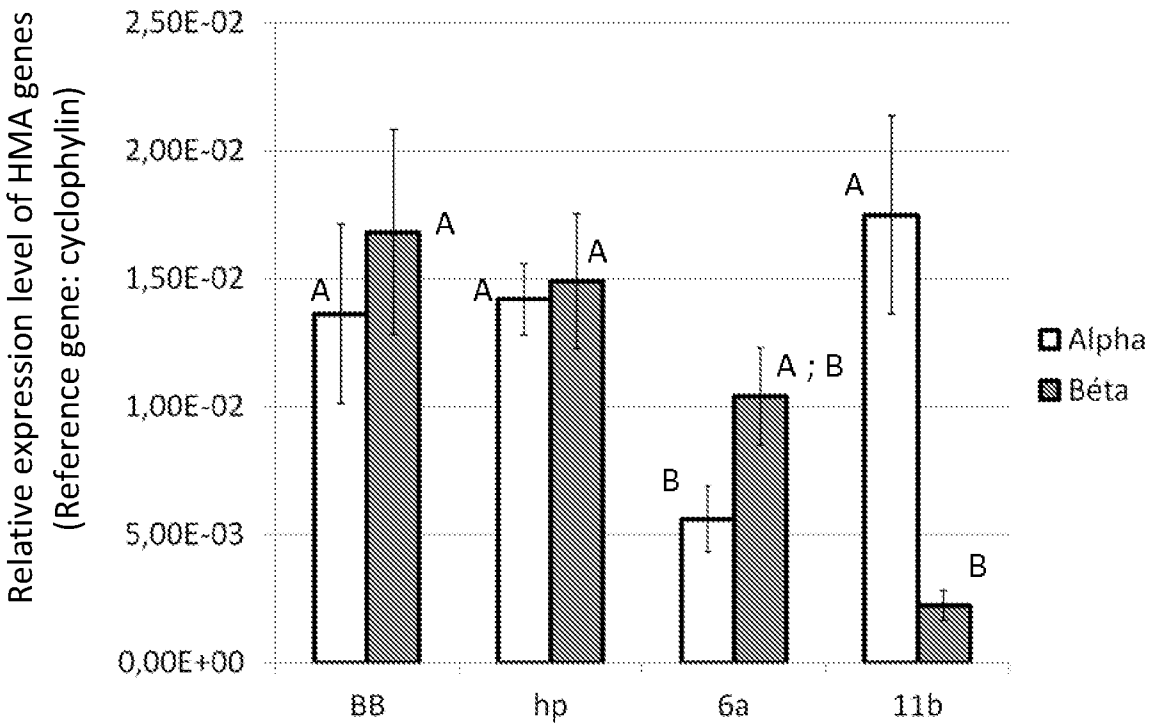

Relative expression level of HMA genes
(Reference gene: cyclophylin)
2,50E-02
2,00E-02
1,50E-02
1,00E-02
5,00E-03
0,00E+00
A
A
A
A
B
A ; B
A
B
BB
hp
6a
11b
Alpha
Béta

Figure 10

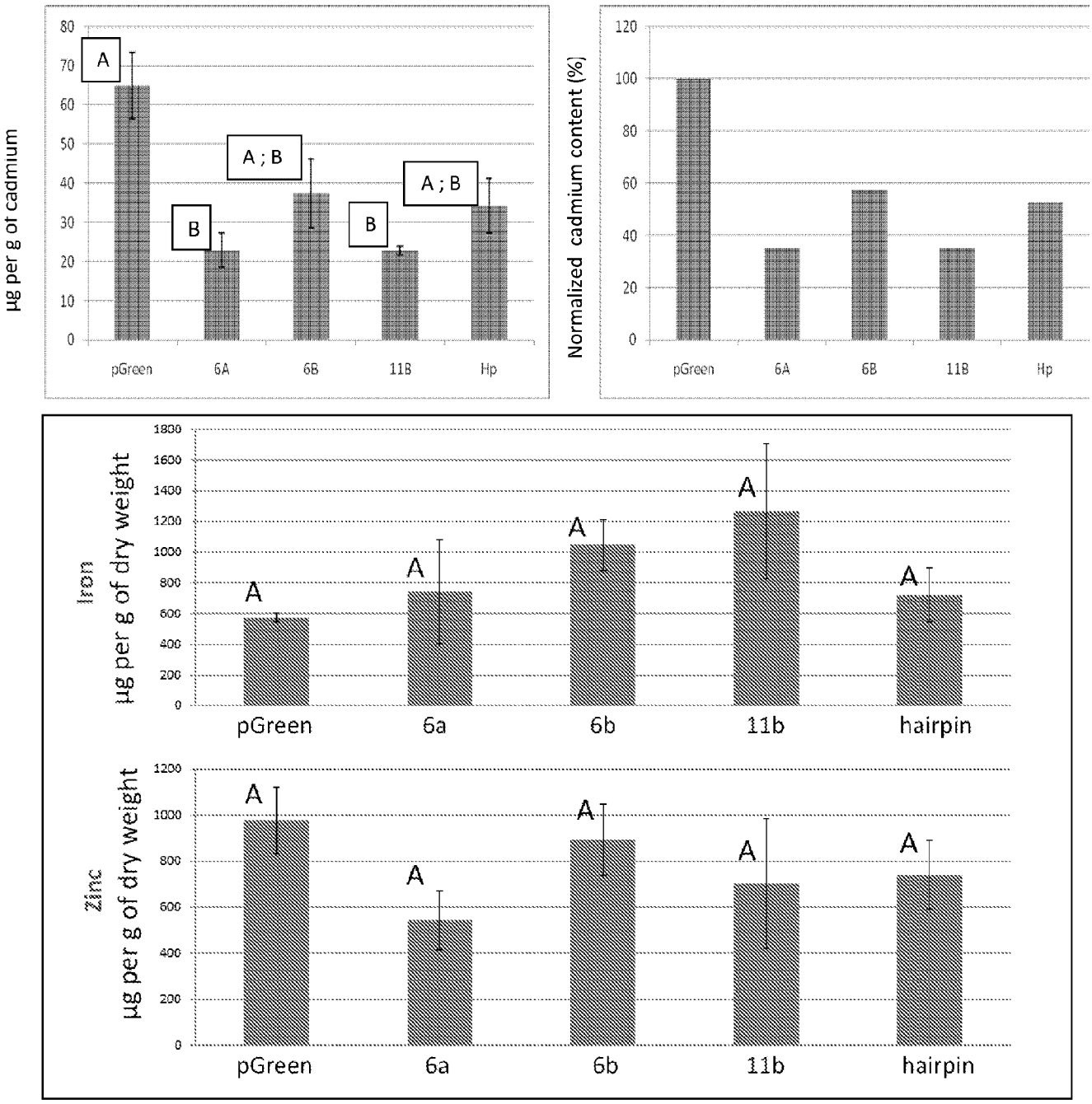

µg per g of cadmium
A
A ; B
B
B
A ; B
pGreen
6A
6B
11B
Hp
Normalized cadmium content (%)
Iron
µg per g of dry weight
Zinc
µg per g of dry weight
pGreen
6a
6b
11b
hairpin

Figure 11

Target 1	Collection	Identification	Amino acid change		Target 2	Collection	Identification	Amino acid change	Cross name
HMA-AS	BB16NN	E2-425	W78*	x	HMA-AT	BY02	390	P68S	F1CD2-425-390
HMA-AS	BB16NN	E2-425	W78*	x	HMA-AT	BY02	769	A40T	F1CD2-425-769
HMA-AS	BB16NN	E2-425	W78*	x	HMA-DT	BB16NN	E3-1370	G333R	F1CD2-425-1370
HMA-AS	BB16NN	E2-425	W78*	x	HMA-AT	BY02	1336	L107F	F1CD2-425-1336
HMA-AS	BB16NN	E2-425	W78*	x	HMA-DT	BY02	835	G328E	F1CD2-425-835
HMA-AS	BB16NN	E2-425	W78*	x	HMA-DT	BB16NN	E3-259	D339N	F1CD2-425-259
HMA-AS	BB16NN	E2-425	W78*	x	HMA-DT	BB16NN	E3-277	R342*	F1CD2-425-277
HMA-AS	BB16NN	E2-425	W78*	x	HMA-DT	BY02	506	A360T	F1CD2-425-506
HMA-AS	BB16NN	E2-425	W78*	x	HMA-AT	BY02	769	A40T	F1CD2-425-769b
HMA-AS	BY02	450	C99Y	x	HMA-DT	BB16NN	E3-277	R342*	F1CD2-450-277
HMA-B	BB16NN	E3-163	C980T	x	HMA-DT	BB16NN	E3-277	R342*	F1CD2-163-277
HMA-B	BY02	399	A186T	x	HMA-DT	BB16NN	E3-277	R342*	F1CD2-399-277
HMA-B	BY02	936	A196T	x	HMA-DT	BB16NN	E3-277	R342*	F1CD2-936-277
HMA-B	BY02	1468	splicing	x	HMA-DT	BB16NN	E3-277	R342*	F1CD2-1468-277
HMA-B	BB16NN	E1-276	E200K	x	HMA-AT	BY02	390	P68S	F1CD2-276-390
HMA-B	BB16NN	E1-276	E200K	x	HMA-AT	BY02	769	A40T	F1CD2-276-769
HMA-B	BB16NN	E1-276	E200K	x	HMA-DT	BB16NN	E3-1370	G333R	F1CD2-276-1370
HMA-B	BB16NN	E1-276	E200K	x	HMA-AT	BY02	1336	L107F	F1CD2-276-1336
HMA-B	BB16NN	E1-276	E200K	x	HMA-DT	BY02	835	G328E	F1CD2-276-835
HMA-B	BB16NN	E1-276	E200K	x	HMA-DT	BB16NN	E3-259	D339N	F1CD2-276-259
HMA-DS	BY02	416	L362F	x	HMA-AT	BY02	390	P68S	F1CD2-416-390
HMA-DS	BY02	416	L362F	x	HMA-AT	BY02	769	A40T	F1CD2-416-769
HMA-DS	BY02	416	L362F	x	HMA-DT	BB16NN	E3-1370	G333R	F1CD2-416-1370
HMA-DS	BY02	416	L362F	x	HMA-AT	BY02	1336	L107F	F1CD2-416-1336
HMA-DS	BY02	416	L362F	x	HMA-DT	BY02	835	G328E	F1CD2-416-835
HMA-DS	BY02	416	L362F	x	HMA-DT	BB16NN	E3-259	D339N	F1CD2-416-259
HMA-AS	BY02	793	E109K	x	HMA-AT	BY02	390	P68S	F1CD2-793-390
HMA-AS	BY02	793	E109K	x	HMA-AT	BY02	769	A40T	F1CD2-793-769
HMA-AS	BY02	793	E109K	x	HMA-DT	BB16NN	E3-1370	G333R	F1CD2-793-1370

Figure 12

Target 1	Collection	Identification	Amino acid change		Target 2	Collection	Identification	Amino acid change	Cross name
HMA-AS	BY02	793	E109K	x	HMA-AT	BY02	1336	L107F	F1CD2-793-1336
HMA-AS	BY02	793	E109K	x	HMA-DT	BY02	835	G328E	F1CD2-793-835
HMA-AS	BY02	793	E109K	x	HMA-DT	BB16NN	E3-259	D339N	F1CD2-793-259
HMA-AS	BY02	450	C99Y	x	HMA-DT	BY02	506	A360T	F1CD2-450-506
HMA-AS	BY02	450	C99Y	x	HMA-AT	BY02	390	P68S	F1CD2-450-390
HMA-AS	BY02	450	C99Y	x	HMA-AT	BY02	769	A40T	F1CD2-450-769
HMAB	BB16NN	E3-163	S137L	x	HMA-DT	BY02	506	A360T	F1CD2-163-506
HMAB	BB16NN	E3-163	S137L	x	HMA-AT	BY02	390	P68S	F1CD2-163-390
HMAB	BB16NN	E3-163	S137L	x	HMA-AT	BY02	769	A40T	F1CD2-163-769
HMA-B	BY02	399	A186T	x	HMA-DT	BY02	506	A360T	F1CD2-399-506
HMA-B	BY02	399	A186T	x	HMA-AT	BY02	390	P68S	F1CD2-399-390
HMA-B	BY02	399	A186T	x	HMA-AT	BY02	769	A40T	F1CD2-399-769
HMA-B	BY02	936	A196T	x	HMA-DT	BY02	506	A360T	F1CD2-936-506
HMA-B	BY02	936	A196T	x	HMA-AT	BY02	390	P68S	F1CD2-936-390
HMA-B	BY02	936	A196T	x	HMA-AT	BY02	769	A40T	F1CD2-936-769
HMA-B	BY02	1445	C167Y	x	HMA-DT	BY02	506	A360T	F1CD2-1445-506
HMA-B	BY02	1445	C167Y	x	HMA-AT	BY02	390	P68S	F1CD2-1445-390
HMA-B	BY02	1445	C167Y	x	HMA-AT	BY02	769	A40T	F1CD2-1445-769
HMA-B	BY02	1468	splicing	x	HMA-DT	BY02	506	A360T	F1CD2-1468-506
HMA-B	BY02	1468	splicing	x	HMA-AT	BY02	390	P68S	F1CD2-1468-390
HMA-B	BY02	1468	splicing	x	HMA-AT	BY02	769	A40T	F1CD2-1468-769

Figure 12 cont'd.

Collection	Identification	Mutated gene	Nucleotide change	Amino acid change	SIFT score
BB16NN	90	Alpha	C576T	L107F	0
BB16NN	425	Alpha	G406A	W78*	0
BY02	1468	Alpha	G964A	splicing	N.D.
BB16NN	277	Beta	C2217T	R342*	0
BB16NN	1275	Beta	C2206T	S338F	0

Figure 13

Name	Targeted gene	Targeted exon	Targeted sequences	Set of primer used	
6a	Alpha	6	TGACTCTTATTATTTCCGCAT	I miR-s	gaTGACTCTTATTATTTCCGCATtctctctttttgtattcc
				II miR-a	gaATGCGGAAATAATAAGAGTCAtcaaagagaatcaatga
				III miR*s	gaATACGGAAATAATTAGAGTCTtcacaggtcgtgatatg
				IV miR*a	gaAGACTCTAATTATTTCCGTATtctacatatatattcct
6b	Beta	6	TGGATCTTATCATTTCCGCAT	I miR-s	gaTGGATCTTATCATTTCCGCATtctctctttttgtattcc
				II miR-a	gaATGCGGAAATGATAAGATCCAtcaaagagaatcaatga
				III miR*s	gaATACGGAAATGATTAGATCCTtcacaggtcgtgatatg
				IV miR*a	gaAGGATCTAATCATTTCCGTATtctacatatatattcct
9	Alpha and Beta	9	TTCATTAGTATAACATGGCCT	I miR-s	gaTTCATTAGTATAACATGGCCTtctctctttttgtattcc
				II miR-a	gaAGGCCATGTTATACTAATGAAtcaaagagaatcaatga
				III miR*s	gaAGACCATGTTATAGTAATGATtcacaggtcgtgatatg
				IV miR*a	gaATCATTACTATAACATGGTCTtctacatatatattcct
11a	Alpha	11	TGATTTTAGCTAGAGTCTCAA	I miR-s	gaTGATTTTAGCTAGAGTCTCAAtctctctttttgtattcc
				II miR-a	gaTTGAGACTCTAGCTAAAATCAtcaaagagaatcaatga
				III miR*s	gaTTAAGACTCTAGCAAAAATCTtcacaggtcgtgatatg
				IV miR*a	gaAGATTTTTGCTAGAGTCTTAAtctacatatatattcct
11b	Beta	11	TGATTTTAGCAAGAGTCTCAA	I miR-s	gaTGATTTTAGCAAGAGTCTCAAtctctctttttgtattcc
				II miR-a	gaTTGAGACTCTTGCTAAAATCAtcaaagagaatcaatga
				III miR*s	gaTTAAGACTCTTGCAAAAATCTtcacaggtcgtgatatg
				IV miR*a	gaAGATTTTTGCAAGAGTCTTAAtctacatatatattcct

Figure 14

A.

Primer pair	Name	Sequence
1	HpEx11FwXhoI	CCCGCTCGAGtgagagcaagtcaggtcatccga
	HpEx11RevKpnI	CGGGGTACCctgtggtacatccagctcttga
2	HpEx11FwXbaI	TGCTCTAGAtgagagcaagtcaggtcatccga
	HpEx11RevClaI	CCCATCGATctgtggtacatccagctcttga

B.

tgagagcaagtcaggtcatccgatggcagccgctctggtggactatgcacaatcaaattccgttgagccaaagcctgatagagttgagcagtttcaaaattttcctggtgaagggatatttggaagaattgatggaatggaaatctatgtcgggaataggaaaatttcttcaagagctggatgtaccacagg

Figure 15

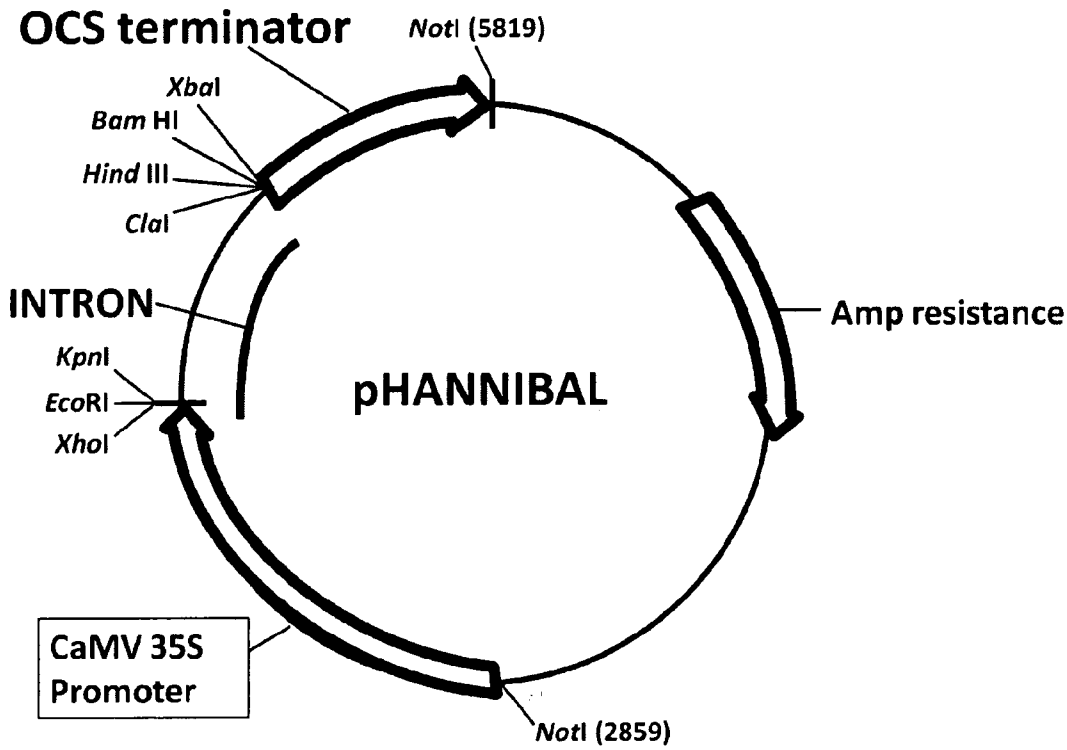

OCS terminator
NotI (5819)
XbaI
Bam HI
Hind III
ClaI
INTRON
Amp resistance
KpnI
EcoRI
XhoI
pHANNIBAL
CaMV 35S
Promoter
NotI (2859)

Figure 16

Genotype	Number of plants	Mean for Cd	Mean for Zn	Mean for Fe	Standard error for Cd	Standard error for Zn	Standard error for Fe
276M	12	37,74	329,99	410,37	9,02	168,84	336,27
276W	10	68,13	492,35	470,04	13,23	198,72	254,38
416M	8	31,07	340,55	256,70	12,01	178,18	149,82
425M	8	31,73	468,43	629,07	7,59	46,77	117,82
425W	10	60,81	374,87	356,70	15,37	185,12	250,46
90M	10	34,87	461,76	534,14	6,52	165,63	265,21
90W	9	47,68	407,53	455,80	9,36	198,59	304,47
amiRNAi	4	20,45	536,55	1045,81	3,48	148,28	530,32
BB16NN	9	55,75	522,82	575,28	9,70	153,98	422,14
BY02	5	45,29	364,86	373,15	5,05	109,60	162,35

Figure 17A

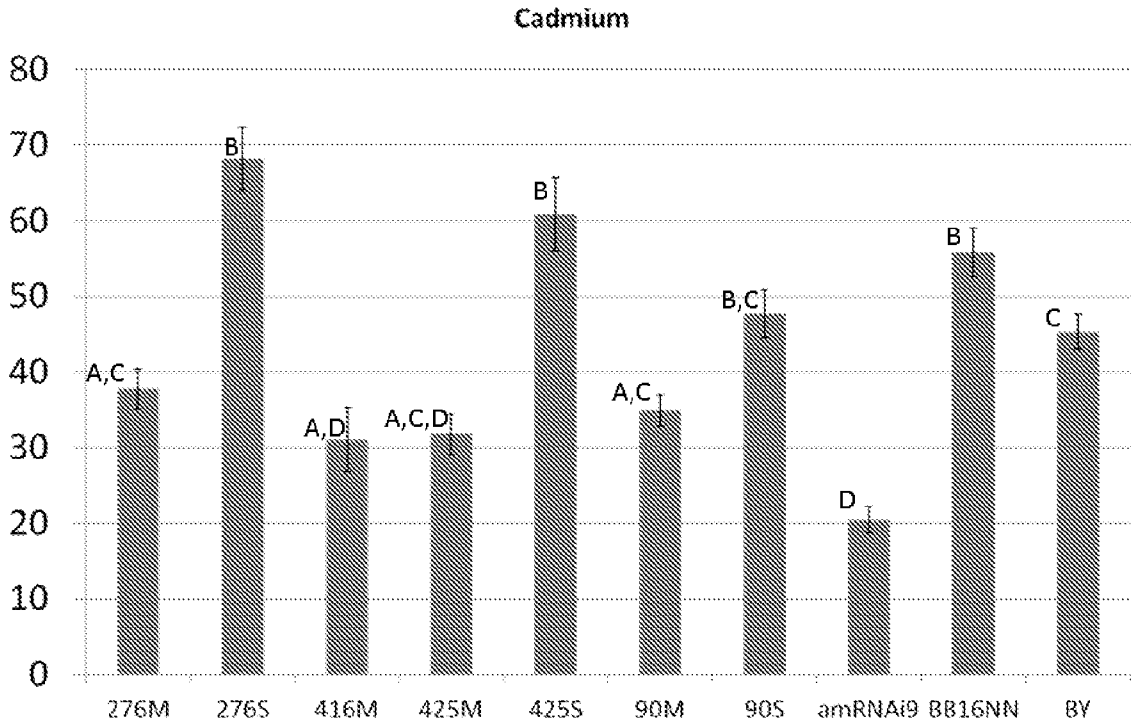


Figure 17B

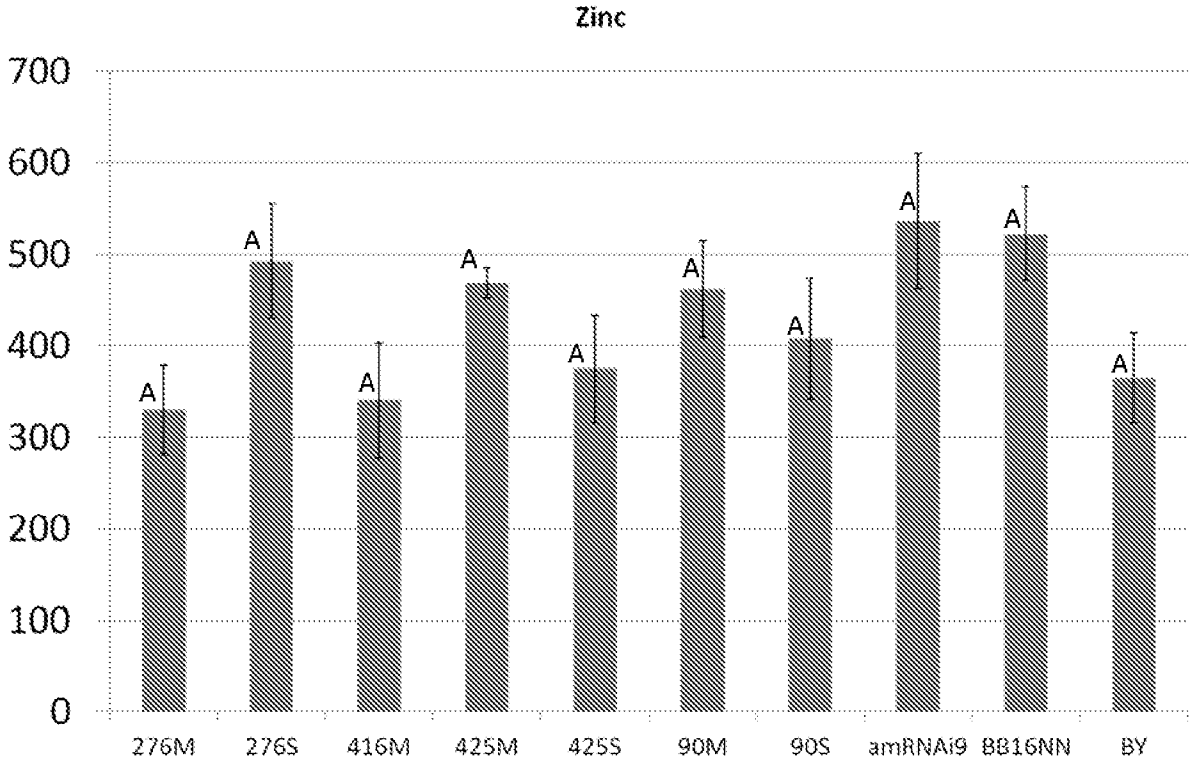


Figure 18A

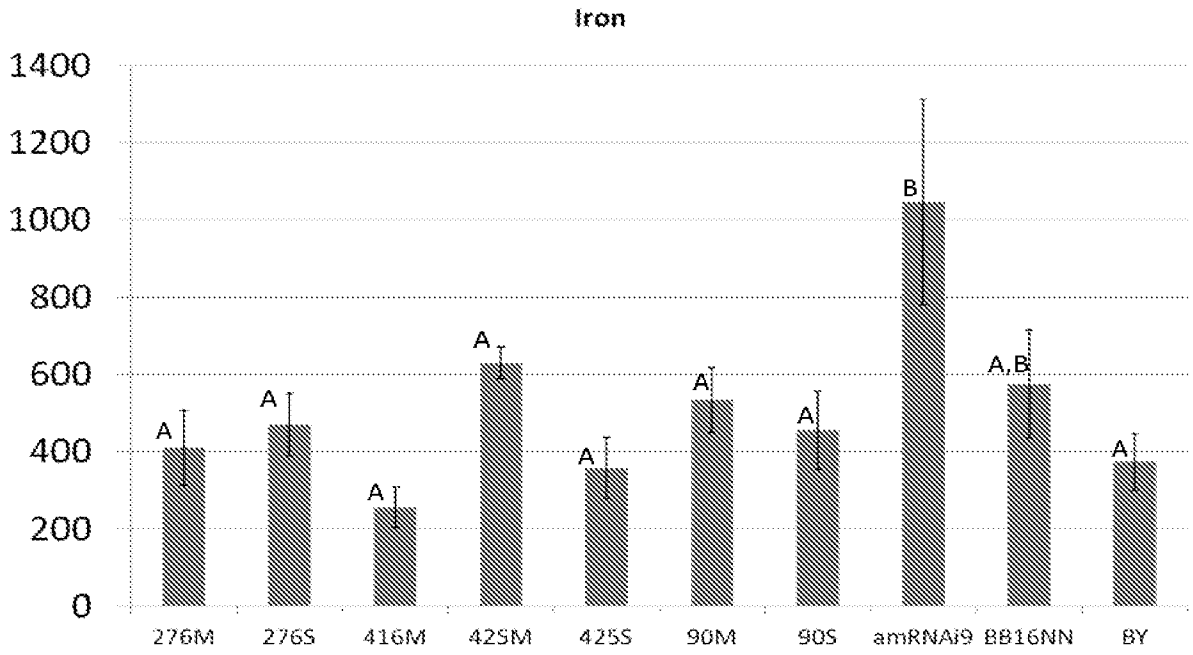


Figure 18B

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2011/066882

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12N15/82 C12N9/14
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C07K C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, BIOSIS, Sequence Search, EMBASE, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2009/074325 A1 (PHILIP MORRIS PROD [CH]) 18 June 2009 (2009-06-18) the whole document	1-18
A	WONG CHONG KUM EDWIN ET AL: "HMA P-type ATPases are the major mechanism for root-to-shoot Cd translocation in Arabidopsis thaliana", NEW PHYTOLOGIST, vol. 181, no. 1, 2009, pages 71-78, XP002624128, ISSN: 0028-646X	1-18

-/--

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
16 November 2011	01/12/2011

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Kania, Thomas

2

Form PCT/ISA/210 (second sheet) (April 2005)

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	VERRET F ET AL: "Overexpression of AtHMA4 enhances root-to-shoot translocation of zinc and cadmium and plant metal tolerance", FEBS LETTERS, vol. 576, no. 3, 22 October 2004 (2004-10-22), pages 306-312, XP004605705, ELSEVIER, AMSTERDAM, NL ISSN: 0014-5793, DOI: 10.1016/J.FEBSLET.2004.09.023	1-18
A	MILLS R F ET AL: "The plant P1B-type ATPase AtHMA4 transports Zn and Cd and plays a role in detoxification of transition metals supplied at elevated levels", FEBS LETTERS, vol. 579, no. 3, 31 January 2005 (2005-01-31), pages 783-791, XP004725196, ELSEVIER, AMSTERDAM, NL ISSN: 0014-5793, DOI: 10.1016/J.FEBSLET.2004.12.040	1-18
A	HANIKENNE MARC ET AL: "Evolution of metal hyperaccumulation required cis-regulatory changes and triplication of HMA4", NATURE (LONDON), vol. 453, no. 7193, May 2008 (2008-05), pages 391-399-METHODS, XP002624129, ISSN: 0028-0836	1-18
A	HUSSAIN DAWAR ET AL: "P-type ATPase heavy metal transporters with roles in essential zinc homeostasis in Arabidopsis", PLANT CELL, vol. 16, no. 5, 1 May 2004 (2004-05-01), pages 1327-1339, XP002495497, AMERICAN SOCIETY OF PLANT PHYSIOLOGISTS, ROCKVILLE, MD, US ISSN: 1040-4651, DOI: 10.1105/TPC.020487	1-18
A	WO 2005/090583 A1 (COMMISSARIAT ENERGIE ATOMIQUE [FR]; RICHAUD PIERRE [FR]; VERRET FREDER) 29 September 2005 (2005-09-29)	1-18
A	MILLS REBECCA F ET AL: "Functional expression of AtHMA4, a P1B-type ATPase of the Zn/Co/Cd/Pb subclass.", PLANT JOURNAL, vol. 35, no. 2, July 2003 (2003-07), pages 164-176, XP002624130, ISSN: 0960-7412	1-18

-/--

2

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2011/066882

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2009/074843 A1 (COMMISSARIAT ENERGIE ATOMIQUE [FR]; RICHAUD PIERRE [FR]; GRAVOT ANTOIN) 18 June 2009 (2009-06-18)	1-18

2

INTERNATIONAL SEARCH REPORT

International application No.
PCT/EP2011/066882

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application and necessary to the claimed invention, the international search was carried out on the basis of:

   a. (means)

   [X] on paper

   [X] in electronic form

   b. (time)

   [X] in the international application as filed

   [X] together with the international application in electronic form

   [ ] subsequently to this Authority for the purpose of search

2. [ ] In addition, in the case that more than one version or copy of a sequence listing and/or table relating thereto has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2011/066882

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2009074325 A1	18-06-2009	CN 101952442 A	19-01-2011
		CO 6280552 A2	20-05-2011
		EP 2231861 A1	29-09-2010
		US 2009183280 A1	16-07-2009
		WO 2009074325 A1	18-06-2009
WO 2005090583 A1	29-09-2005	AT 491796 T	15-01-2011
		CA 2560220 A1	29-09-2005
		DK 1725668 T3	04-04-2011
		EP 1725668 A1	29-11-2006
		JP 2007529217 A	25-10-2007
		US 2009155911 A1	18-06-2009
		WO 2005090583 A1	29-09-2005
		WO 2005093078 A1	06-10-2005
WO 2009074843 A1	18-06-2009	CN 101896608 A	24-11-2010
		EP 2231862 A1	29-09-2010
		US 2011023177 A1	27-01-2011
		WO 2009074843 A1	18-06-2009